The *BABAR* Collaboration

B. Aubert, A. Bazan, A. Boucham, D. Boutigny, I. De Bonis, J. Favier, J.-M. Gaillard, A. Jeremie, Y. Karyotakis, T. Le Flour, J. P. Lees, S. Lieunard, P. Petitpas, P. Robbe, V. Tisserand, K. Zachariadou

Laboratoire de Physique des Particules, F-74941 Annecy-le-Vieux, CEDEX, France

A. Palano

Università di Bari, Dipartimento di Fisica and INFN, I-70126 Bari, Italy

G. P. Chen, J. C. Chen, N. D. Qi, G. Rong, P. Wang, Y. S. Zhu

Institute of High Energy Physics, Beijing 100039, China

G. Eigen, P. L. Reinertsen, B. Stugu

University of Bergen, N-5007 Bergen, Norway

B. Abbott, G. S. Abrams, L. Amerman, A. W. Borgland, A. B. Breon, D. N. Brown, J. Button-Shafer, A. R. Clark, S. Dardin, C. Day, S. F. Dow, Q. Fan, I. Gaponenko, M. S. Gill, F. R. Goozen, S. J. Gowdy, Y. Groysman, C. Hernikl, R. G. Jacobsen, R. C. Jared, R. W. Kadel, J. Kadyk, A. Karcher, L. T. Kerth, I. Kipnis, S. Kluth, J. F. Kral, R. Lafever, C. LeClerc, M. E. Levi, S. A. Lewis, C. Lionberger, T. Liu, M. Long, L. Luo, G. Lynch, P. Luft, E. Mandelli, M. Marino, K. Marks, C. Matuk, A. B. Meyer, R. Minor, A. Mokhtarani, M. Momayezi, M. Nyman, P. J. Oddone, J. Ohnemus, D. Oshatz, S. Patton, M. Pedrali-Noy, A. Perazzo, C. Peters, W. Pope, M. Pripstein, D. R. Quarrie, J. E. Rasson, N. A. Roe, A. Romosan, M. T. Ronan, V. G. Shelkov, R. Stone, P. D. Strother,[1] A. V. Telnov, H. von der Lippe, T. F. Weber, W. A. Wenzel, G. Zizka

Lawrence Berkeley National Laboratory and University of California, Berkeley, CA 94720, USA

P. G. Bright-Thomas, C. M. Hawkes, A. Kirk, D. J. Knowles, S. W. O'Neale, A. T. Watson, N. K. Watson

University of Birmingham, Birmingham, B15 2TT, UK

T. Deppermann, H. Koch, J. Krug, M. Kunze, B. Lewandowski, K. Peters, H. Schmuecker, M. Steinke

Ruhr Universität Bochum, Inst. f. Experimentalphysik 1, D-44780 Bochum, Germany

J. C. Andress, N. R. Barlow, W. Bhimji, N. Chevalier, P. J. Clark, W. N. Cottingham, N. De Groot, N. Dyce, B. Foster, A. Mass, J. D. McFall, D. Wallom, F. F. Wilson

University of Bristol, Bristol BS8 1TH, UK

K. Abe, C. Hearty, J. A. McKenna, D. Thiessen

University of British Columbia, Vancouver, BC, Canada V6T 1Z1

B. Camanzi, T. J. Harrison,[2] A. K. McKemey, J. Tinslay

Brunel University, Uxbridge, Middlesex UB8 3PH, UK

E. I. Antohin, V. E. Blinov, A. D. Bukin, D. A. Bukin, A. R. Buzykaev, M. S. Dubrovin, V. B. Golubev, V. N. Ivanchenko, G. M. Kolachev, A. A. Korol, E. A. Kravchenko, S. F. Mikhailov, A. P. Onuchin, A. A. Salnikov, S. I. Serednyakov, Yu. I. Skovpen, V. I. Telnov, A. N. Yushkov

Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia

J. Booth, A. J. Lankford, M. Mandelkern, S. Pier, D. P. Stoker

University of California at Irvine, Irvine, CA 92697, USA

[1]Now at Queen Mary, University of London, London, E1 4NS, UK
[2]Now at University of Birmingham, Birmingham, B15 2TT, UK

A. Ahsan, K. Arisaka, C. Buchanan, S. Chun

University of California at Los Angeles, Los Angeles, CA 90024, USA

J. G. Branson, R. Faccini,[3] D. B. MacFarlane, S. A. Prell, Sh. Rahatlou, G. Raven, V. Sharma

University of California at San Diego, La Jolla, CA 92093, USA

S. Burke, D. Callahan, C. Campagnari, B. Dahmes, D. Hale, P. A. Hart, N. Kuznetsova, S. Kyre, S. L. Levy, O. Long, A. Lu, J. May, J. D. Richman, W. Verkerke, M. Witherell, S. Yellin

University of California at Santa Barbara, Santa Barbara, CA 93106, USA

J. Beringer, J. DeWitt, D. E. Dorfan, A. M. Eisner, A. Frey, A. A. Grillo, M. Grothe, C. A. Heusch, R. P. Johnson, W. Kroeger, W. S. Lockman, T. Pulliam, W. Rowe, H. Sadrozinski, T. Schalk, R. E. Schmitz, B. A. Schumm, A. Seiden, E. N. Spencer, M. Turri, W. Walkowiak, M. Wilder, D. C. Williams

University of California at Santa Cruz, Santa Cruz, CA 95064, USA

E. Chen, G. P. Dubois-Felsmann, A. Dvoretskii, J. E. Hanson, D. G. Hitlin, Yu. G. Kolomensky,[4] S. Metzler, J. Oyang, F. C. Porter, A. Ryd, A. Samuel, M. Weaver, S. Yang, R. Y. Zhu

California Institute of Technology, Pasadena, CA 91125, USA

S. Devmal, T. L. Geld, S. Jayatilleke, S. M. Jayatilleke, G. Mancinelli, B. T. Meadows, M. D. Sokoloff

University of Cincinnati, Cincinnati, OH 45221, USA

P. Bloom, B. Broomer, E. Erdos, S. Fahey, W. T. Ford, F. Gaede, W. C. van Hoek, D. R. Johnson, A. K. Michael, U. Nauenberg, A. Olivas, H. Park, P. Rankin, J. Roy, S. Sen, J. G. Smith, D. L. Wagner

University of Colorado, Boulder, CO 80309, USA

J. Blouw, J. L. Harton, M. Krishnamurthy, A. Soffer, W. H. Toki, D. W. Warner, R. J. Wilson, J. Zhang

Colorado State University, Fort Collins, CO 80523, USA

T. Brandt, J. Brose, G. Dahlinger, M. Dickopp, R. S. Dubitzky, P. Eckstein, H. Futterschneider, M. L. Kocian, R. Krause, R. Müller-Pfefferkorn, K. R. Schubert, R. Schwierz, B. Spaan, L. Wilden

Technische Universität Dresden, D-01062 Dresden, Germany

L. Behr, D. Bernard, G. R. Bonneaud, F. Brochard, J. Cohen-Tanugi, S. Ferrag, G. Fouque, F. Gastaldi, P. Matricon, P. Mora de Freitas, C. Renard, E. Roussot, S. T'Jampens, C. Thiebaux, G. Vasileiadis, M. Verderi

Ecole Polytechnique, F-91128 Palaiseau, France

A. Anjomshoaa, R. Bernet, F. Di Lodovico, F. Muheim, S. Playfer, J. E. Swain

University of Edinburgh, Edinburgh EH9 3JZ, UK

M. Falbo

Elon College, Elon College, NC 27244

C. Bozzi, S. Dittongo, M. Folegani, L. Piemontese, A. .C. Ramusino

Università di Ferrara, Dipartimento di Fisica and INFN, I-44100 Ferrara, Italy

[3]Jointly appointed with Università di Roma La Sapienza, Dipartimento di Fisica and INFN, I-00185 Roma, Italy
[4]Now at LBNL and University of California, Berkeley, CA 94720, USA

E. Treadwell

Florida A&M University, Tallahassee, FL 32307, USA

F. Anulli,[5] R. Baldini-Ferroli, A. Calcaterra, R. de Sangro, D. Falciai, G. Finocchiaro, P. Patteri, I. M. Peruzzi,[5] M. Piccolo, Y. Xie, A. Zallo

Laboratori Nazionali di Frascati dell'INFN, I-00044 Frascati, Italy

S. Bagnasco, A. Buzzo, R. Contri, G. Crosetti, P. Fabbricatore, S. Farinon, M. Lo Vetere, M. Macri, S. Minutoli, M. R. Monge, R. Musenich, M. Pallavicini, R. Parodi, S. Passaggio, F. C. Pastore, C. Patrignani, M. G. Pia, C. Priano, E. Robutti, A. Santroni

Università di Genova, Dipartimento di Fisica and INFN, I-16146 Genova, Italy

R. Bartoldus, T. Dignan, R. Hamilton, U. Mallik

University of Iowa, Iowa City, IA 52242, USA

J. Cochran, H. B. Crawley, P.-A. Fischer, J. Lamsa, R. McKay, W. T. Meyer, E. I. Rosenberg

Iowa State University, Ames, IA 50011-3160, USA

J. N. Albert, C. Beigbeder, M. Benkebil, D. Breton, R. Cizeron, S. Du, G. Grosdidier, C. Hast, A. Höcker, V. LePeltier, A. M. Lutz, S. Plaszczynski, M. H. Schune, S. Trincaz-Duvoid, K. Truong, A. Valassi, G. Wormser

Laboratoire de l'Accélérateur Linéaire, F-91898 Orsay Cedex, France

O. Alford, D. Behne, R. M. Bionta, J. Bowman, V. Brigljević, A. Brooks, V. A. Dacosta, O. Fackler, D. Fujino, M. Harper, D. J. Lange, M. Mugge, T. G. O'Connor, H. Olson, L. Ott, E. Parker, B. Pedrotti, M. Roeben, X. Shi, K. van Bibber, T. J. Wenaus, D. M. Wright, C. R. Wuest, B. Yamamoto

Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

M. Carroll, P. Cooke, J. R. Fry, E. Gabathuler, R. Gamet, M. George, M. Kay, S. McMahon,[6] A. Muir, D. J. Payne, R. J. Sloane, P. Sutcliffe, C. Touramanis, G. Zioulas

University of Liverpool, Liverpool L69 3BX, UK

M. L. Aspinwall, D. A. Bowerman, P. D. Dauncey, I. Eschrich, N. J. W. Gunawardane, R. Martin, J. A. Nash, D. R. Price, P. Sanders, D. Smith

University of London, Imperial College, London, SW7 2BW, UK

D. E. Azzopardi, J. J. Back, P. Dixon, P. F. Harrison, D. Newman-Coburn,[7] R. J. L. Potter, H. W. Shorthouse, M. I. Williams, P. B. Vidal

Queen Mary, University of London, London, E1 4NS, UK

G. Cowan, S. George, M. G. Green, A. Kurup, C. E. Marker, P. McGrath, T. R. McMahon, F. Salvatore, I. Scott, G. Vaitsas

University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, UK

D. Brown, C. L. Davis, Y. Li, J. Pavlovich

University of Louisville, Louisville, KY 40292, USA

[5]Jointly appointed with Univ. di Perugia, I-06100 Perugia, Italy
[6]Now at University of California at Irvine, Irvine, CA 92697, USA
[7]Deceased

J. Allison, R. J. Barlow, J. T. Boyd, J. Fullwood, F. Jackson, A. Khan,[8] G. D. Lafferty, N. Savvas, E. T. Simopoulos, R. J. Thompson, J. H. Weatherall

University of Manchester, Manchester M13 9PL, UK

R. Bard, C. Dallapiccola,[9] A. Farbin, A. Jawahery, V. Lillard, J. Olsen, D. A. Roberts, J. R. Schieck

University of Maryland, College Park, MD 20742, USA

G. Blaylock, K. T. Flood, S. S. Hertzbach, R. Kofler, C. S. Lin, S. Willocq, J. Wittlin

University of Massachusetts, Amherst, MA 01003, USA

B. Brau, R. Cowan, F. Taylor, R. K. Yamamoto

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

D. I. Britton, R. Fernholz,[10] M. Houde, M. Milek, P. M. Patel, J. Trischuk

McGill University, Montreal, QC, Canada H3A 2T8

F. Lanni, F. Palombo

Università di Milano, Dipartimento di Fisica and INFN, I-20133 Milano, Italy

J. M. Bauer, M. Booke, L. Cremaldi, R. Kroeger, M. Reep, J. Reidy, D. A. Sanders, D. J. Summers

University of Mississippi, University, MS 38677, USA

J. F. Arguin, M. Beaulieu, J. P. Martin, J. Y. Nief, R. Seitz, P. Taras, A. Woch, V. Zacek

Université de Montreal, Montreal, QC, Canada, H3C 3J7

H. Nicholson, C. S. Sutton

Mount Holyoke College, South Hadley, MA 01075, USA

N. Cavallo, G. De Nardo, F. Fabozzi, C. Gatto, L. Lista, D. Piccolo, C. Sciacca

Università di Napoli Federico II, Dipartimento di Fisica and INFN, I-80126 Napoli, Italy

N. M. Cason, J. M. LoSecco

University of Notre Dame, Notre Dame, IN 46556, USA

J. R. G. Alsmiller, T. A. Gabriel, T. Handler, J. Heck

Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

M. Iwasaki, N. B. Sinev,

University of Oregon, Eugene, OR 97403, USA

R. Caracciolo, F. Colecchia, F. Dal Corso, F. Galeazzi, M. Marzolla, G. Michelon, M. Morandin, M. Posocco, M. Rotondo, S. Santi, F. Simonetto, R. Stroili, E. Torassa, C. Voci

Università di Padova, Dipartimento di Fisica and INFN, I-35131 Padova, Italy

[8] Now at University of Edinburgh, Edinburgh EH9 3JZ, UK
[9] Now at University of Massachusetts, Amherst, MA 01003, USA
[10] Now at Princeton University, Princeton, NJ 08544, USA

P. Bailly, M. Benayoun, H. Briand, J. Chauveau, P. David, C. De la Vaissière, L. Del Buono, J.-F. Genat, O. Hamon, Ph. Leruste, F. Le Diberder, H. Lebbolo, J. Lory, L. Martin, F. Martinez-Vidal,[11] L. Roos, J. Stark, S. Versillé, B. Zhang

Université Paris VI et VII, F-75252 Paris, Cedex 05, France

P. F. Manfredi, L. Ratti, V. Re, V. Speziali

Università di Pavia, Dipartimento di Fisica and INFN, I-27100 Pavia, Italy

E. D. Frank, L. Gladney, Q. H. Guo, J. H. Panetta

University of Pennsylvania, Philadelphia, PA 19104, USA

C. Angelini, G. Batignani, S. Bettarini, M. Bondioli, F. Bosi, M. Carpinelli, F. Forti, D. Gagliardi, M. A. Giorgi, A. Lusiani, P. Mammini, M. Morganti, F. Morsani, N. Neri, A. Profeti, E. Paoloni, M. Rama, G. Rizzo, F. Sandrelli, G. Simi, G. Triggiani

Università di Pisa, Scuola Normale Superiore, and INFN, I-56010 Pisa, Italy

M. Haire, D. Judd, K. Paick, L. Turnbull, D. E. Wagoner

Prairie View A&M University, Prairie View, TX 77446, USA

J. Albert, C. Bula, M. H. Kelsey, C. Lu, K. T. McDonald, V. Miftakov, B. Sands, S. F. Schaffner, A. J. S. Smith, A. Tumanov, E. W. Varnes

Princeton University, Princeton, NJ 08544, USA

F. Bronzini, A. Buccheri, C. Bulfon, G. Cavoto, D. del Re, F. Ferrarotto, F. Ferroni, K. Fratini, E. Lamanna, E. Leonardi, M. A. Mazzoni, S. Morganti, G. Piredda, F. Safai Tehrani, M. Serra, C. Voena

Università di Roma La Sapienza, Dipartimento di Fisica and INFN, I-00185 Roma, Italy

R. Waldi

Universität Rostock, D-18051 Rostock, Germany

P. F. Jacques, M. Kalelkar, R. J. Plano

Rutgers University, New Brunswick, NJ 08903, USA

T. Adye, B. Claxton, J. Dowdell, U. Egede, B. Franek, S. Galagedera, N. I. Geddes, G. P. Gopal, J. Kay,[12] J. Lidbury, S. Madani, S. Metcalfe,[12,13] G. Markey,[12] P. Olley, M. Watt, S. M. Xella

Rutherford Appleton Laboratory, Didcot, Oxon., OX11 0QX, UK

R. Aleksan, P. Besson,[7] P. Bourgeois, P. Convert, G. De Domenico, A. de Lesquen, S. Emery, A. Gaidot, S. F. Ganzhur, Z. Georgette, L. Gosset, P. Graffin, G. Hamel de Monchenault, S. Hervé, M. Karolak, W. Kozanecki, M. Langer, G. W. London, V. Marques, B. Mayer, P. Micout, J. P. Mols, J. P. Mouly, Y. Penichot, J. Rolquin, B. Serfass, J. C. Toussaint, M. Usseglio, G. Vasseur, C. Yeche, M. Zito

DAPNIA, Commissariat à l'Energie Atomique/Saclay, F-91191 Gif-sur-Yvette, France

N. Copty, M. V. Purohit, F. X. Yumiceva

University of South Carolina, Columbia, SC 29208, USA

[11] Now at Università di Pisa, I-56010 Pisa, Italy
[12] At CLRC Daresbury Laboratory, Daresbury, Warrington, Cheshire, WA4 4AD, UK
[13] Now at Stanford Linear Accelerator Center, Stanford, CA 94309, USA

I. Adam, A. Adesanya, P. L. Anthony, D. Aston, J. Bartelt, J. Becla, R. Bell, E. Bloom, C. T. Boeheim, A. M. Boyarski, R. F. Boyce, D. Briggs, F. Bulos, W. Burgess, B. Byers, G. Calderini, R. Chestnut, R. Claus, M. R. Convery, R. Coombes, L. Cottrell, D. P. Coupal, D. H. Coward, W. W. Craddock, S. DeBarger, H. DeStaebler, J. Dorfan, M. Doser, W. Dunwoodie, J. E. Dusatko, S. Ecklund, T. H. Fieguth, D. R. Freytag, T. Glanzman, G. L. Godfrey, G. Haller, A. Hanushevsky, J. Harris, A. Hasan, C. Hee, T. Himel, M. E. Huffer, T. Hung, W. R. Innes, C. P. Jessop, H. Kawahara, L. Keller, M. E. King, L. Klaisner, H. J. Krebs, U. Langenegger, W. Langeveld, D. W. G. S. Leith, S. K. Louie, S. Luitz, V. Luth, H. L. Lynch, J. McDonald, G. Manzin, H. Marsiske, T. Mattison,[14] M. McCulloch, M. McDougald, D. McShurley, S. Menke, R. Messner, S. Metcalfe, M. Morii,[15] R. Mount, D. R. Muller, D. Nelson, M. Norby, C. P. O'Grady, L. Olavson, J. Olsen, F. G. O'Neill, G. Oxoby, P. Paolucci,[16] T. Pavel, J. Perl, M. Pertsova, S. Petrak, G. Putallaz, P. E. Raines, B. N. Ratcliff, R. Reif, S. H. Robertson, L. S. Rochester, A. Roodman, J. J. Russel, L. Sapozhnikov, O. H. Saxton, T. Schietinger, R. H. Schindler, J. Schwiening, G.Sciolla,[17] J. T. Seeman, V. V. Serbo, S. Shapiro, K. Skarpass Sr., A. Snyder, E. Soderstrom, A. Soha, S. M. Spanier, A. Stahl, P. Stiles, D. Su, M. K. Sullivan, M. Talby, H. A. Tanaka, J. Va'vra, S. R. Wagner, R. Wang, T. Weber, A. J. R. Weinstein, J. L. White, U. Wienands, W. J. Wisniewski, C. C. Young, N. Yu

Stanford Linear Accelerator Center, Stanford, CA 94309, USA

P. R. Burchat, C. H. Cheng, D. Kirkby, T. I. Meyer, C. Roat

Stanford University, Stanford, CA 94305-4060, USA

R. Henderson, N. Khan

TRIUMF, Vancouver, BC, Canada V6T 2A3

S. Berridge, W. Bugg, H. Cohn, E. Hart, A. W. Weidemann

University of Tennessee, Knoxville, TN 37996, USA

T. Benninger, J. M. Izen, I. Kitayama, X. C. Lou, M. Turcotte

University of Texas at Dallas, Richardson, TX 75083, USA

F. Bianchi, M. Bona, F. Daudo, B. Di Girolamo, D. Gamba, P. Grosso, A. Smol, P. .P. Trapani, D. Zanin

Università di Torino, Dipartimento di Fisica and INFN, I-10125 Torino, Italy

L. Bosisio, G. Della Ricca, L. Lanceri, A. Pompili, P. Poropat, M. Prest, E. Vallazza, G. Vuagnin

Università di Trieste, Dipartimento di Fisica and INFN, I-34127 Trieste, Italy

R. S. Panvini

Vanderbilt University, Nashville, TN 37235, USA

A. De Silva,[18] R. Kowalewski, J. M. Roney

University of Victoria, Victoria, BC, Canada V8W 3P6

H. R. Band, E. Charles, S. Dasu, P. Elmer, J. R. Johnson, J. Nielsen, W. Orejudos, Y. Pan, R. Prepost, I. J. Scott, J. Walsh,[11] S. L. Wu, Z. Yu, H. Zobernig

University of Wisconsin, Madison, WI 53706, USA

[14]Now at University of British Columbia, Vancouver, BC, Canada V6T 1Z1
[15]Now at Harvard University, Cambridge, MA 02138
[16]Now at Università di Napoli Federico II, I-80126 Napoli, Italy
[17]Now at Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[18]Now at TRIUMF, Vancouver, BC, Canada V6T 2A3

T. B. Moore, H. Neal
Yale University, New Haven, CT 06511, USA

The *BABAR* Detector

BABAR is the detector for the SLAC PEP-II asymmetric e^+e^- B Factory operating at the $\Upsilon(4S)$ resonance. *BABAR* was designed to perform comprehensive studies of CP-violation in B-meson decays. The *BABAR* detector has many subsystems configured around a beryllium beam pipe. Charged particle tracks are measured in a multi-layer silicon vertex tracker surrounded by a cylindrical wire drift chamber. Electromagnetic showers from electrons and photons are detected in an array of CsI crystals located just inside the solenoidal coil of a superconducting magnet. Muons and neutral hadrons are identified by arrays of resistive plate chambers inserted into gaps in the iron flux return of the magnet. Charged hadrons are identified by dE/dx measurements in the tracking detectors and in a quartz ring-imaging Cherenkov detector surrounding the drift chamber. The trigger, data acquisition and data-monitoring systems are VME- and network-based, and controlled and operated by custom-designed on-line software. Details of the design and layout of the detector components and their associated electronics and software are presented together with performance data.

1. Introduction

The primary physics goal of the *BABAR* experiment is the systematic study of *CP*-violating asymmetries in the decay of neutral B mesons to *CP* eigenstates. Such measurements are designed to test the Standard Model prediction for *CP*-violation based on the Cabibbo-Kobayashi-Maskawa mechanism [1]; but many other decay modes of the B mesons, charged and neutral, permit searches for this and other sources of *CP*-violation. The secondary goal is to perform precision measurements of decays of bottom and charm mesons and τ leptons, and to search for rare processes that become accessible with the high luminosity of the PEP-II B Factory [2]. The design of the detector is optimized for *CP* studies, but it is also well suited for these other physics studies. The scientific goals of the *BABAR* experiment were first presented in the Letter of Intent [3]; detailed physics studies have been documented in the *BABAR* Physics Book [4] and earlier workshops [5].

The PEP-II B Factory is an asymmetric e^+e^- collider designed to operate at a luminosity of 3×10^{33} cm^{-2}s^{-1}and beyond at a center-of-mass energy of 10.58 GeV, the mass of the $\Upsilon(4S)$ resonance. This resonance decays exclusively to $B^0\overline{B}^0$ and B^+B^- pairs and thus provides an ideal laboratory for the study of B mesons. In PEP-II, the electron beam of 8.98 GeV collides head-on with the positron beam of 3.12 GeV resulting in a boost to the $\Upsilon(4S)$ resonance of $\beta\gamma = 0.55$. This boost makes it possible to reconstruct the decay vertices of the two B mesons and to determine their relative decay times. One can therefore measure the time dependence of decay rates. The crucial test of *CP* invariance is a comparison of these rates for B^0 and $\overline{B}^0$ to a *CP* eigenstate. Experimentally, this requires events in which one B meson decays to a *CP* eigenstate and is fully reconstructed and the other B meson is tagged as B^0 or a $\overline{B}^0$ by its decay products: a charged lepton, a charged kaon, or a D^* decay.

The very small branching ratios of B mesons to *CP* modes, typically 10^{-4}, the need for full reconstruction of final states with two or more charged particles and several π^0, plus the tagging of the second neutral B, place stringent requirements on the detector:

- a large and uniform acceptance, in particular down to small polar angles relative to the boost direction, to avoid particle losses;
- excellent detection efficiency for charged particles down to 60 MeV/c and for photons to 20 MeV;
- very good momentum resolution to separate small signals from background;
- excellent energy and angular resolution for the detection of photons from π^0 and radia-

tive B decays in the range from 20 MeV to 4 GeV;

- very good vertex resolution, both transverse and parallel to the beam;
- identification of electrons and muons, primarily for the detection of semi-leptonic decays used to tag the B flavor, and also for the study of leptonic and rare decays;
- identification of hadrons over a wide range of momenta for B flavor-tagging, for the separation of pions from kaons in decay modes such as $B^0 \rightarrow K^{\pm}\pi^{\mp}$ or $B^0 \rightarrow \pi^+\pi^-$, as well as in charm meson and τ decays;
- a highly redundant, selective trigger system so as to avoid significant signal losses and systematic uncertainties;
- low-noise electronics and a data-acquisition system of high flexibility and operational stability;
- high degree of reliability of components and frequent monitoring and automated calibrations, plus control of the environmental conditions to ensure continuous and stable operation;
- an online computing and network system that can control, process, and store the expected high volume of data; and
- detector components that can tolerate significant doses of radiation and operate under high-background conditions.

To reach the desired sensitivity for the most interesting measurements, data sets of order 10^8 B mesons will be needed. For the peak cross section at the $\Upsilon(4S)$ of 1.1 nb, this will require an integrated luminosity of order 100 fb^{-1} or three years of reliable and highly efficient operation having a detector of state-of-the art capabilities at a storage ring operating at design luminosity or above with few interruptions.

In the following, a brief overview of the principal components of the detector, the trigger, the data-acquisition, and the online computing and control system is given. This is to be followed by brief descriptions of the PEP-II interaction region, the beam characteristics, and the impact of the beam generated background on the design and operation of the detector components. The following chapters contain detailed presentations of the design, construction, and performance for all of the systems that make up the *BABAR* detector.

REFERENCES

1. N. Cabibbo, Phys. Rev. Lett. 10 (1963) 531; M. Kobayashi, T. Maskawa, Prog. Th. Phys. 49 (1973) 652.
2. PEP-II An Asymmetric B Factory, Conceptual Design Report, SLAC-418, LBL-5379 (1993).
3. The *BABAR* Collaboration, Letter of Intent for the Study of *CP* Violation and Heavy Flavor Physics at PEP-II, SLAC-443 (1994).
4. The *BABAR* Physics Book, Physics at an Asymmetric B Factory, P.F. Harrison, H.R. Quinn, Editors, SLAC-R-504, (1998).
5. The Physics Program of a High-Luminosity Asymmetric B Factory at SLAC, SLAC-353 (1989); Proceedings of the Workshop on Physics and Detector Issues for a High Luminosity Asymmetric B Factory at SLAC, SLAC-373 (1991); Proceedings of B Factorys, The State of the Art in Accelerators, Detectors and Physics, SLAC-400 (1992).

2. Detector Overview

The *BABAR* detector was designed and built by a large, international team of scientists and engineers. Details of its original design are documented in the Technical Design Report [1], issued in 1995.

Figure 1 shows a longitudinal section through the detector center, and Figure 2 shows an end-view with the principal dimensions. The detector is configured around the PEP-II interaction region. To maximize the acceptance, while taking into acount the kinematic boost chosen by the asymmetric beam energies, the whole detector is offset by 0.37 m in the direction of the lower energy beam. The inner detector consists of a silicon vertex tracker, a drift chamber, a ring-imaging Cherenkov detector, and a CsI calorimeter. These detector components are surrounded by a superconducting solenoid that is designed for a field of 1.5 T. The iron flux return is instrumented for muon and neutral hadron detection. All these components operate with good performance down to forward angles of 350 mrad and backward angles of 400 mr. For general reference, forward and backward are defined relative to the higher energy beam. As indicated in the two drawings, the right handed coordinate system is anchored on the main tracking system, the drift chamber, with the z-axis coinciding with its principal axis. This axis is offset relative to the beam axis by about 20 mrad. The y-axis points upward and the x-axis points away from the center of the PEP-II storage ring.

The detector is of relatively compact design, its transverse dimension being constrained by the elevation of the beam above the floor, 3.50 m. The solenoid radius was chosen by balancing the physics requirements and capabilities of the drift chamber and calorimeter against the overall cost. As in many similar systems, the calorimeter was the most expensive single system and thus considerable effort was made to minimize its total volume without undue reduction in performance of either the tracking system or the calorimeter. The geometric layout of the vertex detector is constrained by the space available between the components of the PEP-II magnet system, a pair of dipole magnets (B1) followed by a pair of quadrupole magnets (Q1). These magnets are made of SmCo; they extend to 21 cm on either side of the interaction point and the dipoles limit the forward and backward acceptance of the tracking system. The SVT and these two sets of magnets are placed inside a long tube (4.5 m long and 0.217 m inner diameter) that supports the weight of these magnets. The central section of the support tube is fabricated of carbon-fibers composite (CFC) with a thickness to 1.08 % of a radiation length.

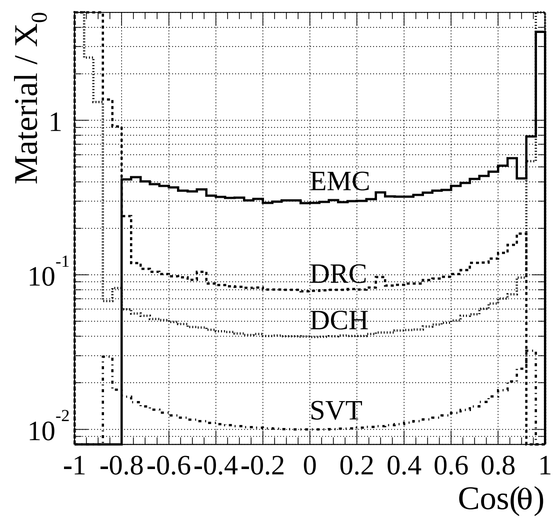

Figure 3. Amount of material surrounding the interaction point.

Since the average momentum of charged particles produced in B meson decay is less than 1 GeV/c, the errors on the measured track parameters are dominated by multiple Coulomb scattering, rather than the intrinsic spatial resolution of the detectors. Similarly, the detection efficiency and energy resolution of low energy photons are severely impacted by material in front of the calorimeter. Thus, special care has been given to keep the material in the active volume of the detector to a minimum. Figure 3 shows the distribution of material in the various detector systems in units of radiation lengths. Specifi-

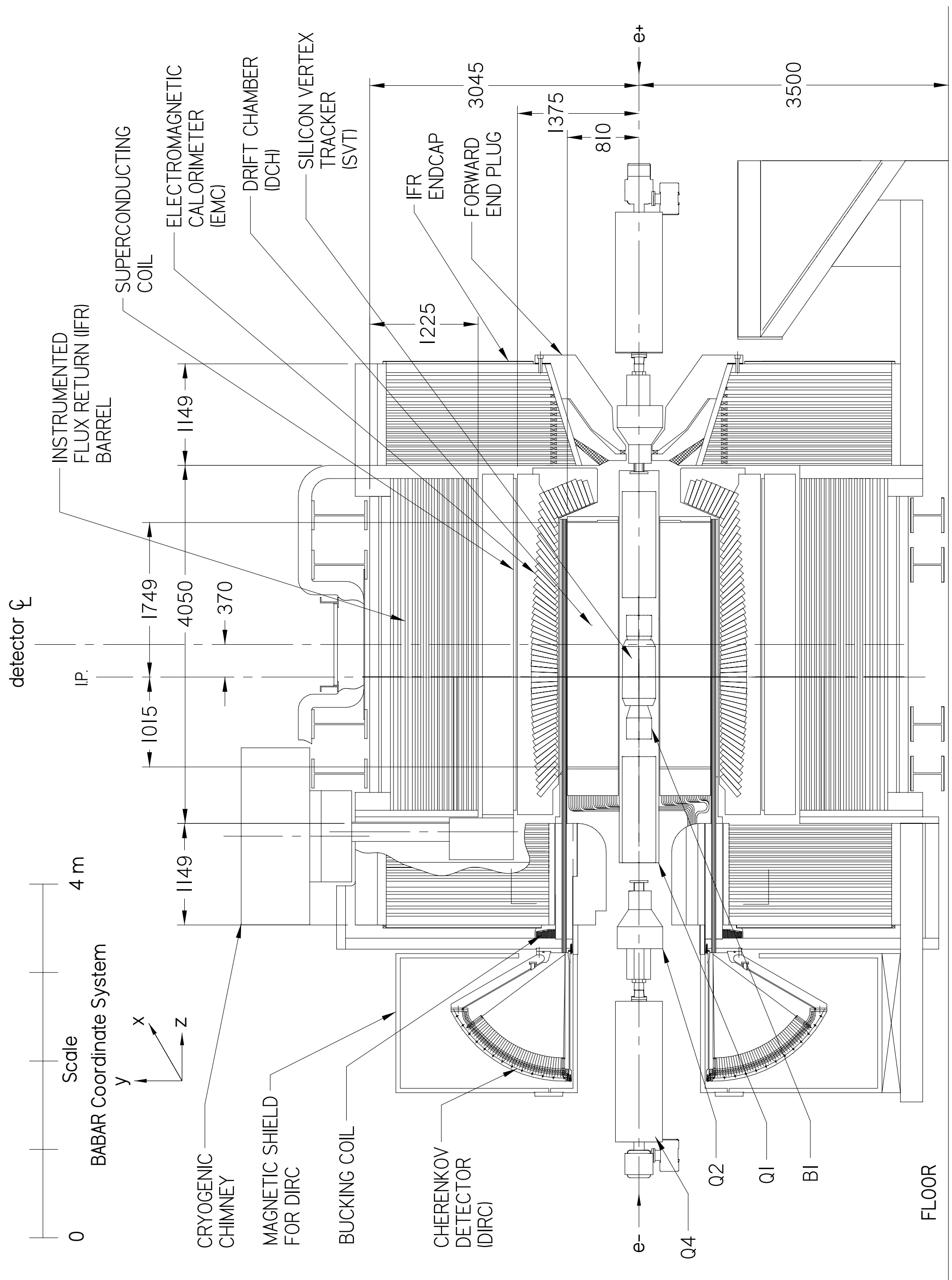

Figure 1. BABAR detector elevation view.

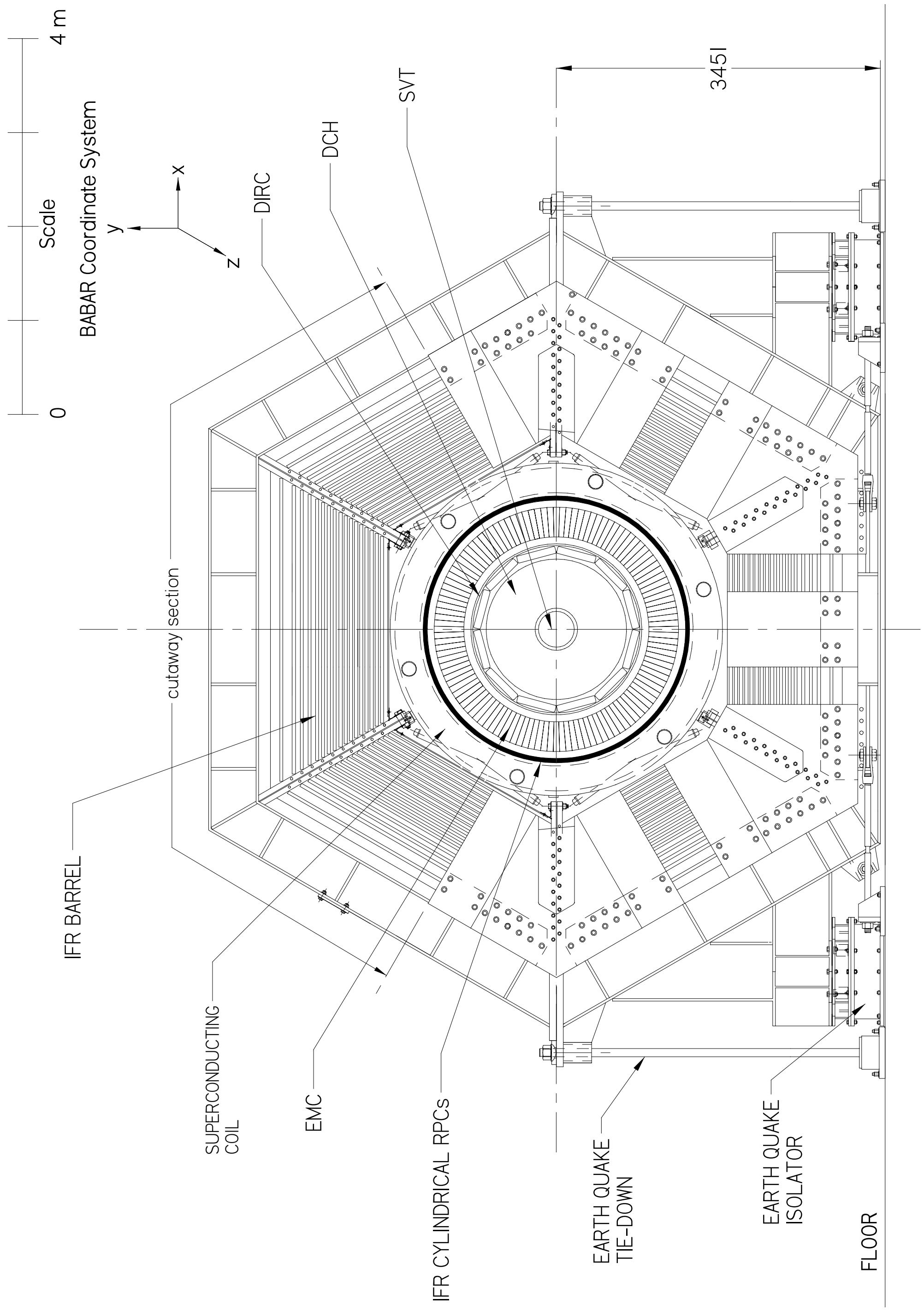

Figure 2. BABAR detector end view.

cally, each curve indicates the material a particle traverses before it reaches the first active element of a specific detector system.

2.1. Detector Components

The charged particle tracking system is made of two components, the Silicon Vertex Tracker (SVT) and the larger cylindrical Drift Chamber (DCH).

The SVT is composed of five layers of double-sided silicon strip detectors that are assembled from modules with read-out at each end, thus avoiding inactive material in the acceptance volume. The inner three layers primarily provide position and angle information for the measurement of the vertex position. They are mounted as close to the water-cooled beryllium beam pipe as practical, thus minimizing the impact of scattering in the material of the beam pipe on the extrapolation to the vertex. The outer two layers layers are placed at much larger radii, thus providing the coordinate and angle measurement needed for linking SVT tracks to DCH tracks.

The principal purpose of the DCH is the momentum and angle measurement for charged particles. It also supplies information for the charged particle trigger and a measurement of $\mathrm{d}E/\mathrm{d}x$ for particle identification. The drift chamber is of compact design, with 40 layers of small, approximately hexagonal cells. Longitudinal information is derived from wires placed at small angles to the principal axis. By chosing low mass wires, and a helium based gas mixture the multiple scattering inside the drift chamber is kept to a minimum. All of the read-out electronics is mounted on the backward endplate of the chamber, minimizing the amount of material in front of the calorimeter endcap.

The DIRC, the Detector of Internally Reflected Cherenkov light, is a novel concept and technology employed primarily for the separation of pions and kaons from about 500 MeV/c to the kinematic limit of 4.5 GeV/c. Cherenkov light is produced in 4.9 m long bars of synthetic fused silica of rectangular cross section, $1.7\,\mathrm{cm} \times 3.5\,\mathrm{cm}$, and transmitted by total internal reflection, while preserving the angle of emission, to an array of photomultipliers. This array forms the backward wall of a donut shaped water tank that is located outside of the backward end of the magnet. Images of the Cherenkov rings are reconstructed from the location and time of arrival of the signals in the photomultipliers.

The Electro-Magnetic Calorimeter (EMC) is designed to detect electromagnetic showers with excellent energy and angular resolution over the full energy range from 20 MeV to 4 GeV. This coverage will allow the detection of low energy π^0 from B decays and higher energy photons and electrons from electromagnetic, weak, or radiative processes. The EMC is a finely segmented array of projective geometry made of thallium doped cesium iodide (CsI(Tl)) crystals. The crystals are arranged in modules that are supported individually from an external support structure. This structure is build from two units, a barrel and a forward endcap. To obtain the desired resolution the amount of material in front and in-between the crystals is held to a minimum. The individual crystals are read-out by pairs of silicon PIN diodes. Low noise analog circuits and frequent, precise calibration of the electronics and energy response over the full dynamic range are crucial for the desired performance.

The Instrumented Flux Return (IFR) is designed to identify muons and to detect neutral hadrons. For this purpose, the magnet flux return steel in the barrel and the two endcaps is segmented into layers, increasing in thickness from 2 cm on the inside to 10 cm on the outside. Between these steel absorbers, single gap Resistive Plate Chambers (RPC) are inserted which detect streamers from ionizing particles via external capacitive read-out strips. There are 19 layers of RPCs in the barrel sectors and 18 layers in the endcaps. Two additional cylindrical layers of RPCs with four read-out planes are placed just in front of the magnet cryostat to detect particles exiting the EMC.

2.2. Electronics, Trigger, Data Acquisition and On-line Computing

The electronics, trigger, data acquisition and on-line computing represent a collection of tightly coupled hardware and software systems. These systems were designed to maximize the physics data acceptance, maintainability and reliability while minimizing complexity, deadtime, downtime and cost.

- *Front-End Electronics* (FEE) assemblies are physically located on the detector and consist of signal processing and digitization electronics along with a standard digital transport mechanism into the data acquisition system.

- A robust and flexible two-level trigger handles the full accelerator collision rate, one in hardware, *Level 1* (L1), the other in software, *Level 3* (L3). A provision is made for an intermediate trigger (Level 2) should unexpected conditions require its implementation.

- A data reduction, transport and event building system, *Online Dataflow* (ODF), handles digitized data from the FEE through the event building. This system also includes hardware elements providing fast control and timing signals.

- A farm of commercial Unix processors and associated software, *Online Event Processing* (OEP), provides the real-time environment within which complete events are processed by the L3 trigger algorithms, partial event reconstruction is performed for monitoring, and event data is logged to an intermediate storage.

- In a second farm of processors, *Online Prompt Reconstruction* (OPR), completely reconstructs on all physics events, performs monitoring and constants generation in near real-time. Physics event data are transferred to an object database and are thus available for further analyses.

- An *Online Run Control* (ORC) system handles the logic for managing the state of the detector systems, starting and stopping runs, and performing calibrations. A user interface provides operator access to detector operation.

- A system to control and monitor the detector and its support systems, *Online Detector Control* (ODC), is based upon the Experimental Physics Industrial Control System (EPICS) toolbox. This system includes communication links with the PEP-II accelerator complex. User interfaces provide operator access and displays.

2.2.1. Electronics

A common electronics architecture is shared by all *BABAR* detector subsystems; common vendors were chosen for many of its components such as VME crates and power supplies. Event data from the detector flows through the FEE circuitry, while monitoring and control signals are handled by a separate, parallel system. All FEE systems, listed in Table 1, are mounted directly on the detector for performance reasons. This solution also minimizes the cable plant and avoids noise pickup and grounding loops generated in long analog signal cables. Electronics for all detector subsystems utilize standard *BABAR* interfaces to the global electronics and software.

Each FEE chain consists of an amplifier, digitizer, a trigger latency buffer for storing data during the L1 trigger processing, and a derandomizing event buffer for storing data between the L1 accept and and subsequent transfer to the data acquisition system (Figure 4). Custom ICs have been developed to perform the signal processing (Table 1). The digital L1 latency buffers function as fixed length data pipelines and are managed by common protocol signals generated by the FCTS. All de-randomizing event buffers function as FIFOs capable of storing a fixed number of events. During normal operation, analog signal processing, digitization, and data readout occur continuously and simultaneously.

Since many of the front-end circuits are totally inaccessible or access requires significant downtime, stringent requirements were placed on reliability. Most components underwent comprehensive MTBF (mean time between failure) studies. A burn-in procedure was followed for all circuits prior to installation with the goal of minimizing initial failure rates.

Table 1
Overview of the Front-End Electronics (PH=pulseheight, ASD=amplifier-shaper-disciminator)

System	Elements	PH (bits)	Time (ns)	No. Channels	Custom ICs
SVT	Si Strips	4	-	150,000	ASD/storage
		-	-		
DCH	Wires	-	-	7,104	ASD
		6	2		ADC/TDC
DIRC	PMTs	-	-	10,752	ASD
		-	0.5		TDC/FIFO
EMC	PIN diodes	-	-	13,160	amplifier/shaper
		17-18	-	6,580	ADC ranging chip
IFR	RPC Strips	1	-	51,584	

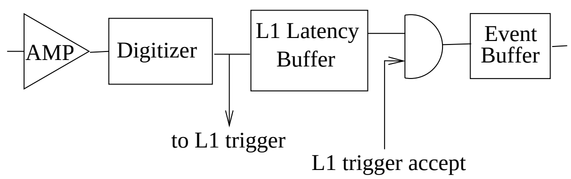

Figure 4. Front-End Electronics. Analog signals arrive from the left, proceed conditionally through the indicated steps and are injected into the remainder of the data acquisition system.

2.2.2. Trigger

The trigger system operates as a sequence of two independent stages, the second conditional upon the first. The Level 1 trigger is performed first at the machine crossing rate. Its goal is to sufficiently reduce that rate to a level acceptable for the Level 3 software trigger which runs on a farm of commercial processors.

The Level 1 trigger is optimized for simplicity and speed. It consists of a pipelined hardware processor. It is designed to provide an output trigger rate of $\lesssim 2$ kHz. The L1 trigger selection is based on a reduced data set from the DCH, EMC and IFR. The maximum L1 response latency for a given collision is 12 μs. Based on both the complete event and L1 trigger information, the Level 3 software algorithms select events of interest stores them for processing. The L3 output rate is administratively limited to 120 Hz so as to not overload the downstream storage and processing capacity. The trigger architecture is designed to accommodate a *Level 2* should unexpected background conditions demand higher performance in the future.

BABAR has no fast counters for triggering purposes, and bunch crossings are nearly continuous at 4 ns. Dedicated Level 1 trigger processors receive data which is continuously clocked in from the DCH, EMC and IFR detector subsystems. These processors produce clocked outputs to the fast control system at 30 MHz, the time granularity of resultant L1 Accept signals. The arrival of an L1 Accept signal by the data acquisition system causes a window of each subsystem's L1 Latency Buffer to be read out. This window size varies from about 500 ns for the silicon vertex tracker to only $4-16$ μs wide for the calorimeter. Absolute timing information for the event, i.e., associating an event with a particular beam crossing, which is determined offline, is calculated using timing distributions for drift chamber hits within a track segment, waveforms from the electromagnetic calorimeter and accelerator timing fiducials.

2.2.3. Data Acquisition and Online Computing

The data acquisition and computing systems are responsible for the transport of event data from the detector FEE circuits all the way to mass storage with a minimum of deadtime and is shown schematically in Figure 5. These systems interface with the trigger and enable specialized modes of operation, such as calibration and testing. Other parts of these systems provide for the control and monitoring of the detector and supporting facilities.

Hardware

The data acquisition system hardware consists of VME crates, specialized VME-based processors called Readout Modules (ROMs) [reference], the Fast Control and Timing System, a UNIX processor farm, various server machines and an Ethernet [3] network. A ROM consists of a commercial single board computer [4], event buffers, an interface to the FCTS, and a custom *Personality Card* which connects with the FEE circuits via 1.2 Gbps fiber optic cables. The ROM provides the standard interface between the detector specific front-end electronics, the FCTS, and the event builder. There are about 157 ROMs in the system which are located in 19 physical VME crates that represent 24 logical crates by virtue of segmented backplanes. The FCTS system consists of a VME crate plus individual Fast Control Distribution Modules in each of the data acquisition VME crates.

Detector monitoring and control is accomplished via a standard set of components, including a VME crate, a Motorola MVME177 single-board computer, and various other VME modules *[this needs to be expanded to a small list of VME modules]*. With the exception of the solenoid magnet and gas systems, which have their own specialized control and monitoring, all other *BABAR* elements plug into this system.

The online computing system relies on a complex of workstation consoles and servers with 0.8 TBytes of attached storage, all interconnected with switched 100 Mbps and 1 Gbps Ethernet networks. Multiple Gbps Ethernet links connect the experimental hall with the SLAC computing center.

Online Dataflow

The Online Dataflow software connects, controls, and manages the flow of data in the data acquisition hardware while incurring almost no dead time. This code is divided between embedded processors in the ROMs which run the VxWorks [5] real-time operating system and UNIX processors running the Solaris operating system. Dataflow provides: a) configuration and read-out of the FEE over fiber links to the ROMs; b) data transport, buffering and event building from the ROMs to the OEP farm; c) masking and prescaling of L1 triggers; and d) partitioning into multiple, independent data acquisition systems for parallel calibrations and diagnostics. Additional *feature extraction* (or FEX) code in the ROMs, supplied by the detector system groups, extracts physical signals from the raw data, performs gain and pedestal correction, data sparsification, and data formatting. Data from electronics calibrations are accumulated in the ROMs, the channel response functions evaluated, the results compared to reference data and subsequently applied in feature extraction. Calibration data is stored in a *Conditions Database*.

Online Event Processing

The On-line Event Processing receives and processes data from the *DataFlow* event builder on each of the 32 UNIX processors. In particular, OEP orchestrates the following tasks: a) L3 trigger algorithms; b) "fast monitoring" to assure data quality; and c) logging the selected events to disk. Logging involves a process to merge the multiple data output streams to a single file.

Online Prompt Reconstruction

Bridging the online and offline worlds is OPR [6]. This system reads raw data recorded to disk by OEP and, operating on farm of 150 Unix processors, selects physics events, performs complete reconstruction, performs a "rolling" calibration, collects extensive monitoring data for quality assurance and writes the result into an object database. A "rolling" calibration is the generation of reconstruction constants during normal event processing which are then applied to the processing of subsequent data.

Online Detector and Run Control

The ODC system controls and extensively monitors the electronics and environment of the

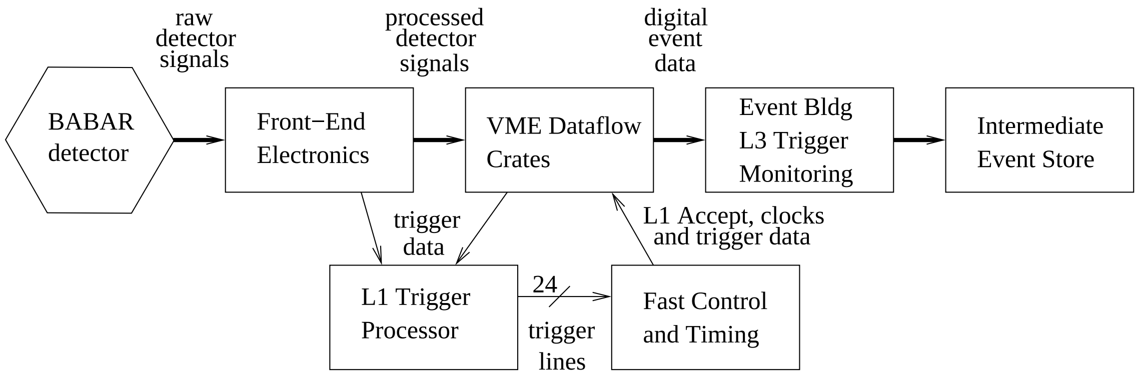

Figure 5. Data acquisition schematic.

detector and assures the safety of the detector. Its implementation is based on EPICS [2], providing detector-wide standardization for the control and monitoring, diagnostics and alarm handling. ODC also provides communication with PEP-II and the magnet control system. Monitoring data are archived in the *Ambient Database*.

The ORC system ties together the various components of the online system and provides the operator with a single graphical interface to control the detector operation. Complex configurations are stored in a *Configuration Database*; keys to the configuration used for any run is stored along with the data. The Event, Ambient, Conditions and Configuration Databases are implemented in an object database [7].

REFERENCES

1. *BABAR* Technical Design Report, SLAC-R-457 (1995)
2. EPICS
3. Any of the IEEE 802 family of standards for computer networking, running at rates of 10 Mbps, 100 Mbps or 1000 Mbps.
4. PowerPC
5. VxWorks Real-time operating system (Tornado??), Wind River Systems, Inc., 500 Wind River Way, Alameda, CA 94501 USA.
6. Francesco Safai Tehrani, "The Babar Prompt Reconstruction Manager: A Real Life Example of a Constructive Approach to Software Development.", submitted to Computer Physics Communications
7. Objectivity/DB, Objectivity, Inc., 301B East Evelyn Avenue, Mountain View, California, 94041, USA.

3. PEP II Storage Rings and Their Impact on the *BABAR* Detector

3.1. PEP-II Storage Rings

PEP-II is an e^+e^- storage ring system designed to operate at a center of mass (c.m.) energy of 10.58 GeV, corresponding to the mass of the $\Upsilon(4S)$ resonance. The parameters of these energy asymmetric storage rings are presented in Table 2. PEP-II has surpassed its design luminosity, both in terms of the instantaneous and the integrated daily luminosity, with significantly fewer bunches than anticipated. A full description of the design and operational experience of PEP-II has recently been published [1] [2].

Table 2
PEP-II beam parameters. Values are given both for design and typical colliding beam operation in the first year. HER and LER refer to the high energy e^- and low energy e^+ ring, respectively. σ_{Lx}, σ_{Ly}, and σ_{Lz} refer to the horizontal, vertical, and longitudinal rms size of the luminous region.

Parameters	Design	Typical
Energy HER/LER (GeV)	9.0/3.1	9.0/3.1
Current HER/LER (A)	0.75/2.15	0.7/1.3
# of bunches	1658	553-829
Bunch spacing (ns)	4.2	6.3-10.5
σ_{Lx} (μm)	110	120
σ_{Ly} (μm)	3.3	5.6
σ_{Lz} (mm)	9	9
Luminosity ($10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	3	2.5
Luminosity (pb^{-1}/d)	135	120

PEP-II typically operates in a series of 40 minute fills during which the colliding beams coast. At the end of each fill, it takes about three minutes to replenish the beams for the next fill. After a loss of the stored beams, it takes approximately 10-15 minutes to refill the two beams. *BABAR* divides the data into runs, defined as periods of three hour duration or less during which beam and detector conditions are judged to be stable. While most of the data are recorded at the peak of the $\Upsilon(4S)$ resonance, some 12% are taken at a c.m. energy 40 MeV below to allow for studies of the non-resonant background.

3.2. Impact of PEP-II on *BABAR* Layout

The large beam currents inherent to the high luminosity of PEP-II and the necessity to separate closely-spaced bunches as close as possible to the interaction point (IP), tightly couple the issues of detector design, interaction region layout, and remediation of machine-induced background. The bunches collide head-on and are separated magnetically in the horizontal plane by a pair of dipole magnets (B1), followed by a series of offset quadrupoles. The tapered B1 dipoles, located at ± 21 cm on either side of the IP, and the Q1 quadrupoles are made of samarium-cobalt and operate inside the field of the *BABAR* solenoid, while Q2, Q4, and Q5, located outside or in the fringe field of the solenoid, are standard iron magnets. The collision axis is off-set from the z-axis of the *BABAR* detector by about 20 mrad in the horizontal plane [3].

The interaction region is enclosed by a water-cooled beam pipe of 27.9 mmouter radius, composed of two thin layers of beryllium (0.83 mm and 0.53 mm) with a 1.48- mm water channel in-between. To attenuate synchrotron radiation, the inner surface of the pipe is covered with a thin layer of gold (approximately 4 μm). In addition, the beam pipe is wrapped with 150 μm of tantalum foil on either side of the IP, i.e., beyond $z = +10.1$ cm and $z = -7.9$ cm. The total thickness of the central beam pipe section at normal incidence corresponds to 1.06% of a radiation length.

The beam pipe, the permanent magnets and the SVT are assembled and aligned and then enclosed in a 4.5 m-long support tube. This rigid structure is inserted into the *BABAR* detector, spanning the IP. The central section of this tube is fabricated from carbon-fiber epoxy composite thus reducing its thickness to 1.08% of a radiation length.

3.3. Luminosity, Beam Energies and Position

The beam parameters that are most critical for the *BABAR* data analysis are the luminosity, the energies of the two beams and the position and size of the luminous region.

3.3.1. Luminosity

While PEP-II uses a high-rate luminosity monitor sampling radiative Bhabha scattering to provide a fast relative measurement of the luminosity for operations, *BABAR* derives the absolute luminosity off-line from QED processes. The best result is obtained from $\mu^+\mu^-$ pairs. For a sample of 1 fb^{-1}, the statistical error is 1.3%, compared to a systematic error of 0.5% on the relative, and 1.5% on the absolute value of the luminosity. This error is currently dominated by uncertainties in the Monte Carlo generator and the simulation of the detector. It is expected that with a better understanding of the efficiency, the overall systematic error on the absolute value of the luminosity can be reduced by a factor of two.

3.3.2. Beam Energies

During operation, the mean energies of the two beams are calculated from the total magnetic bending strength (including the effects of off-axis quadrupole fields, steering magnets, and wigglers) and the average deviations of the accelerating frequencies from their central values. While the systematic uncertainty in the PEP-II calculation of the absolute beam energies is estimated to be 5-10 MeV, the relative energy setting for each beam is accurate and stable to about 1 MeV. The energy spread of the LER and HER is 2.3 MeV and 5.5 MeV, respectively.

To ensure that data are recorded close to the peak of the $\Upsilon(4S)$ resonance, the observed ratio of $B\overline{B}$ enriched hadronic events to lepton pair production is monitored on-line. At the peak of the resonance, a 2.5% change in the $B\overline{B}$ production rate corresponds to a 2 MeV change in the c.m. energy, a value that is close to the tolerance to which the energy of PEP-II can be held. However, a drop in the $B\overline{B}$ rate does not distinguish between energy settings below or above the $\Upsilon(4S)$ peak. The sign of the energy change must be determined from other indicators. The best monitor and absolute calibration of the c.m. energy is derived from the measured c.m. momentum of fully reconstructed B mesons combined with the known B meson mass. An absolute error of 1.1 MeV can be obtained for an integrated luminosity of 1 fb^{-1}. This error is limited by both the knowledge of the B mass [4] and the detector resolution.

The beam energies are necessary input for the calculation of two kinematic variables that are commonly used to separate signal from background in the analysis of exclusive B meson decays. These variables make optimum use of the measured quantities and are largely uncorrelated. They are chosen to be invariant under Lorentz transformations and thus can be evaluated both in the laboratory and c.m. frame.

The first variable, ΔE, can be expressed in a Lorentz invariant form as

$$\Delta E = (2 q_B q_i - s)/2\sqrt{s}, \tag{1}$$

where $q_B = (E_B, \vec{p}_B)$ and $q_i = (E_i, \vec{p}_i)$ are the Lorentz vectors representing the momentum of the B candidate and of the initial state, $q_i = q_{e+} + q_{e-}$. In the c.m. frame, ΔE takes the familiar form

$$\Delta E = E_B^* - E_{beam}^*, \tag{2}$$

where E_B^* is the reconstructed energy of the B meson and E_{beam}^* equals half the c.m. energy. The spread in ΔE receives a sizable contribution from the beam energy spread, but it is generally dominated by detector resolution.

The second variable is the energy-substituted mass, $m_{\rm ES}$, defined as ${m_{\rm ES}}^2 = q_B^2$ with the constraint $\Delta E = 0$. In the laboratory frame, this quantity can be determined from the measured three-momentum $\vec{p}_B$ of the B candidate without explicit knowledge of the masses of the decay products:

$$m_{\rm ES} = \sqrt{(\frac{1}{2}s + \vec{p}_B \vec{p}_i)^2 / E_i^2 - p_B^2}. \tag{3}$$

In the c.m. frame ($\vec{p}_i = 0$) this variable takes the familiar form

$$m_{\rm ES} = \sqrt{E_{beam}^{*2} - p_B^{*2}}, \tag{4}$$

where p_B^* is the c.m. momentum of the B meson, derived from the momenta of its decay products. The B meson energy is substituted by E_{beam}^*. The spread of $m_{\rm ES}$ is dominated by the spread in E_{beam}^*, $\sigma_{E_{beam}^*} = 2.6$ MeV which can be measured from the $m_{\rm ES}$ distribution for samples of fully reconstructed B meson decays.

3.3.3. Beam Direction

The direction of the beams relative to *BABAR* is measured iteratively run-by-run using $e^+e^- \to e^+e^-$ and $e^+e^- \to \mu^+\mu^-$ events. The resultant uncertainty in the direction of the boost from the laboratory to the c.m. frame, $\vec{\beta}$, is about 1 mrad, dominated by alignment errors. This translates into an uncertainty of about 0.3 MeV in m_{ES}. $\vec{\beta}$ is consistent to within 1 mrad with the orientation of the elongated beam spot (see below). It is stable to better than 1 mrad from one run to the next.

3.3.4. Beam Size and Position

The size and position of the luminous region are critical parameters for the decay-time-dependent analyses and their values are monitored continuously on-line and off-line. The design values for the size of the luminous region are presented in Table 2. The vertical size is too small to be measured directly. It is inferred from the measured luminosity, the horizontal size, and the beam currents; it varies typically by 1-2 μm.

The transverse position, size and angles of the luminous region relative to the *BABAR* coordinate system are determined by analyzing the distribution of the distance of closest approach to the z-axis of the tracks in well measured two-track events as a function of the azimuth ϕ. The longitudinal parameters are derived from the longitudinal vertex distribution of the two tracks. A combined fit to 9 parameters (3 average coordinates, 3 widths, and 3 small angles) converges readily, even after significant changes in the beam position. The uncertainties in the average beam position are of the order of a few μm in the transverse plane and 100 μm along the collision axis. Run-by-run variations in the beam position are comparable to these measurement uncertainties, indicating that the beams are stable over the period of a typical run. The fit parameters are stored run by run in the conditions database. These measurements are also checked off-line by measuring the primary vertices in multi-hadron events. The measured horizontal and longitudinal beam sizes, corrected for tracking resolution, are consistent with those measured by PEP-II.

3.4. Beam Background Sources

The primary sources of steady-state accelerator backgrounds are, in order of increasing importance: synchrotron radiation in the vicinity of the interaction region; interactions between the beam particles and the residual gas in either ring; and electromagnetic showers generated by beam-beam collisions [5] [6] [7]. In addition, there are other background sources that fluctuate widely and can lead to very large instantaneous rates, thereby disrupting stable operation.

3.4.1. Synchrotron Radiation

Synchrotron radiation in nearby dipoles, the interaction-region quadrupole doublets and the B1 separation dipoles generates many kW of power and is potentially a severe background. The beam orbits, vacuum-pipe apertures and synchrotron-radiation masks have been designed such that most of these photons are channeled to a distant dump; the remainder are forced to undergo multiple scatters before they can enter the *BABAR* acceptance. The remaining synchrotron radiation background is dominated by x-rays (scattered from the tungsten tip of a mask) generated by dipole fields in HER low-β quadrupoles. This residual background is very low and has not presented a significant problem.

3.4.2. Beam-Gas Scattering

Beam-gas bremsstrahlung and Coulomb scattering off residual gas molecules cause beam particles to escape the acceptance of the ring if their energy loss or scattering angle are sufficiently large. The intrinsic rate of these processes is proportional to the product of the beam current and the residual pressure (which itself increases with current). Their relative importance, as well as the resulting spatial distribution and absolute rate of lost particles impinging the vacuum pipe in the vicinity of the detector, depend on the beam optical functions, the limiting apertures, and the residual-pressure profile around the entire rings. The separation dipoles bend the energy-degraded particles from the two beams in opposite directions and consequently most *BABAR* detector systems exhibit occupancy peaks in the horizontal plane, i.e., the LER background near $\phi = 0°$ and

HER background near $\phi = 180°$.

During the first few months of operation and during the first month after a local venting of the machine, the higher pressures lead to significantly enhanced background from beam-gas scattering. The situation has improved significantly with time due to scrubbing of the vacuum pipe by synchrotron radiation. Towards the end of the first year of data-taking, the dynamic pressure in both rings had dropped below the design goal, and the corresponding background contributions are much reduced. Nevertheless, beam-gas scattering remains the primary source of radiation damage in the SVT and the dominant source of background in all detectors systems, except the DIRC.

3.4.3. Luminosity Background

Radiative Bhabha scattering results in energy-degraded electrons or positrons hitting aperture limitations within a few meters of the IP and spraying *BABAR* with electromagnetic shower debris. This background is directly proportional to the instantaneous luminosity and thus is expected to contribute an increasing fraction of the total background in the future.

3.4.4. Background Fluctuations

In addition to these steady-state background sources, there are instantaneous sources of radiation that fluctuate on diverse time scales:

- beam losses during injection,
- intense bursts of radiation, varying in duration from a few ms to several minutes, currently attributed to very small dust particles, and
- non-Gaussian tails from beam-beam interactions (especially of the e^+ beam) that are highly sensitive to adjustments in collimator settings and ring tunes.

These effects typically lead to short periods of high background and have resulted in a large number of *BABAR*-initiated beam aborts (see below).

3.5. Radiation Protection and Monitoring

A system has been developed to monitor the instantaneous and integrated radiation doses, and to either abort the beams or to halt or limit the rate of injection, if conditions become critical. In addition, DCH and IFR currents, as well as DIRC and IFR counting rates, are monitored; abnormally high rates signal critical conditions.

Radiation monitoring and protection systems are installed for the SVT, the DCH electronics, and the EMC. The radiation doses are measured with silicon photo-diodes. For the SVT, 12 diodes are arranged in three horizontal planes, bottom, middle and top, with four diodes in each plane, placed at $z = +12.1$ cm and $z = -8.5$ cm and at a radial distance of 3 cm from the beam line [8]. The diode leakage current, after correction for temperature and radiation damage effects, is proportional to the dose rate. The four diodes in the middle are exposed to about ten times more radiation than the others. These mid-plane diodes are connected to the beam abort system, while the remaining eight diodes at the top and bottom are used to monitor the radiation dose delivered to the SVT bulk. The accuracy of the measured average dose rate is better than 0.5 mRad/s. The integrated dose, as measured by the SVT diodes, is presented in Figure 6.

The radiation level at the drift chamber and the calorimeter is more than two orders of magnitude lower than at the SVT. To amplify the signal, the PIN diodes are mounted on small CsI(Tl) crystals (about 10 cm^3). These silicon diodes are installed in sets of four. Three sets are mounted on the front face of the endcap calorimeter and one set on the backward endplate of the drift chamber, close to the read-out electronics. The signals of the four diodes in each set are summed, amplified and fed into the radiation protection electronics. Only one of the three diode sets of the EMC is used at any given time. The DCH and the EMC use identical hardware and decision algorithms. They limit injection rates or abort beams, whenever an instantaneous dose equivalent to about 1 Rad/day is exceeded.

The SVT employs a different strategy and circuitry to assess whether the measured radiation levels merit a beam abort or a reduction in single-

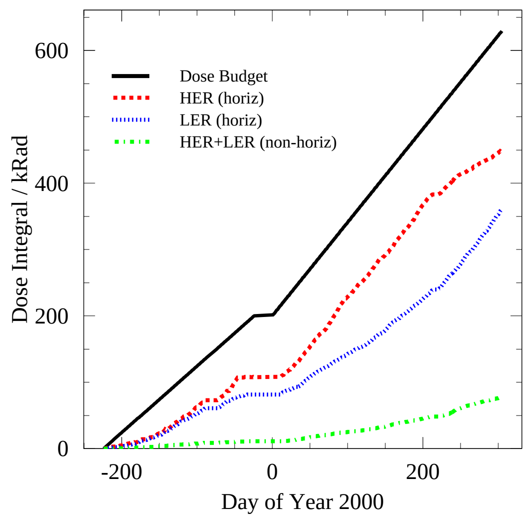

Figure 6. The integrated radiation dose as measured by PIN diodes located at three different positions near the SVT. Also shown is the SVT radiation budget for the first year of operation.

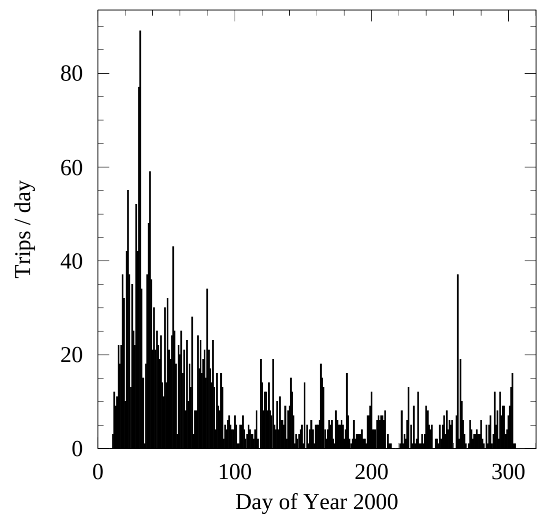

Figure 7. Daily rates of beam aborts initiated by the SVT radiation protection diodes, summed over regular data-taking and PEP-II injection.

beam injection rate. It was realized that every beam dump initiated by *BABAR* would be followed by a 5-10 minute period of injection with significant radiation exposure. Thus to minimize the ratio of the integrated radiation dose to the integrated luminosity, it has been beneficial to tolerate transient high-dose events during data-taking as long as the integrated dose remains less than the typical dose accumulated during injection. To differentiate between very high instantaneous radiation and sustained high dose rates, different trip thresholds are enforced on two different time scales: an instantaneous dose of 1 Rad accumulated over 1 ms, and an average of 50 mRad/s measured over a 5-minute period. In addition, higher thresholds are used during injection, since an aborted injection will delay the return to data-taking.

Figure 7 shows the daily rate of beam aborts initiated by the SVT protection diodes since January 2000. Initially as many as 80 beam aborts were triggered per day, while the average for stable operation is now significantly less than ten. The measures described above, combined with a significant reduction in large background fluctuations, have been very effective in protecting the detector against radiation damage, as well as very important to the effort of increasing the combined live time of the machine and detector to greater than 75%.

3.6. Impact of Beam-Generated Background on *BABAR*

Beam-generated backgrounds affect the detector in multiple ways. They cause radiation damage to the detector components and the electronics and thus may limit the lifetime of the experiment. They may also cause electrical breakdown and damage or generate large numbers of extraneous signals leading to problems with bandwidth limitations of the data acquisition system and with event reconstruction. The event reconstruction is affected both by degraded resolution and by losses in efficiency.

The impact of the beam-generated background on the lifetime and on the operation of the different detector systems varies significantly. Table 3 lists the limits on the instantaneous and integrated background levels in terms of the total dose and instantaneous observables. These limits are estimates derived from tests and/or experience in earlier experiments. For each detector system, an annual radiation allowance has been established taking into account the total estimated lifetime of the components and the expected annual operating conditions. The typical values accumulated for the first year of operation are also stated in the table.

Systematic studies of background rates were performed with stable stored beams. Measurements of the current-dependence of the backgrounds were carried out for single beams, two beams not colliding, and colliding beams with the goal to identify the current background sources, to develop schemes of reducing these sources and to extrapolate to operation at higher luminosity [7]. These experimental studies were complemented by Monte Carlo simulations of beam-gas scattering and of the propagation of showers in the detector. The studies show that the relative importance of the single-beam and luminosity background contributions varies, as illustrated in Figure 8. Data for the IFR are not shown because this system is largely insensitive to beam-generated backgrounds, except for the outer layer of the forward endcap due to insufficient shielding of the external beam line components.

The experience of the first year of operation and the concern for future operation for each of the detectors are summarized as follows.

SVT: The most significant concern for the SVT with regard to machine background is the integrated radiation dose. The instantaneous and integrated dose rates in the radiation protection diodes are representative, to within about a factor of two, of the radiation doses absorbed by the SVT modules. The exposure in the horizontal planes is an order of magnitude larger than elsewhere, averaging 15-25 mRad/s during stable beam operation. The worst integrated dose is 450 kRad, 30% below the allowance, giving confidence that the SVT can be operated several more years (see Figure 6).

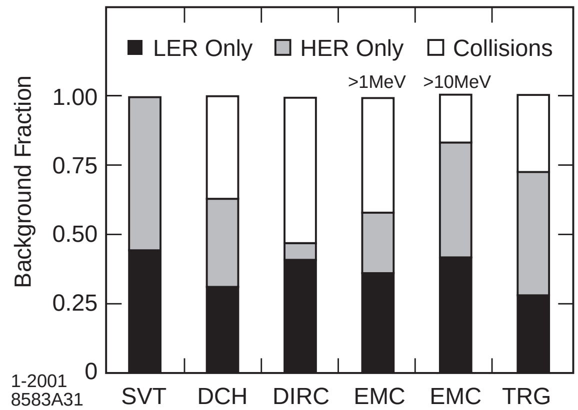

Figure 8. Fractional steady-state background contributions in *BABAR* detector systems, measured for single beams and colliding beams under stable conditions ($I^+ = 1.25\ A$, $I^- = 0.75\ A$, $L = 2.3 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$). The contributions are derived from the measured doses in the horizontal plane for the SVT, the total currents in the DCH, the rates in the DIRC photo-tubes, the occupacany and number of photons above 10 MeV in the EMC, and the L1 triiger rates.

DCH: For the DCH, the currents on the wires are the main concern, both because of the limited capacity of the HV power supplies and the effect of wire aging. The currents drawn are approximately uniformly distributed among the 44 HV supplies, one for each quadrant of superlayers 2-10, and two per quadrant for superlayer 1. Consequently, the total current limit is close to the sum of the limits of the individual supplies. During stable operation the total chamber current is 200-300 μA. However, radiation spikes can lead to currents that occasionally exceed the limit of 1000 μA, causing HV supplies to trip. Other background effects are measured to be well below the estimated lifetime limits and thus not a serious issue at this time. The average wire occupancy has not exceeded 1-2% during stable operation, but the extrapolation to future operation at higher luminosity and currents remains a major concern.

Table 3
BABAR background tolerance. Operational limits are expressed either as lifetime limits (radiation-damage and ageing-related quantities), or in terms of instantaneous observables (DCH current, DIRC and L1-trigger rates).

Detector system	Limiting factor and impact	Operational limit	First-year typical
SVT sensors & electronics	Integrated dose: radiation damage	2 MRad	0.33 MRad (hor.-plane modules) 0.06 MRad (other modules)
SVT sensors	Instantaneous dose: diode shorts	1 Rad/ms	N/A
DCH: electronics	Integrated dose: radiation damage	20 kRad	$\leq$ 100 Rad
DCH: wire current	Accumulated charge: wire aging	100 mC/cm	8 mC/cm
DCH: total current	HV system limitations	1000 μA	250 μA (steady-state)
DIRC photo tubes	Counting rate: TDC dead time	200 kHz	110 kHz (steady-state, well-shielded sector)
EMC crystals	Integrated dose: radiation damage	10 kRad	0.25 kRad (worst case)
L1 trigger	Counting rate: DAQ dead time	2 kHz	0.7 kHz (steady-state)

DIRC: The DIRC quartz bars were tested up to doses of 100 kRad without showing any measurable effects and thus radiation damage is not a concern. The present operational limit of the DIRC is set by the TDC electronics which induces significant dead time at rates above 200-300 kHz, well above the stable beam rate of 110 kHz in well shielded areas. Roughly half the present rate originates from radiative Bhabha scattering. The counting rate is due to debris from electromagnetic showers entering the water-filled stand-off box. Efforts are underway to improve the shielding of the beam pipe nearby.

EMC: The lifetime of the EMC is set by the reduction in light collection in the CsI crystals due to radiation damage. The cumulative dose absorbed by the EMC is measured by a large set of RADFETs [23] distributed over the barrel and endcap. The absorbed dose increases approximately linearly with the integrated luminosity. The highest dose to date was observed in the innermost ring of the endcap, close to 250 Rad, while the barrel crystals accumulated about 80 Rad. The observed reduction in light collection of 10-15% in the worst place, and 4-7% in the barrel, is consistent with expectation (see Chapter 11).

The energy resolution is dependent on the single crystal readout threshold, currently set at 1 MeV. During stable beam conditions the average crystal occupancy for random triggers is 16%, with 10% originating from electronics noise in the absence of any energy deposition. The spectrum of photons observed in the EMC from the LER and HER is presented in Figure 9. The HER produces a somewhat harder spectrum. The average occupancy for a threshold of 1 MeV and the average number of crystals with a deposited energy of more than 10 MeV are shown in Figure 10 as a function of beam currents for both single and colliding beams. The occupancy increases significantly at smaller polar angles, in the forward

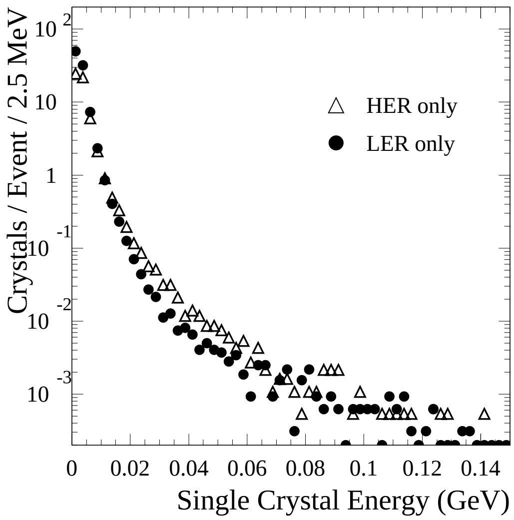

Figure 9. The energy spectrum of photons recorded in the EMC by random triggers with single beams at typical operating currents, LER at 1.1 A and HER at 0.7 A. The electronic noise has been subtracted.

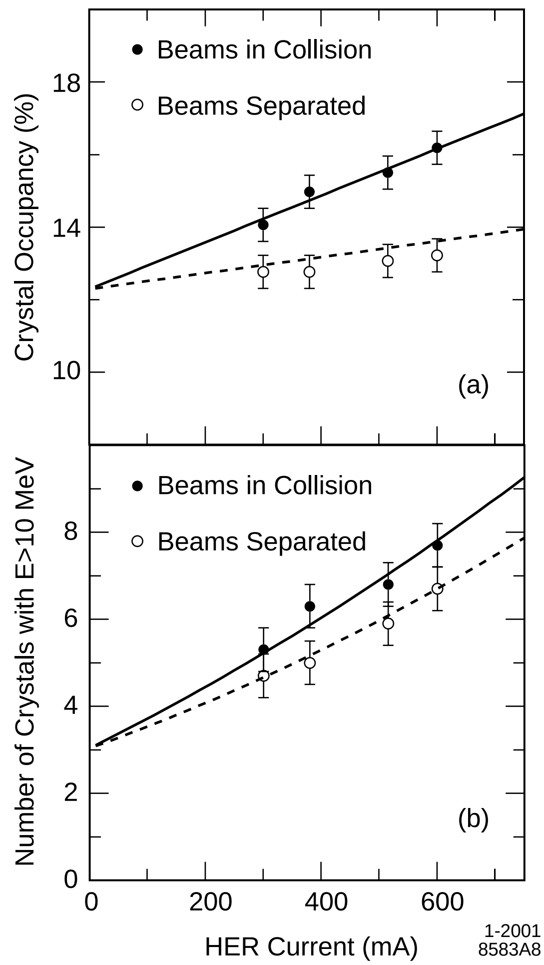

Figure 10. Average rates in the EMC for random triggers as a function of the HER current for a fixed LER current of 1.1 A, both for separated and colliding beams; a) the single crystal occupancy for thresholds of 1 MeV and b) the number of crystals with a deposited energy greater than 10 MeV. The solid curves represent a fit to the colliding beam data, the dashed curves indicate the sum of data recorded for single beams.

endcap and the backward barrel sections, and in the horizontal plane. The rate increase is approximately linear with the single beam currents. The data recorded with separated beams produce backgrounds that are consistent with those produced by single beams. For colliding beams, there is an additional flux of photons originating from small angle radiative Bhabha scattering. This effect is larger for low energy photons and thus it is expected that at higher luminosities the low energy background will raise the occupancy and thereby limit the EMC energy resolution.

L1 Trigger: During stable beam operation the typical L1 trigger rate is below 1 kHz, about a factor of of two below the data acquisition bandwidth limit of about 2-2.5 kHz. Experience shows that background bursts and other rate spikes can raise the data volume by as much as a factor of two and thus it is necessary to aim for steady state rates significant below the stated limit. For the L1 trigger, the dominant sources of DCH triggers

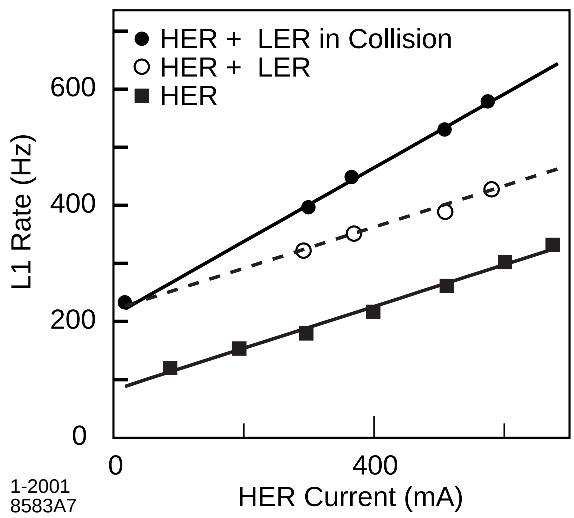

Figure 11. The L1 trigger rate as a function of the HER current for single beam only, for both beams, separated and colliding (with a LER current of 1.1 A).

are particles generated by interactions in vacuum flanges and the B1 magnets (see Figure XX in Chapter ??). This effect is most pronounced in the horizontal plane. At present, the HER background is twice as high as that of the LER, and the luminosity trigger contributes less than half of the combined LER and HER single beam triggers.

3.7. Summary and Outlook

Towards the end of the first year of data-taking, PEP-II routinely delivered beams close to design luminosity. Due to the very close cooperation with the PEP-II operations team, the machine-induced backgrounds have not been a major problem once stable conditions were established. The background monitoring and protection system has become a reliable and useful tool to safeguard the detector operation.

Currently planned upgrades are expected to raise the luminosity to $1.5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ within a few years. The single beam backgrounds will increase with beam currents and the luminosity background is projected to exceed, or at best remain comparable to, the beam-gas contribution. Measures are being prepared to reduce the sources and the impact of machine-related background on *BABAR*, among them upgrades to the DCH power supply system and to the DIRC TDC electronics, the addition of localized shielding against shower debris (especially for the DIRC standoff-off box), new vacuum chambers, adjustable collimators, and additional pumping capacity in critical regions upstream of the interaction point.

With the expected increase in LER current and luminosity both the single-beam and the luminosity-generated L1 trigger rates will increase and are projected to exceed 2 kHz (see Figure 11). Therefore, the DCH trigger is being upgraded to improve the rejection of background tracks originating from outside the luminous region. In addition, the data acquisition and data processing capacity will need to be expanded to meet the demands of higher luminosity.

Overall, the occupancy in all systems, except the IFR, will probably reach levels that are likely to impact the resolution and reconstruction efficiency. For instance, the occupancy in the EMC is expected to more than double. Thus, beyond the relatively straight forward measures currently planned for *BABAR* system upgrades, detailed studies of the impact of higher occupancy will be necessary for all systems.

REFERENCES

1. F. Author et al., The PEP-II Storage Rings submitted to Nucl. Instr. and Methods (2001), SLAC-PUB8xxx.
2. J. Seeman et al., Status Report on PEP-II Performance, in *Proceeedings of the 7th European Particle Accelerator Conference (EPAC 2000)*, Vienna, Austria (2000).
3. M. Sullivan, *B*-Factory Interaction Region Designs, in *Proceedings of the IEEE Particle Accelerator Conference (PAC97)*, Vancouver, B.C., Canada (1997), SLAC-PUB-7563.
4. S.E. Csorna et al., (CLEO Collaboration), Phys. Rev. 61 (2000) 111101.
5. T. Mattison et al., Background Measurements during PEP-II Commissioning in *Proceedings*

of the IEEE Particle Accelerator Conference (PAC99), New York, NY, USA (1999).
6. W. Kozanecki, Nucl. Instr. and Methods A 446 (2000) 59.
7. C. Hast et al., Report of the High-Luminosity Background Task Force, *BABAR* Note-522 (2000).
8. T.I. Meyer et al., contribution to DPF 2000, Meeting of the Division of Particles and Fields of the American Physical Society, Columbus, OH, USA (2000).
9. RADFET

4. The Solenoid Magnet and Flux Return

4.1. Overview

The *BABAR* magnet system consists of a superconducting solenoid [1], a segmented flux return and a field compensating coil. This system provides the magnetic field which enables charged particle momentum measurement, serves as the hadron absorber for hadron/muon separation, and provides the overall structure and support for the detector components. Figures 1 and 2 show key components of the *BABAR* magnet system and some of the nearby PEP-II magnets.

The magnet coil cryostat is mounted inside the hexagonal barrel flux return by four brackets on each end. The end-doors are each split vertically and mounted on skids that rest on the floor. To permit access to the inner detector, the doors can be raised and moved on rollers. At the interface between the barrel and the end-doors, approximately 60% of the area is occupied by structural steel and filler plates; the remaining space is reserved for cables and utilities from the inner detectors. A vertical triangular chase cut into the backward end-doors contains the cryostat chimney. Table 4 lists the principal parameters of the magnet system. The total weight of the flux return is approximately 870 metric tons.

To optimize the detector acceptance for unequal beam energies, the center of the *BABAR* detector is offset by 370 mm in the electron beam direction. The beam collision axis is rotated in the horizontal plane by 20 mrad relative to axis of the solenoid to minimize the perturbation of the beams by the solenoid field. The principal component of the magnetic field, B_z, lies along the $+z$ axis; this is also the approximate direction of the electron beam. The backward end-door is tailored to accommodate the DIRC bar boxes and to allow access to the drift chamber electronics. Both ends allow space and adequate shielding for the PEP-II quadrupoles.

4.2. Magnetic Field Requirements and Design

4.2.1. Field Requirements

A solenoid magnetic field of 1.5 T was specified in order to permit the desired momentum resolu-

Table 4
Magnet Parameters

Field Parameters		
Central Field	1.5	T
Max. Radial Field at Q1 and $r = 200$ mm	<0.25	T
Leakage into PEP-II	<0.01	T
Stored Energy	27	MJ
Steel Parameters		
Overall Barrel Length	4050	mm
Overall Door Thickness (incl. gaps for RPCs)	1149	mm
Overall Height	6545	mm
Plates in Barrel	18	
9	20	mm
4	30	mm
3	50	mm
2	100	mm
Plates in Each Door	18	
9	20	mm
4	30	mm
4	50	mm
1	100	mm
Main Coil Parameters		
Mean Diameter of Current Sheet	3060	mm
Current Sheet Length	3512	mm
Number of layers	2	
Operating Current	4596	A
Conductor Current Density	1.2	kA/mm^2
Inductance	2.57	H
Bucking Coil Parameters		
Inner Diameter	1906	mm
Operating Current	200	A
Number of Turns	140	
Cryostat Parameters		
Inner Diameter	???	
Radial Thickness	???	
Total Length	???	
Total Material	???	

tion for charged particles. To simplify track finding and fast and accurate track fitting, the magnitude of the magnetic field within the tracking volume was required to be constant within a few percent.

The magnet was designed to minimize disturbance of operation of the PEP-II beam elements. The samarium-cobalt B1 dipole and Q1 quadrupole magnets are located inside the solenoid as shown in Figure 1. Although these magnets can sustain the high longitudinal field of 1.5 T, they cannot tolerate a large radial component. Specifically, the field cannot exceed 0.25 T at a radius $r = 200$ mm (assuming a linear dependence of B_r on r) without degrading their field properties due to partial demagnetization. The conventional iron quadrupoles Q2, Q4, and Q5 are exposed to the solenoid stray fields. To avoid excessive induced skew octupole components, the stray field leaking into these beam elements is required to be less than 0.01 T, averaged over their apertures.

4.2.2. Field Design Considerations

Due to unavoidable saturation of the steel near the coil and the gap between the coil and the steel, field non-uniformities arise. To control these non-uniformities the current density of the coil is increased at the ends relative to the center by reducing the thickness of the aluminum stabilizer. While the requirements on the radial field component at Q1 inside the solenoid can be satisfied easily at the forward end, the shape of the backward plug is specifically designed to simultaneously control field uniformity and unwanted radial components.

Sizable leakage of the magnetic flux is a problem, in particular at the backward end. A compensating coil, mounted at the face of the backward door and surrounding the DIRC strong support tube, is designed to reduce the stray field to an acceptable level for the DIRC photomultipliers and the PEP-II quadrupoles.

4.2.3. Magnetic Modeling

Extensive calculations of the magnetic field were performed to develop the detailed design of the flux return, the solenoid coil, and the compensating coil. To cross-check the results of these calculations the fields were modeled in detail in two and three dimensions using commercial software [2]. The shape of the hole in each end-door was designed by optimizing various parameters, such as the minimum steel thickness in areas of saturation. The design of the hole in the forward door was particularly delicate because the highly saturated steel is very close to the Q2 quadrupole. The multiple “finger” design of the hole was chosen to control the saturation of the steel.

Most of the design work was performed in two dimensions, but some three dimensional calculations were necessary to assure the accuracy of modeling the transitions between the end-doors and the barrel, the leakage of field into the PEP-II magnets, and the impact of that leakage on the multipole purity [3,4]. The computations of the leakage field were done for central field of 1.7 T instead of 1.5 T to provide some insurance against uncertainties in the modeling of complex steel shapes and the possible variations of the magnetic properties of the steel.

4.3. Steel Flux Return

4.3.1. Mechanical and Magnetic Forces

The magnet flux return supports the detector components on the inside, but this load is not a major issue. Far greater demands are placed on the structural design by the magnetic forces and the mechanical forces from a potential earthquake.

Magnetic forces are of three kinds. First, there is a symmetric magnetic force on the end-doors which was taken into consideration in their design and construction method. Second, there is an axial force on the solenoid due to the forward-backward asymmetry of the steel. Because the steel is highly saturated in places, the magnitude of the field asymmetry changes when the current is raised from zero and there is no position of the solenoid at which the force remains zero at all currents. Because it is important that this axial force should not change sign, which could cause a quench, the superconducting solenoid was deliberately offset by 30 mm towards the forward door. This offset was chosen to accommodate a worst case scenario, including uncertainties in the calculation. Third, during a quench of the superconducting coil, eddy currents in conducting components inside the magnetic volume could generate sizable forces. These forces were analyzed for components such as the end-plates of the drift chamber and the electromagnetic calorimeter and were found not to be a problem.

4.3.2. Earthquake Considerations

Because SLAC is located in an earthquake zone, considerable attention has been given to protecting the detector against severe damage from such an event. The entire detector is supported on four earthquake isolators, one at each corner, which limit the component acceleration in the horizontal plane to 0.4 g. However, these isolators offer no protection in the vertical direction. Vertical ground accelerations of 0.6 g are considered credible and actual component accelerations may be considerably larger due to resonances. By taking into account resonant frequencies and the expected frequency spectra of earthquakes, the magnet and all detector components have been designed to survive these accelerations without serious damage. Because the magnet is isolated from the ground moving beneath it, worst case clearances to external components, e.g. PEP-II components, are provided. It is expected that even during a major earthquake, damage would be modest.

4.3.3. Fabrication

The flux return was fabricated [5] from drawings prepared by the *BABAR* engineering team. A primary concern was the magnetic properties of the steel. The need for a high saturation field dictated the choice of a low carbon steel, specified by its chemical content (close to AISI 1006). The manufacturer supplied sample steel for critical magnetic measurements and approval. The availability of very large tools at the factory made it possible to machine the entire face of each end of the assembled barrel, thus assuring a good fit of the end-doors. The entire flux return was assembled at the factory, measured mechanically, and inspected before disassembly for shipment.

4.4. Magnet Coils

The design of the superconducting solenoid is conservative and follows previous experience. The superconducting material is composed of niobium-titanium (46.5% by weight Nb) filaments, each less than 40 μm in diameter. These filaments are then wound into 0.8 mm strands, 16 of which are then formed into *Rutherford cable* measuring 1.4 x 6.4 mm. The final conductor [6] consists of Rutherford cable co-extruded with pure aluminum stabilizer measuring 4.93 x 20.0 mm for use on the outer, high current density portion of the solenoid, and 8.49 x 20.0 mm for the central, lower current density portion. The conductor is covered in an insulating dry wrap fiberglass cloth vacuum impregnated with epoxy. The conductor has a total length of 10.3 km.

The solenoid is indirectly cooled to an operating point of 4.5 K using a thermo-syphon technique. Liquid helium [7] is circulated in channels welded to the solenoid support cylinder. Liquid helium and cold gas circulate between the solenoid, its shields, the liquefier/refrigerator and a 4000 ℓ storage dewar via 60 m of coaxial, gas-screened, flexible transfer line. The solenoid coil and its cryostat were fabricated [8] to drawings prepared by the *BABAR* engineering team. Before shipment [9], the fully assembled solenoid was cooled to operating temperature and tested with currents of up to 1000 A, limited by coil forces in the absence of the iron flux return.

To reduce the leakage fields into the PEP-II components and the DIRC photomultipliers, an additional external compensating coil is installed. This is a conventional water cooled copper coil consisting of 10 layers. Although the nominal operating current is 200 A, a current of up to 575 A may be needed to demagnetize the DIRC shield.

To optimally control the stray fields and avoid a magnetization of the DIRC magnetic shield, the currents in the solenoid and the compensating coil are ramped together under computer control. High precision transducers are used to measure the currents and provide the feedback signals to the power supplies. The values of the currents are recorded in the *BABAR* database.

4.5. Magnetic Field Map

The goal of the magnetic field mapping and subsequent corrections was to determine the magnetic field in the tracking volume to a precision of 0.2 mT.

4.5.1. Mapping Procedure

A field mapping device was built specifically for the *BABAR* magnet based on a design concept developed at Fermilab [10]. The magnetic field sensors were mounted on a "propeller" at the end of a long cantilevered spindle which reached through the hole in the forward end-door. The spindle in turn rode on a carriage which moved on precision-aligned rails. The propeller rotated to sample the magnetic volume in ϕ, and the carriage moved along its axis to cover z. Measurements were obtained from five sets of B_r and B_z and two B_ϕ Hall probes, all of which were mounted on a plate at different radial positions. This plate was attached to the propeller and its position could be changed to cover the desired range in the radial distance r from the axis. Precision optical alignment tools were used to determine the position of the sensors transverse to the z axis.

The B_r and B_z probes were two-element devices with a short term precision of 0.01%, the B_ϕ probes were single element devices with a precision of 0.1% [11]. In addition to the Hall probes, an NMR probe [12] was mounted at a radius of 89 mm on the propeller to provide a very precise field reference near the z axis as a function of z for $|z| < 1000$ mm, where $z = 0$ at the magnet center. The NMR measurements set the absolute scale of the magnetic field.

The magnetic field was mapped at the nominal central operating field of 1.5 T, as well as at 1.0 T. Measurements were recorded in 100 mm intervals from -1800 to $+1800$ mm in z, and in 24 azimuthal positions spaced by 15° for each of three different positions of the Hall probe plate. Thus for each z and ϕ position, the components B_r and B_z were measured at 13 radii from 130 mm to 1255 mm and B_ϕ at six radii between 505 mm and 1180 mm.

The field map was parameterized in terms of a polynomial of degree up to 40 in r and z plus additional terms to account for expected perturbations [13]. The fit reproduced the measurements to within an average deviation of 0.2 mT throughout the tracking volume. The fitting procedure also served as a means of detecting and removing questionable measurements.

4.5.2. Perturbations to the Field Map

During the mapping process, the permanent magnet dipoles (B1) and quadrupoles (Q1) were not yet installed. Their presence inside the solenoid results in field perturbations of two kinds. The first is due to the fringe fields of the B1 and Q1 permanent magnets, and of the dipole and quadrupole trim coils mounted on Q1. The B1 field strength reaches a maximum of ≈ 20 mT close to the surface of the B1 casing and decreases rapidly with increasing radius. The fields associated with the trim coils were measured and parameterized prior to installation; they are essentially dipole in character.

The second field perturbation is due to the permeability of the permanent magnet material. Sintered samarium-cobalt has a relative permeability of 1.11 to 1.13 in the z direction, and as a result the solenoid field is modified significantly. Probes between the B1 and Q1 magnets at a radius of about 190 mm measure the effect of the permeability. The field perturbation is obtained from a two-dimensional, finite element analysis which reproduces the r and z dependence of B_r and B_z. The induced magnetization increases B_z by about 9 mT at the interaction point, the effect decreases slowly with increasing radius.

4.5.3. Field Quality

To illustrate the quality of magnetic field, Figure 12 shows the field components B_z and B_r as a function of z for various radial distances r. In the tracking volume the field is very uniform, the B_ϕ component does not exceed 1 mT. The variation of the bend field, i.e. the field transverse to the trajectory, along the path of a high momentum track is at most 2.5% from maximum to minimum within the tracking region, as shown in Figure 13.

4.5.4. Field Computation

In order to reduce the computation of the magnetic field for track reconstruction and momentum determination, the field values averaged over azimuth are stored in a grid of r - z space points spanning the volume interior to the cryostat. Local values are obtained by interpolation. Within the volume of the SVT, a linear interpolation is performed in a 20 mm grid; elsewhere the interpolation is quadratic in a 50 mm grid. Azimuthal dependence is parameterized by means

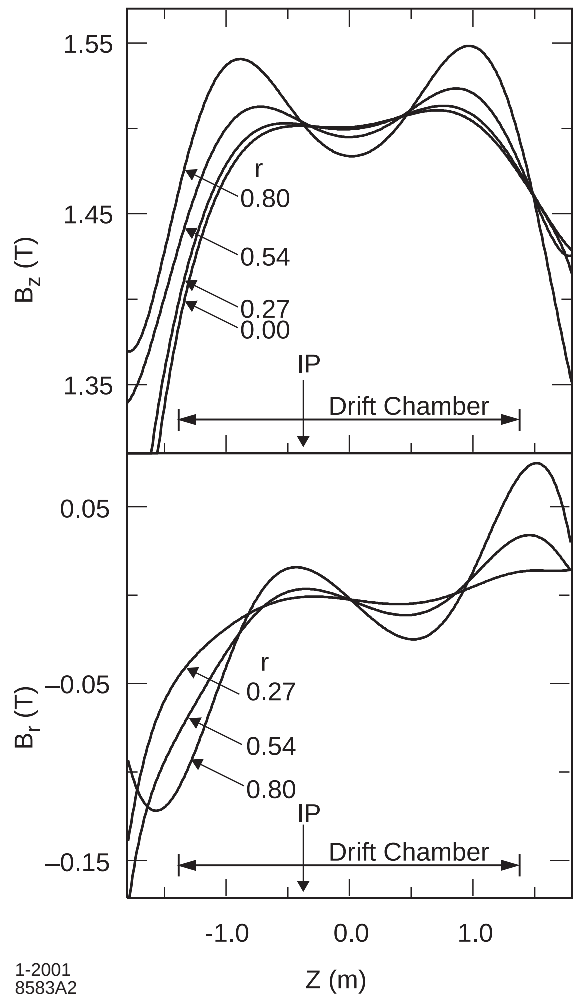

Figure 12. The magnetic field components B_z and B_r as a function of z for various radial distances r (in m). The extent of the drift chamber and the location of the interaction point (IP) are indicated.

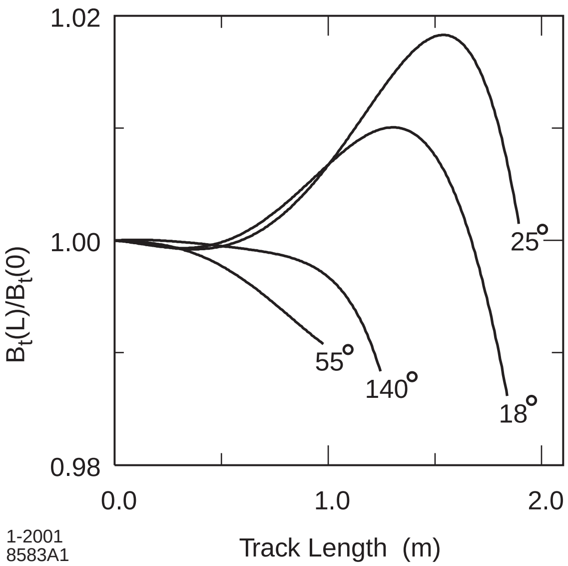

Figure 13. Relative magnitude of magnetic field transverse to a high momentum track as a function of track length from the IP for various polar angles (in degrees). The data are normalized to the field at the origin.

of a Fourier expansion at each $r-z$ point. The Fourier coefficients at the point of interest are obtained by interpolation on the $r-z$ grid, and the average field value is corrected using the resulting Fourier series.

4.6. Summary

Since its successful commissioning, the magnet system has performed without problems. There have been no spontaneous quenches of the superconducting solenoid. In the tracking region the magnetic field meets specifications, both in magnitude and uniformity. The field compensation and magnetic shielding work well for the DIRC photomultiplier array and the external quadrupoles.

REFERENCES

1. IEEE Transactions on Applied Superconductivity, 9 #2 (1999) 847.
2. Opera-2D, Vector Fields, Inc., Aurora, Illinois 60505 (USA).
   ANSYS by ANSYS Inc., Canonsburg, PA 15317 (USA).
   Mermaid (©1994) by Sim Limited, Novosibirsk (Russia).
3. A. Onuchin, et al., Magnetic Field Calculation in the *BABAR* Detector, *BABAR* Note 344 (1996).
4. L. Keller, et al., Magnetic Field Calculation in the *BABAR* Detector, *BABAR* Note 370 (1997).
5. Kawasaki Heavy Industries (KHI), Kobe (Japan).
6. Europa Metali & Cortaillod (Switzerland).
7. Solenoid cool-down and cryogenic Helium is supplied by a modified Linde TCF-200 liquefier/refrigerator.
8. Ansaldo Energia, Genova (Italy).
9. The solenoid was shipped from Genova on a C5-B transport plane of the US Air Force.
10. C. Newman-Holmes, E.E. Schmidt, and R. Yamada, Nucl. Instr. and Meth. A 274 (1989) 443.
11. Sentron model 2MR-4A/3B-14B25-20 (2-element) and, AMR-3B-14B25-20 (1-element) Hall Probes, GMW Associates, San Carlos, CA (USA).
12. Metrolab model PT 2025 Telsameter with model 1060 NMR probe, Metrolab Instruments, SA, CH-1228, Geneva (Switzerland).
13. A. Boyaraki, et al., "Field Measurements in the *BABAR* Solenoid" *BABAR* Note 514 (2000).

5. Silicon Vertex Tracker

5.1. Goals and Design Requirements

The silicon vertex tracker (SVT) has been designed to provide precise reconstruction of charged particle trajectories and decay vertices near the interaction region. It forms one of the two detectors responsible for charged particle tracking in *BABAR*, the other being the drift chamber (DCH). The SVT is the critical detector for the measurement of time-dependent CP asymmetries via the measurement of B^0 meson decay vertices. The design choices were driven primarily by direct requirements from physics measurements and constraints imposed by the PEP-II interaction region and *BABAR* experiment. In this chapter we will discuss the mechanical and electronic design of the SVT, with some discussion of the point resolution per layer and dE/dx performance. We discuss the tracking performance and efficiency of the SVT alone and in combination with the DCH, in Section 9.

5.1.1. Requirements and Constraints

Physics requirements are derived from the decay topologies of B mesons in the laboratory frame. To avoid significant impact on the CP asymmetry measurement ($< 10\%$ with respect to perfect vertex reconstruction) the mean spatial resolution on each B decay vertex along the z-axis must be better than 80 μm [1]. The required resolution in the $x-y$ plane arises from the need to reconstruct final states in B decays as well as in τ and charm decays. For example, in decays of the type $B^0 \to D^+D^-$separating the two D vertices is important. The distance between the two D's in the $x-y$ plane for this decay is typically $\sim 275\,\mu$m. Hence the SVT needs to provide $x-y$ resolution of order $\sim$100 μm.

Many of the decay products of B mesons are at low $p_{\rm t}$. The SVT must provide stand-alone tracking for particles with transverse momentum less than 120 MeV/c, the minimum that can be reliably measured in the DCH alone. This feature is fundamental for the identification of slow pions from D^* meson decays: a tracking efficiency of 70% or more is desirable for tracks with a transverse momentum in the range 50–120 MeV/c. This also means that the material traversed by particles must be minimized.

The stand-alone tracking capability and the need to link SVT tracks to the DCH were crucial in choosing the number of layers. Beyond the stand-alone tracking capability, the SVT provides the best measurement of track angles at the IP, which is required to achieve design resolution for the Cherenkov angle.

Additional constraints are imposed by the storage rings and the environment of the *BABAR* detector. The SVT is located inside the $\sim$4.5 m long support tube, that extends all the way through the detector. To maximize the angular coverage, the SVT must extend down to 350 mrad (17.2°) in polar angle from the beamline in the forward direction. The region at smaller polar angles is occupied by the B1 permanent magnets. In the backward direction, it is sufficient to extend the SVT sensitive area down to 30°.

The SVT must withstand 2 MRad of ionizing radiation. A radiation monitoring system capable of aborting the beams is required. The expected radiation dose is 240 kRad/yr in the horizontal plane immediately outside the beam pipe (where the highest radiation is concentrated), and 33 kRad/yr on average otherwise.

The SVT is inaccessible during normal detector operations. Hence, reliability and robustness are essential: all components of the SVT inside the support tube should have long mean-time-to-failure, because the time needed for any replacement is estimated to be 4-5 months. Redundancies are built in whenever possible and practical.

Because the SVT is physically well insulated from the external environment, the SVT is cooled to remove the heat generated by the electronics or radiated by the water cooled beam pipe. In addition, it operates in the 1.5T magnetic field.

To achieve the position resolution necessary to carry out physics analyses, the relative position of the individual silicon sensors should be stable over long time periods. The assembly allows for relative motion of the support structures with respect to the B1 magnets. This feature is also necessary to sustain an earthquake of moderate intensity.

These requirements and constraints have led to the choice of a SVT made of five layers of double-sided silicon strip sensors. To fulfill our physics requirements, the spatial resolution, for perpendicular tracks, must be 10-15 μm in the three inner layers and about 40 μm in the two outer layers. The inner three layers perform the impact parameter measurements, while the outer layers are necessary for pattern recognition and low $p_{\rm t}$ tracking.

5.2. SVT Layout

The five layers of double-sided silicon strip sensors, which form the SVT detector, are organized in 6, 6, 6, 16, and 18 modules, respectively: a photograph is shown in Figure 14. The strips on the two sides are oriented orthogonally to each other: the ϕ measuring strips ("ϕ strips") run parallel to the beam and the z measuring strips ("z strips") are oriented transversely to the beam axis. The modules of the inner 3 layers are straight, while the modules of layers 4 and 5 are arch-shaped (Figures 15 and 16). This arch design was chosen to minimize the amount of silicon required to cover the solid angle, while increasing the crossing angle for particles near the edges of acceptance. A photograph of an outer layer arch module is shown in Figure 17. The modules are divided electrically into two half-modules, which are read out at the ends.

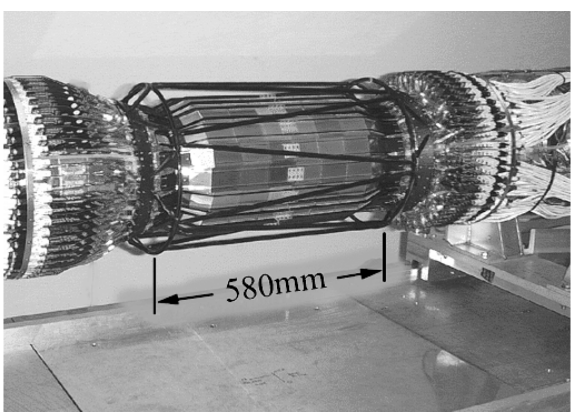

Figure 14. Fully assembled Silicon Vertex Tracker. The silicon sensors of the outer layer are visible, as is the carbon-fiber space frame (black structure) that surrounds the silicon.

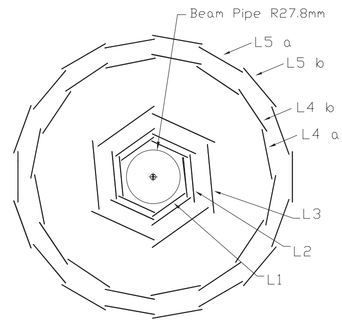

Figure 16. Schematic view of SVT: tranverse section.

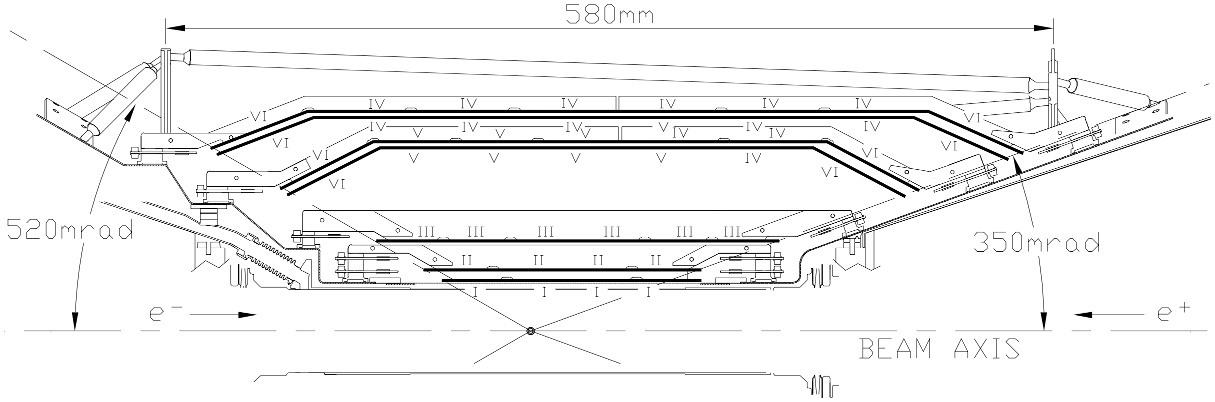

Figure 15. Schematic view of SVT: longitudinal section. The roman numerials label the six different types of sensors.

Figure 17. Photograph of an SVT arch module in an assembly jig.

To satisfy the different geometrical requirements of the five SVT layers, five different sensor shapes are required to assemble the planar sections of the layers. The smallest detectors are 48x40 mm^2, and the largest are 55x52 mm^2. Two identical trapezoidal sensors are added (one each at the forward and backward ends) to form the arch modules. The half-modules are given mechanical stiffness by means of two carbon fiber/kevlar ribs, which are visible in Figure 17. The ϕ strips of sensors in the same half-module are electrically connected with wire bonds to form a single readout strip. This results in a total strip length up to 14 cm (24 cm) in the inner (outer) layers.

The signals from the z strips are brought to the readout electronics using fanout circuits consisting of conducting traces on a thin (50 μm) insulating Upilex [2] substrate. For the innermost three layers, each z strip is connected to its own preamplifier channel, while in layers 4 and 5 the number of z strips on a half module exceeds the number of electronics channels available, requiring that two z strips on different detectors are electrically connected (ganged) to a single electronics channel. The length of a z strip is about 5 cm (no ganging) or 10 cm (two strips connected). The ganging introduces an ambiguity on the z coordinate measurement, which must be resolved in the pattern recognition procedure. The total number of readout channels is approximately 150,000.

The inner modules are tilted in ϕ by 5°, allowing an overlap region between adjacent modules, a feature that is very useful for the alignment. The outer modules cannot be tilted, because of the arch geometry. To avoid gaps and to have a suitable overlap in ϕ, layers 4 and 5 are divided into two sub-layers (4a, 4b, 5a, 5b) placed at slightly different radii. The relevant geometrical parameters of each layer are summarized in Table 5.

Table 5
Geometric parameters for each layer and readout plan of the SVT. "Floating strips" refers to the number of strips between readout strips. Note: parts of the ϕ sides of layers 1 and 2 are bonded at 100 μm and 110 μm pitch, respectively, with one floating strip. Strip length of z-strips for layers 4 and 5 takes includes ganging. The radial range for layers 4 and 5 includes the radial extent of the arched sections.

Layer/ view	Radius (cm)	Readout pitch (μm)	Floating strips	Strip length (mm)
1 z	3.2	100	1	40
1 ϕ	3.2	50-100	0-1	82
2 z	4.0	100	1	48
2 ϕ	4.0	55-110	0-1	88
3 z	5.4	100	1	70
3 ϕ	5.4	110	1	128
4 z	9.1-12.7	210	1	104
4 ϕ	9.1-12.7	100	1	224
5 z	11.4-14.4	210	1	104
5 ϕ	11.4-14.4	100	1	265

In order to minimize the material in the detector acceptance region, the readout electronics are mounted entirely outside the active detector volume. The forward electronics must be mounted in the 1 cm space between the 350 mrad stay-clear space and B1 magnet. This implies that the hybrids carrying the front-end chip must be positioned at an angle of 350 mrad relative to the sensor for the inner layers, and at an even larger angle for the outer ones (Figure 15). In the backward direction the available space is larger and the inner layer electronics can be placed in the sensor plane, allowing a simplified assembly procedure.

The module assembly and the mechanics are quite complicated, especially for the arch modules, and are described in detail elsewhere [3]. The SVT support structure (Figure 14) is a rigid body made from two carbon-fiber cones, connected by a "space-frame", also made of carbon-fiber epoxy laminate.

An optical survey of the SVT on its assembly jig indicated that the global error in placement of the sensors with respect to design was $\sim$200μm, FWHM. Subsequently the detector was disassembled and shipped to SLAC, where it was re-assembled on the IR magnets. The SVT is attached to the B1 magnets by a set of gimbal rings, in such a way as to allow for relative motion of the two B1 magnets while fixing the position of the SVT relative to the forward B1 and the orientation relative to the axis of both B1 dipoles. The support tube structure is mounted on the PEP-II accelerator supports, independently *BABAR*, leading to the possible movements between the SVT and the rest of *BABAR*. Precise position monitoring of the beam interaction point is necessary, this is described in section 5.4.

The total active silicon area is 0.96 m^2 and the material traversed by particles is $\sim$ 4% of a radiation length (see Section 2). The geometrical acceptance of SVT is 92% of the solid angle in the center-of-mass system.

5.3. SVT Components

A block diagram of SVT components is shown in Figure 18. The basic components of the detector are the silicon sensors, the fanout circuits, the front end electronics and the data transmission system. Each of these components is discussed below.

5.3.1. Silicon sensors

The SVT sensors [4] are made of 300 μm thick double-sided silicon strip devices. They were designed at INFN Pisa and Trieste (Italy) and fabricated commercially [5]. They are built on high-

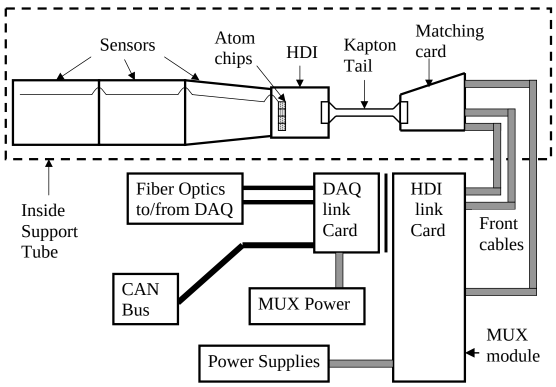

Figure 18. Schematic block diagram showing the different components of the SVT.

resistivity (6-15 kΩ-cm) n-type substrates with p^+ strips and n^+ strips on the two opposite sides. The insulation of the n^+ strips is provided by individual p-stops, so as to achieve an inter-strip resistance greater than 100 MΩ at operating bias voltage, normally about 10 V above the depletion voltage. Typical depletion voltages are in the range 25-35 V. On both sides the strips are biased with polysilicon resistors (4-20 MΩ) to ensure the required radiation hardness, keeping the voltage drop across resistors and the parallel noise as low as possible. Strips are AC-coupled to electronics via integrated decoupling capacitors, whose capacitance depends on the sensor shape, but is always greater than 14 pF/cm. The sensors were designed to maximize the active area, which extends to within 0.7 mm of the physical edges. Another design goal was to control the inter-strip capacitance: values between 0.7 pF/cm and 1.1 pF/cm were obtained for the various sensor shapes. To achieve the required spatial resolution, while keeping the number of readout channels as low as possible, most of the modules have a floating strip between two readout strips.

The leakage currents, because of the excellent performance of the manufacturing process, were as low as 50 nA/cm^2 on average, measured at 10 V above depletion voltage. The silicon sensor parameters have been measured after irradiation with ^{60}Co sources. Apart from an increase in the inter-strip capacitance of about 12% during the first 100 kRad, the main effect was an increase of the leakage current by 0.7 μA/cm^2/MRad. However, in a radiation test performed in a 1 GeV electron beam, an increase in leakage current of about 2 μA/cm^2/MRad, and a significant shift in the depletion voltage, dependent on the initial dopant concentration, were observed. A shift of about 8-10 V was seen for irradiation corresponding to a dose of approximately 1 MRad. These observations indicate significant bulk damage caused by energetic electrons. Due to the change in depletion voltage, the SVT sensors could undergo type inversion after about 1-3 MRad. Preliminary tests show that the sensors continue to function after type inversion [6]. Studies of the behavior of SVT modules as a function of radiation dose continue.

Table 6
Electrical parameters of the SVT, shown for the different layers and views. C_{input} refers to the total input capacitance, R_{series} is the series resistance. The amplifier peaking time is 200 ns for layers 1–3 and 400 ns for layers 4–5.

Layer/ view	C_{input} (pF)	R_{series} (Ω)	Noise, calc. (elec)	meas. (elec)
1 z	6.0	40.	550	880
1 ϕ	17.2	164.	990	1200
2 z	7.2	48.	600	970
2 ϕ	18.4	158.	1030	1240
3 z	10.5	70.	700	1180
3 ϕ	26.8	230.	1470	1440
4 z	16.6	104.	870	1210
4 ϕ	33.6	224.	1380	1350
5 z	16.6	104.	870	1200
5 ϕ	39.7	265.	1580	1600

5.3.2. Fanout circuits

The fanout circuits, which route the signals from the strips to the electronics, have been designed to minimize the series resistance and the inter-strip capacitance. As described in [7], a trace on the fanout has a series resistance about 1.6 Ω/cm, an inter-strip resistance > 20 MΩ, and an inter-strip capacitance < 0.5 pF/cm. The electrical parameters of the final assembly of sensors and fanouts (referred to as Detector Fanout Assemblies or DFAs) are summarized in Table 6. Due to the different strip lengths, there are large differences between the inner and the outer layers. Smaller differences are also present between the forward and backward halves of the module, that are of different lengths.

5.3.3. Front end electronics

The electrical parameters of a DFA and the general *BABAR* requirements are the basic inputs that drove the design of the SVT front-end IC: the ATOM (A Time-Over-Threshold Machine). In particular, the front-end IC had to cope with the following requirements:

- signal to noise ratio greater than 15 for minimum ionizing particle (MIP) signals for all modules;
- signals from all hit strips must be retained, in order to improve the spatial resolution through interpolation, while keeping the number of transmitted hits as low as possible. A 'hit" means a deposited charge greater than 0.95 fC, corresponding to 0.25 MIP (a MIP is the average charge deposited by a minimum ionizing particle) ;
- the amplifier must be sensitive to both negative and positive charge;
- the peaking time must be programmable, with a minimum of 100 ns (in layers 1 and 2, because of the high occupancy), up to 400 ns (outer layers, with high capacitance)
- capability to accept random triggers with a latency up to 11.5 μs and a programmable jitter up to ± 1 μs, without dead time;
- radiation hardness greater than 2.5 MRad;
- small dimensions: 128 channels in a 6.2 mm wide chip.

These requirements are fully satisfied by the ATOM chip [8], which is depicted schematically in Figure 19. The linear analog section consists

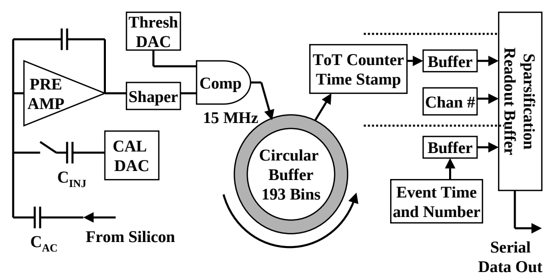

Figure 19. Schematic diagram of the ATOM front end IC.

of a charge-sensitive preamplifier followed by a shaper. Gains of 200 mV/fC (low) or 300 mV/fC (high) may be selected. The channel gains on a IC are uniform to 5 mV/fC. Signals are presented to a programmable-threshold comparator, designed so that the output width of the pulse (time over threshold or ToT) quasi-logarithmic function of collected charge. This output is sampled at 30 MHz and stored in a 193 location buffer. Upon receipt of a first level trigger, the time and ToT is retrieved from this latency buffer, sparcified, and stored in a four event buffer. Upon the receipt of a L1 accept command from the data acquisition system, the output data (the 4 bits for the ToT, 5 bits for the time stamp, and 7 bits for the strip address) are formatted, serialized, and delivered. The IC also contains a test charge injection circuit. The typical noise behavior of the ATOM, as described by the equivalent noise charge (ENC) of the linear analog section is given in Table 7.

The average noise for the various module types is shown in Table 6. Given that shot noise due to sensor leakage current is negligible, the expected noise may be calculated from the parameters of Tables 6 and 7. The results of such a calculation are also shown in Table 6. The maximum average noise is 1,600 electrons, leading to a signal to noise ratio greater than 15.

Table 7
ATOM chip Equivalent Noise Charge (ENC) parameters at different peaking times

ATOM peaking time	Noise Charge at zero Capacitance	Noise slope
100 ns	380 e^-	40.9 e^-/pF
200 ns	280 e^-	33.9 e^-/pF
400 ns	220 e^-	25.4 e^-/pF

The power consumption of the IC is about 4.5 mW/channel. Radiation hardness was checked up to 2.4 MRad of ^{60}Co radiation. At that dose, the gain decreased 20%, and the noise increased less than 15%.

The ATOM IC's are mounted on thick-film double-sided hybrid circuits (known as High Density Interconnects or HDIs) based on an aluminum-nitride substrate with high thermal conductivity. The electronics are powered through a floating power supply system, in such a way as to guarantee a small voltage drop (< 1 V) across the detector decoupling capacitors.

5.3.4. Data transmission

The digitized signals are transmitted from the ATOM chips through a thin kapton cable to the matching cards, from where they are routed to more conventional cables. Just outside the detector signals are multiplexed by the MUX modules, converted into optical signals and transmitted to the ROMs. The MUX modules also receive digital signals from the DAQ via a fiber optical connection. The SVT is connected to the *BABAR* online detector control and monitoring system via the industry standard CAN bus. Details on SVT data transmission system and DAQ can be found in [10,11]. Power to SVT modules (silicon sensor bias voltage and ATOM low voltages) is provided by a CAEN A522 power supply system [12].

5.4. Monitoring and calibration

To identify immediately any operational problems, the SVT is integrated in the control and monitoring system (see Section 14). Major concerns for SVT monitoring are temperature and humidity, mechanical position and radiation dose.

5.4.1. Temperature and humidity monitors

The total power dissipation of the SVT modules is about 350 W, mainly dissipated in the ATOM chip. External cooling is provided by chilled water at 8°C. In addition, humidity is reduced by a stream of dry nitrogen in the support tube.

Since excessive temperature can permanently damage the front-end electronics, temperature monitoring is very important to the safe operation of the SVT. Thermistors are located on the HDIs (for the measurements of front-end electronics temperature), around the SVT, along the cooling systems and in the electronics (MUX) crates. The absolute temperatures are monitored to 0.2°C and relative changes of 0.1°C. Additionally, a series of humidity sensors are employed to monitor the performance of the dry nitrogen system. The temperature and humidity monitors also serve as an interlock to the HDI power supplies.

5.4.2. Position monitoring sensors

The relative alignment of the two tracking systems (DCH and SVT) is monitored with a precision equal to the spatial resolution of the outermost SVT layer (40 μm). This is done on an hourly basis with charged particles from Bhabha and muon pairs from $e^+e^- \to \mu^+\mu^-$ events with a precision of approximately 10 μm. To detect sudden unexpected movements, it is important to monitor the SVT position on a much faster time scale. This is achieved by a system of capacitive sensors that measure the position of the SVT with respect to the PEP-II B1 magnets and the position of the support tube with respect to the drift chamber.

An example of the understanding that can be achieved by the monitoring systems is given in Figure 20. The relative DCH-SVT horizontal position during a period of 6 days, as measured by the capacitive sensors, is shown. This curve is compared with the temperature, measured by the thermistors around the SVT: a correlation is evident, thus explaining the diurnal movement. The displacement measurement performed by the capacitive sensors is also confirmed by the measurements obtained with Bhabha events and $\mu^+\mu^-$

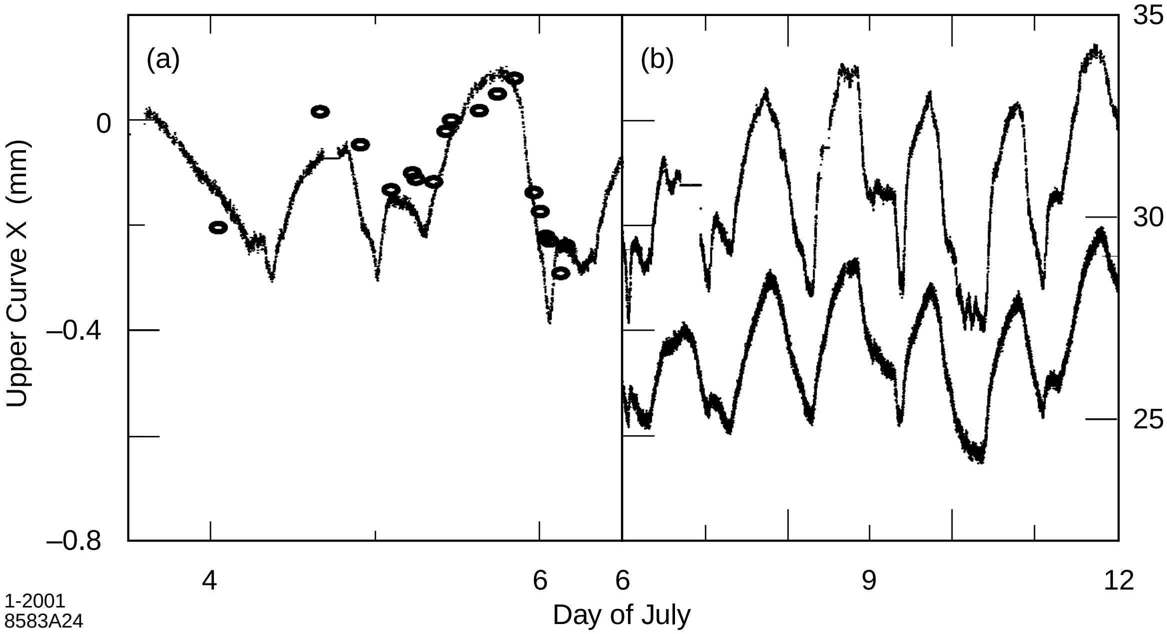

Figure 20. (a) Horizontal motion measured with the capacitive sensors (curve) between the DCH and the support tube and mean x coordinate of the collision point (circles) measured with e^+e^- and $\mu^+\mu^-$ events for a three day period in July 1999. An arbitrary offset and scale has been applied to the beamspot data. (b) Diurnal horizontal motion of the drift chamber relative to the SVT as measured with the capacitive sensors (upper curve) and temperature near the detector (lower curve) as a function of time, for six days in July 1999.

pairs.

5.4.3. Radiation monitors

Radiation monitoring is extremely important to ensure the SVT does not exceed its radiation budget, which could cause permanent damage to the device. To date, the measured radiation absorbed by the SVT is well within the allowed budget.

The monitoring of radiation dose to the SVT is discussed in detail in Section 3.

5.4.4. Calibrations

Once a day, and each time the SVT configuration has changed, calibrations are performed in absence of circulating beams. All electronic channels are tested with pulses through test capacitors, for different values of the injected charge. Gains, thresholds and electronic noise are measured, and defective channels are identified. The calibration results have proved very stable and repeatable. The main variation in time is the occasional discovery of a new defective channel. The calibration procedures have also been very useful for monitoring noise sources external to the SVT.

5.4.5. Defects

Due to a series of minor mishaps incurred during the installation of the SVT, 9 out of 208 readout sections (each corresponding to one side of a half-module) were damaged and are currently not functioning. There is no single failure mode, but several causes: defective connectors, mis-handling during installation and not-fully-understood problems on the front-end electronics hybrid. There has been no module failure due to radiation damage. It should be noted that due to the redundancy afforded by the five layers of the SVT, the presence of the defective modules has minimal impact on physics analyses.

In addition, there are individual channel defects, of various types, at a level of about 1%. Calibrations have revealed an increase in the number of defective channels at a rate of less than 0.2%/year.

5.5. Data analysis and performance

This section describes the clustering of the raw data into hits, the SVT internal and global alignment, single hit efficiency, and resolution and dE/dx performance of the SVT.

5.5.1. Cluster and hit reconstruction

Under normal running conditions the average occupancy of the SVT in a time window of 1 μs is about 3% for the inner layers, with a significant azimuthal variation due to beam-induced backgrounds, and $< 1\%$ for the outer layers, where noise hits dominate. Figure 21 shows the typical occupancy as a function of IC number (equivalent to azimuthal angle, in this case) for layer 1, ϕ side. In the inner layers, the occupancy is dominated by machine backgrounds, which are significantly higher in the horizontal plane, seen in the plot as the peaks near IC numbers 3 and 25, approximately.

The first step of the reconstruction program consists in discarding out-of-time channels. A time correction, the time between the passage of the particle and the time the shaper exceeds threshold, is performed, after which hits with times more than 200 ns from the event time (determined by the drift chamber) are discarded. The loss of real hits from this procedure is negligible. The resulting in-time hits are then passed to the cluster finding algorithm. First, the charge pulse height (Q) of a single pulse is calculated from the ToT value. In a first pass, clusters are formed grouping adjacent strips with consistent times. In a second pass clusters separated by just one strip are merged into one cluster. The two original clusters plus the merged cluster are made available to the pattern recognition algorithm, which chooses among the three.

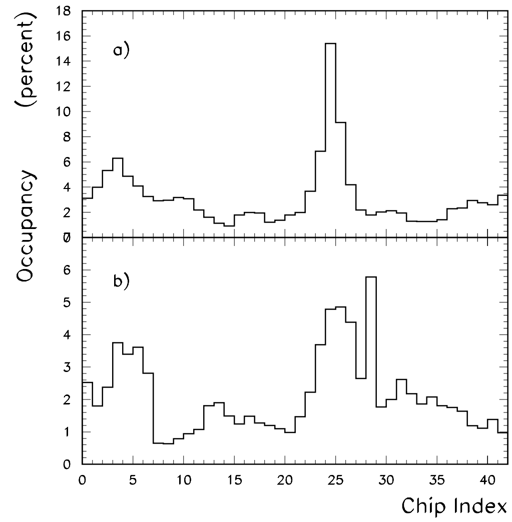

Figure 21. Typical occupancy in percent as a function of IC number in layer 1, ϕ side for a) forward half-modules and b) backward half-modules. IC number corresponds to the azimuthal angle and the increased occupancy in the horizontal plane is visible near ICs 3 and 25.

The position x of a cluster formed by n strips is determined, using the "head-to-tail" algorithm:

$$x = \frac{(x_1 + x_n)}{2} + \frac{p}{2}\frac{(Q_n - Q_1)}{(Q_n + Q_1)},$$

where x_i and Q_i are the position and collected charge of i-th strip, respectively, and p is the readout pitch. This formula results in a cluster position that is always within $p/2$ of the geometrical center of the cluster. The cluster pulse height is simply the sum of the strip charges, while the cluster time is the average of the signal times.

5.5.2. Alignment

The alignment of the SVT is performed in two steps. The first step consists of determining the internal (or local) relative positions of the 340 silicon sensors. Once this is accomplished, the next step is to align the SVT as a whole within the global coordinate system defined by the drift chamber. The primary reason for breaking the alignment process into these two steps is that the local positions are relatively stable in time compared to the global position. Also, the local alignment procedure is considerably more complex than the global alignment procedure. Thus, the global alignment can be updated on a run-by-run basis, while the local alignment constants are changed as needed, typically after magnet quenches or detector access.

The local alignment procedure is performed using tracks from $e^+e^- \rightarrow \mu^+\mu^-$ events and cosmic rays. Well-isolated high momentum tracks from hadronic events are also used to supplement sensors that are not sufficiently illuminated by di-muon and cosmic events. Data samples sufficient to perform the local alignment are collected in 1 to 2 days of typical running conditions.

In $\mu^+\mu^-$ events the two tracks are simultaneously fit using a Kalman filtering technique that utilizes the known beam momentum. The use of tracks from cosmic rays reduces any systematic distortion that may be introduced in the local alignment due to imprecise knowledge of the beam momentum. For all tracks, no information from the DCH is used, effectively decoupling the SVT and DCH.

In addition to the information from tracks, data from an optical survey performed during the assembly of the SVT is included in the alignment procedure. The typical precision of these optical measurements is $4\,\mu$m. This survey information is only used to constrain sensors relative to other sensors in the same module, but not one module to another or one layer to another. Furthermore, only degrees of freedom in the plane of the sensor are constrained as they are expected to be the most stable given the assembly procedure.

Using the hit residuals from the aforementioned set of tracks and the optical survey information, a χ^2 is formed for each sensor and minimized with respect to the sensor's 6 local parameters. Because each sensor's χ^2 is minimized separately, the process must be iterated. The combination of constraints coming from the overlapping regions of the silicon sensors, the di-muon fit, the cosmic rays and the optical survey result in an internally consistent local alignment.

After the internal alignment, the SVT is considered as a rigid body. The second alignment step consists in determining the position of the SVT with respect to the drift chamber. Tracks with sufficient numbers of SVT and DCH hits are fit two times: once using only the DCH information and again using only the SVT hits. The six global alignment parameters, three translations and three rotations, are determined by minimizing the difference between track parameters obtained with the SVT-only and the DCH-only fits. As reported above, because of the diurnal movement of the SVT with respect to the drift chamber, this global alignment needs to be performed once per run ($\sim$ every 2-3 hours). The alignment constants obtained in a given run are then used to reconstruct the data in the subsequent run. This procedure, known as "rolling calibrations", ensures that track reconstruction is always performed with up-to-date global alignment constants.

Figure 22 shows a comparision between the optical alignment made during the SVT assembly in February 1999 and a local alignment using data taken during January, 2000. The alignment from tracking data was made without using cosmics or constraints from the optical survey. The width of the distributions in these plots has four contributions: 1) The SVT was disassembled from its assembly jig and re-assembled on the IR magnets, 2) time dependent motion of the SVT after mounting, 3) statistical errors, and 4) systematic errors. The second set of plots shows the difference in two alignment sets for data taken in January 2000 as compared to March 2000. In general, the stability of the inner three layers is excellent, with slightly larger tails in the outer two layers. The radial coordinate is less tighly constrained in all measurements because the radial location of the charge deposition is not well known, and most of the information about the radial locations comes only from constraints in the overlap region of the sensors.

5.5.3. Performance

The SVT hit efficiency can be calculated for each half-module by comparing the number of associated hits to the number of tracks crossing the active area of the module. As can be seen in

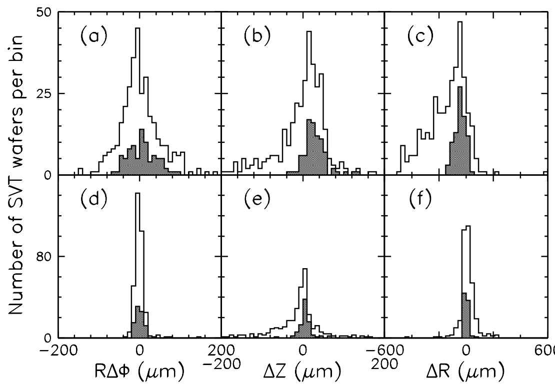

Figure 22. Comparison of a local alignment of all the sensors in the SVT using data from January 2000 with the optical survey of the SVT made during assembly in February 1999 in the (a) $R\Delta\phi$, (b) ΔZ and ΔR coordinates. Plots (d), (e), and (f) show the difference between two local alignments using data from January 15-19 and March 6-7, 2000 for the $R\Delta\phi$, ΔZ, and ΔR coordinates, respectively. In all the plots, the shaded regions correspond to the sensors in the first 3 layers. In comparing the different alignments and optical survey, a six parameter fit (three global translations and three global rotations) has been applied between the data sets.

Figure 23, a combined hardware and software efficiencies of 97% is measured, excluding defective readout sections (9 out of 208), but employing no special treatment for other defects, such as broken AC coupling capacitors or dead channels on front-end chips. Actually, since most of the defects affect a single channel, they do not contribute heavily to inefficiency, because most charge depositions involve two or more strips, due to track crossing angles and, to some degree, charge diffusion.

The spatial resolution of SVT hits is determined by measuring the distance (in the plane of the sensor) between the track trajectory and the hit, using high-momentum tracks in two prong events. The uncertainty due to the track trajectory is subtracted from the width of the residual distribution to obtain the hit resolution. Figure 24 shows the SVT hit resolution for z and ϕ side hits as a function of track incident angle, for each of the five layers. The measured resolutions are in excellent agreement with expectations from Monte Carlo simulations.

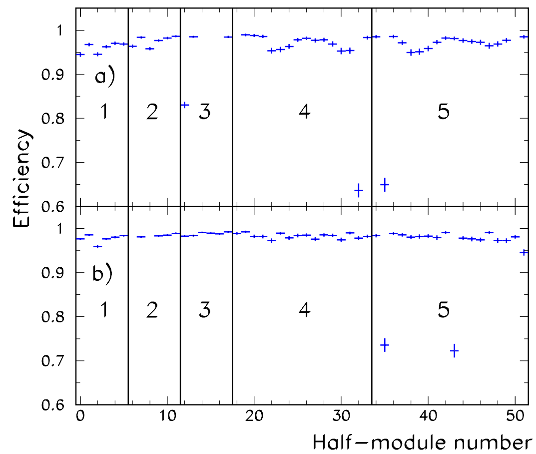

Figure 23. SVT hit reconstruction efficiency, as measured on $\mu^+\mu^-$ events for a) forward half-modules and b) backward half-modules. The plots show the probability of associating both a ϕ and z hit to a track passing through the active part of the detector. The horizontal axis corresponds to the different modules, with the vertical lines separate the different layers as numbered. Missing values correspond to non-functioning half-modules.

Initial studies have shown that hit reconstruction efficiency and spatial resolution are effectively independent of occupancy for the occupancy levels observed so far.

Measurement of the ToT value by the ATOM chips enables one to obtain the pulse height, and hence the ionization dE/dx in the SVT sensor. The values of ToT are converted to pulse height using a lookup table computed from the pulse shapes obtained in the bench measurements. The pulse height is corrected for track length vari-

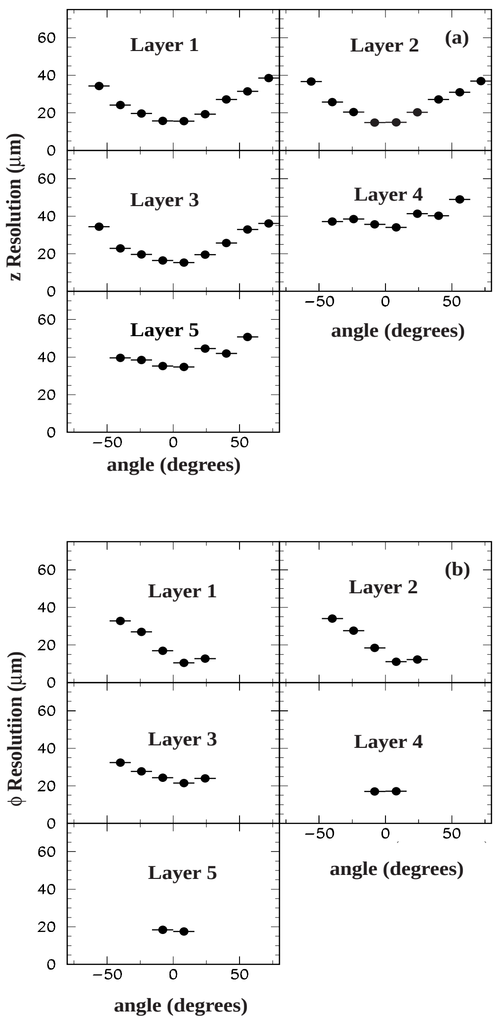

Figure 24. SVT hit resolution in the a) z and b) ϕ coordinate in microns, plotted as a function of track incident angle in degrees. Each plot shows a different layer of the SVT. The plots in the ϕ coordinate for layers 1-3 are asymmetric around $\phi = 0$ because of the "pin-wheel" design of the inner layers. There are fewer points in the ϕ resolution plots for the outer layers as they subtend smaller angles than the inner layers.

ation. The double sided sensors provide up to 10 measurements of dE/dx per track. For every track with signals from at least 4 sensor in the SVT, a 60% truncated mean dE/dx is calculated. The cluster with the smallest dE/dx energy is also removed to reduce sensitivity to electronics noise. For minimum-ionizing particles, the resolution on the truncated mean dE/dx is approximately 14%. We can achieve a 2σ separation between the kaons and pions up to momentum of 500 MeV/c, and between kaons and protons beyond 1 GeV/c.

5.6. Summary and Outlook

The Silicon Vertex tracker has been operating efficiently since its installation in the *BABAR* experiment in May 1999. The five layer device, based on double-sided silicon sensors, has satisfied the original design goals, in particular the targets specified for efficiency, hit resolution and low transverse momentum track reconstruction. The radiation dose during the first $25\,\mathrm{fb}^{-1}$ of integrated luminosity is within the planned budget, and no modules have failed due to radiation damage. The performance of the SVT modules at high radiations dose is currently being studied. Early results indicate that the sensors will continue to function after type inversion (at 1-3 MRad), but further tests with irradiated sensors and ATOM ICs need to be performed. A program of spare module production has commenced, with the goal of replacing modules that are expected to fail due to radiation damage. Beam generated backgrounds are expected to rise with increasing luminosity. Physics studies at five times the current backgrounds levels indicate no change in mass or vertex resolution for the mode $B^0 \rightarrow J/\psi K_S$ and a $\sim$ 20% loss of resolution in the $D^{*+} - D^0$ mass difference. In this study the detector efficiency for both decay modes was lower by 15–20%.

REFERENCES

1. D. Boutigny et al., SLAC-R-0443 (1994) 69.
2. UBE Industries, Japan. see also [3]
3. C. Bozzi et al., *Nucl. Instr. and Methods* **A447** (2000) 20.

4. D. Barbieri et al., *Nuo. Cim.* **A112** (1999) 113.
5. MICRON Semiconductor Ltd., Lancing, U.K.
6. L. Bosisio, private communication of preliminary results.
7. G. Della Ricca et al., *Nucl. Instr. and Methods* **A409** (1998) 258.
8. V. Re at al., *Nucl. Instr. and Methods* **A409** (1998) 354.
9. F. Lanni and F. Palombo, *Nucl. Instr. and Methods* **A379** (1996) 399. The HDIs are manufactured by AUREL, Milano, Italy (?).
10. J. Beringer et al., "The Data Transmission System for the *BABAR* Silicon Vertex Tracker", *BABAR* Note # 518, May 2000.
11. R. Claus et al., SLAC-PUB-8134 (1999).
12. CAEN – Costruzioni Apparecchiature Elettroniche Nucleari – Via Vetraia, 11 – 55049 Viareggio, Italy.

6. Drift Chamber

6.1. Purpose and Design Requirements

The principal purpose of the drift chamber (DCH) is the efficient detection of charged particles and the measurement of their momenta and angles with high precision. These high precision measurements enable the reconstruction of exclusive B and D meson decays with minimal background. The DCH complements the measurements of the impact parameter and the directions of charged tracks provided by the SVT near the IP. At lower momenta, the DCH measurements dominate the errors on the extrapolation of charged tracks to the DIRC, EMC, and IFR. Most critical are the angles at the DIRC, because the uncertainties in the charged particle track parameters add to the uncertainty in the measurement of the Cherenkov angle.

The reconstruction of decay and interaction vertices outside of the SVT volume, for instance the K^0_S decays, relies solely on the DCH. For this purpose, the chamber should be able to measure not only the transverse momenta and positions, but also the longitudinal position of tracks, with a resolution of $\sim$1 mm.

The DCH also needs to supply information for the charged particle trigger with a maximum time jitter of $0.5\,\mu$s (Section **??**).

For low momentum particles, the DCH is required to provide particle identification by measurement of ionization loss ($\mathrm{d}E/\mathrm{d}x$). A resolution of about 7% will allow π/K separation up to 700 MeV/c. This capability is complementary to that of the DIRC in the barrel region, while in the extreme backward and forward directions, the DCH is the only device providing some discrimination of particles of different mass.

Since the average momentum of charged particles produced in B and D meson decays is less than 1 GeV/c, multiple scattering is a significant, if not the dominant limitation on the track parameter resolution. In order to reduce this contribution, material in front of and inside the chamber volume has to be minimized.

Finally, the DCH must be operational in the presence of large beam-generated backgrounds, which were predicted to generate rates of ~5 kHz/cell in the innermost layers.

6.2. Mechanical Design and Assembly

6.2.1. Overview

The DCH is relatively small in diameter, but almost 3 m long, with 40 layers of small hexagonal cells providing up to 40 spatial and ionization loss measurements for charged particles with transverse momentum greater than 180 MeV/c. Longitudinal position information is obtained by placing the wires in 24 of the 40 layers at small angles with respect to the z-axis. By choosing low-mass aluminium field wires and a helium-based gas mixture the multiple scattering inside the drift chamber is held to a minimum, less than 0.2% X_0 of material. The properties of the chosen gas, a 80:20 mixture of helium:isobutane, are presented in Table 8. This mixture has a radiation length that is five times larger than commonly used argon-based gases. The smaller Lorentz angle results in a rather uniform time-distance relationship and thereby improved spatial resolution.

Table 8
Properties of helium-isobutane gas mixture at atmospheric pressure and 20°C. The drift velocity and Lorentz angle are given for an electric field of 600 V/cm with no magenetic field and with 1.5 T, respectively.

Parameter	Values
Mixture He : C_4H_{10}	80:20
Radiation Length	807 m
Primary Ions	21.2 /cm
Drift Velocity	22 μm/ ns
Lorentz Angle	32°
dE/dx Resolution	6.9%

The inner cylindrical wall of the DCH is kept thin to facilitate the matching of the SVT and DCH tracks, to improve the track resolution for high momentum tracks, and to minimize the background from photon conversions and interactions. Material in the outer wall and in the forward direction is also minimized so as not to degrade the performance of the DIRC and the EMC. For this reason, the HV distribution and all of the readout electronics are mounted on the backward endplate of the chamber. This choice also eliminates the need for a massive, heavily shielded cable plant.

A longitudinal cross section and dimensions of the DCH are shown in Figure 25. The DCH is bounded radially by the support tube at its inner radius and the DIRC at its outer radius. The device is asymmetrically located with respect to the interaction point. The forward length of 1749 mm is chosen so that particles emitted at polar angles of 300 mrad traverse at least half of the layers of the chamber before exiting through the front endplate. In the backward direction, the length of 1015 mm means that particles with polar angles down to −475 mrad traverse at least half of the layers. This choice ensures sufficient coverage for forward-going tracks, and thus avoids significant degradation of the invariant mass resolution, while at the same time maintaining a good safety margin on the electrical stability of the chamber. The DCH extends beyond the endplate by 485 mm at the backward end to accommodate the readout electronics, cables, and an rf shield. It extends beyond the forward endplate by 68 mm to provide space for wire feed-throughs and an rf shield.

6.2.2. Structural Components

Details of the DCH mechanical design are presented in Figure 26. The endplates, which carry an axial load of 31,800 kN, are made from aluminium plates of 24 mm thickness. At the forward end, this thickness is reduced to 12 mm beyond a radius of 46.9 cm to minimize the material in front of the calorimeter endcap. For this thickness, the estimated safety margin on the plastic yield point for endplate material (6061T651 aluminium) is not more than a factor of two. The maximum total deflection of the endplates under loading is small, about 2 mm or 28% of the 7 mm wire elongation under tension. During installation of the wires, this small deflection was taken into account by over-tensioning the wires.

The inner and outer cylinders are load bearing to reduce the maximum stress and deflections of the endplates. The stepped forward endplate created a complication during the assembly, because

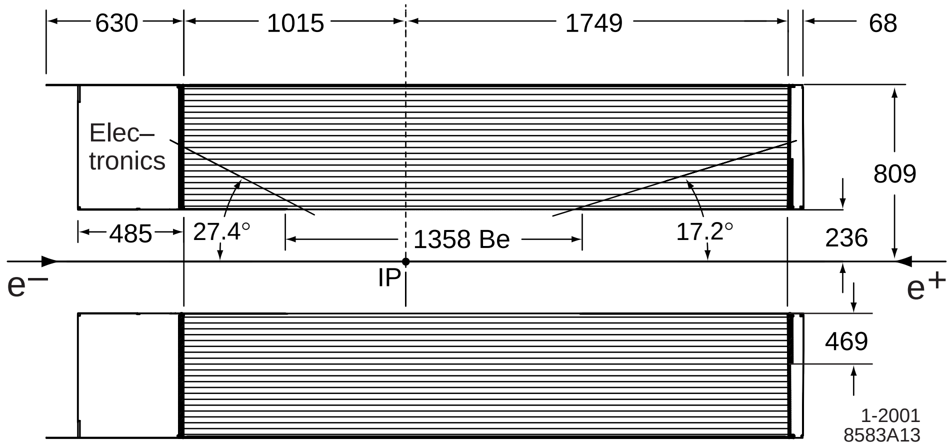

Figure 25. Longitudinal section of the drift chamber with principal dimensions; the chamber center is offset by 370 mm in z from the interaction point.

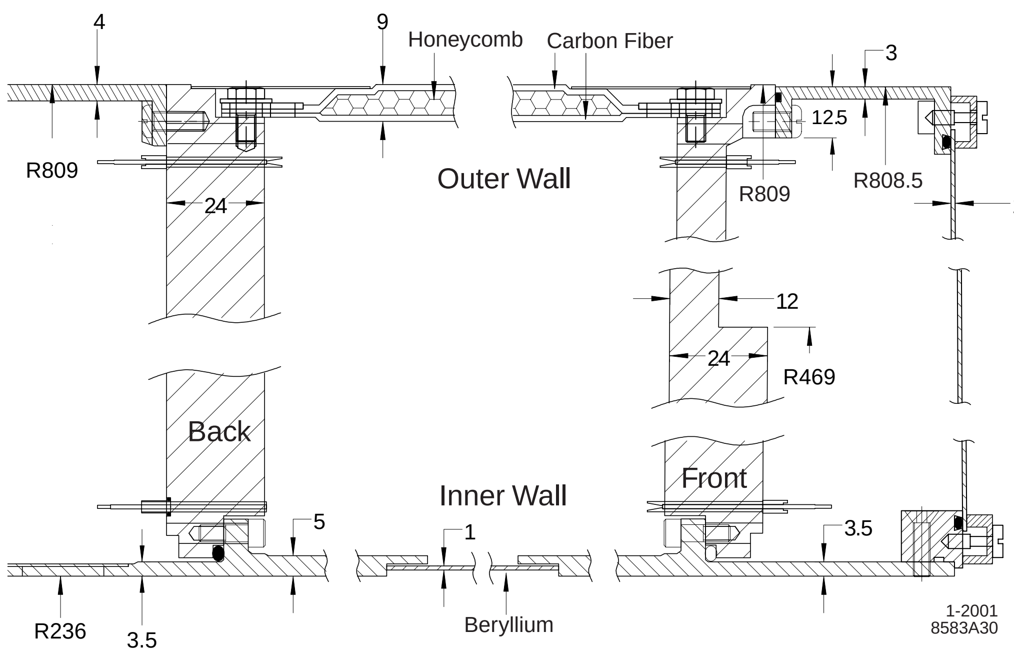

Figure 26. Details of the structural elements of the DCH. All components are made of aluminium, except for the 1 mm-thick inner beryllium wall and the 9 mmthick outer composite wall.

the thinner forward endplate would deflect more than the thicker backward endplate. The outside rim of the forward endplate had to be pre-loaded, i.e., displaced by 2.17 mm in the forward direction, to maintain the inside and outside rims of the rear endplate at the same longitudinal position after the load transfer to the outer cylinder was complete.

Prior to installation on the inner cylinder, the two endplates were inspected on a coordinate-measuring machine. All sense wire holes, as well as 5% of the field and clearing field wire holes, were measured to determine their absolute locations. The achieved concentricity was $38\,\mu$m for both sense and field wires, better than the specification by more than a factor of two. In addition, the diameters of the same sample of endplate holes were checked with precision gauge pins. All holes passed the diameter specification ($4.500\pm^{0.025}_{0.000}$ for sense wires and $2.500\pm^{0.025}_{0.00}$ for the field and guard wires.

The inner cylindrical wall of the DCH, which carries 40% of the wire load, was made from five sections, a central 1 mm-thick beryllium tube with two aluminium extensions which were electron-beam welded to aluminium end flanges to form a 3 m-long cylindrical part. The central section was made from three 120° segments of rolled and brazed beryllium. The end flanges have precision surfaces onto which the endplates were mounted. These surfaces set the angles of the two endplates with respect to the axis and significantly constrain the concentricity of the tube. The inner cylinder also provides a substantial rf shield down to the PEP-II bunch-gap frequency of 136 kHz.

The outer wall bears 60% of axial wire load between the endplates. To simplify its installation, this external wall was constructed from two half-cylinders with longitudinal and circumferential joints. The gas and electrical seals for these joints were made up *in situ*. The main structural element consists of two 1.6 mm-thick ($0.006X_0$) carbon fiber skins laminated to a 6 mm-thick honeycomb core. The outer shell is capable of withstanding a differential pressure of 30 mbar and temperature variations as large as ±20°C, conditions that could be encountered during shipping or installation. Aluminium foil, $25\,\mu$m-thick on the inside surface and $100\,\mu$m on the outside, are in good electrical contact with the endplates, thereby completing the rf shield for the chamber.

The total thickness of the drift chamber at normal incidence is 1.08% X_0, of which the wires and gas mixture contribute 0.2% X_0, and the inner wall 0.28% X_0.

6.2.3. Wire Feed-Throughs

A total of five different types of feed-throughs were required for the chamber to accommodate the sense, field, and clearing field wires, as well as two different endplate thicknesses. The five types are illustrated in Figure 27. They incorporate crimp pins of a simple design which fasten and precisely locate the wires. The choice of pin material (gold-plated copper for the signal wires and gold-plated aluminium for the field wires) and wall thickness in the crimp region was optimized to provide an allowable range of almost $150\,\mu$m in crimp size, as a primary means for avoiding wire breakage.

Crimp pins were either press-fit into an insulator made from a single piece of injection-molded thermoplastic reinforced with 30% silica glass fiber [1], or swaged into a copper jacket for the field wires. The plastic insulates the sense, guard, and clearing field wires from the electrically grounded endplates, while the metal jackets provide good ground contact for field wires ($< 0.1\Omega$) on the backward endplate. The outer diameter of the field and clearing field feed-throughs was maintained at $2.475^{+0.000}_{-0.025}$ mm while the sense wire feed-through had a larger ($4.5^{+0.000}_{-0.025}$ mm) outer diameter and a longer body (41.7 mm). This choice provided both thicker insulating walls and a longer projection into the gas volume to better shield the HV from the grounded endplate.

6.2.4. Assembly and Stringing

Assembly of the chamber components and installation of the wires was carried out in a large clean room (Class 10,000) at TRIUMF in Vancouver. The wires were strung horizontally without the outer cylindrical shell in place. The endplates were mounted and aligned onto the inner cylinder which in turn was supported by a central shaft in

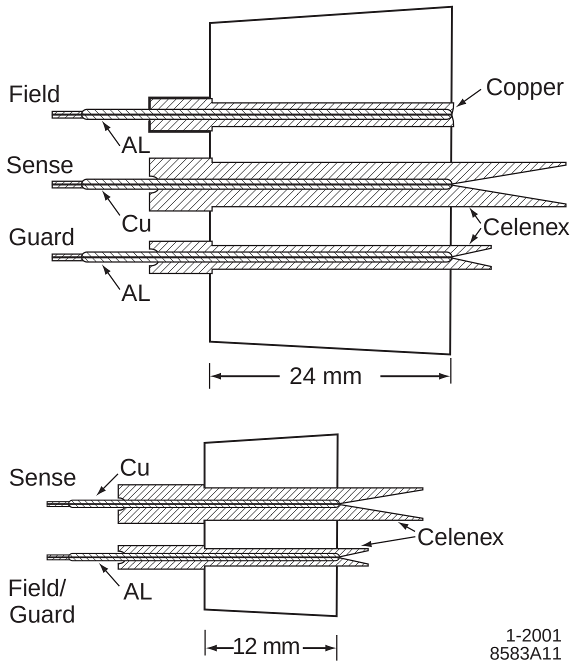

Figure 27. Design of the five wire feed-throughs for the 24 mm-thick endplates and the 12 mm-thick endplate. The copper jacketed feed-through is for grounded field wires, the other four are for sense wires (4.5 mm diameter), and guard and clearing field wires (2.5 mm diameter), all made from a Celenex insulator surrounding the crimp pins.

a mobile fixture. The endplates were mounted on the inner cylinder at the inside rim and and attached to support rings at the outside. These rings were connected by radial *spiders* to the central shaft of the stringing frame.

Two teams of two operators each worked in parallel as the wires were strung from the inner radius outward. The two teams were each assisted by an automated wire transporter [2]. A wire was attached to a needle and inserted through an endplate hole where it was captured magnetically by one of the transporters, and then transported and inserted though the appropriate hole in the other endplate. Once a wire was threaded through the feed-throughs, the feed-throughs were glued into the endplates, and the wire was tensioned and crimped. The automated wire transporters were largely built from industrial components, employing commercial software and hardware. The semi-automatic stringing procedure ensured the correct hole selection, accelerated the stringing rate and greatly improved the cleanliness and quality of the stringing process. The installation of a total of 30,000 wires was completed in less than 15 weeks.

6.3. Drift Cells

6.3.1. Layer Arrangement

The DCH consists of a total of 7,104 small drift cells, arranged in 40 cylindrical layers. The layers are grouped by four into ten superlayers, with the same wire orientation and equal numbers of cells in each layer of a superlayer. Sequential layers are staggered by half a cell. This arrangement enables local segment finding and left-right ambiguity resolution within a superlayer, even if one out of four signals is missing. The stereo angles of the superlayers alternate between axial (A) and stereo (U,V) pairs, in the order AUVAUVAUVA, as shown in Figure 28. The stereo angles vary between ± 45 mrad and ± 76 mrad; they have been chosen such that the drilling patterns are identical for the two endplates. The hole pattern has a 16-fold azimuthal symmetry which is well suited to the modularity of the electronic readout and trigger system. Table 9 summarizes parameters for all superlayers.

6.3.2. Cell Design and Wires

The drift cells are hexagonal in shape, 11.9 mm by approximately 19.0 mm along the radial and azimuthal directions, respectively. The hexagonal cell configuration is desirable because approximate circular symmetry can be achieved over a large portion of the cell. The choice of aspect ratio has the benefit of decreasing the number of wires and electronic channels, while allowing a 40 layer chamber in a confined radial space. Each cell consists of one sense wire surrounded by six field wires, as shown in Figure 28. The properties of the different types of gold-coated wires that make up the drift cells are given in Table 10. The sense wires are made of tungsten-rhenium [3],

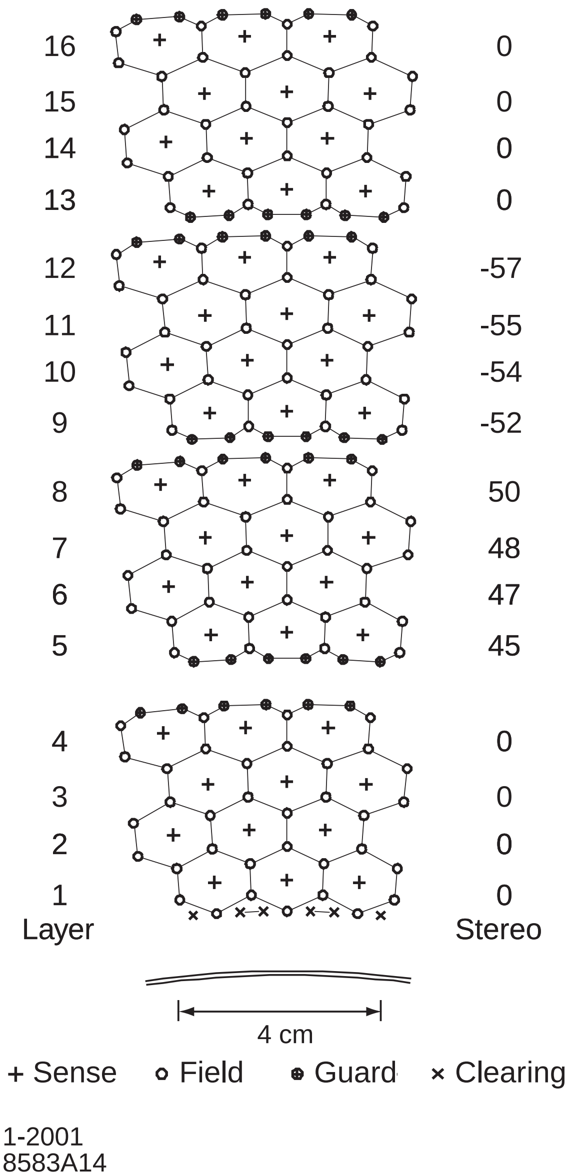

Figure 28. Schematic layout of drift cells for the four innermost superlayers. Lines have been added between field wires to aid in visualization of the cell boundaries. The numbers on the right side give the stereo angles (mrad) of sense wires in each layer. The 1 mm-thick beryllium inner wall is shown inside of the first layer.

Table 9
The DCH superlayer (SL) structure, specifying the number of cells per layer, radius of the innermost sense wire layer, the cell widths, and wire stereo angles, which vary over the four layers in a superlayer as indicated. The radii and widths are specified at the mid-length of the chamber.

SL	# of Cells	Radius (mm)	Width (mm)	Angle (mrad)
1	96	260.4	17.0-19.4	0
2	112	312.4	17.5-19.5	45-50
3	128	363.4	17.8-19.6	-(52-57)
4	144	422.7	18.4-20.0	0
5	176	476.6	16.9-18.2	56-60
6	192	526.1	17.2-18.3	-(63-57)
7	208	585.4	17.7-18.8	0
8	224	636.7	17.8-18.8	65-69
9	240	688.0	18.0-18.9	-(72-76)
10	256	747.2	18.3-19.2	0

20 μm in diameter and tensioned with a weight of 30 g. The deflection due to gravity is 200 μm at mid-length. Tungsten-rhenium has a substantially higher linear resistivity (290 Ω/m), compared to pure tungsten (160 Ω/m), but it is considerably stronger and has better surface quality. While the field wires are at ground potential, a positive high voltage is applied to the sense wires. An avalanche gain of approximately 5×10^4 is obtained at a typical operating voltage of 1960 V and a 80:20 helium:isobutane gas mixture.

Table 10
DCH wire specifications (all wires are gold plated).

Type	Material	Diameter (μm)	Voltage (V)	Tension (g)
Sense	W-Re	20	1960	30
Field	Al	120	0	155
Guard	Al	80	340	74
Clearing	Al	120	825	155

The relatively low tension on the sense wires was chosen so that the aluminium field wires have matching gravitational sag and are tensioned well below the elastic limit. A simulation of the electrostatic forces shows that the cell configuration has no instability problems. At the nominal operating voltage of 1960 V, the wires deflect by less then 60 μm.

The field wires [4] are tensioned with 155 g to match the gravitational sag of the sense wires to within 20 μm. This tension is less than one-half the tensile yield strength of the aluminium wire. For cells at the inner or outer boundary of a superlayer, two guard wires are added to improve the electrostatic performance of the cell and to match the gain of the boundary cells to those of the cells in the inner layers. At the innermost boundary of layer 1 and the outermost boundary of layer 40, two clearing wires have been added per cell to collect charges created through photon conversions in the material of the walls.

6.3.3. Drift Isochrones

The calculated isochrones and drift paths for ions in adjacent cells of layer 3 and 4 of an axial superlayer are presented in Figure 29. The isochrones are circular near the sense wires, but deviate greatly from circles near the field wires. Ions originating in the gap between superlayers are collected by cells in the edge layers after a delay of several μs. These lagging ions do not affect the drift times measurements, but they contribute to the dE/dx measurement.

6.3.4. Cross Talk

A signal on one sense wire produces oppositely-charged signals on neighboring wires due to capacitive coupling. The cross talk is largest between adjacent cells of adjacent layers, ranging from -0.5% at a superlayer boundary to -2.7% for internal layers within superlayers. For adjacent cells in the same layer, the cross talk ranges from -0.8 to -1.8%, while for cells separated by two layers it is less than 0.5%.

6.4. Electronics

6.4.1. Design Requirements and Overview

The DCH electronic system is designed to provide a measurement of the drift time and the integrated charge, as well as a bit for every hit wire to the trigger system [5]. In the 80:20 helium-isobutane gas mixture, there are on average some 22 (44) primary (total) ionization clusters produced per cm. The position of the primary ionization clusters is derived from timing of the leading edge of the amplified signal. The design goal was to achieve a position resolution of 140 μm, averaged over the cells. To reduce the time jitter in the signal arrival and at the same time maintain a good signal-to-noise ratio, the signal threshold was set at about 2.5 primary electrons. For the dE/dx measurement, a resolution of 7% was projected for a 40-layer chamber.

The small cell size and the difficult access through the DIRC strong support tube require a very high density of electronics components. As a consequence, a compact and highly modular design was chosen. The readout is installed in well shielded assemblies that are plugged into the end-

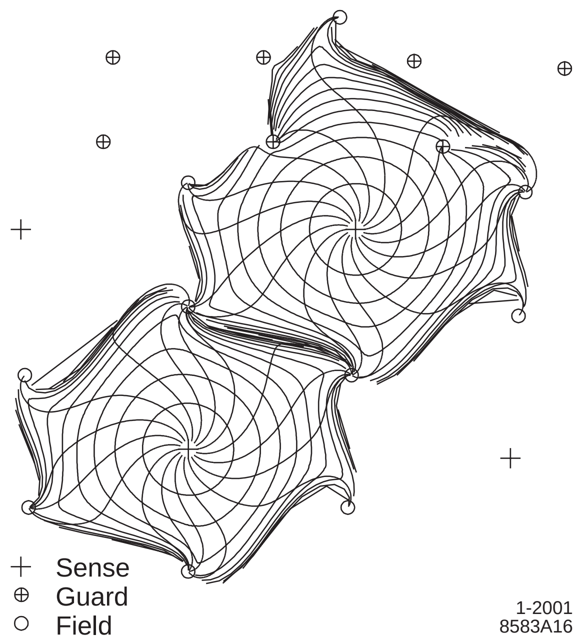

Figure 29. Drift cell isochrones, i.e., contours of equal drift times of ions in cells of layers 3 and 4 of an axial superlayer. The isochrones are spaced by 100 ns. They are circular near the sense wires, but become irregular near the field wires, and extend into the gap between superlayers.

plate and are easily removable for maintenance.

A schematic overview of the DCH electronics is presented in Figure 30 [6]. The 16-fold azimuthal symmetry of the cell pattern is reflected in the readout segmentation. The DCH amplifier and digitizer electronics are installed in 3×16 electronics front-end assemblies (FEAs) that are mounted directly onto the rear endplate. There are 16 radial bars that extend from the inner to the outer chamber walls. These bars provide mechanical support and water cooling for the FEAs. The assemblies connect to the sense wires through service boards, which route the signals and HV distribution. A readout interface board in each FEA organizes readout of the digitized data. Data I/O and trigger I/O modules multiplex serial data from the FEAs to high-speed optical fibers for transfer to the readout modules that are located in the electronics building.

6.4.2. Service Boards

Service boards provide the electrostatic potentials for signal, guard, and clearing wires, and pass signals and ground to the front-end readout electronics. A side view of a service board is shown in Figure 31. The lower board contains the high voltage buses and filtering, current limiting resistors, and blocking capacitors. Jumpers connect adjacent boards. The stored energy is minimized by 220 pF HV blocking capacitors.

The series resistors for the protection circuits carry the signals to the upper signal board which contains the protection diodes and standard output connectors. Mounting posts, anchored into the rear endplate, also serve as ground connections.

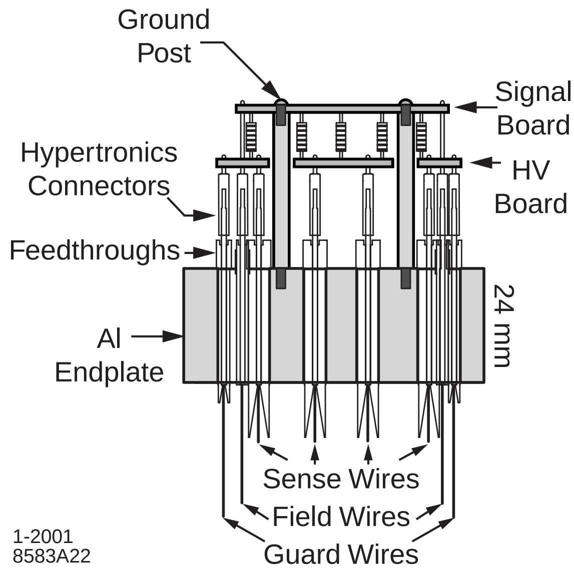

Figure 31. Side view of service boards showing two-tiered structure for HV distribution and signal collection.

6.4.3. Front-End Assemblies

The front-end assemblies (FEAs) plug into connectors on the back side of the service boards. These custom wedge-shaped crates are aluminium boxes that contain a single readout interface board (ROIB) and 2, 3, or 4 amplifier/digitizer boards (ADB) for superlayers 1–4, 5–7, and 8–10, respectively, as shown in Figure 32. The crates are mounted with good thermal contact to the water cooled radial support bars. The total heat load generated by the FEAs is 1.3 kW.

The ADBs are built from basic building blocks consisting of two 4-channel amplifier IC [7] feeding a single 8-channel digitizer custom ASICs [8]. The number of channels serviced by an ADB is 60, 48, or 45, for the inner, middle, and outer FEA modules, respectively.

The amplifier IC receives the input signal from the sense wire and produces a discriminator output signal for the drift time measurement and a shaped analog signal for the $\mathrm{d}E/\mathrm{d}x$ measurement. Both outputs are fully differential. The discriminator has gain and bandwidth control, and a voltage controlled threshold. The analog circuit has integrator and gain control.

The digitizer IC incorporates a 4-bit TDC for time measurement and a 6-bit 15 MHz FADC to measure the total deposited charge. The TDC is a phase-locked digital delay linear vernier on the sample clock of 15 MHz, which achieves a 1 ns precision for leading edge timing. The FADC design is based on a resistor-divider comparator ladder that operates in bi-linear mode to cover the

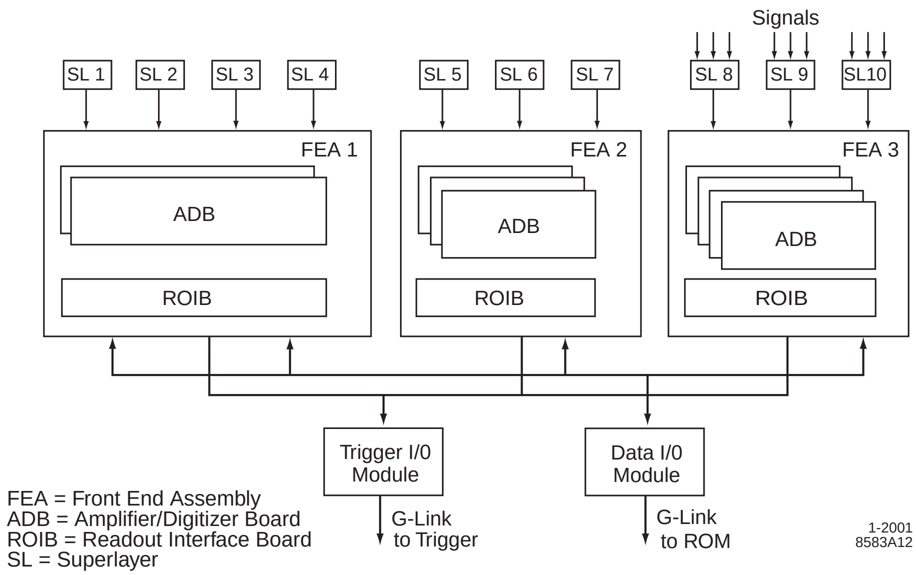

Figure 30. Block diagram for a $1/16^{th}$ wedge of the readout system, showing logical organization of the three front-end assemblies and their connections to the trigger and data I/O modules

full dynamic range. The digitized output signals are stored in a trigger latency buffer for $12.9\,\mu$s, after which a L1 Accept initiates the transfer of a $2.2\,\mu$s block of data to the readout buffer. In addition, trigger information is supplied for every channel, based either on the presence of a TDC hit during the sample period or FADC differential pulse height information, should a higher discriminator level be desirable.

The ROIB interprets FCTS commands to control the flow of data and trigger information. Data are moved to FIFOs on the ROIBs, and then to data and trigger I/O modules via 59.5 MHz serial links. A total of four such links are required per $1/16^{th}$ wedge, one for each of the outer two FEAs and two for the innermost of the FEA. Each data I/O module services all FEAs one quadrant and transmits the data to a single ROM via one optical fiber link. The trigger stream is first multiplexed onto a total of 30 serial lines per wedge for transmission to the trigger I/O module. Trigger data from two wedges of FEAs are then transmitted to the trigger system via three optical links. Thus, a total of 28 optical fibers, four for the data and 24 for the charged particle trigger, are required to transfer the DCH data to the electronics building.

6.4.4. Data Acquisition

The data stream is received and controlled by four *BABAR* standard readout modules. Drift chamber-specific feature extraction algorithms convert the raw FADC and TDC information into drift times, total charge, and a status word. The time and charge are corrected channel-by-channel for time offsets, pedestals, and gain constants. These algorithms take about $1\,\mu$s per channel, and reduce the data volume by roughly a factor of four.

6.4.5. High Voltage System

The HV bias lines on the chamber are daisy-chained together so that each superlayer requires only four power supplies, except for superlayer 1 which has eight. The high voltages are supplied to the sense, guard and clearing wires by a

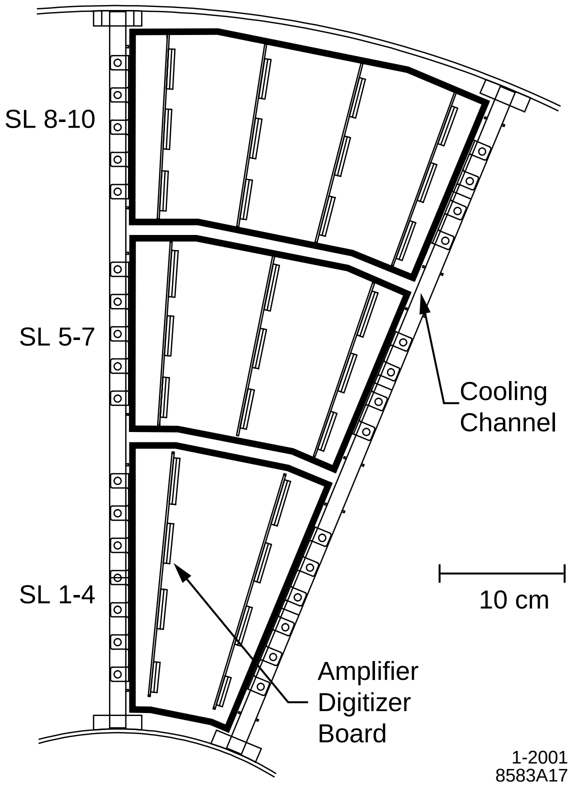

Figure 32. Layout of $1/16^{th}$ of the rear endplate, showing 3 FEA boxes between water cooled channels.

CAEN SY527 HV mainframe [9], equipped with 24-channel plug-in modules. The sense wires are supplied by 44 HV channels providing up to 40 μA of current each that can be monitored with a resolution of 0.1%.

6.5. Gas System

The gas system has been designed to provide a stable 80:20 helium-isobutane mixture at a constant over pressure of 4 mbar [10]. Gas mixing and recirculation is controlled by precise mass flow controllers; the total flow is tuned to 15 ℓ/min, of which 2.5 ℓ/min are fresh gas. During normal operation, the complete DCH gas volume is re-circulated in six hours, and one full volume of fresh gas is added every 36 hours. The pressure in the DCH is measured by two independent pressure gauges, one of which is connected to a regulator controlling the speed of the compressor. The relative pressure in the chamber is controlled to better than ± 0.05 mbar.

Oxygen is removed from the gas mixture using a palladium catalytic filter. The water content is maintained at 3500 ± 200 ppm by passing an adjustable fraction of the gas through a water bubbler. This relatively high level of water vapor is maintained to prevent electrical discharge. In addition to various sensors to monitor pressure, temperature, and flow at several points of the system, a small wire chamber with an ^{55}Fe source continuously monitors gain of the gas mixture.

6.6. Calibrations and Monitoring

6.7. Electronics Calibration

The front-end electronics are calibrated daily to determine the channel-by-channel correction constants and thresholds. Calibration pulses are produced internally and input to the preamplifier at a rate of about 160 Hz. The calibration signals are processed in the ROM to minimize the data transfer and fully exploit the available processing power. The results are stored for subsequent feature extraction. The entire online calibration procedure takes less than two minutes.

6.7.1. Environmental Monitoring

The operating conditions of the DCH are monitored in real time by a variety of sensors and read out by the detector-wide CAN bus system. These sensors monitor the flow rate, pressure, and gas mixture; the voltages and currents applied to the wires in the chamber; the voltages and currents distributed to the electronics from power supplies and regulators; instantaneous and cumulative radiation doses; temperature and humidity around the chamber electronics and in the equipment racks. Additional sensors monitor the atmosphere in and around the detector for excess isobutane, which could pose a flammability or explosive hazard in the event of a leak.

Many of the sensors are connected to hardware interlocks, which ensure that the chamber is automatically put into a safe state in response to an unsafe condition. All of these systems have performed reliably. In addition, automated software monitors raw data quality, chamber occupancies and efficiencies to sense variations in electronics performance that might indicate more subtle operational problems.

6.7.2. Operational Stability

The design of the DCH specifies a voltage of 1960 V on the sense wires to achieve the desired gain and resolution. The chamber voltage was lowered for part of the run to 1900 V out of concern for a small region of the chamber that was damaged during the commissioning phase by inadvertently applying 2 kV to the guard wires. Wires in this region were disconnected when continuous discharge was observed over extended periods of time.

6.8. Performance

The DCH was first operated with full magnetic field immediately after the installation into *BABAR*. Cosmic ray data were recorded and extensive studies of the basic cell performance were performed to develop calibration algorithms for the time-to-distance and $\mathrm{d}E/\mathrm{d}x$ measurements. These algorithms were then implemented as described below for colliding beam data. Calibrations are monitored continuously to provide feedback to the operation; some time varying parameters are updated continuously as part of OPR. For charge particle tracking the DCH and SVT information is combined; the performance of the combined tracking system is described in Chapter 9.

6.8.1. Time-to-Distance Relation

The precise relation between the measured drift time and drift distance is determined from samples of e^+e^- and $\mu^+\mu^-$ events. For each signal, the drift distance is estimated by computing the distance of closest approach between the track and the wire. To avoid bias, the fit does not use the hit on the wire under consideration. The estimated drift distances and measured drift times are averaged over all wires in a layer, but the data are accumulated separately for tracks passing on the left of a sense wire and on the right. The time-distance relation is fit to a sixth-order Chebychev polynomial. An example of such a fit is shown in Figure 33.

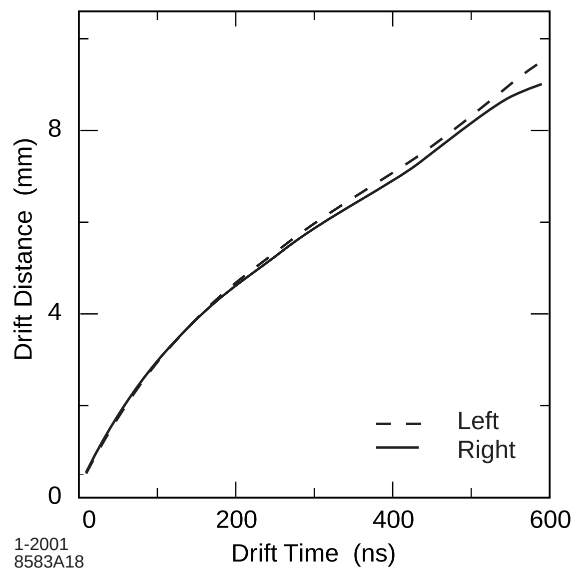

Figure 33. The drift time versus distance relation for left and right half of a cell, obtained from a fit to data averaged over all cells in layer 18 of the DCH.

An additional correction is made for tracks with varying entrance angle into the drift cell. This angle is defined relative to the radial vector from the IP to the sense wire. The correction is applied as a scale factor to the drift distance and was determined layer-by-layer from a Garfield [11] simulation. The entrance angle correction is implemented as a tenth-order Chebychev polynomial of the drift distance, with coefficients which are functions of the entrance angle.

Figure 34 shows the position resolution as a function of the drift distance, separately for the left and the right side of the sense wire. The resolution is taken from Gaussian fits to the distributions of residuals obtained from unbiased track fits. The results are based on multi-hadron events, for data averaged over all cells in layer 18.

6.8.2. Charge Measurement

The specific energy loss, $\mathrm{d}E/\mathrm{d}x$, for charged particles traversing the DCH is derived from measurement of total charge deposited in each drift

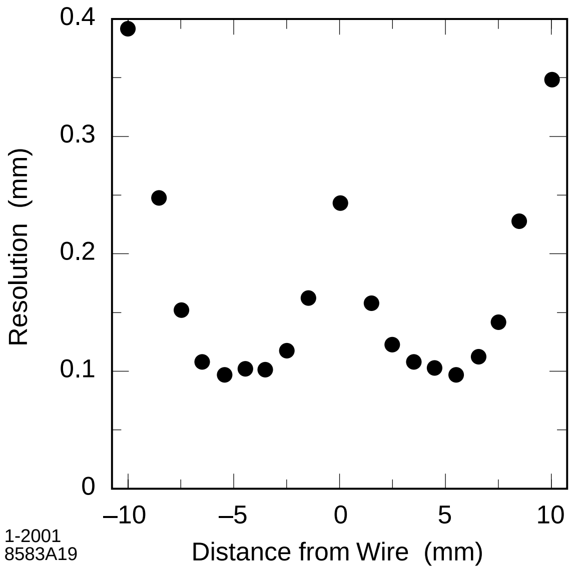

Figure 34. Position resolution as a function of the drift distance, separately for tracks on the left and right of the sense wire. The data are averaged over all cells in layer 18 of the DCH.

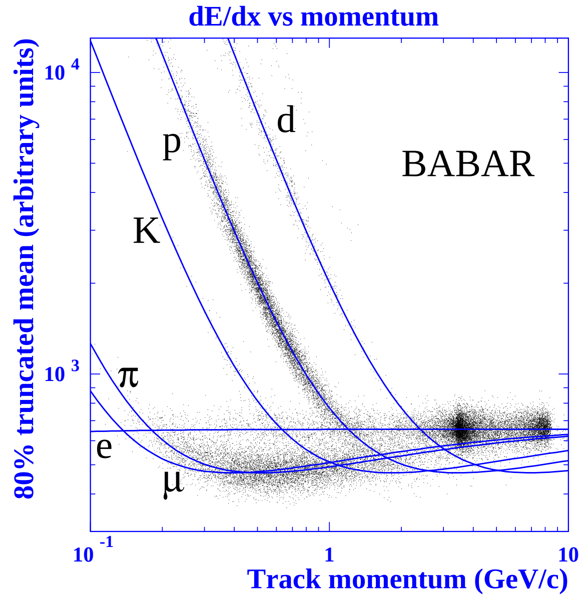

Figure 35. Measurement of $\mathrm{d}E/\mathrm{d}x$ in the DCH as a function of track momenta. The data include large samples of beam background triggers, as evident from the high rate of protons.

cell. The charge collected per signal cell is measured as part of the feature extraction algorithm in the ROM. Individual measurements are corrected for gain variations, pedestal-subtracted and integrated over a period of approximately $1.8\,\mu$s.

The truncated mean of the specific energy loss per track is computed from the lowest 80% of the individual measurements. Various corrections are applied to remove sources of bias that degrade the accuracy of the primary ionization measurement. These corrections include the following:

- changes in gas pressure and/or temperature, leading to $\pm 9\%$ variation in $\mathrm{d}E/\mathrm{d}x$, corrected by a single overall multiplicative constant;
- differences in cell geometry and charge collection ($\pm 8\%$ variation), corrected by a set of multiplicative constants for each wire;
- signal saturation due to space charge build-up ($\pm 11\%$ variation), corrected by a second-order polynomial in the dip angle, λ, of the form $1/\sqrt{\sin^2\lambda + const}$;
- non-linearities in the most probable energy loss at large dip angles ($\pm 2.5\%$ variation), corrected with a fourth-order Chebychev polynomial as a function of λ; and
- variation of cell charge collection as a function entrance angle ($\pm 2.5\%$ variation), corrected using a sixth-order Chebychev polynomial in the entrance angle.

The overall gas gain is updated continuously based on calibrations derived as part of prompt reconstruction of the colliding beam data; the remaining corrections are determined once for a given HV voltage setting and gas mixture.

Corrections applied at the single cell level can be large compared to the single-cell $\mathrm{d}E/\mathrm{d}x$ resolution, but have only a modest impact on the average resolution of the ensemble of hits. Global

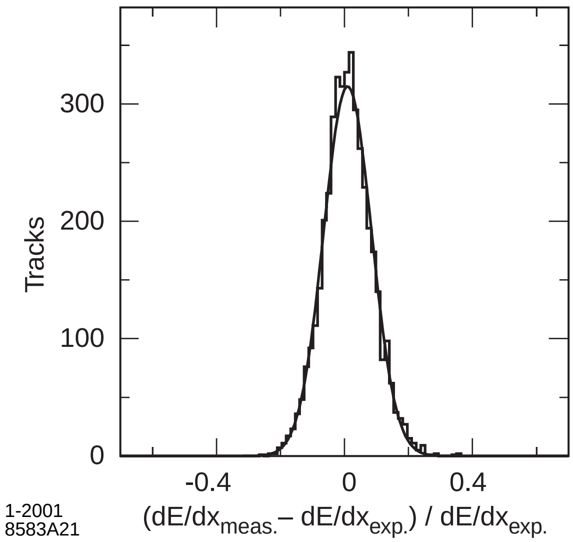

Figure 36. Difference between the measured and expected energy loss dE/dx for $e^{\pm}$ from Bhabha scattering, measured in the DCH at an operating voltage of 1900 V.

corrections applied to all hits on a track are therefore the most important for the resolution.

Figure 35 shows the distribution of the corrected dE/dx measurements as a function of track momenta. The superimposed Bethe-Bloch predictions for particles of different masses have been determined from selected control samples.

The measured dE/dx resolution for Bhabha events is shown in Figure 36. The performance achieved to date is typically 7.5%, limited by the number of samples and Landau fluctuations. This value is close to the expected resolution of 7%. Further refinements and additional corrections are being considered to improve performance.

6.9. Conclusions

The drift chamber has been performing close to design expectations from the start of operations. With the exception of a small number of wires that were damaged by an unfortunate HV incident during the commissioning phase, all cells are fully operational. The DCH performance has proven very stable over time. The design goal for the intrinsic position and dE/dx resolution have been met. Backgrounds are acceptable at present beam currents, but there is concern for rising occupancies and DAQ capacity at the high end of the planned luminosity upgrades.

REFERENCES

1. Celenex 3300-2 polyester thermoplastic with 30% silica glass fiber
2. E. Borsato et al., Nucl. Instr. and Methods A451 (2000) 414.
3. Luma Metall AB, Kalmar, Sweden.
4. California Fine Wire, Grover Beach, CA, USA.
5. A. Berenyi et al., IEEE Trans. Nucl. Sci. 46 (1999), 348; *ibid.* 46 (1999) 928.
6. J. Albert et al., IEEE Trans. Nucl. Sci. 46 (1999) 2027; A. Bouchan et al., Nucl. Instr. and Methods A409 (1998) 46; G. Sciolla et al., Nucl. Instr. and Methods A419 (1998) 310.
7. D. Dorfan et al., Nucl. Instr. and Methods A409 (1998) 310.
8. S.F. Dow et al., IEEE Trans. Nucl. Sci. 46 (1999) 785.
9. CAEN SY527 HV mainframe by CAEN, Viareggio, Italy.
10. Y. Karyotakis, D. Boutigny, LAPP Annecy, *private communication.*
11. Garfield: Simulation of Gaseous Detectors, CERN Program Library (1992).

7. Drift Chamber

8. Charged Particle Tracking

The principal purpose of the *BABAR* charged particle tracking systems, the SVT and the DCH, is the efficient detection of charged particles and the measurement of their momentum and angles with high precision. Among many applications, these high precision measurements allow for the reconstruction of exclusive B and D meson decays with high resolution and thus minimal background.

The reconstruction of multiple decay vertices of weakly decaying B and D mesons, is of prime importance to the physics goals of *BABAR*. The design goal was to achieve a mean spatial resolution for each B decay vertex along the z-axis of better than 80 μm.

Track measurements are also important for the extrapolation to the DIRC, EMC, and IFR. At lower momenta the DCH measurements are more important, while at higher momenta the SVT measurements dominate. Most critical are the angles at the DIRC, because the uncertainties in the charged particle track parameters add to the uncertainty in the measurement of the Cherenkov angle. Thus the track errors from the combined SVT and DCH measurements should be small compared to the average DIRC angle measurements, i.e., of order of 1 mrad, particularly at the highest momenta.

Charged particle tracking has been studied with large samples of cosmic ray muons, e^+e^-, $\mu^+\mu^-$, and $\tau^+\tau^-$ events, as well as multi-hadrons. At this time, these studies are far from complete and the results are still preliminary. In particular, many issues related to the intrinsic alignment of the SVT, the variation with time of the relative alignment of the SVT and the DCH, and movement of the beam position relative to *BABAR* remain under study.

8.1. Track Reconstruction

The reconstruction of charged particle tracks relies on data from both tracking system, the SVT and the DCH. Charged tracks are defined by five parameters $(d_0, \phi_0, \omega, z_0, \tan\lambda)$ and their associated error matrix. These parameters are measured at the point of closest approach to the z-axis; d_0 and z_0 are the distances of this point from the origin of the coordinate system in the $x-y$ plane and along the z-axis, respectively. The angle ϕ_0 is the azimuth of the track, λ the dip angle relative to the transverse plane, and $\omega = 1/p_t$ is its curvature. d_0 and ω are signed variables; their sign depends on the charge of the track. The track finding and the fitting procedures make use of Kalman filter algorithm [1] that takes into account the detailed distribution of material in the detector and the full map of the magnetic field.

The off-line charged particle track reconstruction builds on information available from the L3 trigger and tracking algorithm. It begins with an improvement of the event start time t_0, obtained from a fit to the parameters d_0, ϕ_0, and t_0 based on the four-hit track segments in the DCH superlayers. The next step is to select tracks in the DCH by performing helix fits to the hits found by the L3 track finding algorithm. A search for additional hits in the DCH that may belong on these tracks is performed, while t_0 is further improved by using only hits associated with tracks. Two more sophisticated tracking procedures are applied which are designed to find tracks that either pass through fewer than the full ten DCH superlayers or do not originate from the interaction point. These algorithms use primarily track segments that have not already been assigned to other tracks, and thus benefit from a progressively cleaner tracking environment with a constantly improving t_0. At the end of this process, tracks are fit again using a Kalman filter fit.

The resulting tracks are then extrapolated into the SVT, and SVT track segments are added provided they are consistent with the expected error in the extrapolation through the intervening material and inhomogeneous magnetic field. Among the possible SVT segments, those with the smallest residuals and the largest number of SVT layers are retained and a Kalman fit is performed to the full set of DCH and SVT hits.

Any remaining SVT hits are then passed to two complementary stand-alone track finding algorithms. The first reconstructs tracks starting with triplets of space points (matched ϕ and z hits) in layers 1, 3 and 5 of the SVT, and adding consistent space points from the other layers. A minimum of four space points are required to form a good track. This algorithm is efficient over a wide range of impact parameters and z_0 values. The second algorithm starts with circle trajectories from ϕ hits and then adds z hits to form helices. This algorithm is less sensitive to large combinatorics and to missing z information for some of the SVT modules.

Finally, an attempt is made to combine tracks that are only found by one of the two tracking systems and thus recover pattern recognition problems caused by multiple scattering at the support tube. The relative alignment of the two tracking devices and the internal alignment of the SVT modules are discussed in Chapter 5.

8.2. Tracking Efficiency

The efficiency for reconstructing tracks in the DCH has been measured as a function of transverse momentum, polar and azimuthal angles in multi-track events, as well as the charged track multiplicity. These measurements rely on specific final states and exploit the fact that the track reconstruction can be performed independently in the two tracking devices, the SVT and the DCH.

A comparison of the tracking efficiency in data and Monte Carlo simulation has been performed for a sample of $\tau^+\tau^-$ pairs with 1-prong plus 3-prong decay topology. The events are selected without any requirements on the third track of the 3-prong decay, and the fraction of events for which all tracks are successfully reconstructed provides a measurement of the overall reconstruction efficiency for tracks in the range $0.2 < p_{\rm t} < 3.0\,{\rm GeV}/c$. Initial results indicated that the measured tracking efficiency was lower than expected from the simulations. Since then, a number of improvements in the track finding software have been made, and the DCH operating voltage was raised by 60 V to the design value of 1960 V. These measures have restored the efficiency for data close to the expected level.

The absolute DCH tracking efficiency is determined as the ratio of reconstructed DCH tracks to the tracks detected in the SVT with the requirement that they fall within the acceptance of the DCH. Corrections are made for fake tracks in the SVT and DCH as well as for a bias in the

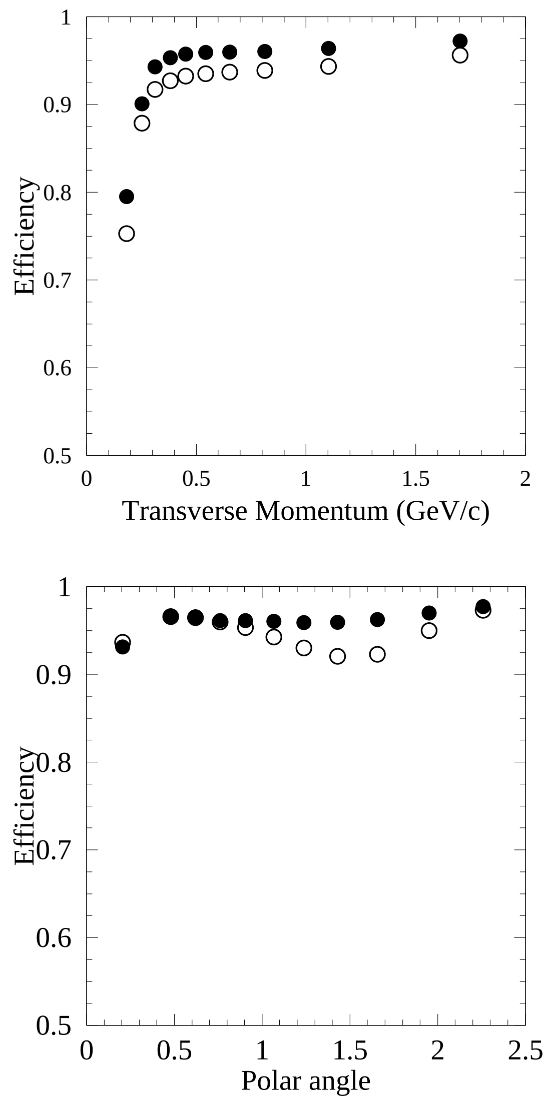

Figure 37. DCH track reconstruction efficiency at operating voltages of 1900 V and 1960 V, as a function of a) transverse momentum, and b) polar angle. The efficiency is measured in multi-hadron events as a fraction all tracks detected in the SVT for which the DCH portion is also reconstructed.

track fnding algoriths. Such studies have been performed for different samples of multi-hadron events. Figure 37 shows the result of one such study for the two voltage settings. The results of the different samples agree well within the estimated error. For 1900 V, the data show a reduction in efficiency by about 5% at close to normal incidence, indicating that the cells are not fully efficient at this voltage. At the design voltage of 1960 V, the efficiency averages $(98\pm1)\%$ per track of more than $200\,\mathrm{MeV}/c$ and polar angle $\theta > 0.5$. The errors of the measurements are dominated by the uncertainty in the contamination of SVT tracks.

As a cross-check and measure of systematic uncertainties in the above method, an analysis is performed using a sample of $D^0 \to K^-\pi^+\pi^+\pi^-$ decays, where the D^0 is reconstructed from the decay $D^{*+} \to D^0\pi^+$. All of the tracks from the D^0 decay are required to have SVT hits and the fraction of D^0 for which all four tracks have DCH tracks is taken as the relative DCH tracking efficiency. The results agree with the more inclusive method to within the estimated error of 1%.

The stand-alone SVT tracking algorithms have a high efficiency for tracks with low transverse momentum, which do not reach the drift chamber, as illustrated in Figure 38. This feature is very important for the detection of D^* decays which is used for a variety of inclusive studies of B and D mesons. To study the efficiency, the decays $D^{*+} \to D^0\pi^+$ are selected by reconstructing events of the type $\bar{B} \to D^{*+}X$ followed by $D^{*+} \to D^0\pi^+ \to K^-\pi^+\pi^+$, where the pion from the D^* decay has a soft momentum spectrum. The majority of these low momentum tracks do not reach the inner layers of the DCH, their momentum resolution is in most cases limited by multiple scattering, but the production angle can be determined from the hits in innermost layers of the SVT. Figure 38 shows the momentum of the slow pions and the mass difference $\Delta M = M(K^-\pi^+\pi^+) - M(K^-\pi^+)$, for the total sample and the subsample of events in which the slow pion has been reconstructed in both the SVT and the DCH. To derive an estimate of the tracking efficiency for these low momentum tracks, a detailed Monte Carlo simulation was performed. In particular, the spectrum of the pions was derived from simulation of the inclusive D^* production in $B\overline{B}$ events, and the events were selected in the same way as for the data. Contributions for non-$B\overline{B}$ events were subtracted. A comparison of the detected slow pions spectra with the Monte Carlo prediction is presented in Figure 39. Also shown is the detection efficiency derived from the Monte Carlo simulation, which is taken as an estimate, based on the good agreement between the spectra. While there is some remaining concern about adequate details in the simulation of the SVT, one can conclude that the SVT significantly extends the capability of the charged particle detection of tracks down to transverse momenta of $\sim$50 MeV/c.

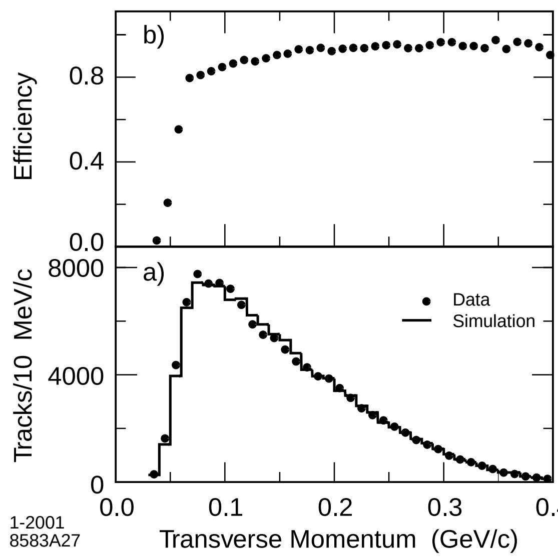

Figure 39. Monte Carlo studies of low momentum tracking in the SVT a) comparison of data (non-resonant data have been subtracted) and simulation of the transverse momentum spectrum of pions from $D^{*+} \to D^0\pi^+$ in $B\overline{B}$ events, and b) efficiency for pion detection derived from simulated events.

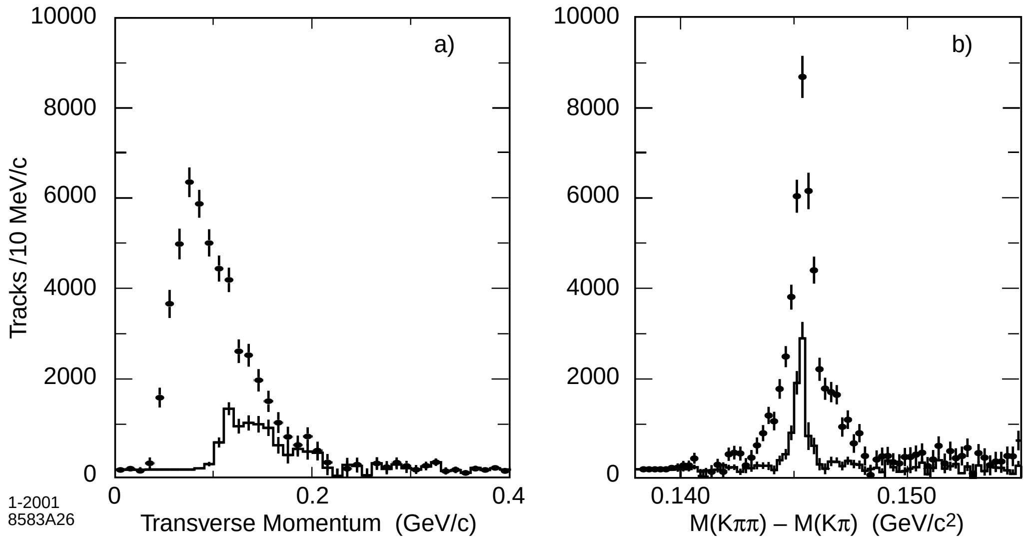

Figure 38. Reconstruction of low momentum tracks in the SVT: a) Transverse momentum spectrum of pions from $D^{*+} \to D^0\pi^+$ in $B\overline{B}$ events, and b) the mass difference, $\Delta M = M(K^-\pi^+\pi^+) - M(K^-\pi^+)$, both for all detected events and for events in which the slow pion track is reconstructed both in the SVT and DCH. Non-resonant data have been subtracted.

8.3. Track Parameter Resolutions

The resolution in the five track parameters is monitored in OPR using e^+e^- and $\mu^+\mu^-$ pair events. It is further investigated off-line for tracks in multi-hadron events and cosmic ray muons.

Cosmic rays that are recorded during normal data-taking offer a simple way of measuring the resolution of track resolution. The upper and lower halves of the muon tracks traversing the DCH are fit as two separate tracks, and the resolution measurements are taken from the difference of the measured parameters for the two track halves. To assure that the tracks pass close to the beam interaction point, cuts are applied on the d_0, z_0, and $\tan\lambda$. The results of this comparison for the impact parameters and the angles are shown in Figure 40 for tracks with momenta above p_t of 3 GeV/c. The distributions for the impact parameters and the angles are symmetric; the non-Gaussian tails are small. The z_0 and $\tan\lambda$ distributions show a clear off-set, attributed to residual problems with the internal alignment of the SVT. Based on the full width at half maximum of these distributions the resolutions for single tracks are

$$\sigma_{d_0} = 23\mu m \qquad \sigma_{z_0} = 29\mu m$$
$$\sigma_{\phi_0} = 0.43\,\mathrm{mrad} \qquad \sigma_{\tan\lambda} = 0.53 \cdot 10^{-3}.$$

The dependence of the resolution in the impact parameter, d_0 and z_0, on the transverse momentum p_t is presented in Figure 41. The measurement is based on tracks in multi-hadron events. The resolution is determined as the width of the distribution of the difference between the measured impact parameters, d_0 and z_0, and the coordinates of the vertex reconstructed from the remaining tracks in the event. These distributions peak at zero, but have a tail for positive values due to the effect of weak decays. Consequently, only the negative part of the distributions reflects the measurement error and is used in the Gaussian fit. Event shape cuts and a cut on the χ^2 of the vertex fit are applied to further reduce the effect of weak decay on this measurement. The contribution from the vertex errors are removed from the measured resolutions in quadrature. The d_0

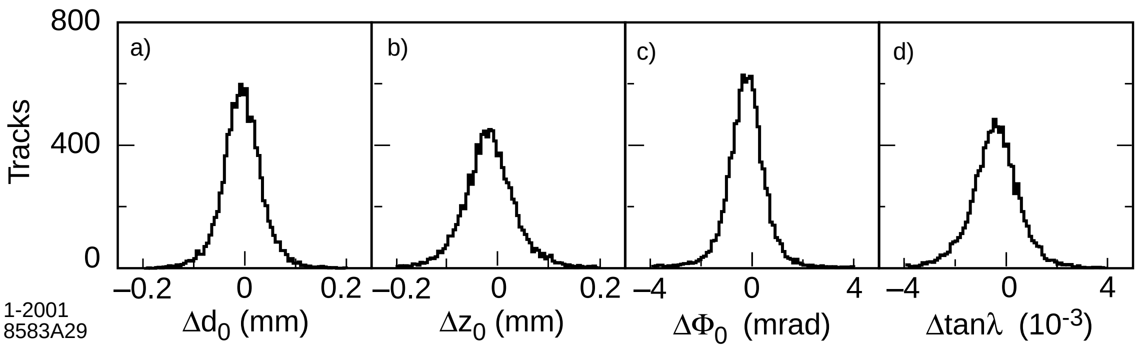

Figure 40. Resolution for track parameters of cosmic ray muons reconstructed in the SVT and DCH. The resolution is measured as the difference between the fitted track parameters of the two halves of cosmic ray tracks; a) d_0, b) z_0, c) ϕ_0, and d) $\tan\lambda$.

and z_0 resolutions so measured are about 25 μm and 40 μm respectively at $p_t = 3\,\mathrm{GeV}/c$ and above. These values agree well with expectations and Monte Carlo simulation, and are also in reasonable agreement with the results obtained from cosmic rays.

While the impact parameter and angle measurements near the IP are dominated by the SVT measurements, the DCH contributes primarily to the p_t measurement. Figure 42 shows the resolution in the transverse momentum derived from cosmic muons. The data are well represented by a linear, rather than the usual quadratic function,

$$\frac{\sigma_{p_t}}{p_t} = (0.13 \pm 0.01)\% \cdot p_t + (0.45 \pm 0.03)\%,$$

where the transverse momentum p_t is measured in GeV/c. These values for the resolution parameters are very close to the initial estimates and are well reproduced by Monte Carlo simulations. More sophisticated treatment of the DCH time-to-distance relations and overall resolution function are presently under study.

8.4. Conclusions

The two tracking devices, the SVT and DCH, have been performing close to design expectations from the start of operations. Studies of track resolution at lower momenta and as a function of polar and azimuthal angles are still under way. Likewise, the position and angular resolution at the entrance to the DIRC or EMC are still being studied. Such measurements are very sensitive to internal alignment of the SVT and relative place-

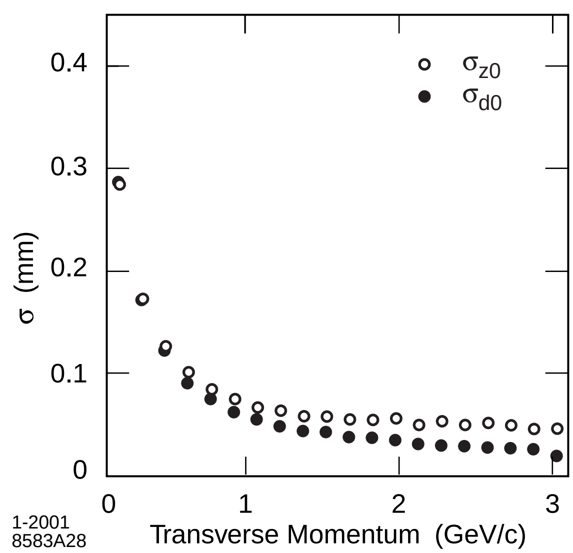

Figure 41. Impact parameter resolution for tracks in multi-hadron events as a function of the transverse momentum, in the $x - y$ plane and for the projection on the z-axis. The data are corrected for the impact of particle decays and vertexing errors.

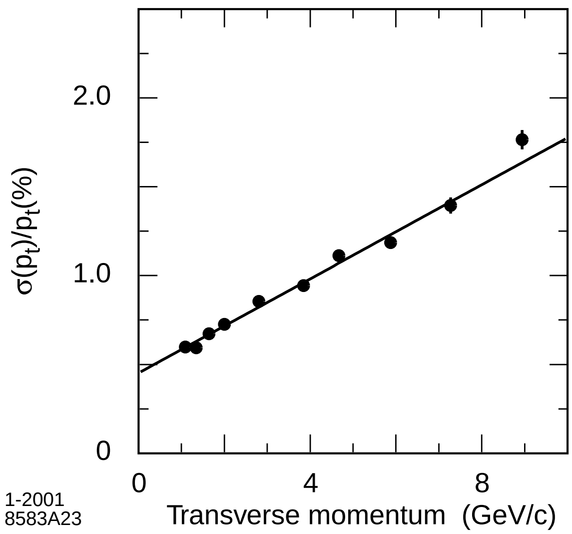

Figure 42. Resolution in the transverse momentum $p_{\rm t}$ determined from cosmic ray muons traversing the DCH and SVT.

ment of the SVT and the DCH. A better understanding will not only reduce the mass resolution for the reconstruction of exclusive states, it will also be particularly important for improvement of the performance of the DIRC.

REFERENCES

1. Kalman Filter reference needed here

9. Tracking

10. DIRC

10.1. Purpose and Design Requirements

The study of CP-violation using hadronic final states of the $B\overline{B}$ meson system requires the ability to tag the flavor of one of the B mesons via the cascade decay $b \to c \to s$, while fully reconstructing the final state of the other over a large region of solid angle and momentum. The momenta of the kaons used for flavor tagging extend up to about 2 GeV/c, with most below 1 GeV/c. On the other hand, pions from the rare two-body decays $B^0 \to \pi^+\pi^-(K^-\pi^+)$ must be well-separated from kaons, and have momenta between 1.5 and 4.5 GeV/c with a strong momentum-polar angle correlation between the tracks (higher momenta occur at the more forward angles because of the c.m. system boost) [1].

The particle identification (PID) system inside the calorimeter volume should be thin and uniform in radiation lengths (to minimize degradation of the calorimeter energy resolution) and thin in the radial dimension to reduce the volume, hence, the cost of the calorimeter. Finally, for high-luminosity running conditions, the PID system must have fast signal response, and be able to tolerate high backgrounds.

The PID system being used in *BABAR* is a new kind of ring imaging detector called the DIRC [2] (the acronym DIRC stands for detection of internally reflected Cherenkov light). It is expected to be able to provide π/K separation of $\sim 4\sigma$ or greater, for all tracks from B meson decays with momentum greater than 600 MeV/c. Particle identification below 700 MeV/c is provided by dE/dx measurements in the DCH and SVT.

The DIRC is a ring imaging Cherenkov detector, based on the principle that the magnitudes of angles are maintained upon reflection from a flat surface. Figure 43 shows a schematic of the DIRC geometry that illustrates the principles of light production, transport, and imaging. The radiator material of the DIRC is synthetic fused silica in the form of long, thin bars with rectangular cross section. These bars serve both as radiators and as light pipes for the portion of the light trapped in the radiator by total internal reflection. Fused

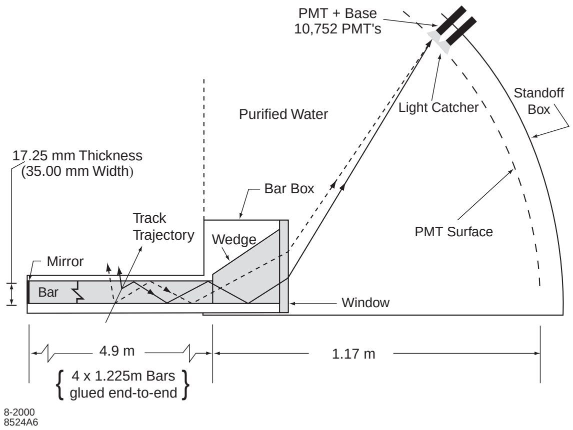

Figure 43. Schematics of the DIRC fused silica radiator bar and imaging region. Not shown is a 6 mrad angle on the bottom surface of the wedge (see text).

silica (Spectrosil [3]) is chosen because of its resistance to ionizing radiation, its large index of refraction, low chromatic dispersion, long attenuation length, and because it allows an excellent optical finish on the surfaces of the bars [4].

The variable θ_c is used to designate the Cherenkov angle, ϕ_c denotes the azimuthal angle of the Cherenkov photon around the track direction, and n represents the mean index of refraction of fused silica ($n = 1.473$), with the familiar relation $\cos\theta_c = 1/n\beta$.

For particles with $\beta \approx 1$, some photons will always lie within the total internal reflection limit, and will be transported to either one or both ends of the bar, depending on the particle incident angle. To avoid having to instrument both bar ends with photon detectors, a mirror is placed at the forward end, perpendicular to the bar axis, to reflect incident photons to the backward (instrumented) bar end.

Once photons arrive at the instrumented end, most of them emerge into a water-filled expansion region, called the standoff box. A fused silica wedge at the exit of the bar reflects photons at large angles and thereby reduces the size of the required detection surface. The photons are detected by an array of densely packed photomultiplier tubes (PMTs), each surrounded by reflecting "light catcher" cones [5] to capture light which would otherwise miss the PMT active area. The PMTs are placed about 1.2 m from the bar end. The expected Cherenkov light pattern at this surface is essentially a conic section, whose cone opening-angle is the Cherenkov production angle modified by refraction at the exit from the fused silica window.

The DIRC is intrinsically a three-dimensional imaging device. Photons are focused onto the phototube detection surface via a "pin-hole" defined by the exit aperture of the bar, so that the photon propagation angles can be measured in two-dimensional space (α_x, α_y). The time taken for the photon to travel down the bar is also related to the photon propagation angle (α_z) with respect to the bar axis. As the track position and angles are known from the tracking system, these three α angles can be used to (over)-determine the two Cherenkov angles (θ_c, ϕ_c). Imaging in the DIRC occurs in all three of these dimensions, by recording the time at which a given PMT is hit. This over-constraint on the angles is particularly useful in dealing with ambiguities (see below) and high background rates.

10.2. Mechanical Design/Physical Description

The DIRC bars are arranged in a 12-sided polygonal barrel. Because of the beam energy asymmetry, particles are produced preferentially forward in the laboratory. To minimize interference with other detector subsystems in the forward region, the DIRC photon detector is placed at the backward end.

The principal components of the DIRC are shown schematically in Figs. 44,45. The bars are placed into 12 hermetically sealed containers, called bar boxes, made of very thin aluminum-hexcel panels. Dry nitrogen gas flows through each box, and is monitored for humidity to ensure that the bar box/water interface remains tightly sealed. Each bar box, shown in Figure 45, in turn contains 12 long bars, for a total of 144 long bars. Within a bar box the twelve bars are optically isolated by a $\sim 150\,\mu$m air gap between neighboring bars, enforced by custom shims made from aluminum foil.

The bars are 17 mm thick, 35 mm wide, and 4.9 m long. Each long bar is assembled from four 1.225 m "short" bars that are glued end-to-end; that length being the longest high-quality bar currently obtainable from industry [4].

The bars are supported at 600 mm intervals on small nylon buttons for optical isolation from the bar box. Each long bar has a fused silica wedge glued to it at the readout end. The wedge is made of the same material as the bar, 91 mm long with very nearly the same width as the bars (33 mm) and a trapezoidal profile (27 mm high at bar end, and 79 mm at the light exit end). The bottom of the wedge has a slight ($\sim$ 6 mrad) upward slope to minimize the displacement of the downward reflected image due to the finite bar thickness. The wedges are glued to a 10 mm thick fused silica window, which provides the interface and seal to the purified water in the standoff box.

The mechanical support of the DIRC, shown in Figure 44, is cantilevered from the iron of the instrumented flux return (IFR). The strong support tube (SST) is a steel cylinder located inside the end-doors of the IFR and provides the basic support for the entire DIRC. It, in turn, is supported by an iron support gusset that fixes the SST to the Barrel magnet iron. It also minimizes the magnetic flux gap caused by the DIRC bars extending through the IFR, and supports the axial load of the inner magnetic plug surrounding the beam in this region.

The bar boxes are supported in the active region by an aluminum tube, the central support tube (CST), attached to the SST via an aluminum transition flange. The CST is a thin, double-walled, cylindrical shell, using aircraft-type construction with stressed aluminum skins and bulkheads having riveted or glued joints. The CST also provides the support for the drift chamber.

The standoff box is made of stainless steel, consisting of a cone, cylinder, and 12 sectors of PMTs. It contains about 6,000 liters of purified water. Water is used to fill this region because it is inexpensive, has an index of refraction ($n \sim 1.346$) reasonably close to that of fused silica, thus minimizing the total internal reflection at their interface, and its chromaticity index is a close match to that of fused silica, effectively eliminating dispersion at the silica-water interface. The iron gusset supports the standoff box. An iron shield, supplemented by a bucking coil,

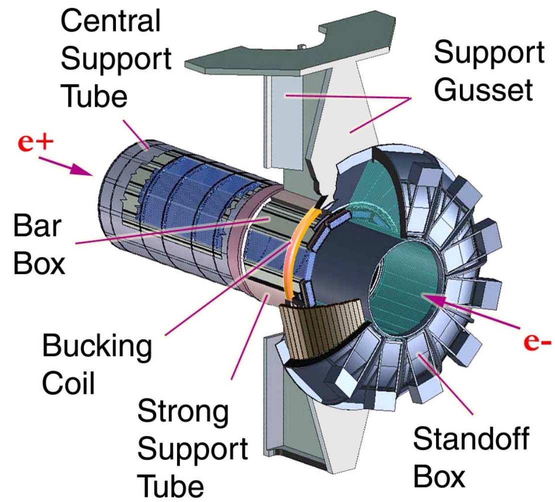

Figure 44. Exploded view of the DIRC mechanical support structure.

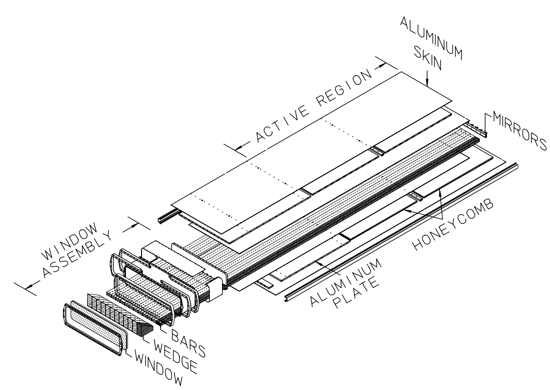

Figure 45. Schematics of the bar box assembly.

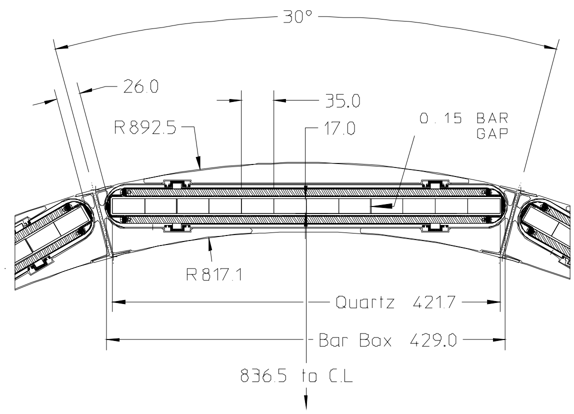

Figure 47. 1/6 cross-section view of the nominal DIRC system geometry. All dimensions are given in mm.

surrounds the standoff box to reduce the field in the PMT region to below 1 Gauss [7].

The PMTs at the rear of the standoff box lie on a surface that is approximately toroidal. Each of the 12 PMT sectors contains 896 PMTs (ETL model 9125 [8,9]) with 29-mm diameter, closely packed, inside the water volume. Each PMT is mounted from the inside of the standoff box and is connected via a feed-through to a base mounted outside. A hexagonal "light catcher" cone is mounted in front of the photocathode of each PMT, which results in an effective active surface area light collection fraction of about 90%. The geometry of the DIRC is shown in an elevation view and a cross-section in Figures 46 and 47.

The DIRC occupies 80 mm of radial space in the central detector volume including supports and construction tolerances, with a total radiation length thickness of about 19% at normal incidence. The radiator bars subtend about 94% of the azimuthal angle and 87% of the center-of-mass polar angle cosine.

The distance from the end of the bar to the PMTs is 1.174 m, which together with the size of the bars and PMTs, gives a geometric contribution to the single photon Cherenkov angle resolution of ~ 7 mrad. This is a bit larger than the resolution contribution from the photon production (dominated by a ~ 5.4 mrad chromatic term) and transmission dispersions. The overall single photon resolution expected is about 9 mrad.

10.2.1. Cherenkov Photon Detection Efficiency

Figure 48 shows the contribution of various optical and electronic components of the DIRC to the Cherenkov photon detection efficiency (number of detected photoelectrons per Cherenkov photon) as a function of wavelength. The data pertain to a particle entering the center of the bar at 90°. A typical design goal for the photon transport in the bar was that no single component specification should be responsible for more than 10–20% loss of detection efficiency. To satisfy this requirement implied an extremely high internal reflection coefficient of the bar surfaces (greater than 0.9992 per bounce), so that about 80% of the light is maintained after multiple bounces down the bars: 365 bounces in the example of Figure 48. The ultraviolet cut-off is ~ 300 nm, determined by the epoxy (Epotek 301-2 [10]) used to glue the fused silica bars together. The dominant contributor to the overall detection efficiency is the quantum efficiency of the photomultiplier tube, taken from the manufacturer's data. On the basis of these data, and taking into account wavelength independent factors such as the PMT packing fraction and the geometrical efficiency for trapping Cherenkov photons in the fused silica bars via total internal reflection, the number of expected photoelectrons (N_{pe}) is ~ 28 for a $\beta = 1$ particle entering normal to the surface at the center of a bar.

10.2.2. DIRC Water System

The DIRC water system is designed to maintain good transparency at wavelengths as small as 300 nm. The only sure way to do this is to use ultra-pure, de-ionized water, close to the theoretical limit of 18 MΩ/cm resistivity. In addition, the water must be de-gassed and the entire system kept free of bacteria. To maintain the necessary level of water quality, all components are made of stainless steel or of polyvinylidene fluoride.

The system contains an input line with six me-

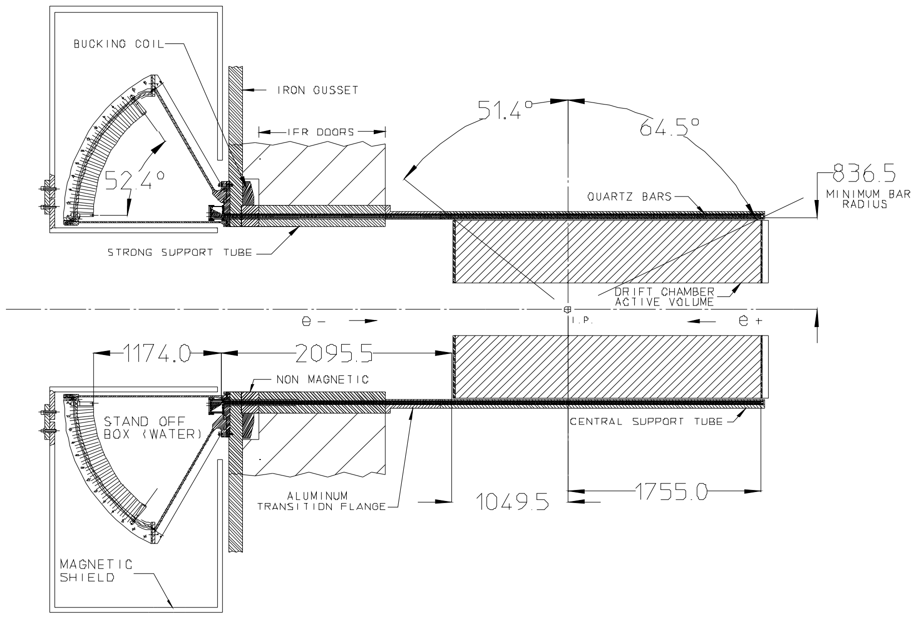

Figure 46. Elevation view of the nominal DIRC system geometry. For clarity, the end plug is not shown. All dimensions are given in mm.

chanical filters (three 10 μm, two 5 μm, and one 1 μm), a reverse osmosis unit, de-ionization beds, a Teflon microtube de-gasser and various pumps and valves. To prevent bacteria growth, it is equipped with a UV lamp (254 nm wavelength) and filters (two 1 μm, two 0.2 μm, and charcoal filters). Sampling ports are provided to check the water quality and to monitor resistivity, temperature, and flow. A gravity feed return system prevents overpressure. The water volume can be recirculated up to four times a day.

The operating experience with the water system so far has been very good. The water transparency is routinely measured using lasers of three different wavelengths. The transmission is better than 92% per meter at 266 nm and exceeds 98% per meter at 325 nm and 442 nm.

Potential leaks from the water seals between the bar boxes and the standoff box are detected by a water leak detection system of 20 custom water sensors in and about the bar box slots, two commercial ultrasonic flow sensors to monitor water flow in two (normally dry) drain lines in addition to the twelve humidity sensors on a nitrogen gas output line from each bar box (see below). Should water be detected, a valve in a 100 mm diameter drain line is opened, and all the water is drained in about 12 minutes.

10.2.3. DIRC Gas System

Nitrogen gas from boil-off is used to prevent humidity from condensing on the bars and to detect water leaks. The gas flows through each bar box at the rate of $100 - 200$ cm^3/min, and is monitored for humidity to ensure that the seal around the bar box/water interface remains tight. The gas is filtered through a molecular sieve and three mechanical filters to remove particulates (7 μm, 0.5 μm, and 0.01 μm). Dew points of the gas returned from the bar boxes are typically -40°C. About one third of the input nitrogen gas leaks from the bar boxes and keeps the bar box slots in the mechanical support structure free of condensation.

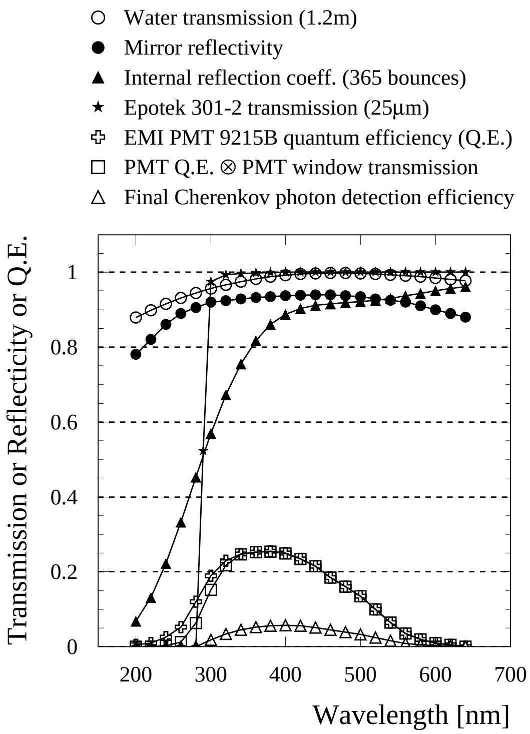

Figure 48. Transmission, reflectivity and quantum efficiency for various components of the DIRC as a function of wavelength for a $\beta = 1$ particle at normal incidence to the center of a bar [11].

10.3. Electronics

10.3.1. DIRC PMT Electronics

The DIRC PMT base system contains a single printed circuit board, equipped with surface mounted components. The operating high voltage (HV) of the PMTs is typically 1.1 kV, with a range between 900 and 1.3 kV.

Groups of 16 tubes are selected for uniformity of gain to allow their operation at a common HV provided from a single distribution board.

The HV is provided by a CAEN SY-527 high voltage distribution system. Each of the 12 sectors receives HV through 56 high voltage channels, distributed through a single cable bundle. Each voltage can be set between 0 and 1.6 kV. A total of 672 high voltage channels are needed for the entire DIRC system.

10.3.2. DIRC Front-End Electronics

The DIRC front-end electronics (FEE) is designed to measure the arrival time of each Cherenkov photon detected by the phototube array [12] to an accuracy that is limited by the intrinsic 1.5 ns transit time spread of the PMTs. In addition, the electronics also allow the photoelectron charge spectra to be measured in calibration mode to ensure that the PMTs operate on the HV plateau. However, because the ADC information is not needed to reconstruct events, 64 PMT are multiplexed onto a single ADC for monitoring and calibration. The design contains a pipeline to deal with the level one trigger latency of 12 μs, and can handle random background rates of up to 200 kHz/PMT with zero deadtime.

The DIRC FEE is mounted on the outside of the standoff box and highly integrated in order to minimize cable lengths and to retain the required single photo-electron sensitivity. Each of the 168 DIRC front-end boards (DFB) processes 64 PMT inputs, housing eight custom analog chips along with their associated level translators, four custom TDC chips, one 8-bit flash ADC (FADC), two digitally controlled calibration generators, multi-event buffers and test hardware.

The PMT signals are amplified, and pulse shaped by an eight channel analog chip. A digital pulse timed with the peak of the input pulse is output by a zero-crossing discriminator, as well as a pulse shaped by a CR-RC filter with 80 ns peaking time. The 80 ns shaping time was chosen to allow the ADC multiplexing mechanism time to take place. The multiplexer selects the analog channel to be digitized by the FADC for calibration.

The TDC chip [13] is a 16-channel TDC with 0.5 ns binning, input buffering and selective readout of the data in time with the trigger. To cope with the L1 maximum trigger latency of 12 μs and jitter of 1 μs, the selective readout process extracts data in time with the trigger within a programmable time window. The acceptance window width (latency) is programmable between 64 ns and 2 μs (16 μs) and is typically set at 600 ns. The six read-out modules (ROMs) are connected by 1.2 Gbit/s optical fibers to twelve DIRC crate controllers (DCCs) that form the interface to the DIRC VME front-end crates.

10.3.3. DAQ Feature Extraction

Raw data from the DFBs are processed in the ROMs by a feature extraction algorithm before being transmitted to the segment and event builder. This software algorithm reduces the data volume by roughly 50% under typical background conditions. DFB data that contain errors are flagged and discarded. The only data errors seen to date have been traced to damaged DFBs. Because the dataflow system can reliably transmit at most 32 kBytes/crate, the feature extraction must sometimes truncate data to limit the event size. Hit data are replaced with a per-DFB occupancy summary when a ROM's hit occupancy exceeds 56%, which occurs in roughly one in 10^4 events. An appropriate flag is inserted into the feature extraction output whenever truncation or deletion occurs. Errors, truncation, and feature extraction performance are continuously monitored online.

10.3.4. DIRC Calibration

The DIRC uses two independent approaches for a calibration of the *a priori* unknown PMT time response and the delays introduced by the front-end electronics chain and the fast control system. The first is a conventional pulser calibration, and the second uses reconstructed tracks from collision data.

The pulser calibration is performed online using a light pulser system, which generates precisely timed 1 ns duration pulses of blue LED light. This light is transmitted through approximately 47 m long optical fibers to diffusers mounted on the standoff box wall opposite the PMT detector surface of each of the 12 DIRC sectors. The pulser produces roughly 10% photoelectron occupancy nearly uniformly throughout the standoff box. The LEDs are triggered by the global fast control calibration strobe command sent to the DCCs. The DCC triggers an individual LED for each sector upon receipt of calibration strobe. Pulses in adjacent sectors are staggered by 50 ns to prevent light crosstalk between sectors. The pulser is run at roughly 2 kHz for the time delay calibration. Histograms of TDC times for each PMT are accumulated in parallel in the ROMs, and then fit to an asymmetric peaked function. About 65,000 light pulses are used to determine the mean time delay of each of the PMTs in the standoff box to a statistical accuracy of better than 0.1 ns. The LED pulser is also used to monitor the phototube gains using the ADC readout. As with the TDC calibration, histograms and fits of the ADC spectrum are accumulated and fit in the ROM. A calibration run including both TDC and ADC information for all PMTs completes in a few minutes, and is run once per day. Daily calibrations not only verify the time delays, but allow defective hardware to be fixed quickly.

The data stream calibration uses reconstructed tracks from the collision data. For calibration of the global time delay, the values of uncalibrated differences of observed and expected arrival times, δt_γ, are collected during the online prompt reconstruction processing using all DIRC channels. To calculate the individual channel calibrations, δt_γ values for each DIRC channel are accumulated until sufficient statistics of about 100,000 tracks are achieved. The collected distribution for each channel is fitted to extract the global time offset calibration.

The data stream and online pulser calibrations of the electronic delays and the PMT time response and gain yield fully consistent results, although the data stream results in 15% better timing resolution than the pulser calibration. The time delay values per channel are typically stable to an *rms* of less than 0.1 ns over more than one year of daily calibrations.

10.3.5. DIRC Environmental Monitoring System

The DIRC environmental monitoring system is divided into three parts, corresponding to three separate tasks. The first deals with the control and monitoring of the high voltage system for the photomultipliers. The second is devoted to monitoring low voltages related to the front-end electronics. The third controls a variety of other detector parameter settings. An interlock system, based on a standard VME module (SIAM), is provided. For the purposes of the DIRC, three dedicated VME CPUs are running the application code. The communication between the HV mainframes and the monitoring crate is achieved by a CAENET controller (V288). The high voltage monitor task controls the step sizes for ramping the HV up or down as well as the communication of alarm conditions, and the values and limits for the HV and current of each channel.

The purpose of FEE monitoring is to control and monitor parameters related to the front-end electronics. For each DIRC sector, a custom multi-purpose board, the DCC, equipped with a micro-controller [14] incorporating the appropriate communication protocol (CANbus) is situated in the same crate as the DFB. All monitoring and control tasks are implemented on this card. The parameters monitored are the low voltages for the DFBs and DCCs, the status of the optical link (Finisar), the temperature on supply boards and the VME crate status.

The third part of the monitoring system is based on a custom ADC VME board (VSAM) used to monitor various type of sensors: magnetic field mappers, an ensemble of 12 beam monitoring scalers, 16 CsI radiation monitors, the level of the water in the standoff box as well as the pH-value, resistivity and temperature of the water.

10.4. Operational Issues

The DIRC was successfully commissioned and attained performance close to that expected from Monte Carlo simulation. The DIRC has been robust and stable, and indeed, serves also as a background detector for PEP-II tuning. Figure 49 shows a typical di-muon ($e^+e^- \rightarrow \mu^+\mu^-$) event. In addition to the hits caused by the Cherenkov light from the two tracks, about 500 backgrounds hits can be seen in the readout window of ± 300 ns. This background is dominated by low energy photons from the PEP-II machine hitting the standoff box. Some care in machine tuning is required to stay under a noise limit of about 200 kHz/tube imposed by limited DAQ throughput. Lead shielding has been installed around the beamline components just outside the backward endcap, and has substantially reduced this background.

After about 2 years of running, about 99.7% of all PMTs and electronic channels are still operating with nominal performance.

Corrosion of the PMT glass face plates that are immersed in the pure water of the standoff box has been observed. For most of the tubes, the observable effect is typically a slight cloudiness, but for ~ 50 of the tubes, it is much more pronounced. Extensive R&D has demonstrated that the corrosion is associated with a loss of sodium and boron from the surface of the glass. For most tubes, the corrosion rate is a few microns per year, and is expected to be acceptable for the full projected 10 year lifetime of the experiment. However, for the ~ 50 tubes, the wrong glass was used by the PMT manufacturer. This glass did not contain zinc, which makes it much more susceptible to rapid leaching. This leaching may eventually lead to either a loss of performance, or some risk of mechanical failure of the face plates, for these tubes. Direct measurements of the number of Cherenkov photons observed in di-muon events as a function of time suggest that the total loss of photons from all sources is less than 2%/year, although the accuracy of this number is limited by a number of systematic effects at this time.

10.5. Data Analysis and Performance

Figure 49 shows the pattern of Cherenkov photons in a di-muon event, before and after reconstruction. The time distribution of real Cherenkov photons from a single event is of order ~ 50 ns wide, and during normal data taking they are accompanied by hundreds of random photons in a flat background within the trigger acceptance window. Given a track pointing at a particular fused silica bar and a candidate signal in a PMT within the optical phase space of that bar, the Cherenkov angle is uniquely determined up to a sixteen fold ambiguity: top/bottom, left/right, forward/backward and wedge/no-wedge reflections. The goal of the reconstruction program is to associate the correct track with the candidate PMT signal, with the requirement that the transit time of the photon from its creation in the bar to its detection at the PMT is consistent with the measurement error of ~ 1.5 ns.

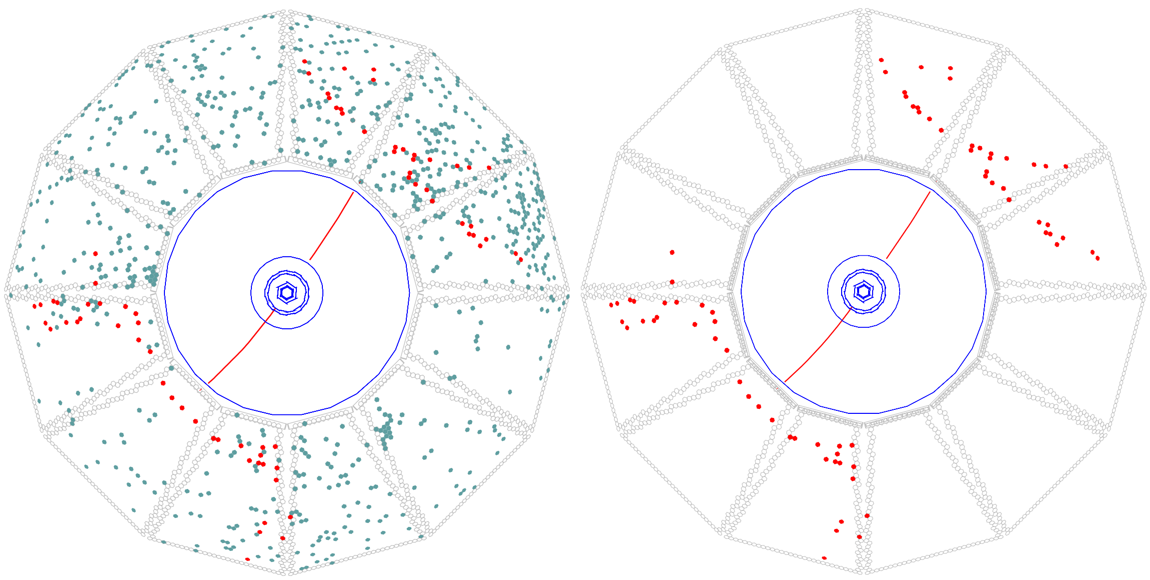

Figure 49. Display of one $e^+e^- \to \mu^+\mu^-$ event reconstructed in *BABAR* with two different time cuts. On the left, all DIRC PMTs that were hit within the ±300 ns trigger window are shown. On the right, only those PMTs that were hit within 8 ns of the expected Cherenkov photon arrival time are displayed.

10.5.1. Reconstruction

We use an unbinned maximum likelihood formalism to incorporate all information provided by the space and time measurements from the DIRC.

The emission angle and the arrival time of the Cherenkov photons are reconstructed from the observed space-time coordinates of the PMT hits, transformed into the Cherenkov coordinate system. The known spatial position of the bar through which the track passed and the PMTs hit within the readout window of ±300 ns of the trigger signal is used to calculate the three-dimensional vector pointing from the center of the end of the bar to the center of each tube. This vector is then extrapolated into the synthetic fused silica bar (using Snell's law). This procedure defines, up to the 16 fold ambiguity described above, the Cherenkov angles θ_c and ϕ_c of a photon.

The DIRC time measurement represents the third dimension of the photomultiplier hit reconstruction. The timing resolution is not competitive with the positional information for Cherenkov angle reconstruction, but timing information is used to suppress background hits from the accelerator and, more importantly, exclude other tracks in the same event as the source of the photon. Timing information is also used to resolve the forward-backward and wedge ambiguities in the hit-to-track association.

The relevant observable to distinguish between signal and background photons is the difference between the measured and expected photon arrival time, δt_γ. It is calculated for each photon using the track time-of-flight (assuming it to be a charged pion), the measured time of the candidate signal in the PMT and the photon propagation time within the bar and the water filled standoff box. The time information and the requirement of using only physically possible photon propagation paths that are within the region of total internal reflection reduces the number of ambiguities from 16 to typically 3. Applying the time information also substantially improves the correct matching of photons with tracks and reduces the number of accelerator induced background hits by approximately a factor 40, as can be seen in Figure 49.

The reconstruction routine currently provides a likelihood value for each of the five stable particle types (e,μ,π,K,p) if the track passes through the active volume of the DIRC. These likelihoods are calculated in an iterative process by maximizing the likelihood value for the entire event while testing different hypotheses for each track. If enough photons are found, the result of a fit of θ_c and the number of observed signal and background photons are calculated for each track.

10.5.2. Results

The parameters of expected DIRC performance were derived from extensive studies with a variety of prototypes, culminating with a full-size prototype in a series of test beam runs at CERN [15]. The results were well-described by Monte Carlo simulations of the detector. The present results are close to expectations, and additional offline work, particularly on geometrical alignment, is expected to lead to further improvements.

In the absence of correlated systematic errors, the resolution ($\sigma_{C,track}$) on the track Cherenkov angle should scale as

$$\sigma_{C,track} = \sigma_{C,\gamma}/\sqrt{N_{pe}} \ , \qquad (5)$$

where $\sigma_{C,\gamma}$ is the single photon Cherenkov angle resolution, and N_{pe} is the number of photons detected. Figure 50(a) shows the single photon resolution obtained for photoelectrons from di-muon events, $e^+e^- \to \mu^+\mu^-$. The average single photon resolution obtained is about 10.2 mrad, about 10% worse than the expected value of about 9 mrad. There is a broad background of less than 10% relative height under the peak, that originates mostly from track-associated sources. The time resolution obtained, shown in Figure 50(b), is 1.7 ns, close to the 1.5 ns value expected from the single-photon resolution of the PMTs.

The number of photoelectrons shown in Figure 51 varies from a minimum of about 20 for small polar angles at the center of the barrel to well over 50 at large polar angles. This is in good agreement with the value expected from the Monte Carlo simulation at all angles. The shape of the distribution can be understood as follows:

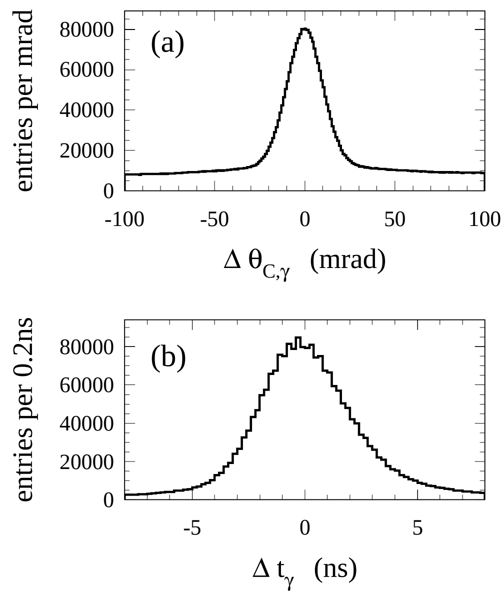

Figure 50. Resolution of (a) the reconstructed Cherenkov polar angle for single photons and (b) the difference between measured and expected arrival time.

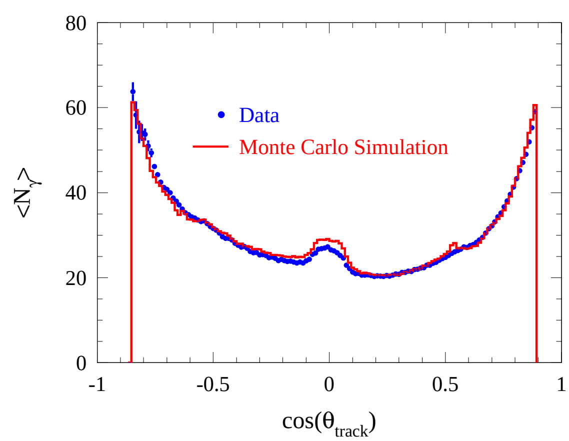

Figure 51. Number of detected photoelectrons *vs.* track polar angle for reconstructed di-muon events in data and simulation.

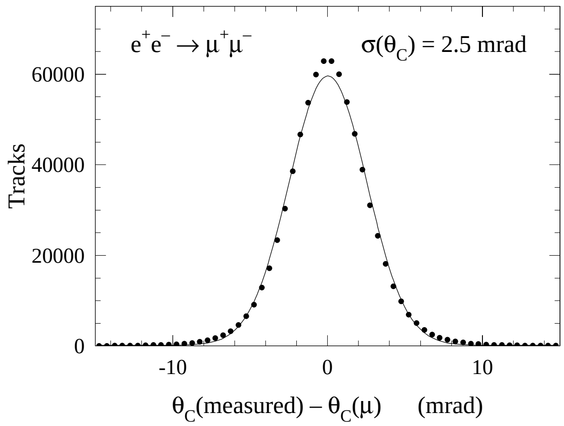

Figure 52. Resolution of the reconstructed Cherenkov polar angle per track for di-muons.

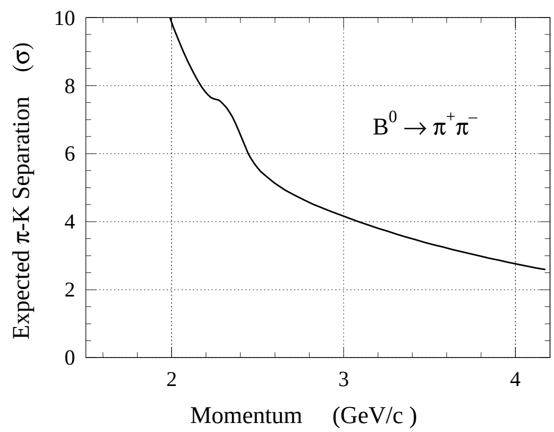

Figure 53. π-K separation in $B^0 \rightarrow \pi^+\pi^-$ events *vs.* track momentum inferred from the measured Cherenkov angle resolution and number of Cherenkov photons per track in di-muon events.

Using tracks at $\cos\theta_c = 0$ as a reference, the number of photons initially decreases as the track angle moves in the forward (or backward) direction because fewer Cherenkov photons are trapped by total internal reflection in the bar. At larger values of $\mid \cos\theta \mid$, the number of photons increases for two reasons: 1) the path length of the track, and therefore the number of Cherenkov photons, increases in the fused silica, and 2) the fraction of photons trapped in the silica by total internal reflection increases.

This spectrum also demonstrates a very useful feature of the DIRC in the *BABAR* environment, namely, the performance improves (see equation 1) in the forward direction, as is needed to cope with the angle-momentum correlation of particles from the boost.

With the present alignment, the average track Cherenkov angle resolution for di-muon events is shown in Figure 52. The width of the fitted Gaussian is 2.5 mrad. This is about 15% worse than the 2.2 mrad expected from simulation. From the measured single track resolution *vs.* momentum in di-muon events and the difference between the expected Cherenkov angles of charged pions and kaons, the pion-kaon separation power of the DIRC can be inferred. As shown in Figure 53, the separation between kaons and pions at 3 GeV/c is about 4.2 σ, approximately 15% worse than the design goal.

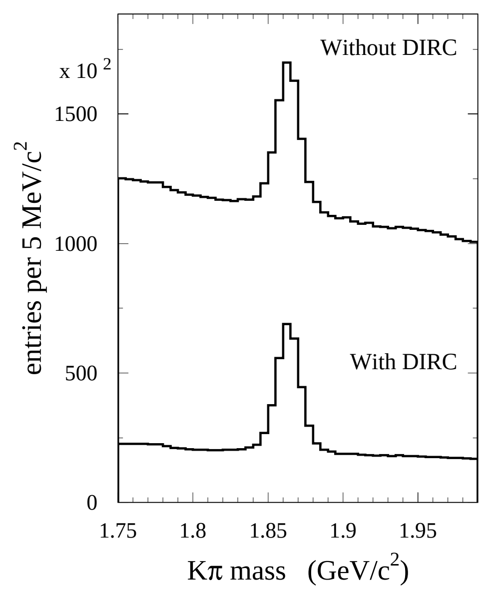

Figure 54. Invariant Kπ inclusive mass spectrum without and with the use of the DIRC for kaon identification. The mass peak corresponds to the decay of the D^0 particle.

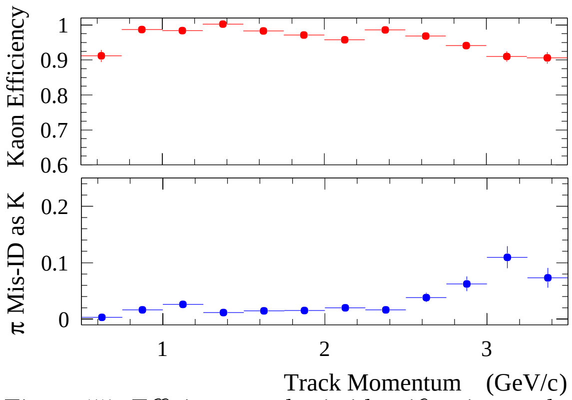

Figure 55. Efficiency and mis-identification probability for the selection of charged kaons as a function of track momentum determined using D^0 decays selected kinematically from inclusive D^* production.

Figure 54 shows an example of the use of the DIRC for sample selection. The $K\pi$ invariant two body mass spectra are shown without and with the use of the DIRC for kaon identification. The mass peak corresponds to the decay of the D^0 particle.

The efficiency for correctly identifying a charged kaon that hit a radiator bar and the probability to wrongly identify a pion as kaon are determined using D^0 decays selected kinematically from inclusive D^* production and are shown as a function of the track momentum in Figure 55 for a particular choice of particle selection criteria. The mean kaon selection efficiency and pion mis-identification are $96.2 \pm 0.2\%$ (stat.) and $2.1 \pm 0.1\%$ (stat.), respectively.

10.6. Conclusions

The DIRC is a novel ring imaging Cherenkov detector that is well-matched to the hadronic particle identification requirements of *BABAR*. The DIRC has been robust and stable and, two years after installation, about 99.7% of all PMTs and electronic channels are still operating with nominal performance. Additional shielding in the standoff box tunnel region should reduce the sensitivity to beam related backgrounds, as should faster front-end electronics, both installed during the winter 2000-2001 shutdown. At luminosities around $1 \times 10^{34}/cm^2 sec$, the TDC chip will have to be replaced with a faster version and deeper buffering. The design process for this is underway.

The initial detector performance obtained is already rather close to that predicted by the Monte Carlo simulations. Alignment and further code developments are underway which are expected to further improve performance.

REFERENCES

1. The *BABAR* Collaboration, The *BABAR* Physics Book, SLAC-R-504 (1998).
2. B. N. Ratcliff, SLAC-PUB-5946 (1992); B. N. Ratcliff, SLAC-PUB-6067 (1993); P. Coyle et al., *Nucl. Instr. and Methods* **A343** (1994) 292.
3. Spectrosil is a trademark of TSL Group PCL, Wallsend, Tyne on Wear, NE28 6DG, England; Sold in the USA by Quartz Products Co., 160 W. Lee Street, Louisville, Kentucky 40201.
4. I. Adam et al., IEEE Trans. Nucl. Sci., Vol. 45, No. 3 (June 1998) 657; I. Adam et al., *ibid* 450; J. Cohen-Tanugi, M. C. Convery, B. N. Ratcliff, X. Sarazin, J. Schwiening, J. Va'vra, SLAC-JOURNAL-ICFA-21, *ICFA Instrumentation Bulletin,* Fall 2000 Issue.
5. M. Benkebil *et al.*, *Nucl. Instr. and Methods* **A442**, 364 (2000).
6. Boeing Optical Fabrication, Albuquerque, NM, USA.
7. E. Antokhin et al., *Nucl. Instr. and Methods* **A432**, 24 (1999).
8. Electron Tubes Limited (formerly: Thorn EMI Electron Tubes), Ruislip, Middlesex HA47TA, England.
9. P. Bourgeois, M. Karolak and G. Vasseur, *Nucl. Instr. and Methods* **A442**, 105 (2000).
10. Epoxy Technology, Inc., MA, USA.
11. J. Va'vra, *Nucl. Instr. and Methods* **A453**, 262 (2000).
12. P. Bailly *et al.*, *Nucl. Instr. and Methods* **A433**, 450 (1999).
13. P. Bailly *et al.*, *Nucl. Instr. and Methods* **A433**, 432 (1999).
14. The micro-controller is a MC68HC05x32, Motorola Inc., Schaumburg, IL, USA.

15. R. Aleksan et al., *Nucl. Instr. and Methods* **A397** (1997) 261.

11. Electromagnetic Calorimeter

11.1. Requirements and Design

11.1.1. Requirements

The electromagnetic calorimeter (EMC) is designed to measure electromagnetic showers with excellent efficiency, and energy and angular resolution over the energy range from 20 MeV to 9 GeV. This capability allows the detection of photons from π^0 and η decays as well as from electromagnetic and radiative processes. By identifying electrons, the EMC contributes to the flavor tagging of neutral B mesons via semi-leptonic decays, to the reconstruction of vector mesons like J/ψ, and the study of semi-leptonic and rare decays B and D mesons and τ leptons. The upper bound of the energy range is set by the need to measure QED processes, like $e^+e^- \rightarrow e^+e^-(\gamma)$ and $e^+e^- \rightarrow \gamma\gamma$, for calibration and luminosity determination. The lower bound is set by the need for highly efficient reconstruction of B meson decays containing multiple π^0 and η.

The measurement of extremely rare decays of B mesons containing π^0s (e.g. $B^0 \rightarrow \pi^0\pi^0$) poses the most stringent requirements on energy resolution, namely of order 1% to 2%. Below energies of 2 GeV, the π^0 mass resolution is dominated by the energy resolution. At higher energies, the angular resolution becomes dominant, and therefore it is required to be of the order of a few mrad.

Furthermore, the EMC has to be compatible with the 1.5 T field of the solenoid and operate reliably over the anticipated ten-year lifetime of the experiment. To achieve excellent resolution, stable operating conditions have to be maintained. Temperatures and the radiation exposure must be closely monitored, and precise calibrations of the electronics and energy response over the full dynamic range must be performed frequently.

11.1.2. Design Considerations

The requirements stated above lead to the choice of a hermetic, total-absorption calorimeter, composed of a finely segmented array of thallium-iodide-doped cesium iodide (CsI(Tl)) crystals. The crystals are read out with silicon photodiodes that are matched to the spectrum of scintillation light. Recent experience at CLEO [1] has demonstrated the suitability of this choice for physics at the $\Upsilon(4S)$ resonance.

The energy resolution of a homogeneous crystal calorimeter can be described empirically in terms of a sum of two terms added in quadrature

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus b, \tag{6}$$

where E and σ_E refer to the energy of a photon and its rms error, measured in GeV. The energy dependent term a arises primarily from the fluctuations in photon statistics, but it is also impacted by electronic noise of the photon detector and electronics. Furthermore, beam-generated background will lead to large numbers of additional photons that add to the noise. This term is dominant at low energies. The constant term, b, is dominant at higher energies ($> 1\,\mathrm{GeV}$) and it arises from non-uniformity in light collection, from leakage or absorption in the material between and in front of the crystals, and from uncertainties in the calibrations. Most of these effects can be influenced by design choices, and they are stable with time. Others will be impacted by changes in the operating conditions, like variations in temperature, electronics gain, and noise, as well as by radiation damage caused by beam-generated radiation.

The angular resolution is determined by the transverse crystal size and the distance from the interaction point. It can also be empirically parameterized as a sum of an energy dependent and a constant term,

$$\sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E(\,\mathrm{GeV})}} \oplus d, \tag{7}$$

where the energy E is measureed in GeV. The design of the EMC required a careful optimization of a wide range of choices, from the crystal material and dimensions to the choice of the photon detector and readout electronics to the design of a calibration and monitoring system. These choices were made on the basis of extensive studies, prototyping and beam tests [2], and Monte Carlo simulation, taking into account limitations of space and the impact of other *BABAR* detector systems.

Under ideal conditions, values for the energy resolution parameters a and b close to 1% to 2% could be obtained. A position resolution of a mm will translate into an angular resolution of a few mrad, i.e., values of parameters $c \approx 3\,\mathrm{mrad}$ and $d \approx 1\,\mathrm{mrad}$.

However in practice, for a large system with a small, but unavoidable amount of inert material and gaps, limitations of electronics, and background in multi-particle events, plus contributions from beam-generated background, such performance is very difficult to achieve.

Though in CsI(Tl) the intrinsic efficiency for the detection of photons is close to 100% down to a few MeV, the minimum measurable energy in colliding beam data is expected to be about 20 MeV, a limit that is largely determined by beam- and event-related background and the amount of material in front of the calorimeter. Because of the sensitivity of the π^0 efficiency to the minimum detectable photon energy it is extremely important to keep the amount of material in front of the EMC to the lowest possible level.

11.1.3. CsI(Tl) Crystals

Thallium-doped CsI meets the needs of *BABAR* in several ways. Its properties are listed in Table 11. The high light yield and small Molière radius allow for excellent energy and angular resolution, while the short radiation length allows for shower containment at *BABAR* energies with a relatively compact design. Furthermore, the high light yield and the emission spectrum permit efficient use of a silicon photo-diodes which operate well in high magnetic fields. The transverse size of the crystals is chosen to be comparable to the Molière radius achieving the required angular resolution at low energies while appropriately limiting the total number of crystals (and readout channels).

11.2. Layout and Assembly

11.2.1. Overall Layout

The EMC consists of a cylindrical barrel and a conical forward endcap. It has full coverage in azimuth and extends in polar angle from $\cos\theta = 0.96$ to $\cos\theta = -0.77$ for a solid-angle coverage of 90% in the c.m. system (Figure 56

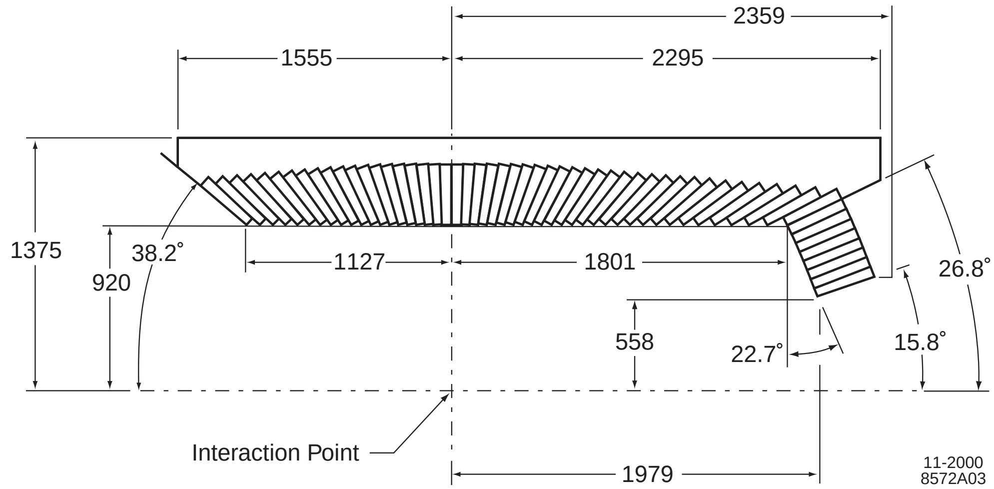

Figure 56. A longitudinal cross-section of the EMC (only the top half is shown) indicating the arrangement of the 56 crystal rings. The detector is axially symmetric around the z axis. All dimensions are given in mm.

Table 11
Properties of CsI(Tl) .

Parameter	Values
Radiation Length	1.85 cm
Molière Radius	3.8 cm
Density	4.53 g/cm^3
Light Yield	50,000 γ/ MeV
Light Yield Temp. Coeff.	0.28%/°C
Peak Emission λ_{max}	565 nm
Refractive Index (λ_{max})	1.80
Signal Decay Time	680 ns (64%)
	3.34 μs (36%)

Table 12
Layout of the EMC, composed of 56 axially symmetric rings, each consisting of CsI crystals of identical dimensions.

$\cos\theta$ Interval	Length (X_0)	# Rings	Crystals /Ring
	Barrel		
−0.774 – 0.349	16.0	27	120
0.350 – 0.620	16.5	7	120
0.621 – 0.793	17.0	7	120
0.794 – 0.890	17.5	7	120
	Endcap		
0.892 – 0.822	17.5	3	120
0.923 – 0.947	17.5	3	100
0.948 – 0.955	17.5	1	80
0.956 – 0.962	16.5	1	80

and Table 12). The barrel contains 5,760 crystals arranged in 48 distinct rings with 120 identical crystals each. The endcap holds 820 crystals arranged in 8 rings, adding up to a total of 6,580 crystals. The crystals have a tapered trapezoidal cross-section. The length of the crystals increases from 29.6 cm in the backward to 32.4 cm in the forward direction to limit the effects of shower leakage from the increasing average particle energy.

To minimize the probability of pre-showering, the crystals are supported at the outer radius, with only a thin gas seal at the front. The barrel and outer five rings of the endcap have less than $0.3-0.6X_o$ of material in front of the crystal faces. The SVT support structure and electronics, as well as the B1 dipole shadow the inner three rings of the endcap, resulting in up to 3.0 X_o for the innermost ring. The principal purpose of the two innermost rings is to enhance shower containment for particles incident at small polar angles.

11.2.2. Crystal Fabrication and Assembly

The crystals were grown in boules from a melt of CsI salt doped with 0.1% Thallium [3]. They were cut from the boules, machined into tapered trapezoids (Figure 57) to a tolerance of ± 150 μm and then polished [4]. The transverse dimensions of the crystals for each of the 56 rings vary to achieve the required hermetic coverage. The typical area of the front face is 4.7×4.7 cm, while the back face area is typically 6.1×6.0 cm. The crystals act not only as a total-absorption scintillating medium, but also as a light guide to collect light at the photo-diodes that are mounted on the rear surface. At the polished crystal surface light is internally reflected, and a small fraction is transmitted. The transmitted light is recovered in part by wrapping the crystal with two layers of diffuse white reflector [5] [6], each 165 μm thick . The uniformity of light yield along the wrapped crystal was measured by recording the signal from a highly collimated radioactive source at 20 points along the length of the crystal. The light yield was required to be uniform to within $\pm 2\%$ in the front half of the crystal; the limit increased linearly up to a maximum of $\pm 5\%$ at the rear face. Adjustments were made on individual crystals to meet these criteria by selectively roughing or polishing the crystal surface to reduce or increase its reflectivity.

Following these checks, the crystals were further wrapped in 25 μm thick aluminum foil which was electrically connected to the metal housing of the photo-diodes/preamplifier assembly to provide a Faraday shield. The crystals were covered on the outside with a 13 μm thick layer of mylar to assure electrical isolation from the external support.

11.2.3. Photo-Diodes and Preamplifier Assembly

The photon detector consists of two $2 \times 1\,\mathrm{cm}^2$ silicon PIN diodes glued to a transparent 1.2 mm-thick substrate of polysterene that in turn is glued to the center of the rear face of the crystal by an optical epoxy [7] to maximize light transmission [8]. The surrounding area of the crystal face is covered by a plastic plate coated with white reflective paint [9]. The plate has two 3-mm-diameter penetrations for the fibers of the light pulser monitoring system.

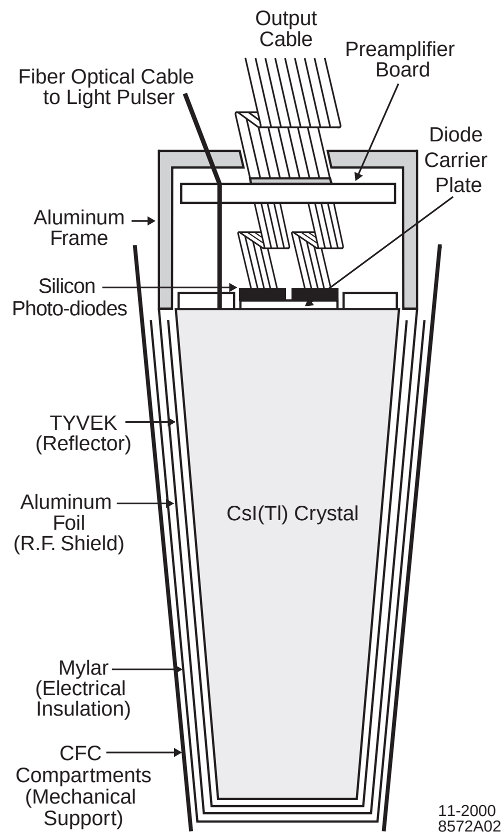

Figure 57. A schematic of the wrapped CsI(Tl) crystal and the front-end readout package mounted on the rear face. Also indicated is the tapered, trapezoidal CFC compartment, which is open at the front. This drawing is not to scale.

As part of the quality control process, the 1.836 MeV photon line from a ^{88}Y radioactive source was used to measure the light yield of every crystal-diode assembly, employing a Canberra

2003T preamplifier with $2\,\mu$s Gaussian shaping. The resulting signal distribution has a mean and rms of 7300 photo-electrons/ MeV and 890 photo-electrons/ MeV, respectively, with no crystal below 4600 photo-electrons/ MeV [8]. [19]

Each of the diodes is directly connected to a low-noise preamplifier. The entire assembly is enclosed by an aluminium fixture as shown in Figure 57. This fixture is electrically coupled to the aluminium foil wrapped around the crystal and thermally coupled to the support frame to dissipate the heat load from the preamplifiers.

Extensive aging tests were performed to ascertain that the diodes and the preamplifiers met the ten-year life time requirements. In addition, daily thermal cycles of $\pm 5^o C$ were run for many months to assure that the diode-crystal epoxy joint could sustain modest temperature variations.

11.2.4. Crystal Support Structure

The crystals are inserted into modules that are supported individually from an external support structure. This structure is built in three sections, a cylinder for the barrel and two semi-circular structures for the forward endcap. The barrel support cylinder carries the load of the barrel modules plus the forward endcap to the magnet iron through four flexible supports. These supports decouple and dampen any acceleration induced by movements of the magnet iron during a potential earthquake.

The modules are built from tapered, trapezoidal compartments made from carbon-fiber-epoxy composite (CFC) with 300 μm thick walls (Figure 58). Each compartment loosely holds a single wrapped and instrumented crystal and thus assures that the forces on the crystal surfaces never exceed its own weight. Each module is surrounded by an additional layer of 300 μm CFC to provide additional strength. The modules are bonded to an aluminum strong-back that is mounted on the external support. This scheme minimizes inter-crystal materials while exerting minimal force on the crystal surfaces; this prevents deformations and surface degradation that could compromise performance. By supporting the modules at the back, the material in front of the crystals is kept to a minimum.

[19]The calibration procedure employed in this measurement introduces a dependency of the light yield on the shaping time of the preamplifier. When connected to the actual front-end electronics in the *BABAR* detector, the signal is reduced by a factor 1.29.

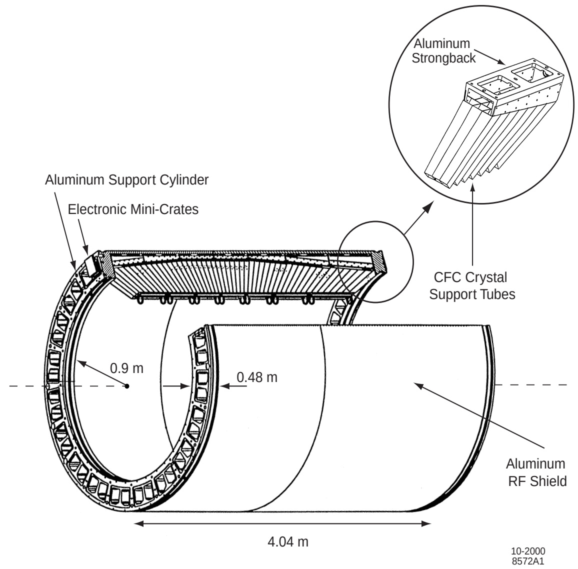

Figure 58. The support structure and assembly of the EMC barrel.

The barrel section is divided into 280 separate modules, each holding 21 crystals (7×3 in $\theta \times \phi$). After the insertion of the crystals, the aluminum readout frames, which also stiffen the module, are attached with thermally-conducting epoxy to each of the CFC compartments. The entire 100 kg-module is then bolted and again thermally epoxied to an aluminum strong-back. The strong-back contains alignment features as well as channels to couple into the cooling system. Each module was installed into the 2.5 cm-thick, 4 m-long aluminum support cylinder, and subsequently aligned. On each of the thick annular end-flanges this cylinder contains access ports for digitizing electronics crates with associated cooling channels, as well as mounting features and alignment dowels for the forward endcap.

The endcap is constructed from 20 identical

CFC modules (each with 41 crystals), individually aligned and bolted to one of two semi-circular support structures. This vertical split into two halves was necessary to facilitate access to the central detector components.

The entire calorimeter is surrounded by a double Faraday shield composed of two 1 mm-thick aluminum sheets so that the diodes and preamplifiers are further shielded from external noise. This cage also serves as the environmental barrier, allowing the slightly hygroscopic crystals to reside in a dry, temperature controlled nitrogen atmosphere.

11.2.5. Cooling System

The EMC is maintained at constant, accurately monitored temperature. Of particular concern are the stability of the photo-diode leakage current which rises exponentially with temperature, and the large number of diode-crystal epoxy joints that could experience stress due to differential thermal expansion. In addition, the light yield of CsI(Tl) is weakly temperature dependent.

The primary heat sources internal to the calorimeter are the preamplifiers (2×50 mW/crystal) and the digitizing electronics (3 kW per end-flange). In the barrel, the preamplifier heat is removed by conduction to the module strong backs which are directly cooled by Fluorinert (polychlorotrifluoro-ethylene) [10]. The digitizing electronics are housed in 80 mini-crates, each in contact with the end-flanges of the cylindrical support structure. These crates are indirectly cooled by chilled water pumped through channels milled into the end flanges close to the inner and outer radii. A separate Fluorinert system in the endcap cools both the 20 mini-crates of digitizing electronics and the preamplifiers.

11.3. Electronics

The EMC electronics system, shown schematically in Figure 59, is required to have negligible impact on the energy resolution of electromagnetic showers from 20 MeV to 9 GeV, while accommodating the use of a 6.13 MeV radioactive source for calibration. These requirements set a limit of less than 250 keV equivalent noise energy (ENE) per crystal and define an 18-bit effective dynamic range of the digitization scheme. For source calibrations, the least significant bit is set to 50 keV, while for colliding beam data it is set to 200 keV. To reach the required energy resolution at high energies, the coherent component has to be significantly smaller than the incoherent noise component. In addition, it is important that the impact of high rates of low energy (<5 MeV) beam-induced photon background be minimized.

11.3.1. Photo-Diode Readout and Preamplifiers

The ENE is minimized by maximizing the light yield and collection, employing a highly efficient photon detector, and a low-noise electronic readout. The PIN silicon photo-diodes [11] have a quantum efficiency of 85% for the CsI(Tl) scintillation light [12]. At a depletion voltage of 70 V, their typical dark current wre measured to be 4 nA for an average capacitance of 85 pF; the diodes are operated at a votage of 50 V. The input capacitance to the preamplifier is minimized by connecting the diodes to the preamplifier with a very short cable. The preamplifier is a low-noise charge-sensitive amplifier implemented as a custom application specific integrated circuit (ASIC) [13]. It shapes the signal and acts as a band-pass filter to remove high- and low-frequency noise components. The optimum shaping time for the CsI(Tl)-photodiode readout is $2-3$ μs, but a shorter time was chosen to reduce the probability of overlap with low-energy photons from beam background. The commensurate degradation in noise performance is recovered by implementing a real-time digital signal-processing algorithm following digitization.

To achieve the required operational reliability [14] for the inaccessible front-end readout components, two photo-diodes were installed, each connected to an preamplifier. In addition, all components were carefully selected and subjected to rigorous tests, including a 72-hour *burn-in* of the preamplifiers at 70 °C to avoid *infant mortality*. The dual signals are combined in the postamplification/digitization circuits, installed in mini-crates at the end-flanges, a location that is acces-

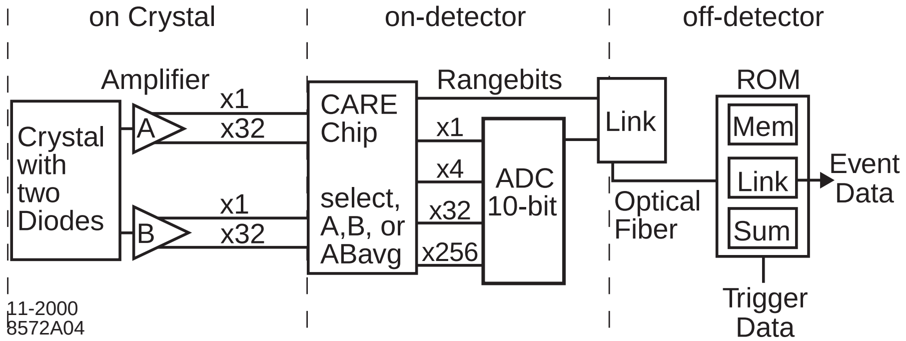

Figure 59. Schematic diagram of the EMC readout electronics.

sible for maintenance.

11.3.2. Postamplification, Digitization and Readout

The two preamplifiers on each crystal, A and B, each provide amplification factors of 1 and 32 and thus reduce the dynamic range of the signal that is transmitted to the mini-crates to 13-bits. A custom auto-range encoding (CARE) circuit [13] further amplifies the signal to arrive at a total gain of 256, 32, 4 or 1 for four energy ranges, 0-50 MeV, 50-400 MeV, 0.4-3.2 GeV, and 3.2-13.0 GeV, respectively. The appropriate range is identified by a comparator and the signal is digitized by a 10-bit, 3.7 MHz ADC. Data from 24 crystals are multiplexed onto a fiber-optic driver and sent serially at a rate of 1.5 Gbits/s across a 30 m-long optical fiber to the ROM. In the ROM, the continuous data stream is entered into a digital pipeline. A correction for pedestal and gain is applied to each sample. The pipeline is then tapped to extract the input to the calorimeter trigger.

Upon receipt of the L1 Accept signal, data samples within a time window of $\pm 1\,\mu$s are selected for the feature extraction. Up to now, the calorimeter feature extraction algorithm performs a parabolic fit to the peak of the signal waveform to derive its energy and time. In the future, it is planned to employ a digital filter prior to the signal fit to further reduce noise. For this filter algorithm, the frequency decomposition of an average signal pulse and the typical noise spectrum are measured for all channels and subsequently used to derive an optimum set of weights that maximizes the signal-to-noise ratio. These weights are then applied to individual samples to obtain a filtered waveform.

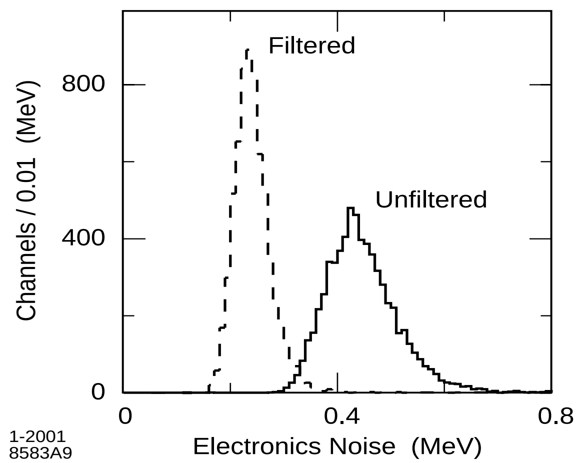

Figure 60. The distribution of equivalent noise energy for all channels of the EMC with/without digital filtering. The data were recorded in the absence of beams by a random trigger.

The magnitude of the electronic noise is measured as the rms width of the pedestal distribution as shown in Figure 60. The observed distribution for all channels translates to an ENE of 230 keV and 440 keV with and without digital filtering; this result is comparable to design expectations. Measurements of the auto-correlation function indicate that the coherent noise component is negligible compared to the incoherent noise, except for regions where the preamplifiers saturate (see below).

During data taking, the data acquisition imposes a single-crystal readout threshold in order to keep the data volume at an acceptable level. This energy threshold is currently set to 1 MeV and during stable colliding beam conditions on average 1,000 crystals are read out (measured with 600 mA of e^- and 1100 mA of e^+ and a random clock trigger), corresponding to an average occupancy of 16%. The electronic noise accounts for about 10%, while the remaining signals originate from beam-generated background (see Chapter 3). In addition, a typical hadronic event contributes signals in xxx crystals.

11.3.3. Electronics Calibration and Linearity

To measure pedestal offsets, determine the overall gain, and to remove non-linearities the front-end electronics are calibrated by precision charge injection into the preamplifier input. Initially, residual non-linearities of up to 12% in limited regions near each of the range changes were observed and corrected for offline [15]. These non-linearities were traced to oscillations on the ADC cards that have since been corrected. The correction resulted in markedly improved energy resolution at high energies. Residual non-linearities (typically $2-4$%) arise primarily from cross-talk, impacting both the electronics calibrations and the colliding-beam data. The effect is largest at about 630 MeV (950 MeV) in a high (low) gain preamplifier channel. The implementation of an energy dependent correction is expected to significantly reduce this small, remaining effect, and lead to a further improvement of the energy resolution.

11.3.4. Electronics Reliability

With the exception of minor cable damage during installation (leaving two channels inoperative), the system of 13,160 readout channels has met its reliability requirements. After the replacement of a batch of failing optical-fiber drivers, the reliability of the digitizing electronics improved substantially, averaging channel losses of less than 0.1%.

11.4. Energy Calibration

The energy calibration of the EMC proceeds in two steps: First, the measured pulse height in each crystal has to be translated to the actual energy deposited. Second, the energy deposited in a shower spreading over several adjacent crystals has to be related to the energy of the incident photon or electron by correcting for energy loss mostly due to leakage at the front and the rear, and absorption in the material between and in front of the crystals, as well as shower energy not associated with the cluster.

The offline pattern recognition algorithm that groups adjacent crystals into "clusters" is described in detail in section 11.6.

11.4.1. Individual Crystal Calibration

In spite of the careful selection and tuning of the individual crystals, their light yield varies significantly and is generally non-uniform. It also changes with time under the impact of beam-generated radiation. The absorbed dose results is largest at the front of the crystal and results in increased attenuation of the transmitted scintillation light. The light yield must therefore be calibrated at different energies, corresponding to different average shower penetration, to track the effects of the radiation damage.

The calibration of the deposited energies is performed at two energies at opposite ends of the dynamic range, and these two measurements are combined by a logarithmic interpolation. A 6.13 MeV radioactive photon source [17] provides an absolute calibration at low energy, while at higher energies (3 - 9 GeV) the relation between polar angle and energy of $e^{\pm}$ from Bhabha events is exploited [18].

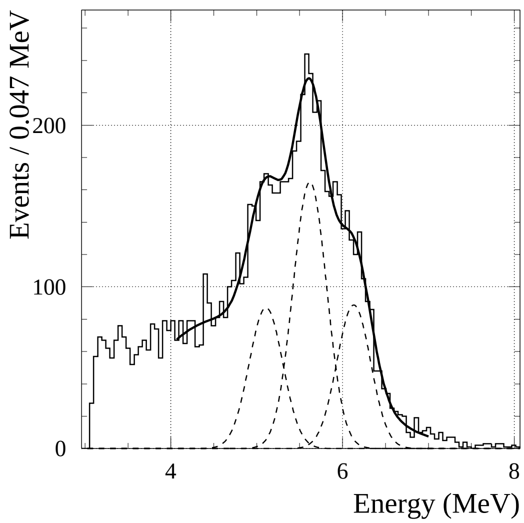

Figure 61. A typical pulse-height spectrum recorded with the radioactive source to calibrate the single-crystal energy scale of the EMC. The spectrum shows the primary 6.13 MeV peak and two associated escape peaks at 0.511 MeV and 1.022 MeV below.

A flux of low-energy neutrons ($4\ \times 10^8/s$) is used to irradiate Fluorinert [10] to produce photons of 6.13 MeV via the reaction $^{19}F + n \rightarrow ^{16}N + \alpha$, $^{16}N \rightarrow ^{16}O^* + \beta$, $^{16}O^* \rightarrow ^{16}O + \gamma$. The activated ^{16}N has a half-life of 7 seconds and thus does not cause radiation damage or long-term activation. The fluid is pumped at a rate of 125 ℓ/s from the neutron generator to a manifold of thin-walled (0.5 mm) aluminum pipes that are mounted immediately in front of the crystals. At this location, the typical rate of photons is 40 Hz/crystal.

Figure 61 shows a typical source spectrum that was derived from the raw data by employing a digital filter algorithm. For a 30-minute exposure, a statistical error of 0.35% is obtained, compared to a systematic uncertainty of less than 0.1%. This calibration is performed weekly.

At high energies, single crystal calibration is performed with a pure sample of Bhabha events [18]. As a function of the polar angle of the $e^{\pm}$, the deposited cluster energy is constrained to equal the prediction of a GEANT [19] based Monte Carlo simulation. For a large number of energy clusters, a set of simultaneous linear equations relates the measured to the expected energy and thus permits the determination of a gain constant for each crystal. In a 12-hour run at a luminosity of $3\ \times 10^{33}\mathrm{cm}^{-1}\mathrm{sec}^{-2}$ some 200 hits per crystal can be accumulated, leading to a statistical error of 0.35%. This calibration has been performed about once per month, and will be fully automated in the future.

11.4.2. Cluster Energy Correction

The correction for energy loss due to shower leakage and absorption is performed as a function of cluster energy and polar angle. At low energy ($\mathrm{E} < 0.8\,\mathrm{GeV}$), it is derived from π^0 decays [20]. The true energy of the photon is expressed as a product of the measured deposited energy and a correction function which depends on $\ln E$ and $\cos\theta$. The algorithm constrains the two-photon mass to the nominal π^0 mass and iteratively finds the coefficients of the correction function. The typical corrections are of order $6 \pm 1\%$. The uncertainty in the correction is due to systematic uncertainties in the background estimation and the fitting technique.

At higher energy ($0.8\,\mathrm{GeV} < \mathrm{E} < 9\,\mathrm{GeV}$) the correction is estimated from single-photon Monte Carlo simulations. A second technique using radiative Bhabha events [21] is being developed. The beam energy and the precise track momenta of the e^+ and e^-, together with the direction of the radiative photon, are used to fit the photon energy. This fitted value is compared to the measured photon energy to extract correction coefficients, again as a function of $\ln E$ and $\cos\theta$.

11.5. Monitoring

11.5.1. Environmental Monitoring

The temperature is monitored by 256 thermal sensors that are distributed over the calorimeter, and has been maintained at $20 \pm 0.5°$C. Dry nitrogen is circulated throughout the detector to stabilize the relative humidity at $1 \pm 0.5\%$.

11.5.2. Light-Pulser System

The light response of the individual crystals is measured daily using a light-pulser system [16] [22]. Spectrally filtered light from a xenon flash lamp is transmitted through optical fibers to the rear of each crystal. The light pulse is similar in spectrum, rise-time and shape to the scintillation light in the CsI(Tl) crystals. The pulses are varied in intensity by neutral-density filters, allowing a precise measurement of the linearity of light collection, conversion to charge, amplification and digitization. The intensity is monitored pulse-to-pulse by comparison to a reference system with two radioactive sources, ^{241}Am and ^{148}Gd, that are attached to a small CsI(Tl) crystal that is read out by both a photo-diode and a photo-multiplier tube. The system is stable to 0.15% over a period of one week and has proven to be very valuable in diagnosing problems. For example, the ability to accurately vary the light intensity has lead to the detection of non-linearities in the electronics [16].

11.5.3. Radiation Monitoring and Damage

The radiation exposure is monitored by 60/56 RadFETs placed in front of the barrel/endcap crystals. RadFETs [23] are real-time integrating dosimeters based on solid-state Metal Oxide Semiconductor (MOS) technology. In Figure 62 the accumulated dose is compared to the observed loss in scintillation light, separately for the endcap, the forward, and the backward barrel. The dose appears to follow the integrated luminosity, approximately linearly. The light loss is greatest in the forward region corresponding to the area of highest integrated radiation dose. The size of the observed light loss is close to expectations, based on extensive irradiation tests.

11.6. Reconstruction Algorithms

A typical electromagnetic shower spreads over many adjacent crystals, forming a *cluster* of energy deposits. Pattern recognition algorithms have been developed to efficiently identify these clusters and to differentiate single clusters with one energy maximum from merged clusters with more than one local energy maximum, referred to as a *bumps*. Furthermore, the algorithms have

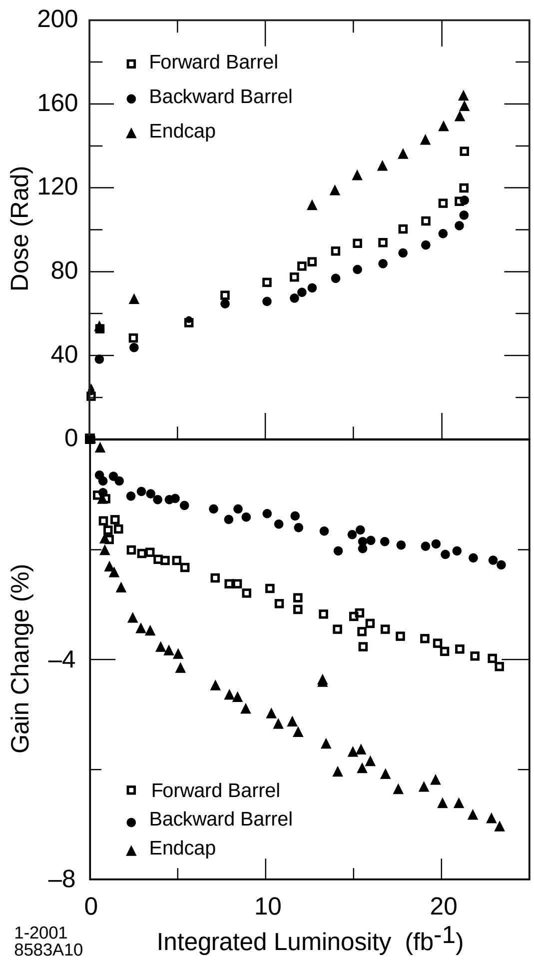

Figure 62. Impact of beam-generated radiation on the CsI(Tl) crystals: a) the integrated dose measured with RadFETs placed in front of the crystals, b) the degradation in light yield measured with the radioactive-source calibration system.

to determine whether a bump is generated by a charged or a neutral particle.

Clusters are required to contain at least one seed crystal with an energy above 10 MeV. Surrounding crystals are considered as part of the cluster if their energy exceeds a threshold of 1 MeV, or if they are contiguous neighbors (including corners) of a crystal with at least 3 MeV. The value of the single crystal threshold is set by the data acquisition system in order to keep the data volume at an acceptable level, given the current level of electronics noise and beam-generated background. It is highly desirable to reduce this threshold since fluctuations in the effective energy loss at the edges of a shower cause a degradation in resolution, particularly at low energies.

Local energy maxima are identified within a cluster by requiring that the candidate crystal have an energy, $E_{LocalMax}$, which exceeds the energy of each of its neighbors, and satisfy the following condition: $0.5(N-2.5) > E_{NMax}/E_{LocalMax}$, where E_{NMax} is the highest energy of the neighboring N crystals with an energy above 2 MeV.

Clusters are divided into as many bumps as there are local maxima. An iterative algorithm is used to determine the energy of the bumps. Each crystal is given a weight, w_i, and the bump energy is defined as $E_{bump} = \sum_i w_i E_i$, where the sum runs over all crystals in the cluster. For a cluster with a single bump, the result is $w_i \equiv 1$. For a cluster with multiple bumps, the crystal weight for each bump is calculated as

$$w_i = E_i \frac{\exp(-2.5 r_i/r_M)}{\sum_j E_j \exp(-2.5 r_j/r_M)},$$

where the index j runs over all crystals in the cluster. r_M refers to the Molière radius, and r_i is the distance of the ith crystal from the centroid of the bump. At the outset, all weights are set to one. The process is then iterated, whereby the centroid position used in calculating r_i is determined from the weights of the previous iteration, until the bump centroid position is stable to within a tolerance of 1 mm.

The position of a bump is calculated using a center-of-gravity method with logarithmic, rather than linear weights [24] [25], $W_i = 4.0 + \ln E_i/E_{bump}$, where only crystals with positive weights (i.e., $E_i > 0.0184 \times E_{bump}$, are used in the calculation. This procedure emphasizes lower-energy crystals, while utilizing only those crystals that make up the core of the cluster. A systematic bias of the calculated polar angle originates from the non-projectivity of the crystals. This bias is corrected by a simple offset of -2.6 mrad for $\cos\theta < 0$ and $+2.6$ mrad for $\cos\theta > 0$.

A bump is associated with a charged particle by projecting a track to the inner face of the calorimeter. The distance between the track impact point and the bump centroid is calculated and if it is consistent with the angle and momentum of the track, the bump is associated with this charged particle. Otherwise, it is assumed to originate from a neutral particle.

On average, xxx clusters are detected per hadronic event, of which yyy are identified as neutral particles. At current operating conditions, beam-induced background contributes on average on 1.4 neutral clusters with energies above 20 MeV. This number is significantly smaller than the average number of crystals with energies above 10 MeV(see Chapter 3.

11.7. Performance

11.7.1. Energy Resolution

At low energy, the energy resolution of the EMC is measured directly with the radioactive source yielding $\sigma_E/E = (5.0 \pm 0.8)\%$ at 6.13 MeV (see Figure 61). At high energy, the resolution is derived from Bhabha scattering, where the energy of the detected shower can be predicted from the polar angle of the $e^{\pm}$. The measured resolution is $\sigma_E/E = (1.9 \pm 0.07)\%$ at 7.5 GeV (Figure 63). Figure 64 shows the energy resolution extracted from a variety of processes as a function of energy. Below 2 GeV, the mass resolution of π^0 and η mesons decaying into two photons of approximately equal energy is used to infer the EMC energy resolution. The decay $\chi_{c1} \to J/\psi\,\gamma$ provides a measurement at an average energy of $\approx$ 500 MeV, and measurements at high energy are derived from Bhabha scattering. A fit to the

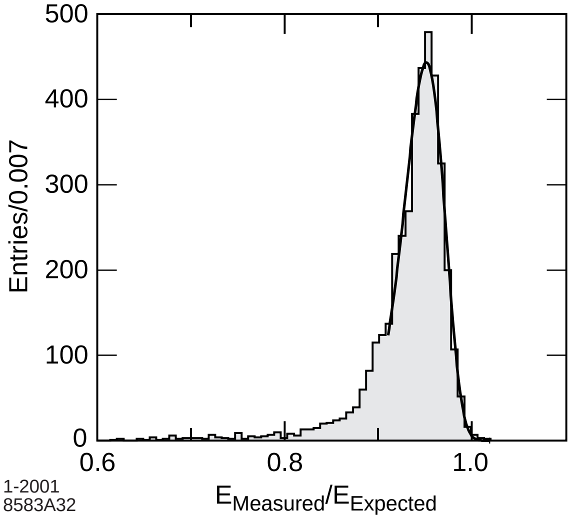

Figure 63. The ratio of the measured energy to the expected energy for electrons from Bhabha scattering. The solid line indicates a fit to a Gaussian distribution.

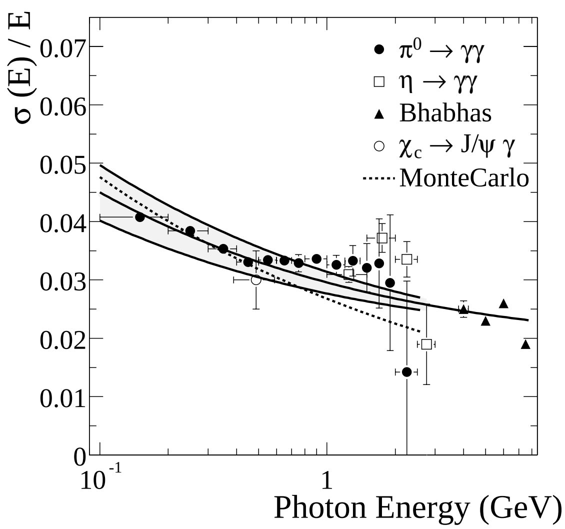

Figure 64. The energy resolution for the electromagnetic calorimeter measured for photons and electrons from various processes. The solid curve is a fit to Equation 6and the shaded area denotes the r.m.s. error on the fit.

energy dependence results in

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus (1.85 \pm 0.12)\%, \tag{8}$$

These values of these fitted parameters are higher than the somewhat optimistic earlier estimates, but they agree with detailed Monte Carlo simulations which include the contributions from electronic noise and beam background, as well as the impact of the material and the energy thresholds.

11.7.2. Angular Resolution

The measurement of the angular resolution is based on the analysis of π^0 and η decays to two photons of approximately equal energy. The result is presented in Figure 65. The resolution varies between about 12 mrad at low energies and 3 mrad at high energies. A fit to the empirical parameterization of the energy dependence results in

$$\sigma_\theta = \sigma_\phi = (\frac{3.87 \pm 0.07}{\sqrt{E(\,\mathrm{GeV})}} \oplus 0.00 \pm 0.04) \ \mathrm{mrad}).(9)$$

These fitted values are slightly lower than one would expect from detailed Monte Carlo simulations.

11.7.3. π^0 Mass and Width

Figure 66 shows the two-photon invariant mass in $B\overline{B}$ events. The reconstructed π^0 mass is measured to be 135.1 MeV/c^2 and is stable to better than 1% over the full photon energy range. The width of 6.9 MeV/c^2 agrees well with the prediction obtained from detailed Monte-Carlo simulations. In low-occupancy $\tau^+\tau^-$ events the width is slightly smaller, 6.5 MeV/c^2, for π^0 energies below 1 GeV. A similar improvement is also observed in analyses using selected isolated photons in hadronic events.

11.7.4. Electron Identification

Electrons are separated from charged hadrons primarily on the basis of the shower energy, lateral shower moments, and track momentum [26].

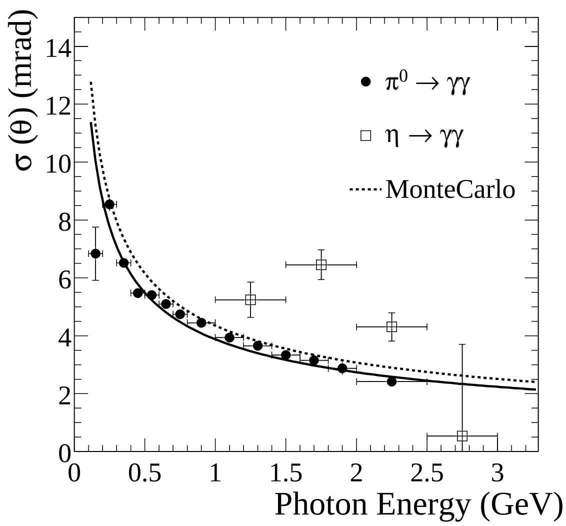

Figure 65. The angular resolution of the EMC for photons from π^0 and η decays. The solid curve is a fit to Equation 7.

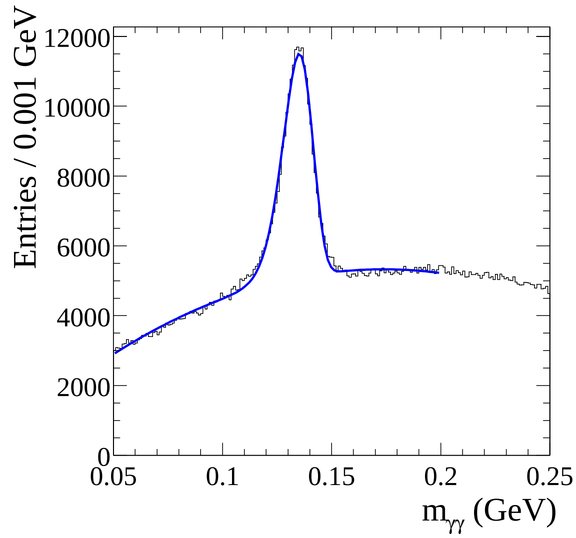

Figure 66. Invariant mass of two photons in $B\overline{B}$ events. The energies of the photons and the π^0 are required to exceed 30 MeV and 300 MeV, respectively. The solid line is a fit to the data [20].

In addition, the dE/dx energy loss in the DCH and the DIRC Cherenkov angle are required to be consistent with an electron. The most important variable for the discrimination of hadrons is the ratio of the shower energy to the track momentum (E/p). Figure 67 shows the efficiency for electron identification and the pion misidentification probability as a function of momentum for two sets of selection criteria. The electron efficiency is measured using radiative Bhabhas and $e^+e^- \rightarrow e^+e^-e^+e^-$ events. The pion misidentification probability is measured for selected charged pions from K_S^0 decays and three-prong τ decays. A tight (very tight) selector results in an efficiency plateau at 94.8% (88.1%) in the momentum range $0.5\,\mathrm{GeV}/c < p < 2\,\mathrm{GeV}/c$. The pion misidentification probability is of order 0.3% for the very tight criteria. The selection criteria can, of course, be tailored to meet the needs of specific physics analyses.

11.8. Summary

The EMC is presently performing close to design expectations. Improvements in the energy resolution are expected from the optimization of the feature-extraction algorithms designed to further reduce the electronics noise. Modifications to the electronics should allow for more precise calibrations. The expected noise reduction should permit a lower single-crystal readout threshold. However, this decrease in noise might be off-set by an increase in the beam background that is expected for higher luminosities and beam currents.

REFERENCES

1. T. Swarnicki, *Performance of the CLEO-II* CsI(Tl)Calorimeter in *Proceedings of Workshop on B Factories*, Stanford, CA, USA (1992).
2. R.J. Barlow *et al.*, *Results from the BABAR Electromagnetic Calorimeter Beam Test*,

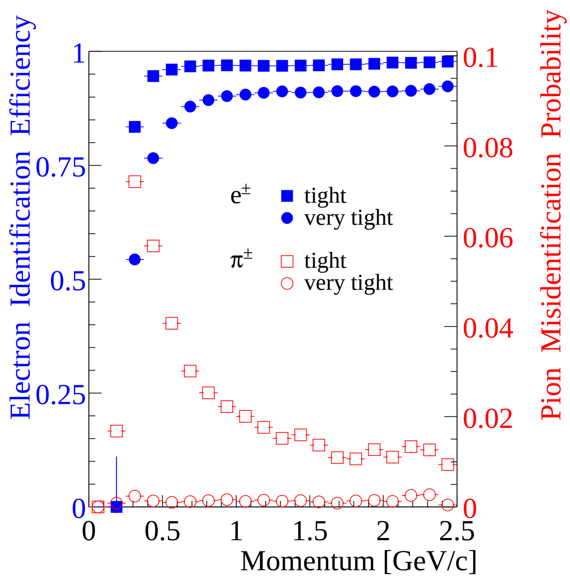

Figure 67. The electron efficiency and pion misidentification probability as a function of the particle momentum.

Nucl. Instr. and Methods **A**420 (1999) 162.

3. Chemetall and Aldrich-APL.
4. Shanghai Institute of Ceramics, Shanghai, P.R.China;
   Beijing Glass Research Institute, Beijing, P.R.China;
   Hilger Analytical, United Kingdom;
   Crismatec, France;
   Amcrys-H, Ukraine.
5. TYVEK, registered trademark of E.I. DuPont de Nemours & Co.
6. G. Dahlinger, *Aufbau und Test eines Kalorimeter-Prototyps aus CsI(Tl) zur Energie- und Ortsmessung hochenergetischer Photonen*, PhD Thesis, Technische Universität Dresden, Germany (1998).
7. A17-01 manufactured by EPILOX, Germany.
8. J. Brose, G. Dahlinger, K.R. Schubert, Nucl. Instr. and Methods **A**417 (1998) 311; C. Jessop *et al.*, *Development of Front End Readout for the BABAR Electromagnetic Calorimeter*, BABAR Note-216 (199x).
9. NE-561 manufactured by Nuclear Enterprises, USA.
10. Fluorinert (polychlorotrifluoro-ethylene) is manufactured by 3M Corporation.
11. S-2744-08 PIN diode by Hamamatsu Photonics, K. K., Hamamatsu City, Japan.
12. C. Jessop *et al. Performance Tests of Hamamatsu 2774-08 Photo-Diodes for the BABAR Electromagnetic Calorimeter Front End Readout*, BABAR Note-236 (199x).
13. G. Haller, D. Freytag, IEEE Trans.Nucl.Sci. 43, 1610 (1996)
14. C. Jessop, *Reliability Issues for the Front-End Readout of the BaBar Electromagnetic Calorimeter*, BABAR Note-217 (199x).
15. S. Menke, *Offline Correction of Non-Linearities in the BaBar Electromagnetic Calorimeter*, BABAR Note-527 (199x).
16. M. Kocian, *Das Lichtpulsersystem des elektromagnetischen CsI(Tl)-Kalorimeters des BABAR-Detektors*, PhD Thesis, TUD-IKTP/00-03, Technische Universität Dresden, Germany (2000).
17. F. Gaede, D. Hitlin, M. Weaver, *The Radioactive Source Calibration of the BABAR Electromagnetic Calorimeter*, BABAR Note-531 199x.
    J. Button-Shafer *et al.*, *Use of Radioactive Sources for Calibrating and Monitoring the BABAR Electromagnetic Calorimeter*, BABAR Note-322 (199x).
18. R. Müller-Pfefferkorn, *Die Kalibration des elektromagnetischen CsI(Tl)-Kalorimeters des Babar-Detektors mit Ereignissen der Bhabha-Streuung*, PhD Thesis, TUD-IKTP/01-01, Technische Universität Dresden, Germany 2001.
19. GEANT Detector Description and Simulation tool, CERN Program Library, Long Writeup W5013, (1994.
20. S. Menke *et al. Calibration of the BABAR Electromagnetic Calorimeter with $\pi^0 s$* , BABAR Note-528 (199x).
21. J. Bauer, *Radiative Bhabha Calibration for the BABAR Electromagnetic Calorimeter*, BABAR Note-521 (2000).
22. P.J. Clark, *The BaBar Light Pulser System*,

PhD Thesis, University of Edinburgh, UK (2000);
B. Lewandowski, *Entwicklung und Aufbau eines Lichtpulsersystems fr das Kalorimeter des BABAR-Detektors*, PhD Thesis, Ruhr-Universität Bochum, Germany (2000).
23. RADFET
24. B. Brabson *et al.* Nucl. Instr. and Methods **A**332 (1993) 419.
25. S. Otto, *Untersuchungen zur Ortsrekonstruktion electromagnetischer Schauer*, Diplomarbeit, Technische Universität Dresden, Germany 2000.
26. U. Langenegger, *Electron Identification with BABAR*, *BABAR* Note-530 (2000).

12. Detector for Muons and Neutral Hadrons

12.1. Physics Requirements and Goals

The Instrumented Flux Return (IFR) was designed to identify muons with high efficiency and good purity, and to detect neutral hadrons (primarily K_L^0 and neutrons) over a wide range of momenta and angles. Muons are important for tagging the flavor of neutral B mesons via semi-leptonic decays, for the reconstruction of vector mesons, like the J/ψ, and the study of semi-leptonic and rare decays involving leptons of B and D mesons and τ leptons. K_L^0 detection allows for the study of exclusive B decays, in particular CP eigenstates. The IFR could also help in vetoing charm decays and improve the reconstruction of neutrinos.

The principal requirements for IFR are large solid angle coverage, good efficiency, and high background rejection for muons down to momenta below 1 GeV/c. For neutral hadrons, high efficiency and good angular resolution are most important. Because this system is very large and difficult to access, high reliability and extensive monitoring of the detector performance and the associated electronics plus the voltage distribution are necessary.

12.2. Overview and RPC Concept

The IFR uses the steel flux return of the magnet as muon filter and hadron absorber. Single gap resistive plate chambers (RPCs) [1] with two-coordinate readout have been chosen as detectors.

The RPCs are installed in the gaps in the finely segmented steel (see Chapter 4) of the six barrel sectors and the two end-doors of the flux return, as illustrated in Figure 68. The steel segmentation has been optimized on the basis of Monte Carlo studies of muon penetration and charged and neutral hadron interactions. The steel is segmented into 18 plates, increasing in thickness from 2 cm of the inner nine plates to 10 cm of outermost plate(s). The nominal gap width is 3.5 cm in the inner layers of the barrel and 3.2 cm elsewhere. There are 19 RPC layers in the barrel and 18 in the endcaps. In addition, two layers of cylindrical RPCs are installed between the EMC

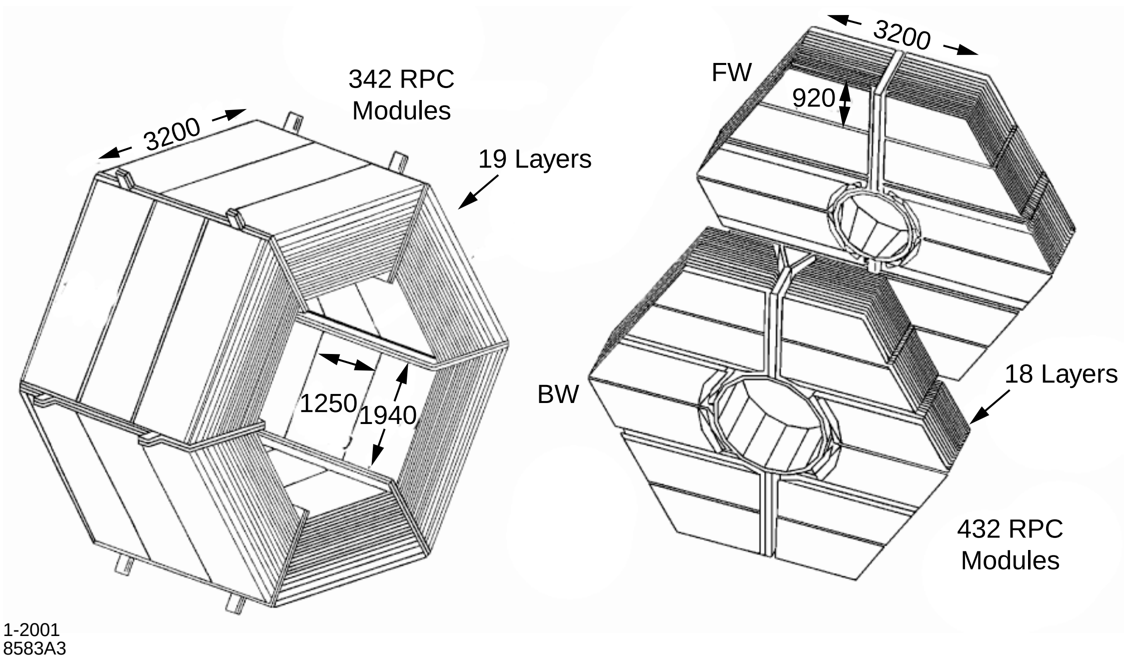

Figure 68. Overview of the IFR: Barrel sectors and forward (FW) and backward (BW) end-doors; the shape of the RPC modules and their dimensions are indicated.

and the magnet cryostat to detect particles exiting the EMC.

RPCs detect streamers from ionizing particles via capacitive readout strips. They offer several advantages: simple, low cost construction and the possibility of covering odd shapes with minimum dead space. Further benefits are large signals and fast response allowing for simple and robust front end electronics and good time resolution, typically 1-2 ns. The position resolution depends on the segmentation of the readout, a few mm are achievable.

A cross section of an RPC is shown schematically in Figure 69. The construction of the planar and the cylindrical RPCs differ in detail, but are based on the same concept.

The planar RPCs consist of two bakelite (phenolic polymer) sheets, 2 mmthick and separated by a gap of 2 mm. The gap is enclosed at the edge by a 7 mm wide frame. The gap width is kept uniform by polycarbonate spacers (0.8 cm^2) that are

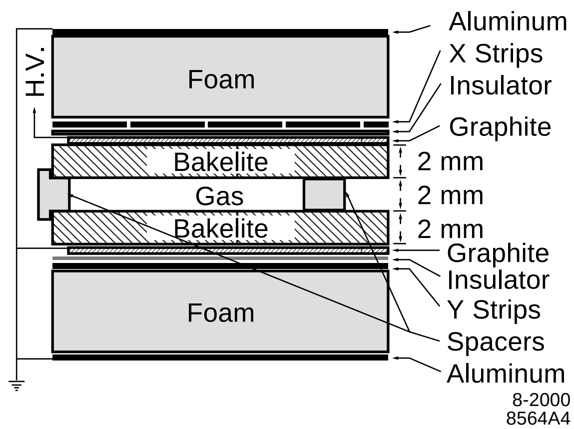

Figure 69. Cross section of a planar RPC with the schematics of the high voltage (HV) connection.

glued to the bakelite, spaced at distances of about 10 cm. The bulk resistivity of the bakelite sheets has been especially tuned to $10^{11} - 10^{12}$ Ωcm. The external surfaces are coated with graphite to achieve a surface resistivity of ≈ 100 kΩ/square. These two graphite surfaces are connected to high voltage (≈ 8 kV) and ground, and protected by an insulating mylar film. The bakelite surfaces facing the gap are treated with linseed oil to improve performance. The modules are operated in limited streamer mode and the signals are read out capacitively, on both sides of the gap, by external electrodes made of aluminium strips on a mylar substrate.

The cylindrical RPCs have resistive electrodes made of a special plastic composed of a conducting polymer and ABS. The gap thickness and the spacers are identical to the planar RPCs. No linseed oil or any other surface treatments have been applied. The very thin and flexible electrodes are laminated to fiberglass boards and foam to form a rigid structure. The copper readout strips are attached to the fiberglass boards.

12.3. RPC Design and Construction

The IFR detectors cover a total active area of about 2,000 m^2 There are a total of 806 RPC modules, 57 in each of the 6 barrel sectors, 108 in each of the four half end-doors, and 32 in the two cylindrical layers. The size and the shape of the modules are matched to the steel dimensions with very little dead space. More than 25 different shapes and sizes were built. Because the size of a module is limited by the maximum size of the material available, i.e. for the bakelite sheets (320×130 cm^2), two or three RPC modules are joined to form a gap-size chamber. The modules of each chamber are connected to the gas system in series, the gas enters in two corners on the same side a short side and exit from two outlets on the opposite side. The high voltage is supplied separately to each module.

In the barrel sectors, the gaps between the steel plates extend 375 cm in the z direction and vary in width from 180 cm to 320 cm. Three modules are needed to cover the whole area of the gap, as shown in Figure 68. Each barrel module has 32 strips running perpendicular to the beam axis to measure the z coordinate and 96 strips in the orthogonal direction extending over three modules to measure ϕ.

Each of the four half end-doors is divided into three sections by steel spacers that are needed for mechanical strength. Each of these sections is covered by two RPC modules that are joined to form a larger chamber with orthogonal readout strips.

The readout strips are separated from the ground aluminium plane by 4 mm thick foam sheet and form strip lines of 33 Ω impedance. The strips are connected to the readout electronics at one end and terminated with a 2 kΩ resistor at the other. Even and odd numbered strips are connected to different front-end cards, so that the failure of a card does not result in a total loss of signal, since a particle crossing the gap typically generates signals in two or more adjacent strips.

The cylindrical RPC is divided into four sections, each covering a quarter of the circumference. Each of these sections has four sets of two single gap RPCs with orthogonal readout strips, the inner with helical $u-v$ strips that run parallel to the diagonals, and the outer with ϕ and z strips. Within each section, the strips in a given readout plane from different modules are connected to form long strips extending over the whole chamber. Details of the segmentation and dimensions can be found in Table 13.

Prior to shipment to SLAC, all RPC modules (equipped with only one readout plane) were tested with cosmic rays. The single rates, dark currents and efficiency were measured as a function of HV. In addition, detailed studies of the efficiency, spatial resolution, and strip multiplicity were performed [2],[3].

After the assembly of RPC modules into gap size chambers, a new series of cosmic rays tests was performed to assure stable and efficient operation. Before the installation of the iron flux return, the planar chambers were then inserted horizontally into the gaps. The cylindrical chambers were inserted after the installation of the solenoid and the EMC.

For each module, test results and conditions are retained in a database, together with records of the critical parameters of the components, the

Table 13
IFR Readout segmentation. The total number of channels is close to 53,000. There are a few deviations from the numbers quoted here, in particular, in layer 19 (due to the external steel structure), and in layer 18 (due to the reduced length). The central chambers in the forward end-door have more vertical strips because of the central hole.

section	# of sectors	coordinate	# of layers	# strips layer/sect	strip length (cm)	strip width (mm)	total # channels
barrel	6	ϕ	19	96	350	19.7-32.8	$\approx 11,000$
		z	19	96	190-318	38.5	$\approx 11,000$
endcap	4	y	18	6x32	124-262	28.3	13,824
		x	18	3x64	10-180	38.0	$\approx 15,000$
cylinder	4	ϕ	1	128	370	16.0	512
		z	1	128	211	29.0	512
		u	1	128	10-422	29.0	512
		v	1	128	10-423	29.0	512

assembly and cabling. In addition, operational data are stored, such as the results of the weekly efficiency measurements that are used in the reconstruction and simulation software.

12.4. Power and Utilities

Once the return flux assembly was completed, the front-end cards (FECs) [4], were installed and services were connected, the low (LV) and high voltage (HV), and the gas system. There are approximately 3,300 FECs, most of them were placed right on top of the RPC modules, inside the steel gap. The remainder were installed in custom crates mounted on the outside of the steel.

Each FEC is individually connected to the LV power distribution. The total power required by the entire system is about 8 kW at +7.0 V and 2.5 kW at -5.2 V. The LV power is supplied by custom built switching devices with load and line regulation to better than 1%. Additional features are precision shunts to measure output currents and TTL logic to inhibit output.

The HV power system is custom adaptation by CAEN [5]. Each HV mainframe can hold up to 10 pods, each carrying two independent 10 kV outputs at 1 mA and 2 mA. The modules are connected to the HV supply which is located in the electronics building via a distribution box. Each distribution box services six RPC modules and up to six distribution boxes are daisy-chained to one pod output. Provisions are made for monitoring the currents drawn by each module. To reduce noise, the RPC ground plane is decoupled from the HV power supply ground by a 100 kΩ resistor.

The RPCs operate with a non-flammable gas mixture containing approximately 56.7% Argon, 38.8% Freon 134a (1,1,1,2 tetrafluoroethane), and 4.5% isobutane. This mixture is drawn from a 760-liter tank that is maintained at an absolute pressure of $1500 - 1600$ Torr. The mixing tank is filled on demand with the three component gases under control of mass-flow meters, each adjusted to provide the desired amount to the mixture. Samples are extracted from the mixing tank periodically and analyzed to verify the correct mixture.

The mixed gas is distributed at a gauge pressure of approximately 6.5 Torr through a parallel manifold system of 12.7-mm-diameter copper tubing. Each chamber is connected to the manifold through several meters of 6 mm-diameter plastic tubing (polyamide or Teflon). The flow to each of these is adjusted individually with a small multi-turn metering valve. Protection against overpressure is provided by an oil bubbler to atmosphere in parallel with each chamber (after the valve), limiting the gauge pressure in the chamber to a maximum of about 1 Torr. Return flow of gas from each chamber is monitored by a second

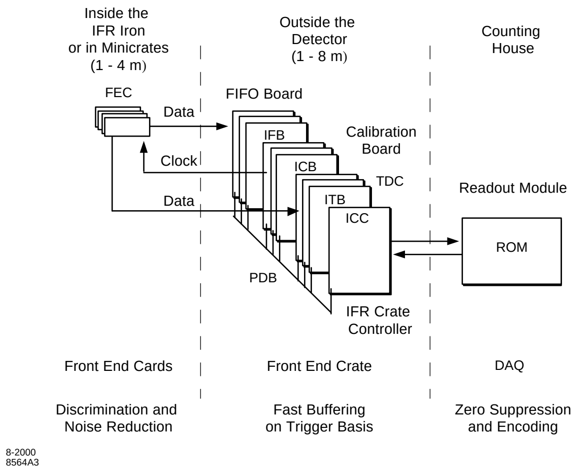

Figure 70. Block diagram of the IFR electronics.

oil bubbler which creates a back pressure of about 0.2 Torr. The total flow through the entire system is approximately 5 ℓ/minute and corresponds on average to two gas exchanges per day.

12.5. Electronics

A block diagram of the IFR electronics system [6] is shown in Figure 70. It includes the front-end cards, the data acquisition, and the trigger.

The detector-mounted FECs service 16 channels. They shape and discriminate the input signals and set a bit for each strip with a signal above a fixed threshold. The input stage operates continuously and is connected directly to the strips which act as transmission lines. A fast OR of all FEC input signals provides time information and is also used for diagnostic purposes. Two types of FECs are employed to handle the inputs of different polarity for signals from the opposite side of the gap. Because of the very low occupancy there is no provision for buffering during the trigger latency [4].

Signals from 3,300 FECs are transmitted to eight custom IFR front-end crates that are located near the detector and from there via a standard G link to four ROMs in the electronics building. Each front-end crate houses up to 16 data handling cards, four trigger cards and a crate controller card (ICC) that collects data from the DAQ cards and forwards them to a ROM. There are three kinds of data cards: the FIFO boards (IFBs) that buffer strip hits, the TDC boards (ITBs) that provide time information, and the calibration boards (ICBs) that inject test pulses into the FECs. To deliver the data and clock signals within the jitter limit to all the boards in the front-end crate, a custom backplane (PDB) for a standard 6U Eurocard crate was designed using 9-layers “strip line” technology. Each board is connected to the ICC via 3 point-to point lines for three single-end signals (data-in, data-out and clock), all of the same length and impedance (50 Ω).

The IFB reads the digital hit patterns from the FECs in less than 2.2 ms, stores the data into FIFOs and transfers FIFO contents into one of the ROMs. Each IFB handles 64 FECs acting as an acquisition master: it receives commands via the PDB, and transmits and receives data patterns from the ROM (via GLINK and ICC). This card operates with the system clock frequency of 59.5 MHz.

The ICB is used for front-end test and calibrations. A signal with programmable amplitude and width is injected into the FEC input stage. To provide timing calibration and to determine the correct readout delay, the board is also used together with the TDCs.

The ICC interfaces the crate backplane with the G link. The physical interface is the FINISAR transceiver, a low cost and highly reliable data link for applications up to 1.5 Gbit/s.

The TDC board exploits the excellent time resolution of the RPCs. Each board has 96 ECL differential input channels, reading fast OR signals from the FECs. Time digitization is achieved by three custom TDCs, designed at CERN [7]. Upon receipt of a L1 trigger, data are selected and stored until readout by the ROM. The intrinsic resolution of the board is better than 200 ps. The 59.5 MHz clock signal is synchronized with the data and distributed to the 16 boards. High performance drivers are used to provide a reliable clock distribution with a jitter of less than 0.5 ns.

12.6. Slow Controls and On-line Monitoring

The IFR is a system with a large number of components and electronics distributed all over the *BABAR* detector. To assure safe and stable operation, an extensive monitoring and control system was installed and has been in operation from the start. The IFR On-line Detector Control IODC) monitors the performance of the RPCs by measuring the singles counting rate and the dark current of every module. It also controls and monitors the operation of the electronics, the DAQ and trigger, as well as the LV, the HV, and the gas system. The total number of hardware channels is close to 2,500, and the hardware is installed in 8 custom slave DAQ crates [8]).

The system has been easy to operate. HV trips are rare. Temperature monitoring in the steel structure and the electronics crates has proven very useful for the diagnosis of operational problems. The occupancy is extremely low everywhere but in layer 18 of the forward end-door which lacks adequate shielding from machine generated background. On average, there are about 100 strip hits per event.

12.7. Efficiency Measurements and Performance

The efficiency of the RPCs is evaluated for both normal collision data and cosmic ray muons recorded with the IFR trigger. Every week cosmic ray data are recorded at different voltage settings and the efficiency is measured chamber by chamber as a function of the applied voltage. The absolute efficiency at the nominal working voltage (typically 7.6 kV) is stored in the database for use in the event reconstruction software.

To calculate the efficiency in a given chamber, nearby hits in a given layer and hits in different layers are combined to form clusters. Two different algorithms are used: The first is based solely on the IFR information and uses data recorded with a dedicated FIR trigger; the second matches the FIR clusters with the tracks reconstructed in the drift chamber. Both these algorithms start from one-dimensional IFR clusters defined as a group of adjacent hits in one of the two readout coordinates. The cluster position is defined as the centroid of the strips in the cluster. In the first algorithm, two-dimensional clusters are formed by joining one-dimensional clusters (of the same readout coordinate) in different layers, provided the distance between their coordinate centroids is less then a given value. In each sector, two-dimensional clusters in different coordinates are combined to three-dimensional clusters as long as there are less than three layers missing one of the two coordinates. The second algorithm extrapolates charged tracks reconstructed by the drift chamber into the FIR. FIR clusters which are less than 12 cm from the extrapolated track are combined to form three-dimensional or two-dimensional clusters. A detailed discussion of the clustering algorithm can be found elsewhere [9].

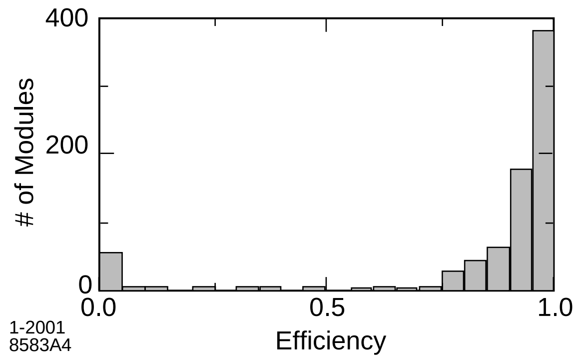

Figure 71. Distribution of the efficiency for all RPC modules measured with cosmic rays in June 1999. Some 50 modules were not operational at that time.

The residual distributions from straight line fits to two-dimensional clusters typically have an rms width of less than 1 cm. An RPC is considered efficient if a signal is detected at a distance of less than 10 cm from the fitted straight line in either of the two readout planes. Following the installation and commissioning of the IFR system, all the PC modules were tested with cosmic rays and their efficiency was measured. The results are presented in Figure 71. Of the active RPC modules, 75% exceed an efficiency of 90%.

Early tests indicated that the RPC dark current was very temperature dependent, specifi-

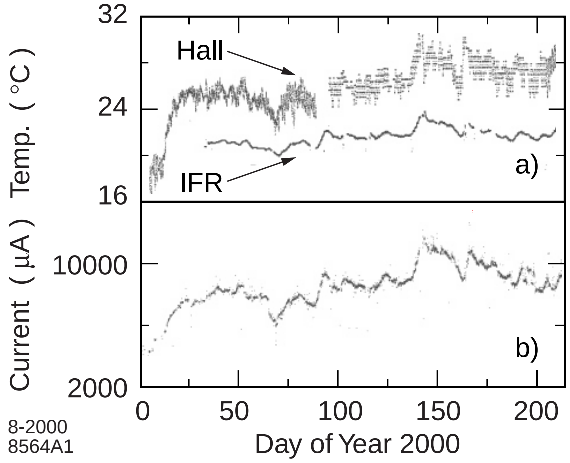

Figure 72. History of the temperature and dark current in the RPC modules since January 2000. top: temperature in the IR-2 hall and in the backward end-door; bottom: total dark current in the 216 modules of the backward end-door.

cally, the current increases $14 - 20\%$ per °C. Because the IR experimental hall does not have temperatures regulation this presents a serious problem. The FECs that are installed in the steel gaps each produce 3 W, adding up to a total power dissipation of 3.3 kW in the barrel and 1.3 kW in the forward end-door.

During the first summer of operation the daily average temperature in the IR hall was 28°C and the maximum hall temperature frequently exceeded 31°C. The temperature inside the steel rose to more than 37°C and the dark currents in many modules exceeded the capabilities of the HV system and some RPCs had to be temporarily disconnected.

To overcome this problem, water cooling was installed on the barrel and end-door steel, removing ≈ 10 kW of heat and stabilizing the temperature to $20 - 21$°C in the barrel, 22°C in the backward and 24°C in the forward end-doors. Figure 72 shows the history of temperature in the hall and temperature and total dark current in the backward end-door. While the current closely follow the temperature variations, the range of

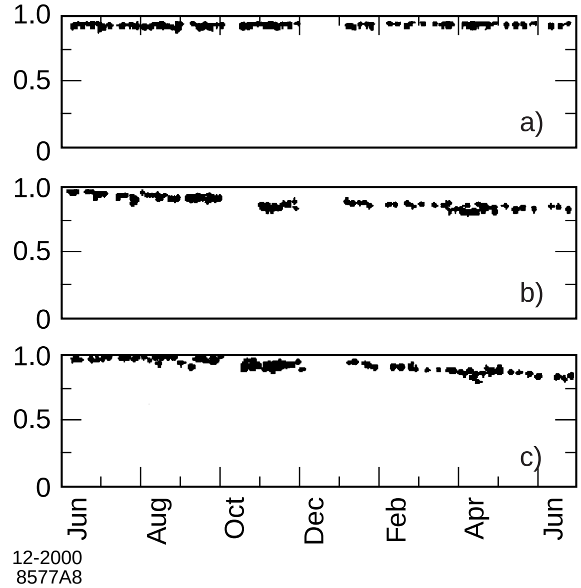

Figure 73. Efficiency history for 12 months starting in June 1999 for RPC modules showing different performance: a) highly efficient and stable; b) continuous slow decrease in efficiency; c) more recent, faster decrease in efficiency.

change is now limited to a few degrees.

During the operation at high temperatures a large fraction of the RPCs ($> 50\%$) showed not only very high dark currents but also a some reduction in efficiency compared to earlier measurements [10]. After the cooling was installed and the RPCs were reconnected, some of them continued to deteriorate while others remained stable, some of them ($> 30\%$) at full efficiency. (see Figure 73). Detailed studies revealed large regions of very low efficiency in these modules, but no clear pattern was identified.

The cause of the efficiency loss remains under investigation. Several possible causes have been excluded as the primary source of the problem, such as the change in the bakelite bulk resistivity, loosened spacers, gas flow or composition. A number of prototype RPCs developed similar efficiency problems after being operated above a temperature 36°C for a period of two weeks. In some of these modules, evidence was found that the linseed oil had failed to cure and had accumulated at various spots under the impact of the high voltage.

12.8. Muon Identification

While muon identification relies almost entirely on the IFR, other detector systems provide complementary information. Charged particles are reconstructed in the SVT and DCH and muon candidates are required to meet the criteria for minimum ionizing particles in the EMC, i.e. tracks depositing large amounts of energy are rejected. Charged tracks that are reconstructed in the tracking systems are extrapolated to the IFR taking into account the non-uniform magnetic field, multiple scattering and the average energy loss. The projected intersections with the RPC planes are computed and for each readout plane all clusters detected within a predefined distance from the predicted intersection are associated with the track.

For each cluster in the IFR associated with a charged track a number of variables are defined to discriminate muons from charged hadrons: 1) the total number of interaction lengths traversed from the IP to the last RPC layer with an associated cluster, 2) the difference between this measured number of interaction lengths and the number of interaction lengths predicted for a muon of the same momentum and angle, 3) the average number and the rms of the distribution of RPC strips per layer, 4) the χ^2 for the geometric match between the projected track and the centroids of clusters in different RPC layers, and 5) the χ^2 of a polynomial fit to the 2-dimensional IFR clusters. These variables will be combined into global likelihood analysis to optimize the efficiency and purity of muons. At present, cuts on individual variables are employed.

The performance of muon selection has been tested on samples of muons from $\mu\mu ee$ and $\mu\mu\gamma$ final states and pions from 3-prong τ decays and $K_S \rightarrow \pi^+\pi^-$ decays. The selection of these control samples is based on kinematic variables, and not on variables used for muon selection. As illustrated in Figure 74, a muon detection efficiency of close to 90% has been achieved in the momen-

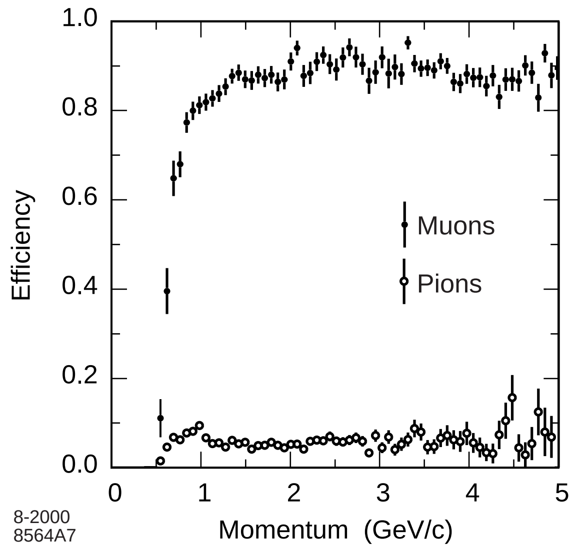

Figure 74. Muon efficiency and pion misidentification probability as a function of the track momentum, otained with loose selection criteria.

tum range of $1 < p < 3$ GeV/c with a fake rate for pions of about 5%. The pion misidentification can be reduced by a factor of two by tighter selection cuts which lower the detection efficiency to about 80%. Above 1 GeV/c decays in flight contribution 1.5% to the pion misidentification.

12.9. K_L^0 and Neutral Hadron Detection

K_L^0 and other neutral hadrons interact in the iron of the IFR and can be identified as clusters that are not associated with a charged track. Monte Carlo simulations predict that about 64% of K_L's above a momentum of 1 GeV/c produce a cluster in the cylindrical RPC, and/or a cluster with hits in two or more planar RPC layers.

Unassociated clusters that have an angular separation of ≤ 0.3 rad are combined into a composite cluster, joining clusters that originate from showers that spread into adjacent sectors of the barrel, several sections of the end-doors and/or the cylindrical RPC. This procedure also combines multiple clusters from large fluctuations in the hadronic showers. The direction of the neutral hadron is determined from the event vertex

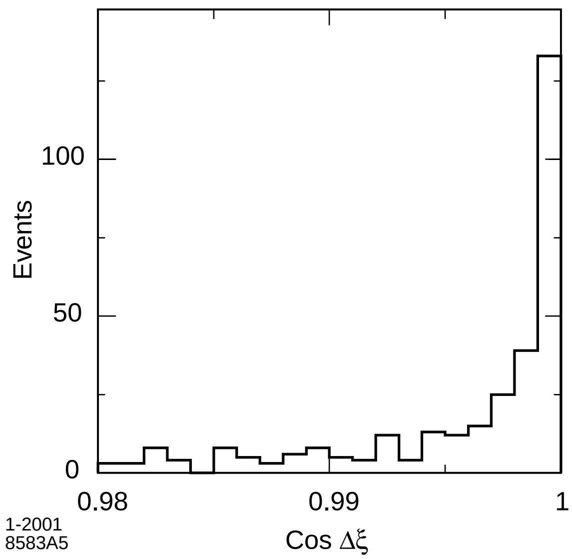

Figure 75. Angular difference, $\cos\Delta\xi$, between the direction of the missing momentum and the closest neutral IFR cluster for a sample of ϕ mesons produced in the reaction $e^+e^- \to \phi\gamma$ with $\phi \to K_L^0 K_S^0$.

and the centroid of the neutral cluster. No information on the energy of the cluster can be obtained.

Since a significant fraction of the hadrons interact before reaching the IFR, information from the EMC and the cylindrical RPCs is combined with the IFR cluster information. Neutral showers in the EMC are associated with the neutral hadrons detected in the IFR, based on a match in production angles. For a good match a χ^2 probability of $\geq 1\%$ is required.

An estimate of the angular resolution of the neutral hadron cluster can be derived from a sample of K_L^0 produced in the reaction $e^+e^- \to \phi\gamma \to K_L^0 K_S^0 \gamma$. The K_L^0 direction is infered from the missing momentum computed from the measured particles in the final state, γ and K_S^0. The data in Figure 75 indicate that the angular resolution is of the order of xxx °.

For multi-hadron events with a reconstructed J/Ψ decay, Figure 76 shows the angular difference, $\Delta\phi$, between the missing momentum and

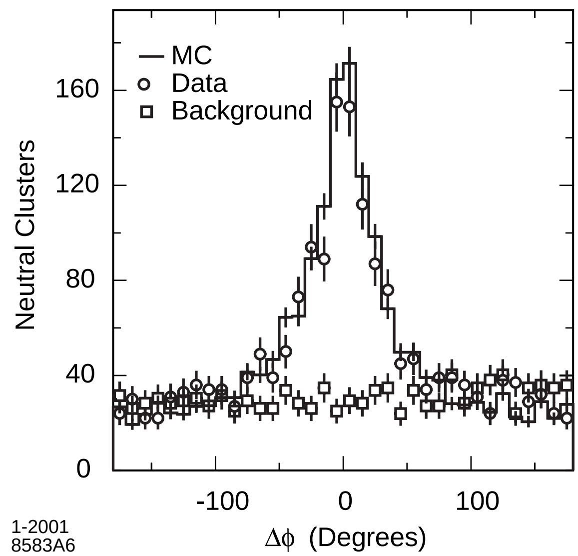

Figure 76. Difference between the direction of the reconstructed neutral hadron cluster and the transverse missing momentum of the event. The Monte Carlo simulation is nomalized to the luminosity of the data; the background is obtained using neutral hadrons and the missing momentum from different events.

the direction of the nearest neutral hadron cluster. The observed peak demonstrates clearly that the missing momentum can be associated with a neutral hadron, assumed to be a K_L^0.

????? estimate of efficiency and resolution for KL ???

12.10. Summary and Outlook

The IFR is the largest RPC system built so far. It provides efficient muon identification and allows for the detection of K_L^0 interacting in the steel and the calorimeter. During the first year of operation a large fraction of the RPC modules have suffered significant losses in efficiency. This effect appears to be correlated with high temperatures, but the full extent of the problem and its cause remain under study. Thanks to the large number of RPC layers, this problem has not yet impacted the overall performance too severely. But current extraplolations, even after installation of water cooling on the steel, indicate a severe problem for the future operation. Recently, 24 end-door modules have been replaced by new RPCs with improved treatment of the bakelite surfaces that should lead to full curing of the linseed oil coating. Results from this new production and other tests will need to be evaluated before decisions on future improvements of the IFR can be made.

REFERENCES

1. R. Santonico, R. Cardarelli, Nucl. Instr. and Methods **A**187 (1981) 377.
2. F. Anulli *et al.* Nucl. Instr. and Methods **A**409 (1998) 542.
3. A. Calcaterra, *et al.* Performance of the *BABAR* RPCs in a Cosmic Ray Test in *Proceedings of the International Workshop on Resistive Plate Chambers and Related Detectors*, Naples, Italy (1997).
4. N.Cavallo, *et al.* Nucl. Phys. **B(**Proc. Suppl.) 61B (1998) 545; N. Cavallo, *et al.*, Nucl. Instr. and Methods **A 409** (1998) 297.
5. HV system SY-127, Pod Models A300-P and A300-N, by CAEN, Viareggio, Italy.
6. G. Crosetti *et al.* Data Acquisition System for the RPC Detector of *BABAR* Experiment in *Proceedings of the International Workshop on Resistive Plate Chambers and Related Detectors*, Naples, Italy (1997).
7. TDC custom chip INFO NEEDED!
8. P. Paolucci, The IFR Online Detector Control System, SLAC PUB-8167 (1999).
9. L. Lista, Object Oriented Reconstruction Software for the IFR Detector of *BABAR* Experiment in *Proceedings of the Conference on Computing in High Energy Physics*, Padova, Italy (2000).
10. A. Zallo *et al.* The *BABAR* RPC System in *Proceedings of the 5th International Workshop on Resistive Plate Chambers and Related Detectors*, Bari, Italy (1999).

13. Trigger

13.1. Trigger design requirements

The most basic requirements for the trigger system are to select the physics events of interest (Table 14) with a high, stable and well-understood efficiency while keeping the total event output rate to permanent storage under 120 Hz at the design luminosity of 3×10^{33} cm^{-2}s^{-1}. This total includes some prescaled samples of events, such as ones failing the trigger conditions or from random beam crossings, which are required for diagnostic and background studies.

The actual efficiency requirement depends on the physics channel being considered. The total trigger efficiency is required to be at least 99% for all $B\overline{B}$ events and at least 95% for continuum events, which are required for background subtraction. Less stringent requirements are put on the efficiencies of other physics channels of interest, e.g. the efficiencies for fiducial τ pair events are required to be 90-95%, depending on decay channels.

The trigger is required to be robust and flexible in order to achieve stable operation even under unexpected situations. It is designed to be able to operate at up to ten times the expected nominal PEP-II background rates and to degrade slowly for backgrounds above that level. It is also required to be able to operate when some of the detector channels are dead or noisy. Additionally, the trigger should not cause more than 1% deadtime.

Table 14
Physics event production and trigger rates at $\Upsilon(4S)$ and a luminosity of $\mathcal{L} = 3 \times 10^{33}\,\mathrm{cm^{-2}s^{-1}}$. The e^+e^- cross section refers to fiducial events with either e^+ or e^- inside EMC detection volume. [Numbers need check]

Event type	Cross-section (nb)	Production Rate (Hz)	Level 1 Trigger Rate (Hz)
$B\overline{B}$	1.05	3.2	3.2
$u\overline{u}+d\overline{d}+c\overline{c}+s\overline{s}$	3.39	10.2	10.1
e^+e^-	˜ 53	159	156
$\mu^+\mu^-$	1.16	3.5	3.1
$\tau^+\tau^-$	0.94	2.8	2.4

The trigger is implemented in two levels; the Level 1 (L1) hardware and Level 3 (L3) software triggers. The underlying concept is that the Level 1 is as open as possible to physics, after which Level 3 selects the events of most interest, consistent with the desired output rate.

As an indication for the level of background the Level 1 need to reduce from, the expected rate of background interactions producing at least one track in the drift chamber with $p_{\rm t} > 120\,{\rm MeV}/c$ or at least one cluster in the calorimeter with $E > 100\,{\rm MeV}$ is $\sim 10\,{\rm kHz}$ each at a luminosity of 3×10^{33}. The Level 1 is required to have an output rate of less than 2 kHz. It needs to operate with a latency fixed within a defined range; for all events which are triggered, 99% of the trigger signals have to be delivered to the Fast Control and Timing System (FCTS) within the time window between 11 and $12\,\mu$s after the event occured. All parts of the Level 1 are also required to read out sufficient trigger data for each event to allow offline calculation of the efficiencies from the data themselves.

The Level 3 performs the second stage of rate reduction from the maximum Level 1 output rate to a maximum final rate of 120 Hz. It must reject events from a number of background processes which are not of physics interest. In addition, it must flag events needed for luminosity, diagnostics and calibration purposes. It is required that only 90 Hz of the output rate is used by physics events, with the remaining 30 Hz used for this other information. Finally, Level 3 software is required to comply with the *BABAR* software conventions and standards; in particular, the code has to usable in both online and offline systems to reduce duplication of effort and also provide consistency.

13.2. Level 1 trigger system

The Level 1 trigger selects events and achieves its rate reduction based on a fraction of the detector data processed by specialized hardware. Hits in the drift chamber, energy deposition in the electromagnetic calorimeter and hit occupancy in the instrumented flux return systems are used to quickly identify particle tracks, the presence of deposited energy, and the presence of muons, respectively. The drift chamber trigger (DCT) and electromagnetic calorimeter trigger (EMT) are both designed to independently satisfy all triggering requirements with high efficiency, thereby providing the needed degree of redundancy. The instrumented flux return trigger (IFT) is used for triggering μ-pair and cosmic ray events, and for diagnostic purposes.

Each of the three Level 1 trigger processors generates trigger "primitives" (summary data on the position and energy of particles) that is sent to the global trigger (GLT) which then forms 24 discreet triggers (which may be "true" or "false"). The FCTS, receiving these 24 signals, can optionally select a subset of triggers with a mask, prescale some of the triggers and — if a valid trigger remains — issue a Level 1 Accept to initiate event readout. The trigger definition logic, masks and prescale values are all configurable on a per run basis.

The overall structure of the Level 1 system is illustrated in Figure 77. All the Level 1 hardware shown in this diagram is hosted in five 9u VME crates. The Level 1 system operates in a continuous sampling mode of processing input data and generating output trigger information at fixed time intervals. The DCH frontend electronics and the EMC UPCs send raw data to the DCT and EMT $\sim 2\,\mu$s after the event time. The DCT and EMT event processing time are 4–5 μs and it is followed by another $\sim 3\,\mu$s in the GLT to issue an Level 1 trigger. The Level 1 trigger takes $\sim 1\,\mu$s to propagate through the fast control system and the ROMs to generate a readout just before the event data reaching the end of the $12.8\,\mu$s detector frontend buffer.

The basic trigger primitives generated by the DCT, EMT and IFT are mostly only the azimuthal angle (ϕ) projections of the particle raw trigger signals, with only a few cases of very coarse polar angle (θ) information, such as barrel and endcap distinctions. The trigger primitives are processed in the GLT to generate up to 24 Level 1 trigger lines which are passed on to the FCTS. The DCT and EMT primitives sent to the GLT are ϕ maps signalling tracks or energy deposits in the various ϕ regions. The IFT primitive is a three-bit pattern representing sex-

Table 15
Trigger primitives. Most energy thresholds are adjustable and the listed thresholds are typical values. The detailed defintion for the IFT U pattern can be found in Table 16.

	Description	Origin	No. of bits	Threshold
B	Short track reaching DCH superlayer 5	BLT	16	120 MeV/c
A	Long track reaching DCH superlayer 10	BLT	16	180 MeV/c
A′	High p_{t} track	PTD	16	800 MeV/c
M	All-θ MIP energy	TPB	20	100 MeV
G	All-θ intermediate energy	TPB	20	250 MeV
E	All-θ high energy	TPB	20	700 MeV
X	Forward endcap MIP	TPB	20	100 MeV
Y	Backward barrel high energy	TPB	10	1 GeV
U	Muon IFR sextant hit pattern	IFS	3	

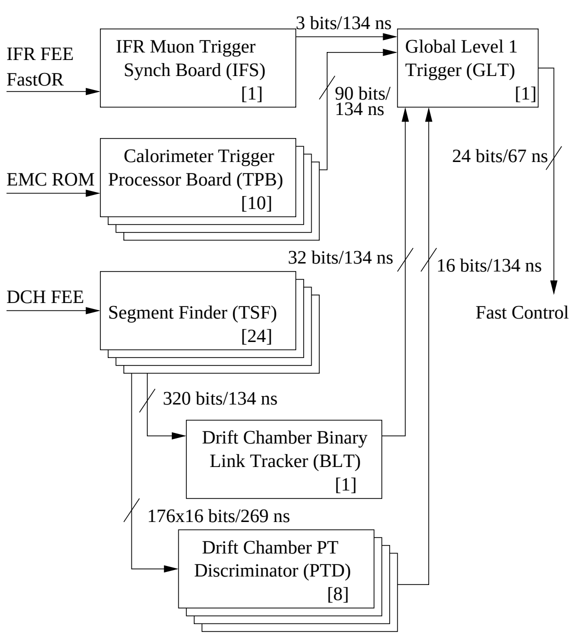

Figure 77. Level 1 system schematics for the major L1 components. The numbers on the links between different components are the tansmission rates of total signal bits.

tant hit topology in the IFR. The meanings of the various trigger primitive inputs to the GLT are summarized in Table 15. The DCT, EMT and IFT primitives are all time-stretched to reflect the uncertainty in their time determination before being sent to the GLT.

The DCT, EMT and GLT contain 4-event buffer DAQ readout capabilities similar to other detector system front-end DAQ designs. The DAQ data contain information from various stages of the trigger for each event and are used for monitoring and determining trigger efficiencies, as well as providing seed information for Level 3.

13.2.1. Level 1 drift chamber trigger

The input data from the DCH front-end electronics (FEE) to the DCT consist of one bit, updated every 269 ns, for each of the 7104 DCH cells (see section 6.4 for details). These bits convey time information from amplitude discriminators whose input is the wire signal for that cell. The DCT outputs primitives consisting of three ϕ maps of 16 bits each, as listed in Table 15, which are sent to the GLT every 134 ns.

From the input signal bits, the DCT continuously links wire hits into "segments", a series of associated hits in one DCH superlayer. It then chains segments into tracks, and finds the ϕ position at superlayer 5 (10) for a short (long) track. It also determines whether the tracks represent particles having transverse momentum (p_{t})

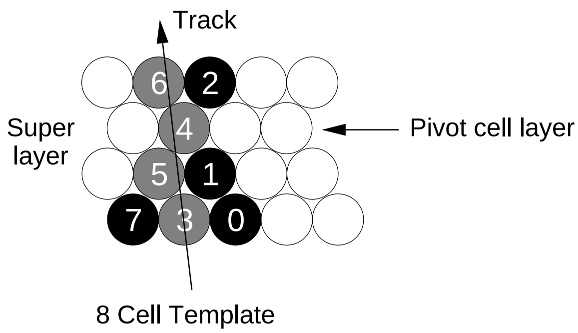

Figure 78. Track Segment Finder pivot cell group.

greater than a preset value. The algorithms of the DCT are executed in three types of modules. First, the track segments and ϕ positions are found using a set of 24 Track Segment Finder (TSF) modules [1]. The segments are then passed to the Binary Link Tracker (BLT) module [2], where they are linked into complete tracks. In parallel, detailed ϕ information for segments found in axial superlayers is transmitted to the eight PT Discriminator (PTD) modules [3], which determine if the segments are consistent with tracks of particles having a p_{t} greater than a user-specified minimum.

Each TSF is responsible for processing a subset of the DCH input data and extracting track segments from a set of contiguous hits within a "pivot group", see Figure 78. For any pivot group, the cells are numbered 0 through 7, with cell 4 being the pivot cell. The shape of a pivot group was chosen such that only reasonably straight tracks originating from the interaction point can produce a segment. Each of the 24 TSF modules consists of 72-75 track segment engines, one for each pivot group, which amounts to a total of 1776 pivot groups. The engine processes the data from the eight cells in its assigned pivot group to determine whether a grouping of hit cells can produce a track segment. Depending on where a track passes through a particular cell, the resulting ionization will take one to four 269 ns clock ticks to drift to the signal wire. It is this time delay, or drift time, that the TSF uses to more accurately establish the position of the track as well as the event time. Typically, valid segment patterns consist of hits, close together in time and in at least three out of four layers within the superlayer (to account for cell inefficiencies). At the hardware level, this is implemented using a self-triggered two-bit counter for each of the eight cells in the pivot group. A counter is enabled when a hit is registered. By incrementing the counter every 269 ns, a 16-bit address, corresponding to 65,536 possibilities, is obtained at any given clock tick. Each non-zero address is then translated by the pre-loaded look-up table (LUT) into a two-bit weight indicating whether there is a four-layer segment, a three-layer segment, or a low-quality (un-calibrated) segment. A three-layer segment with the pivot cell hit missing is also allowed. The pivot group is monitored to determine which of the three subsequent clock ticks produces the highest weight, or "best" pattern. This time information is used to align each segment in time and to create a coincidence window for all segments related to the same event. The contents of the LUT are derived from an offline calibration and consist of high-precision position and event time information. The segment arrival time is estimated mainly based on the variation of weight in time. The resulting time-adjusted weight and position are then passed on to the BLT and PTD's. The data sent to the BLT are the 'coarse-ϕ' data with just the wire address, while the 'fine-ϕ' data sent to the PTD's contain the high resolution position information. Since the TSF segment arrival time estimate combining information of several hits is considerably more precise than the 269 ns input time interval, the TSF data for BLT is transmitted at 134 ns intervals to gain on timing resolution.

The position resolution as measured from the data after calibration, is $\sim$600 μm for a 4-layer segment and $\sim$900 μm for a 3-layer segment typically. For tracks from the IP, the efficiency for all TSF segments is 97%, and the efficiency for calibrated 3-layer or 4-layer TSF segments is 94%.

The BLT receives segment hit information from all 24 TSF's, corresponding to the entire DCH, at a rate of 320 bits every 134 ns and links them into complete tracks. The segment hits are mapped onto the DCH geometry in terms of "supercells", with 10 radial superlayers and 32 sectors in ϕ, and each bit indicates whether a segment is found in that supercell or not. The input data to the BLT are combined using a logical OR with a programmable mask pattern. The masking capability allows the system to activate track segments that correspond to dead or highly inefficient cells so that the track efficiency does not degrade. The linking algorithm uses an extension of a simple method, developed for the CLEO-II trigger [5], to link the segments into a continuous track. It starts from the innermost superlayer, A1, and moves radially outward. A track is found if there is a segment hit in every layer and if the segments in two consecutive superlayers are within a certain number of supercells (three or five depending on the superlayer type) of each other. This allows for track curvature in the magnetic field and dip angle variations. Up to two superlayers are allowed to be missing. Tracks that reach the outer layer of the DCH (superlayer A10) are classified as type A. Tracks that reach the middle layer (superlayer U5) are classified as type B. The data are then compressed and output to the GLT in the form of two words of 16 bits each corresponding to A and B tracks. Each bit in a word represents the "track hit" state of a supercell in the designated superlayer.

The eight PTD's receive the fine-ϕ information on track segments from axial superlayers only, and determine if the segments are consistent with tracks having a $p_{\rm t}$ greater than some configurable minimum. An envelope for tracks above the minumum $p_{\rm t}$ is defined using the interaction point (IP) and a track segment position in one of the "seed" superlayers, A7 or A10. A high $p_{\rm t}$ A$'$ candidate is identified when there are sufficient track segments with accurate ϕ information lie inside this envelope.

The processing on each PTD is subdivided in eight processing engines, one for each supercell in each of the two superlayers A7 and A10. The principal components in each engine are an algorithmic processor and LUT's containing the limits for each individual seed position. The contents of the LUT's thus specify the allowed track segment positions for each of the three other axial superlayers and consequently define the effective $p_{\rm t}$ discrimination threshold.

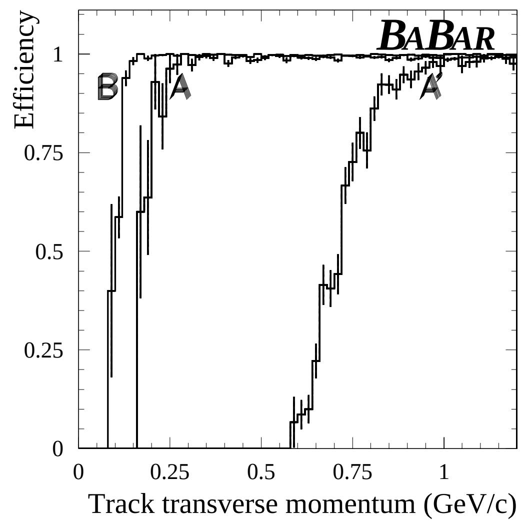

Figure 79. DCT track efficiency vs. $p_{\rm t}$, where the A$'$ threshold is set to 800 MeV/c.

The resulting $p_{\rm t}$ turn-on for the PTD A' tracks is shown in Figure 79 together with the BLT A,B track efficiency.

Each of the three main boards relies heavily on multiple FPGA's [4] which perform the on-board control and algorithmic functions. All cabling is handled by a small (6u) back-of-crate interface behind each main board.

13.2.2. Level 1 calorimeter trigger

The input data for the EMT are 280 "tower sums" in the calorimeter. The EMC barrel consists of an array of 48×120 $(\theta \times \phi)$ crystals. These are grouped into "towers" of 8×3, giving 6 towers in θ by 40 in ϕ. Each barrel tower therefore corresponds to 24 crystals. The endcap is divided into 40 wedges in ϕ, making a further 40 towers, each containing between 19 and 22 crystals. The towers thus form an array of 7×40 covering the whole of the EMC. The energies of crystals in each tower, if above a threshold of 10 MeV, are summed and sent to the EMT as an unsigned 16-

bit word every 269 ns on dedicated cables from the EMC UPC's.

The EMT primitives consist of five types of ϕ maps, as listed Table 15, which are sent to the GLT every 134 ns. Further details of the EMT system can be found in [6].

For each of the 40 ϕ positions, the EMT algorithm sums the seven tower energies in θ and sends the resulting sum to a eight-tap finite impulse response (FIR) filter *[reference needed]*. The FIR output zero-crossing time is used as the estimate of the time of the deposit. The energy is determined from this zero-crossing by a time offset, tuned to coincide with the peak of the shaped EMC pulse. This energy is compared with five thresholds, one for each of the ϕ maps, and the corresponding primitive bit is set if the energy is above the threshold.

The conversion of the tower data into the GLT ϕ bits is performed by ten Trigger Processor Boards (TPB). Each TPB receives data from 28 towers, corresponding to a strip of 7×4 in $\theta \times \phi$, and so performs the algorithm for four ϕ positions. In addition, data from the towers bordering this strip are needed. Each TPB receives summary data from another TPB for the seven towers in θ which are at the ϕ position immediately next to its 28 towers. Each TPB also outputs the equivalent data to another TPB for the seven towers in θ at the highest ϕ position of its 28 towers. These "nearest neighbour" data allow energy deposits which overlap two TPB's to be accurately reconstructed. The complete algorithm is implemented in one FPGA [7] for each ϕ position, with four identical components per TPB.

The basic performance of the EMT can be expressed through the efficiency and timing jitter of the trigger primitives. The efficiency of the primitives can be measured by the number of times a trigger bit is set as a function of the actual energy reconstructed offline in the EMC. Figure 80 shows this efficiency for energies around the M threshold. The trigger bit changes from 10% to 90% efficiency over a range from 110 to 145 MeV and is at the full efficiency of 99% by the average minimum ionizing particle (MIP) energy of 180 MeV.

The EMT time jitter with respect to the energy deposit time is measured by comparing the FIR output time offline with the DCH track start time determination, t_0. The difference has an RMS of 150 ns, and 97% are within the ± 500 ns window. The event timing is considerably more precise when averaging over many hits.

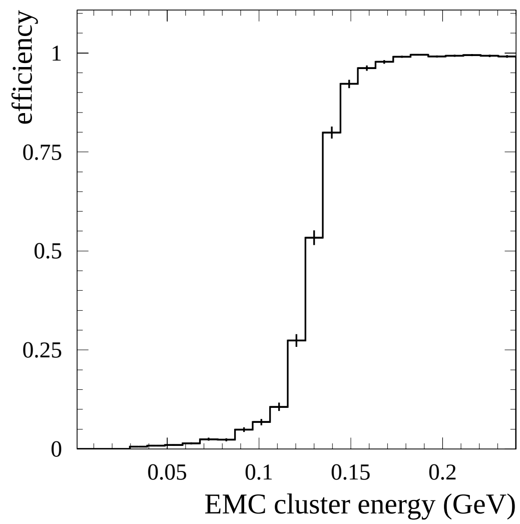

Figure 80. EMT M efficiency vs. EMC cluster energy for an M threshold setting of 120 MeV.

13.2.3. Level 1 IFR trigger

The IFT is used for triggering on μ pairs and for diagnostics. For the purposes of the trigger, the IFR is divided in ten sectors, namely the six barrel sextants and the four half-endcap doors. The inputs to the IFT are the "Fast OR" signals which cover all ϕ strips from eight selected layers in each sector. The output to the GLT consists of three bits every 134 ns.

A majority logic algorithm defines trigger objects for every sector if at least four of the eight trigger layers have hits inside a time window of 134 ns. The IFR Trigger Synchronization (IFS) module processes the trigger objects from the ten sectors and generates the three-bit trigger word (U) encoding seven possible exclusive trigger conditions, as defined in Table 16. The trigger $U \geq 5$, for example, covers all the 2-prong μ-pair topologies.

Table 16
IFR trigger pattern (U) definition.

U	Trigger condition
1	other $\geq$2 μ topologies not in U=5-7
2	1 μ in backward endcap
3	1 μ in forward endcap
4	1 μ in barrel
5	2 back-back μ's in barrel + 1 forward μ
6	1 μ in barrel + 1 forward μ
7	2 back-back μ's in barrel

The efficiency of the IFT has been evaluated with cosmic rays triggered by the DCT and crossing the detector close to the interaction point. For these events, 98% have been triggered by the IFT as events with at least one track, and 73% as events with two tracks, inside the geometrical region of the IFR. Most of the inefficiency is concentrated in the angular regions between two adjacent sextants and between the barrel and endcaps.

13.2.4. Level 1 global trigger

The GLT receives inputs consisting of the primitives from the DCT, EMT and IFT, as listed in Table 15, every 134 ns. Due to the different latencies of the primitives, the GLT first aligns the input signals using configurable delays. It then forms some additional combined ϕ maps from the DCT and EMT data. A total of 16 ϕ maps are input to a LUT which treats each ϕ map as an input address and a three-bit output for each trigger object which counts, e.g., the number of B tracks or number of M clusters. To count as distinct trigger objects, the map bits are typically required to have a separation of more than one ϕ bin.

These 16 counts not only include the standard inputs of A, B, A$'$, M, G, E and Y, but also matched objects such as BM for B tracks matched to an M cluster in ϕ. B and M are also used to derive additional back-to-back objects, B* and M*, where the count requires a pair of ϕ bits separated by a configurable angle of typically $\sim 120°$. E and M are also used to make an EM* object for back-to-back EM pairs.

The 16 counts plus the IFT hit pattern are then used to form 24 raw trigger lines. The final stage logic selects the highest priority trigger to determine the Level 1 trigger time and latches the other triggers compatible in time for the 24 line trigger output. The trigger decision logic for each trigger line is a logical AND of the 17 outputs from operations applied to the object counts. The decision operations on each object can be: always-pass, or $\geq$, = or < to a configurable cut parameter. The trigger logic evaluates the time-centroid of the highest priority line with a 67 ns sampling period based on the last $\sim$1 μs of history. This faster sampling rate can improve output trigger timing resolution in the case of sending output at the time-centroid of a trigger which is on for an odd number of 134 ns intervals. If a higher priority trigger line starts within this time, the process resets to follow the higher priority line instead. There is an optional configurable output delay for each line to control the Level 1 trigger to be set as late as allowed by its timing jitter so as to minimize the chance of early background hits poisoning real signals in the SVT single hit electronics.

The 24 bit GLT output signal is sent to the fast control system every 67 ns, which handles trigger mask selection and prescaling. The achieved timing resolution for hadronic data events has an RMS of 52 ns and 99% of the events are within 77 ns.

The GLT consists of a single 9u VME module. Most of the logic, including diagnostic and DAQ memories, are implemented in FPGA's [4]. The LUT section is an array of 16 memory chips with 8 Mbytes of VME downloadable configuration data.

13.3. Level 1 trigger performance and operational experience

The Level 1 trigger configuration consists of pure DCT, pure EMT, mixed and prescaled triggers, not only aimed for maximum efficiency and background suppression, but also for the convenience of trigger efficiency determination. The trigger lines are designed to collectively preserve the many types of physics interactions. Although most trigger lines have as their primary target a specific physics source, they are not strictly classified to only serve that source. For example, some of the two-particle triggers are not only efficient for Bhabha, $\mu\mu$ and $\tau\tau$ events, but are also useful for selecting jetlike hadronic events and some rare B decays.

The efficiencies and rates of some selected Level 1 trigger lines are listed in Table 17 and for various physics processes. Although triggering on generic $B\overline{B}$ events is relatively easy, it is essential to make sure high efficiencies are also maintained for the important rare B decay processes which are typically less distinctive in decay multiplicity. We therefore include efficiencies for events containing two bench mark processes of $B^0 \to \pi^0\pi^0$ and $B^- \to \tau\overline{\nu}$ (while the other B in the event decay generically) in Table 17.

The efficiencies for the hadronic events are absolute efficiencies including acceptance losses, based on MC simulation including local inefficiency effects calibrated using single particle triggering efficiencies from data. The efficiencies for τ-pair events are for MC "fiducial" events, having two or more tracks with $p_{\rm t} > 120\,{\rm MeV}/c$ originating close to the IP and with θ in the range to reach at least DCH superlayer U5. The Bhabha and μ-pair efficiencies are determined from the data, for events with 2 high momentum particles back to back in e^+e^- center of mass frame, and within the EMC fiducial volume. It can be seen that with the exception of $\tau\tau$ events, there are efficient orthogonal triggers using only DCT or only EMT/IFT for these processes. The efficiencies predicted by the MC are generally in good agreement with data when tested using events passing typical analysis selections and based on orthogonal triggers. Prescaled triggers with a very open acceptance of physics events, such as (B$\geq$2 & A$\geq$1) or (M$\geq$2) are also used to measure the trigger efficiencies.

The trigger rates listed in Table 17 are for a typical run with HER (LER) currents at 650 mA (1350 mA) and a luminosity of 2.2×10^{33} $\rm cm^{-2}s^{-1}$. These are stable to within 20% for the same PEP-II setup. Longer term they can be 50% higher after a major shutdown, when starting with a poor vacuum, and gradually improve in the following months. There are also occasional background spikes which can double the Level 1 rate. However, even many of these spikes cannot induce significant dead time due to the Level 1 2 kHz capability. The expected Level 1 trigger rate for the design luminosity of $3\times10^{33}\,\rm cm^{-2}s^{-1}$, based on current observations, will be well below 2 kHz.

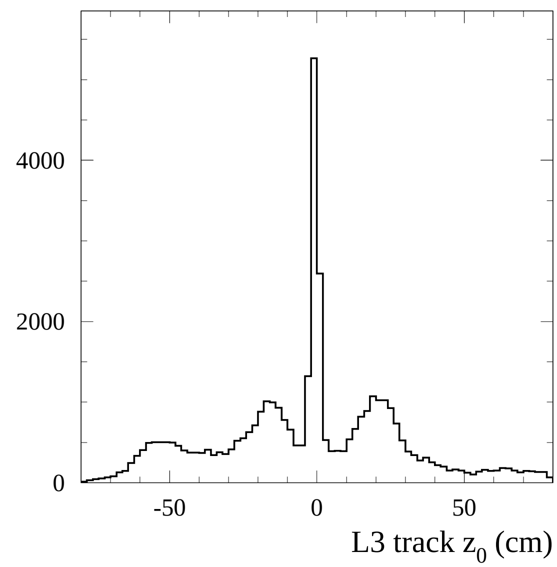

Figure 81. Track z_0 for all Level 1 triggers.

Within the typical Level 1 rate of 970 Hz, Bhabha and annihilation physics events amount to $\sim$130 Hz. There is also 100 Hz of cosmics and 20 Hz of random beam crossing triggers. The dominant source of background which causes the remaining triggers is due to lost particles interacting with the beam line components. The distribution of track z_0 (as reconstructed by Level 3) for all input Level 1 triggers are shown in Figure 81.

The most prominent peaks at $z = \pm 20\,\rm cm$ correspond to the end flange of the beam pipe. The peak at $z = -55\,\rm cm$ corresponds to a step in the synchrotron mask.

The Level 1 trigger hardware operation has been very stable. For the first one and half years of operation, there have been only 4 hardware failures in the whole Level 1 system, mainly auxilliary or communication boards. The only occa-

Table 17
Level 1 Trigger efficiency (%) and rates (Hz) at a luminosity of trigger 2.2×10^{33} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for some selected triggers lines applied to various physics processes. The symbols refer to the counts for each object. The unfilled "–" entries are cases where the line is not intended for this physics source.

Level 1 Trigger	$\epsilon_{B\overline{B}}$	$\epsilon_{B\to\pi^0\pi^0}$	$\epsilon_{B\to\tau\nu}$	$\epsilon_{c\overline{c}}$	ϵ_{uds}	ϵ_{ee}	$\epsilon_{\mu\mu}$	$\epsilon_{\tau\tau}$	Rate
A≥3 & B* ≥1	97.1	66.4	81.8	88.9	81.1	–	–	17.7	180
A≥1 & B* ≥1 & A'≥1	95.0	63.0	83.2	89.2	85.2	98.6	99.1	79.9	410
All pure DCT OR	99.1	79.7	92.2	95.3	90.6	98.9	99.1	80.6	560
M≥3 & M* ≥1	99.7	98.6	93.7	98.5	94.7	–	–	53.7	160
EM* ≥1	71.4	94.9	55.5	77.1	79.5	97.8	–	65.8	150
All pure EMT OR	99.8	99.2	95.5	98.8	95.6	99.2	–	77.6	340
B≥3 & A≥2 & M≥2	99.4	81.2	90.3	94.8	87.8	–	–	19.7	170
M* ≥1 & A≥1 & A'≥1	95.1	68.8	83.7	90.1	87.0	97.8	95.9	78.2	250
E≥1 & B≥2 & A≥1	72.1	92.4	60.2	77.7	79.2	99.3	–	72.8	140
M* ≥1 & U≥5 (μ-pair)	–	–	–	–	–	–	60.3	–	70
All Level 1 triggers	>99.9	99.8	99.7	99.9	98.2	>99.9	99.6	94.5	970

sional online adjustment needed was for the EMT tower mask to temporarily suppress noisy channels in the EMC electronics.

13.4. Level 3 trigger system

The Level 3 trigger makes its selection based upon the complete event data from *BABAR* processed by a farm of 32 commercial Unix processors. The Level 3 also has access to the output of the Level 1 trigger processors and FCTS trigger scalers. Level 3 operates by improving upon the themes defined in Level 1. For example, better DCH tracking (vertex resolution) and EMC clustering allow for greater rejection of beam backgrounds and more effective prescaling of Bhabhas. Filters to select specific rates of luminosity and calibration events are also implemented.

The Level 3 system runs within the Online Event Processing (OEP) framework (see section ??). Along with other responsibilities, OEP hands events to the Level 3 algorithm, then prescales and logs those which trigger. The execution time budget for the complete Level 3 code is 10 ms per event on a 333 MHz Sun Ultra5 node.

To provide optimum flexibility under different running conditions, Level 3 was designed according to a general logic model that can be configured to support an unlimited variety of event selection mechanisms. The logic model provides for a number of different, independent classification tests, called "paths", that are executed in parallel, together with a mechanism for combining these tests into the final set of classification decisions. Thus, modified or new algorithms can be handled with no changes to the underlying logic.

The Level 3 design is based on a three-phase logic model. In the first phase, events are classified on the basis of the Level 1 and FCTS decisions (the 32 FCTS output lines). These classification decisions are referred to as Level 3 "input lines". Any number of Level 3 input lines may be defined, but it is required that each Level 1 output line must be used in defining at least one Level 3 input line, thereby ensuring that any interesting physics events selected by Level 1 are viewed by Level 3.

In the second phase, events are tested using a series of "scripts", each of which requires the logical AND of various event properties. The decision of which scripts to execute is determined by the values of the Level 3 input lines. Each script constructs objects of interest from the raw event data and tests the event for the logical AND of various properties, e.g., charged particle multiplicity, EMC energy, etc. Each script returns a pass/fail decision in the form of a flag.

In the final phase, the Level 3 output lines are formed by considering combinations of the results of the scripts. Each output line is defined as the logical OR of a subset of the script flags. The Level 3 design allows output lines to consider certain script flags as vetoes, rejecting events that otherwise satisfy the line's requirement if the veto is true. As an example, Bhabha-scattering events, a particularly high-rate background for scripts which select events based on charged tracks or EMC energy deposits, are rejected by a Bhabha veto from lines that correspond to interesting physics and therefore may not be prescaled by OEP.

The Level 3 implementation depends crucially on several aspects of Framework. Any code in the form of "modules" can be included and runtime configuration using a standard Tcl interface is allowed. Framework supports multiple execution paths and Level 3 exploits this feature. When multiple paths are used, the same instance of a module may be included in multiple paths. Framework ensures that the event processing for such modules is only invoked once per event. Hence calculations done by a module are not repeated even if the module is called more than once. Thus, the same module may be used on multiple paths without incurring significant CPU overhead. Finally, Framework provides the ability to terminate processing of an event before all modules have been executed. Thus, a "filter module" can deliver a pass/fail indicator to Framework for a given event. If the filter module fails the event, Framework terminates the path and proceeds immediately to either the next path, if any paths remain to be executed, or to the next event.

13.4.1. Level 3 drift chamber tracking algorithm

A large fraction of the events which passes Level 1 but must be rejected by Level 3, is background events in which charged particles are produced in material located close to the IP. Level 1 does not currently have sufficient tracking resolution to identify these background tracks but Level 3 can do full three-dimensional track finding and fitting. The DCH-based algorithm, L3Dch, consists of fast pattern recognition, that is efficient for tracks coming from the IP, and a track fitter, which determines the five helix track parameters for tracks with p_t above 250 MeV/c.

To speed up the process of pattern recognition, L3Dch starts with the track segments from the TSF system. The TSF provides an address that can be decoded to give information about which wires were hit in the pivot cell and an estimated ϕ position. To improve the resolution L3Dch makes use of the actual DCH hit information for hits which were reported as used on the TSF segments.

For those TSF segments that have a staggered pattern indicating a simple solution to the left-right ambiguity for each layer, a track t_0 is determined. The t_0 values for each segment in an event are binned and the average value in the most populated bin is used as the estimated event t_0. Almost all events, physics or background, which pass Level 1 have enough segments to form a t_0 estimate. The rms resolution on this estimate is 1.8 ns for Bhabha events and 3.8 ns for hadronic events.

The actual pattern recognition for L3Dch is done with a look up table. This track table is made by considering the DCH as divided into 120 ϕ sectors, corresponding to the number of cells in the innermost layers. The track table is made up of the ϕ sectors for all tracks originating within 2 cm of the IP in the xy plane and within 10 cm in z and having a p_t above 250 MeV/c. The pattern recognition consists of looping over the table entries looking for matches to segments found by the TSF's. The located set of segments for a given track is then passed to the track fitter. In order to account for inefficiencies, the track table allows for up to two missing DCH layers.

The track fitter is provided with both the track segments and the "seed" track which made the entry in the lookup table. From this information, the track fitter fits the five helix parameters, adding segments not found in the original pattern recognition, if needed. Hits which have large residuals are dropped from the fit. The fit is iterated so that segments close to the initial track fit can be added to the track before a new fit is performed. The final fit does not demand that the track originate from the IP.

The two-track miss distances near the IP for

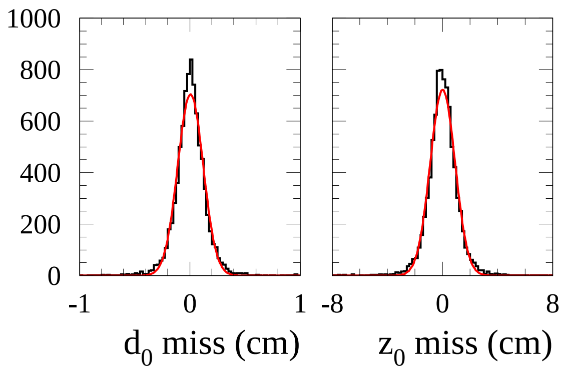

Figure 82. Level 3 track d_0 and z_0 miss distance for Bhabha events.

Bhabha events are plotted in Figure 82. The individual track impact parameter resolution can be derived from the width of these distributions by multiplying by $1/\sqrt{2}$. This gives an Level 3 track d_0 resolution of 0.83 mm and and an average z_0 resolution of 6.9 mm. Similarly, fitting the $1/p_{\mathrm{t}}$ difference between the 2 tracks in μ-pair events yields a p_{t} resolution of $\delta p_{\mathrm{t}}/p_{\mathrm{t}} \sim 0.019 \cdot p_{\mathrm{t}}/\,\mathrm{GeV}/c$ for Level 3 tracks.

13.4.2. Level 3 calorimeter clustering algorithm

A number of events which are interesting for CP physics, calibration or luminosity measurements consist essentially of only neutral particles. Hence an all-neutral trigger for Level 3 based on information from the EMC is essential. In addition, calorimeter information is vital to complement the drift chamber data for the identification of Bhabha events, in particular for the veto algorithm. Finally, partial orthogonality in triggering based on information from the DCH and the EMC is necessary to obtain a thorough understanding of the efficiency of triggers based on either of these systems.

The Level 3 EMC-based trigger algorithm, L3Emc, operates by finding EMC clusters at energies down to the MIP level. The clustering is done in two steps, the processing of the EMC data to assemble lists of adjacent crystals with significant energy deposits and the formation of clusters and computation of integrated quantities, such as the total energy, centroid position, and cluster moments.

Out of the total 6580 crystals in the calorimeter, the EMC at present sends typically 1900 so-called “EMC digi’s” per event, the vast majority of these being caused by electronics noise. An EMC digi encodes the peak energy and time of the crystal waveform as determined by the EMC feature extraction. To filter out noise, L3Emc applies an energy threshold of currently 30 MeV and and a time-window cut of $t > 5.7\,\mu$s on the raw data on a per crystal basis. For the remaining crystals, raw energies and times are converted into physical units to create “L3Emc digi’s”.

The L3Emc digis are filled into a linear list and fed into a fast clustering algorithm that performs a single iteration to split it into sub-lists of neighboring crystals. The neighboring information is not buried in the algorithm but is configured for each crystal by a lookup table. The lookup table addresses are mapped to detector coordinates by a hash algorithm using module, fiber, channel id’s. From the resulting lists of adjacent hits, clusters are formed by summing contiguous crystal energies. If the total energy is above 100 MeV, the Level 3 clusters are stored with an energy weighted centroid and average time, the number of crystals and a lateral and Zernicke moment *[reference needed]* describing the shower shape for particle identification.

13.4.3. Level 3 filters

Based on the Level 3 tracks and clusters that are computed with the tools described above, a variety of filter algorithms is implemented, which perform the event classification and background reduction.

The logging decision is essentially represented by two orthogonal sets of filters which form the open physics trigger lines and which are based on pure drift chamber and pure calorimeter information, respectively.

The drift chamber algorithms comprise two IP track filters which select events with one “tight” (high p_{t}) track or two “loose” tracks originating from the interaction point, respectively. The high p_{t} track is required to have a transverse momentum of $p_{\rm t}$>800 MeV/c and to satisfy tight vertex cuts of $|d_0|$ <1.0 cm, $|z_0|$ <7.0 cm. Two tracks are accepted with $p_{\rm t}$>250 MeV/c and somewhat looser vertex cuts of $|d_0|$ <1.5 cm, $|z_0|$ <10.0 cm.

The calorimeter algorithm uses two cluster filters which select events with either high energy deposits or high cluster multiplicity and with a high (pseudo) event mass. The event mass is calculated from the cluster energy sums and the energy weighted centroid positions assuming massless particles. The filters require either 2 clusters of $E_{CM} > 350$ MeV in the center-of-mass or 4 clusters, which form an event mass greater than 1.5 GeV in both cases.

At current luminosities the output of these open filters is dominated by Bhabha events, which therefore need to be rejected from the respective physics lines. This is accomplished by a Bhabha veto algorithm, which makes use of a very pure identification based on tight track-cluster matching criteria. The veto algorithm is separated into the 1-prong (with only a positron in the backward part of the detector) and 2-prong category (with both particles in the detector). Both algorithms impose stringent criteria on EMC energy deposits that are consistent with Bhabha events, while relying on both the track momenta and on E/p to allow for final state radiation. The 2-prong veto requires either a small acolinearity between the tracks in the center-of-mass or one that is consistent with initial state radiation (ISR), where a photon is radiated along the beam direction leading to a missing energy close to the observed missing track momentum.

For purposes of calibration and offline luminosity measurements, a fraction of Bhabha events are flagged on a separate trigger line by an efficiency-oriented algorithm. The output event θ distribution is 'flattened' using a binned prescaling mechanism in order to obtain similar calibration statistics for all EMC crystals. In addition, calibration samples include radiative Bhabha events, $\gamma\gamma$ final state events and cosmic rays. Radiative Bhabha events are identified by selecting two-prong events with missing energy and requiring a calorimeter cluster in a cone around the missing momentum vector with no requirement on the cluster energy. $e^+e^- \rightarrow \gamma\gamma$ events are selected with no tracks and two high energy clusters back-to-back in the center-of-mass. The cosmics selection is drift-chamber-based and requires two tracks back-to-back in the laboratory frame with a small difference in impact parameters and curvature, where a significant background from ISR Bhabha events faking this topology is removed based on the relation between the visible energy and the track momenta (as used by the 2-prong veto).

As an important part of online monitoring and diagnostics, there are filters for the online luminosity measurement, which is served by a track-based lepton-pair selection with a well known efficiency, and hadronic filters for selection of continuum and $B\overline{B}$-enriched samples. The latter two are distinguished by an event shape cut using a ratio of Fox-Wolfram moments [8]. Both selections are combined for an online determination of the hadron to luminosity ratio, which provides sensitivity to the $\Upsilon(4S)$ line shape and is used to monitor the center-of-mass energy of PEP-II.

13.5. Level 3 performance and operational experience

The Level 3 trigger efficiencies for Level 1 accepted events are tabulated in Table 18, for various physics processes with similar definitions as in Table 17, based on production Level 3 processing applied to MC simulation events. High efficiencies are achieved for the DCH and EMC based filters independently for hadronic events. The comparison between data and MC Level 3 trigger pass fractions for the various filters also show good agreement when requiring tracking and EMC based hadronic event selections in turn. An example of the actual Level 3 event display used for online monitoring with Level 3 recontructed tracks and Level 3 EMC clusters is shown in Fig. 83, together with the Level 1 and Level 3 trigger line status for the event.

For a typical run on the $\Upsilon(4s)$ peak with an average luminosity of $2.6\times10^{33}\,{\rm cm}^{-2}{\rm s}^{-1}$, the Level 3 total output rate is 120 Hz and the event composition is tabulated in table 19. The desired physics sources contribute 13% of the total output and the calibration and diagnostic samples take up 40% of the total. An improved version of the Level 3 IP track filter to be deployed in the 2001

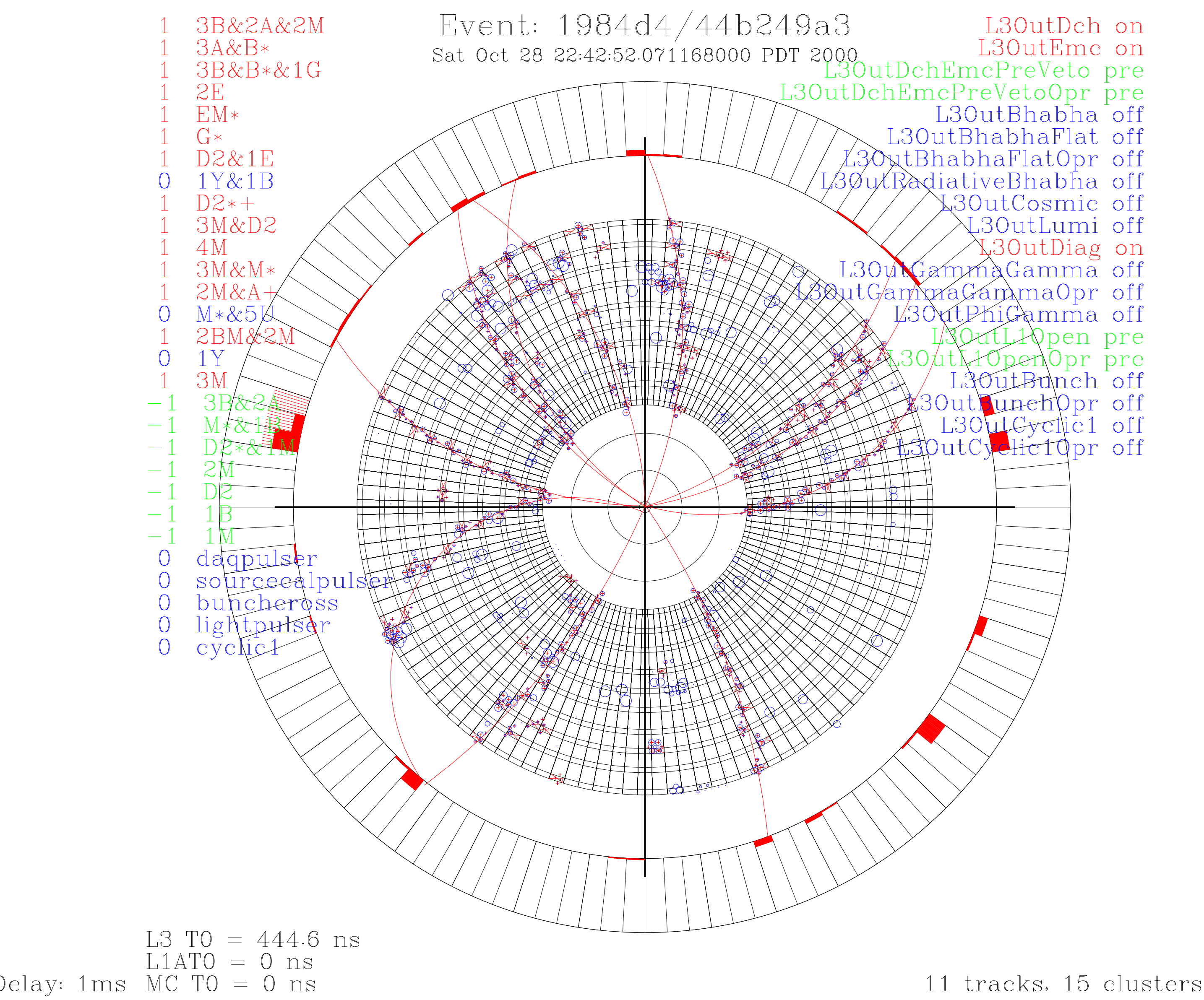

Figure 83. An Level 3 event display. The small circles and small crosses in the DCH volume are DCH hits and TSF segment hit wires respectively. The filled EMC crystals represent energy deposit (full crystal depth=2 GeV) from Level 3 EMC clusters.

Table 18
Level 3 Trigger efficiency (%) for various physics processes.

Level 3 Trigger	$\epsilon_{B\overline{B}}$	$\epsilon_{B\to\pi^0\pi^0}$	$\epsilon_{B\to\tau\nu}$	$\epsilon_{c\overline{c}}$	ϵ_{uds}	$\epsilon_{\tau\tau}$
1 track filter	89.9	69.9	86.5	89.2	88.2	94.1
2 track filter	98.9	84.1	94.5	96.1	93.2	87.6
Combined DCH filters	99.4	89.1	96.6	97.1	95.4	95.5
2 cluster filter	25.8	91.2	14.5	39.2	48.7	34.3
4 cluster filter	93.5	95.2	62.3	87.4	85.5	37.8
Combined EMC filters	93.5	95.7	62.3	87.4	85.6	46.3
Combined DCH+EMC filters	>99.9	99.3	98.1	99.0	97.6	97.3
Combined L1+L3	>99.9	99.1	97.8	98.9	95.8	92.0

Table 19
Level 3 output event composition at a luminosity of $2.6\times10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

Event type	Rate (Hz)
L1,L3 pass through diagnostics	7
Random triggers, cosmics	2
Calibration Bhabhas	30
$\gamma\gamma$, Radiative Bhabha	10
Total calibration/diagnostics	49
Hadrons and $\tau\tau,\mu\mu$	16
Unidentified Bhabhas	18
Other QED, 2-photon	13
Beam wall interactions	26
Total physics accept	73

run is expected to reduce the beam wall background by a factor of 2, which will be sufficient to meet the 120Hz maximum total output rate for the design luminosity of $3\times10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, despite the somewhat larger calibration sample demand than the original design.

At the current performance, the Level 3 executable takes an average processing time of 8.5 ms per event on a current 333 MHz Sun Ultra-5 farm node. At the highest Level 1 input rates, the Level 3 process runs at $\sim$72 % CPU usage, while the rest is spent in OEP (including the network event builder) and in the operating system kernel. This results in an estimated maximum Level 1 input rate of about 2700 Hz with 32 nodes, which is well above the design value of 2 kHz.

13.6. Summary and outlook

Both the Level 1 and Level 3 trigger systems have met the efficiency and the maximum rate requirements for the original design luminosity of $3\times10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The triggering efficiencies for $B\overline{B}$ events are generally well over the 99% design goal for both Level 1 and Level 3. The orthogonal trigger setup based on DCH only and EMC only information, for both Level 1 and Level 3, has successfully delivered stability and measurability of the overall trigger efficiency. The current system also provides a solid foundation with a clear upgrade path to higher luminosities of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ or more.

Future Level 1 trigger improvements will mostly come from further background rejection capability afforded by algorithm refinements and upgrades of the DCT. This is essential for reducing the load on the DAQ and Level 3 at much higher luminosities. Refined algorithms for the PTD and BLT are expected to be deployed in the 2001 run. The new PTD algorithm will effectively narrow the track d_0 acceptance window, while new BLT algorithm will narrow the track z_0 acceptance. For the longer term, a major DCT upgrade design is underway for a new set of trigger boards with the capability of selecting tracks with a narrow z_0 window around the IP. The TSF segment spatial resolution for the stereo layers can be utilized to provide segment z coordinates at $\sim$2 cm resolution. The implied track z_0 resolution of $\sim$4 cm show good promise for rejecting majority of the beam wall background events at $z=\pm20\,\mathrm{cm}$.

The future improvement for Level 3 will also emphasize background rejection. The current physics filter algorithms are still rather simple and there is room for better efficiency at the same background rejection rate. The remaining beam-wall background event fraction are expected to be significantly reduced at higher luminosities. The main tasks will be more aggressive rejection of Bhabha events and suppression of uninteresting QED and two-photon events. To enhance the physics topology recognition capability, a major source of improvement is expected to be L3Dch tracking for low momentum tracks at $p_t < 250\,\mathrm{MeV}/c$. Given the rapid growth of computing technology, a rather moderate CPU upgrade for the Level 3 online farm in the near future will be sufficient to keep up with higher luminosities of greater than $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, as well as having the CPU capacity for improved Level 3 algorithms.

REFERENCES

1. A. Berenyi *et al.*, "Continuously Live Image Processor for Drift Chamber Track Segment Triggering", IEEE Trans. Nucl. Sci. 46 (1999) 348.
2. A. Berenyi *et al.*, "A Binary Link Tracker for the *BABAR* Level 1 Trigger System", IEEE Trans. Nucl. Sci. 46 (1999) 928.
3. A. Berenyi *et al.*, "A Real-Time Transverse Momentum Discriminator for the *BABAR* Level 1 Trigger System", submitted to IEEE Transactions on Nuclear Science.
4. The FPGA's used in the DCT and GLT are from ORCA 2C series manufactured by Lucent Technologies.
5. K. Kinoshita, Nucl. Instr. and Methods A276 (1989) 242.
6. P. D. Dauncey *et al.*, "Design and Performance of the Level 1 Calorimeter Trigger for the BaBar Detector", submitted to IEEE Transactions on Nuclear Science.
7. The FPGA's used for EMT algorithm logic are Xilinx 4020E.
8. G. C. Fox and S. Wolfram, Nucl. Phys. **B1**49 (1979) 413.

14. The Online Computing System

14.1. Overview

The *BABAR* online computing system comprises the data acquisition chain from the common front-end electronics, through the embedded processors in the data acquisition system and the Level 3 trigger, to the logging of event data. It also includes those components required for detector and data acquisition control and monitoring, immediate data quality monitoring, and online calibration.

14.1.1. Design requirements

The data acquisition chain was designed to meet the following basic performance requirements. It must support a Level 1 trigger accept rate of up to ∼2000 Hz, with an average event size of ∼32 kbytesand a maximum output (Level 3 trigger accept) rate of 120 Hz. While performing these functions it must not contribute more than a time-averaged 3% to dead time during normal data acquisition.

The online system is also required to be capable of performing data acquisition simultaneously on independent *partitions* — sets of detector system components — to support calibrations and diagnostics.

The final design embodies many other basic requirements to ensure its effective functioning as described in the following paragraphs.

Normal detector operation, data acquisition and routine calibrations are performed efficiently and under the control of a simple user interface with facilities for detecting, diagnosing, and recovering from common error conditions.

Following standard practice, the event data acquired by the system is subjected to monitoring. Such monitoring is configurable by experts and designed to detect anomalies in the detector systems which, if present, are reported to operators for rapid assessment and, if necessary, corrective action.

Environmental conditions of the detector, such as the state of low and high voltage power, high purity gas supplies, and of the accelerator, such as beam luminosity and currents, are measured and recorded in a fashion that permits its association with the event data logged. Conditions relevant to data quality are monitored for consistency with specified standards. Operators are alerted if these are not met. Data-taking is inhibited or otherwise flagged if conditions are incompatible with maintaining the quality of the data.

Operational configurations, calibration results, software versions in use, and routine and error messages are also recorded and correlatable, in support of the reconstruction of the conditions of operation when analyzing data or diagnosing problems.

14.1.2. System Components

The online computing system is designed as a set of subsystems using elements of a common software infrastructure running on a dedicated collection of hardware.

The major subsystems are:

- Online Dataflow (ODF) – responsible for communication with and control of the detector systems' front-end electronics, and the acquisition and building of event data from them
- Online Event Processing (OEP) – responsible for processing of complete events, including Level 3 (software) triggering, data quality monitoring, and the final stages of calibrations;
- Logging Manager – responsible for receiving selected events sent from OEP and writing them to disk files for use as input to the "Prompt Reconstruction" processing
- Detector Control – responsible for the control and monitoring of environmental conditions of the detector systems
- Run Control – which ties together all the other components, and is responsible for sequencing their operations, interlocking them as appropriate, and providing a graphical user interface for operator control

Each of these components, as well as a selection of the common tools which tie them together are described below.

The entire system is coded primarily in the C++ language, with some use of Java for graphical user interfaces. Object-oriented analysis and design techniques have been used throughout. This has been an important factor in the success of the project, having produced benefits in development speed, maintainability, and extensibility.

14.1.3. Hardware infrastructure

The hardware infrastructure for the online system is shown schematically in Figure 84.

The data from the front-end electronics (FEE) of the various detector systems are routed via optical fiber links to a set of 157 custom VME *Readout Modules* (ROMs). These ROMs are grouped by detector system and contained within 23 crates controlled by the ODF software described below. One ROM in each crate aggregates the data and forwards the result for event building into a farm of commercial Unix workstations [1], 32 of which are used during normal operation. The crates and farm machines communicate via full-duplex 100 Mbps Ethernet, linked by a network switch—the *event builder switch* [2]. The ROMs are supported by a *boot server* providing core and system-specific code and configuration information [3].

The farm machines host the OEP and Level 3 trigger software. The events accepted by the trigger are logged via TCP/IP to a *logging server* [3] and written to a disk buffer for later reconstruction and archival storage. Various data quality monitoring processes run on farm machines not used for data acquisition.

Several additional file servers hold the online databases and production software releases. A further set of *application servers* host the central functions of the various online subsystems. Operator displays are supported by a group of ten *console servers* [4].

An additional set of fifteen VME crates, each with an embedded processor, contain the data acquisition hardware for the detector control subsystem.

These, the online farm, and all the application and console servers are connected via a switched 100 Mbps Ethernet network distinct from that used for event building, with 1 Gbps fiber Ethernet used for the file servers and inter-switch links.

14.1.4. User interaction

Operator control of the online system is achieved primarily through a custom Motif graphical user interface (GUI) for run control and an extensive hierarchy of displays for detector control, including control panels, strip charts and an alarm handler. An electronic logbook is made available through a Web browser interface. These and other GUIs are organized across seventeen displays for the use of the experiment's operators. This operator environment provides for basic control of data acquisition, the overall state of the detector, and certain calibration tasks.

Each detector system has developed a set of specialized calibration and diagnostic applications using the tools provided in the online system. A subset of these calibrations has been specified to be run once per day, during a ten minute scheduled beam-off period. The run control logic, combined with the capability for creating partitions, allows all detector systems' calibrations to be run in parallel and provides the operator with basic feedback on the success or failure of each.

14.2. Online Dataflow

Online Dataflow (ODF) handles data acquisition and processing from the detector systems' FEE through the delivery of complete events in the online farm [5]. The ODF subsystem receives the Level 1 trigger outputs, filters and distributes them to the FEE, reads back the resulting data and hierarchically assembles it into events. It provides interfaces for control of data acquisition, processing and calibration of detector system data, and FEE configuration. Multiple independent *partitions* of the detector may be operated simultaneously.

Event data acquisition proceeds from a trigger decision formed in the Fast Control and Timing System (FCTS) hardware [6], a part of ODF, based on inputs from the Level 1 trigger. Trigger decisions are distributed, in the full detector configuration, to the 133 ROMs connected via optical fibers to the detector system FEE. These ROMs read and process the data from the FEE. One to ten such ROMs from a single detector system are located in each of the 23 data acquisition VME crates. ODF builds complete events from them, first collecting the data in each crate into an additional dedicated ROM, and then collecting the data from the 23 of these, across the event builder network switch, into the online farm Unix workstations.

The operation of the system is controlled by ODF software running on one of the application servers, under the direction of run control. A single ROM in the VME crate containing the central FCTS hardware supports the software interface to ODF. The distribution of ROMs by detector system is shown in Table 20. The numbers of ROMs connected directly to the detector FEE and of those used for event building are shown separately.

Table 20
Online Dataflow Hardware Components

Detector System	VME Crates	Readout Modules
SVT	5	14+5
DCH	2	4+2
DRC	2	6+2
EMC	10	100+10
IFR	1	4+1
EMT	1	1+1
DCT	1	3+1
GLT	1	1+1
FCTS	1	1
TOTAL	24	157

All of the ROM CPUs boot via NFS over the event building network from the boot server described above. The entire system takes about 40 seconds to load and boot 1.5 MB of core ODF code and, typically, 4 MB of detector-specific code.

The ODF software allows all the components of this heterogeneous system to be represented as entities in a uniform object-oriented application framework. The components are organized into five *levels* which map closely onto the system's physical structure.

For each component at each level, its behavior is abstracted as a finite state machine. The com-

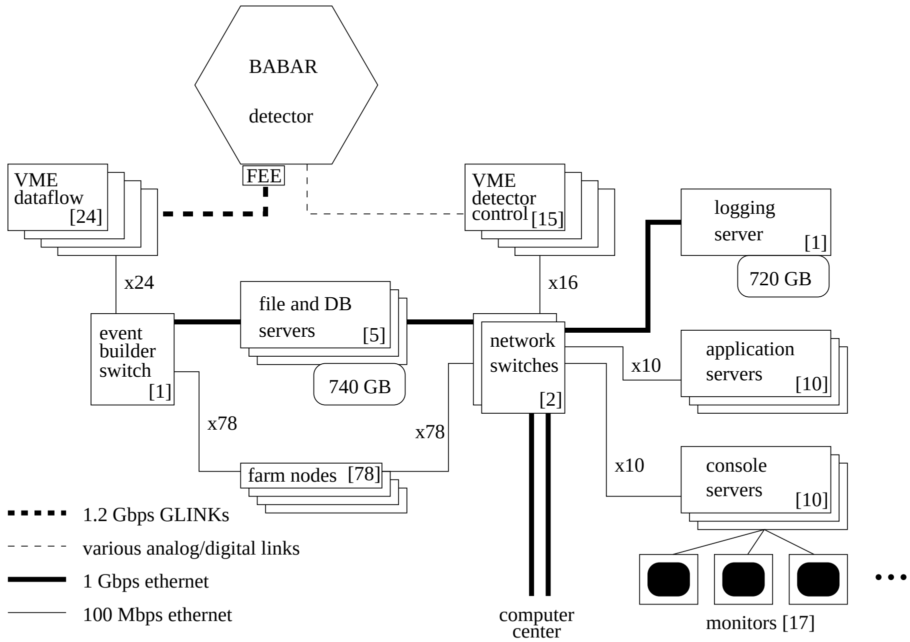

Figure 84. Physical infrastructure of the *BABAR* online system

plete set of these machines is kept coherent by passing messages and data regarding state transitions along the chain of levels. The basic flow of control and data is shown in Figure 85. The mapping of levels to components is as follows:

Control – the Unix-based process controlling the operation of each partition and the source of all state transitions except for L1 Accept. It transmits state transition messages over the network to the source level, waiting for acknowledgement of their successful processing by all levels.

Source – the FCTS hardware and the software running in the ROM located in the FCTS VME crate. For each partition in existence, its source level receives control level transitions and Level 1 trigger outputs and distributes them via the FCTS hardware to all ROMs in the VME crates that are included in the partition. Level 1 triggers are modeled in the subsystem as an additional, idempotent state transition, *L1 Accept,* and are treated uniformly with the others wherever possible.

Segment – the ROMs connected to the detector FEE, with their ODF and detector system-specific software. Each segment level ROM receives state transition messages from the source level and runs appropriate core and detector system-specific tasks in response. These tasks include the acquisition of raw data from the FEE in response to L1 Accepts, and *feature extraction* — taking this data, eliminating uninteresting hits and applying calibration corrections, and returning results in a reduced form. Output data resulting from this processing is attached to the transition messages, which are then forwarded over the VME backplane to the fragment level ROM in each crate.

Fragment – the per-crate event builder ROMs and software. The single fragment level ROM in each crate aggregates the messages from the crate's segment level ROMs — the first stage of event building — and forwards the combined message to one of the event level Unix nodes.

Event – the processes on the online farm nodes receiving complete events and handing them over to OEP for filtering and logging. The ODF event level code aggregates messages, with their attached data, from all the crates in a partition – the second and final stage of event building. The resulting data may be further processed by user code in the event level, but is normally just passed on to OEP. The control level is notified of the completion of processing of all transitions other than L1 Accept. Both the fragment and event level event builders use a data-driven "push" model, with a back pressure mechanism to signal when they are unable to accept more data.

Test stands of varying complexity are supported. The simplest possible consists of a single Unix machine which runs both control and event level code, with two FCTS modules and a single ROM, running source, segment, and fragment level code, in one VME crate. Subsystem configuration is detected at run-time, so the same code that runs in the full subsystem can also run in test stand systems.

14.2.1. Control and Source Levels

The control level sends state transition messages for a partition over the network, using the UDP datagram protocol [10], to the source level in the single ROM inside the FCTS crate. In the source level, the transition message is sent over VME to an FCTS module which forwards it as a 104-bit 59.5 MHz serial word to all VME crates in the relevant partition. This serial word contains a 56-bit event time stamp (counting at 59.5 MHz), a 32-bit transition-specific word and additional control bits. L1 Accept transitions and calibration sequences originate in the source level, but then the same mechanism is used to transmit them through the system.

The FCTS hardware receives the 24 Level 1 trigger output lines and eight additional external trigger lines. The FCTS crate is a 9U VME crate, with a custom P3 backplane on which all the trigger lines are bussed. For each partition, an FCTS module receives these lines. It is configurable with a bit mask specifying the trigger lines enabled for its partition, and an optional prescale factor for each line. A trigger decision is formed for the partition by taking the logical OR of the enabled prescaled lines. Twelve of these modules are installed in the full system, thus setting its maximum number of partitions.

The FCTS crate receives two timing signals from the accelerator: a 476 MHz clock tied to the RF bucket structure of PEP-II and a 136 kHz fiducial that counts at its revolution frequency. The former is divided by eight to create a 59.5 MHz system clock. The fiducial is used to start timing counters and to check that the clocks have no problems.

There are two types of deadtime in the ODF subsystem. The first arises due to the minimum spacing of 2.7 μs required between L1 Accept transitions. This restriction simplified the logic design of the FEE readout, because each datum in the silicon tracker and drift chamber is thus associated with only one L1 Accept. The FCTS hardware enforces this by ensuring that all transitions have at least the minimum separation. The command spacing introduces an irreducible, yet minimal deadtime: 0.54% at 2 kHz.

The second type of deadtime arises when all FEE buffers are full and thus unable to accept another event. In a time required to be less than the inter-command spacing, each VME crate in a partition may send back a FULL signal indicating that it is no longer able to process further L1 Accept transitions. The FCTS hardware detects these signals and disables triggering until the FEE are once again prepared to accept data.

An actual L1 Accept signal is only generated from a partition's trigger decision when neither form of deadtime is asserted.

14.2.2. Segment and Fragment Levels

The segment and fragment levels reside in the 23 detector system VME crates. These are standard 9U crates with a custom P3 backplane.

The 104 bit serial transition messages that leave the source level are received by a FCTS module in each VME crate in a partition. This module in turn forwards these messages to the ROMs in the crate over the custom backplane, along with the 59.5 MHz system clock.

A ROM consists of four components (Figure 86): a commercial single-board computer

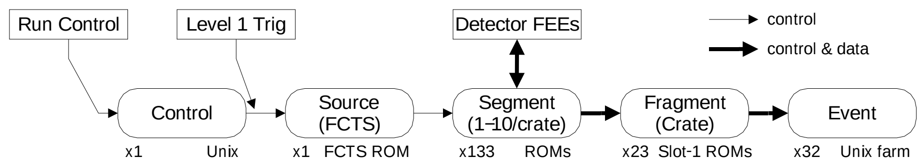

Figure 85. Schematic of the ODF *levels*, their mapping onto physical components, and the flow of control and data between them.

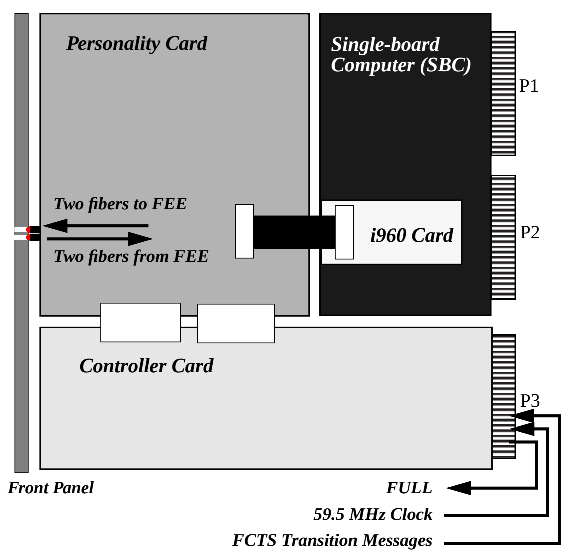

Figure 86. A ROM with a triggered personality card (TPC)

(SBC) [7] and three custom boards: a *controller card* for receiving FCTS commands and supporting FEE reads and writes; a *personality card* that transmits commands to and receives data from the FEE; and a PCI mezzanine card with an Intel i960 I/O processor. The SBCs run the VxWorks [9] operating system with custom code written in C++ and assembly language.

There are two styles of personality cards in the system: triggered *(TPC)* and untriggered *(UPC)*. UPCs are used only in the EMC system. These accept data continuously from the FEE into a buffer pipeline, at a rate of 3.7 MHz. From these samples EMC trigger information is derived and sent over a dedicated serial link to the trigger hardware, providing it with a continuous data stream. An L1 Accept causes up to 256 samples of the raw data stream to be saved to an intermediate memory store on the UPC.

A TPC (used in all other systems) reads out FEE data only when an L1 Accept is received, again saving it into an intermediate memory store. Each detector reads out data in a window of time around the trigger signal, large enough to allow for trigger jitter (see section *[ref?]* above) and detector time resolution. For instance, this window is about 500 ns wide for the silicon vertex tracker. The actual event time within the window is determined approximately in the Level 3 trigger software in OEP and then refined off-line in the course of full event reconstruction.

FEE commands (such as for initiating event reading) are sent and data received by the personality cards over unidirectional 1.2 Gbps serial optical fiber links [8]. All systems' FEE provide zero suppression in hardware except in the EMC and IFR. Data are transferred from the personality card to SBC memory using the i960 as a direct memory access (DMA) engine. This DMA runs at nearly the ideal 133 MB/s rate of the PCI bus.

The FEE for various systems are able to buffer data for three to five L1 Accept transitions. The ROM keeps track of the buffer occupancy and sends, when necessary, a FULL signal (within the required 2.7 μs interval) back to the FCTS to suppress further triggers. The FULL condition is removed when event reads by the ROM free sufficient buffer space. This mechanism handles back pressure from any stage of the data acquisition through to logging by OEP; when downstream processes back up for any reason, back pressure is applied all the way up the chain until FEE buffers are filled and the assertion of FULL throttles the L1 Accept rate.

The ODF application framework provides uniform software entry points for the insertion of user code at each level of the system. This capability is used primarily at the segment level, for FEE configuration and feature extraction. Table 21 shows detector output sizes, bytes per hit, and overhead (e.g. addressing).

Table 21
Typical Event Sizes from Detector Systems

Detector	Hit Size	Total Size	Overhead
SVT	2 bytes	4.9 kB	0.4 kB
DCH	10 bytes	4.8 kB	0.2 kB
DRC	4 bytes	3.1 kB	0.3 kB
EMC	4 bytes	9.1 kB	3.0 kB
IFR	??	1.2 kB	0.2 kB
EMT	–	1.2 kB	<0.1 kB
DCT	–	2.7 kB	0.1 kB
GLT	–	0.9 kB	<0.1 kB
TOTAL		27.9 kB	4.2 kB

Data from the segment level ROMs in a crate are gathered by the fragment level ROM using a *chained* sequence of DMA operations. The maximum throughput of the fragment level event builder is observed to be about 31 Mbps.

In calibrations, ODF may be operated in a mode in which L1 Accept data are not transferred out of the segment level ROMs. This allows calibration data accumulation to occur at high rates inside the ROMs, not limited by the throughput of the event builders or any downstream consumers. Completed calibration results are computed, read out of the ROMs, and written to a database on a non-L1 Accept state transition following the accumulation.

14.2.3. Event Level

For each L1 Accept transition passing through the ODF subsystem, all fragment ROM data are sent to one of the Unix farm machines. The destination is chosen by a deterministic calculation based on the L1 Accept's 56-bit time stamp, available from the FCTS in each ROM. This technique has been demonstrated to produce a uniform quasi-random distribution and to introduce no detectable inefficiency in processing. Events sent to a farm machine still busy with a previous event are held in a buffer to await processing.

All fragment data for an event are sent over the switched 100 Mbps Ethernet event building network to the selected farm machine. The connectionless User Datagram Protocol (UDP/IP) [10] was chosen as the data transport protocol [11], allowing a flow control mechanism to be tailored specifically to this application. Dropped packets are minimized by the network's purely point-to-point, full duplex switched architecture, and by careful tuning of the buffering in the network switch and other parameters. In the rare instance when a packet is lost this is detected by the event builder and the resulting incomplete event flagged.

The event level provides the standard software entry points for user code. In normal operation, these are used only to transfer events via shared memory to the OEP subsystem for Level 3 triggering, monitoring, and logging.

14.2.4. System Monitoring

It is important that the clocks of the FEE stay synchronized with the rest of the system. Each FEE module maintains a time counter which is compared to the time stamp of each L1 Accept in order to ensure that the system remains synchronized. If it becomes unsynchronized, a special *synch* command can be sent through the FCTS, causing all systems to reset their clocks.

To ensure that the data from the correct event is retrieved from the FEE, a five-bit number is incremented and sent down from the FCTS to the FEE with each L1 Accept. These bits are stored in the FEE along with the data and are compared on read-back. If they disagree, a special *clear-readout* command is sent which resynchronizes ROM buffer pointers with FEE buffer pointers.

All transitions, including L1 Accept, are logged in a 4 kB deep by 20 byte wide FIFO as they pass through the FCTS crate. The transition type, the event time stamp, a bit list of the trigger lines contributing to the decision, and the current FULL bit list from all VME crates are recorded in this FIFO. There are also scalers which record delivered and accepted luminosity, deadtime due to the 2.7 μs minimum inter-command spacing, deadtime caused by VME crates being FULL and triggers on each line. These FIFOs and scalers are read out over VME by the FCTS ROM, which then transmits the data to monitoring programs that calculate quantities such as luminosity, deadtime and trigger rates. The UDP *multicast* protocol [12] is used to allow efficient simultaneous transmission of data to multiple clients.

To provide diagnostics, a system which multicasts additional performance information on demand from each CPU, typically at 1 Hz, is used. This information is currently received by a single client on one of the Unix application servers and archived. It can be retrieved subsequently to investigate any unusual behaviour observed in the system.

14.3. Online Event Processing

The online event processing (OEP) subsystem provides a framework for the processing of complete events delivered from the ODF event builder [13]. The Level 3 (software) trigger described in Section [trg:l3] above operates in this framework, along with event-based data quality monitoring and the final stages of online calibrations.

The OEP subsystem serves as an adapter between the ODF event builder interface and the application framework originally developed for the off-line computing system. Raw data delivered from the ODF subsystem are put into an object-oriented form and made available through the standard event data analysis interface.

The use of this technique permitted the Level 3 trigger and most of the data quality monitoring software to be written and debugged within the off-line environment. This software is decomposed into small, reusable units — *modules*, pluggable software components in the framework — many of which are shared among multiple applications.

The OEP interfaces allow user applications to append new data blocks to the original raw data from ODF. The results of Level 3 event analysis are stored in this manner so that the trigger decision and the tracks and calorimeter clusters on which it is based may be used in later processing, such as reconstruction and trigger performance studies.

In order to provide sufficient CPU capacity, with headroom, the Level 3 trigger task within OEP is typically run on 32 of the nodes in the online farm described above.

Histograms and other monitoring data are accumulated across the farm. In order to be able to monitor the full resulting statistics, a "distributed histograming package" (DHP) [14] was developed which provides to networked clients a single view of histograms and time history data summed across all nodes, communicating via CORBA-based protocols [15,16].

In addition to the primary triggering and monitoring functions carried out on the 32 nodes, OEP provides a "trickle stream" protocol that allows networked clients to subscribe to a sampling of the event data. This scheme provides support for event displays and additional detailed data quality monitoring ("Fast Monitoring"). Figure 87 shows the basic flow of data in the OEP subsystem.

A system for automated comparisons of monitoring data against defined references was developed. Statistical comparisons of live histograms, or the results of fits to them, may be performed at configurable time intervals to reference histograms, analytic spectra, or specified nominal values of fit parameters.

Comparison failures, tagged with configurable severity levels based on the confidence levels of the comparisons, are displayed to operators and logged in the common occurrence database, described below.

The Java Analysis Studio package [17] previously developed at SLAC has been enhanced with the ability to serve as a DHP client, and it is used for operator viewing of monitoring data accumulated in OEP. This was implemented by devising a Java "middleware server" which adapts the DHP

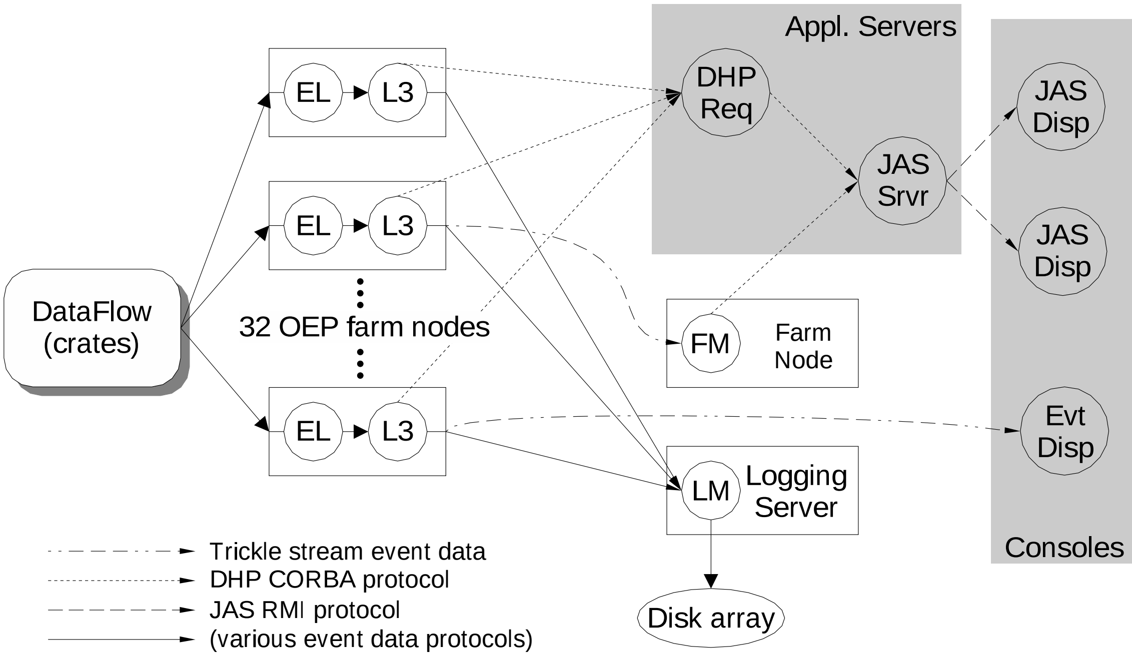

Figure 87. Flow of data in the OEP subsystem: ODF event level (EL) and Level 3 trigger processes on each OEP node; the Logging Manager (LM) on the logging server; the DHP "requestor" process that combines histograms from all 32 Level 3 processes; one instance of a Fast Monitoring (FM) process with DHP histograms; the Java server that makes DHP histograms available to JAS clients; two such clients, and one event display for the Level 3 trigger. OEP-specific data transport protocols are identified.

protocol to the native JAS data protocol.

A dedicated operator console supports the data quality monitoring system, providing for its control and for the display of the histograms from Fast Monitoring and the Level 3 trigger processes, using JAS, along with any error conditions detected by the automatic histogram analysis facility.

14.4. Data logging

Events selected by the Level 3 trigger algorithms in OEP are retained for subsequent full reconstruction. The events are sent from the 32 OEP nodes via TCP/IP to a single multithreaded process, the *Logging Manager* (LM), running on the logging server described above. The LM writes them to the server's RAID arrays in a format specific to OEP. Data from all 32 nodes are combined into a single file for each data-taking *run* (typically two to three hours of data acquisition, resulting in files of about 15–20 GB in size).

Completed files are copied to the SLAC HPSS (High Performance Storage System) [18] system for archiving to tape. These files are retrieved from HPSS for reconstruction, which is typically completed within less than eight hours of data acquisition. The data files are also retrievable for other tasks such as detector system hardware diagnostics and offline tests of the Level 3 trigger algorithms.

14.5. Detector Control

14.5.1. Design Principles

The Experimental Physics and Industrial Control System, EPICS [19], was selected to provide the "engineering layer" for the online detector control (ODC) subsystem: the direct connection to the electrical signals of the power supplies and other hardware, with sufficient monitoring and control to allow commissioning, fault diagnosis and testing. A summary of monitoring and control points is presented in Table 22.

Table 22
Detector Control monitor points *this is a place holder table awaiting data from Vera*

Detector System	Monitor Points	Control Points
SVT	?	?
DCH	?	?
DRC	?	?
EMC	?	?
IFR	?	?
PEPII	?	?
General	?	?
TOTAL	?	?

Beyond the writing of custom drivers, only minor additions or changes were required to EPICS as distributed. EPICS and the additional software specializing it for *BABAR* are written in the C language.

Detector-wide standard hardware was adopted to ease development and maintenance. The standard ODC crate is a 6U VME chassis containing a single board computer [20] serving as an EPICS input/output controller (IOC). Fifteen such crates are used in the experiment. EPICS is fully distributed: each IOC supplies its own naming service, notify-by-exception semantics, and processing. The IOCs boot from a dedicated server.

Analog data are either digitized by modules within the crates or, more commonly, on digitizer boards located directly on the detector. In the latter case, the CANbus standard [21] is used for the transport of signals to and from the detector. A custom "general monitoring board" (GMB) [22] was developed to interface CANbus to the on-detector electronics. The GMB contains a microcontroller, an ADC, multiplexors and operational amplifiers. It can digitize up to 32 signals.

14.5.2. User interface

The operator view of this part of the control system is via screens controlled by the EPICS display manager, DM.

Dedicated control and display panels were developed using DM for each of the detector subsystems, using common "color rules" to show the status of devices, aggregated up the hierarchy of detector system components. A top-level panel for ODC summarizes all subsystems' status and provides access to their specialized panels.

The EPICS alarm handler (ALH) with some *BABAR*-specific modifications is used to provide operators with audible and color-coded alarms and warnings in a hierarchical view of all the subsystems and components. Conditions directly relevant to personnel or detector safety are further enforced by hardware interlocks, the states of which are themselves reflected in a set of uniform EPICS displays and in the ALH, and an alarm annunciator.

14.5.3. Interfaces to other *BABAR* software

A custom C++ layer above EPICS consisting of *Component Proxies* and *Archivers* provides for device-oriented state management and archival data collection. This is ODC's interface to the rest of the online system.

The 27 component proxies (CPs), running on a Unix application server, each define a logical component representing some aspect of a detector system or the experiment's central support systems, aggregated from the 10^5 individual EPICS channels.

The CPs present a simple finite state machine model as their interface to Run Control. The most important actions available are *Configure*, on which the CP accesses the configuration database described below, retrieving set points for its component's channels, and *Begin Run*, which puts the CP into the *Running* state, in which setpoint changes are prohibited and readbacks are required to match settings. While in the *Running* state, the CP maintains a *Runnable* flag which reflects that requirement and allows Run Control to ensure that data acquisition is performed only under satisfactory conditions. When the flag is negated, indicating that readback values are out of tolerance, Run Control will automatically pause data-taking.

The CP's other principal function is to provide an interface for the rest of the online system to the data collected by ODC on the state of the detector hardware and its environment — dubbed "ambient data."

The recording of this data so that it may be analyzed and correlated with event data is essential to the full understanding of the detector. It is the task of the *Archiver* processes, each paired with a CP, to collect the ambient data, aggregate them and write out time histories approximately every hour to the Ambient Database described below. The data are recorded associated with times in order that they may be correlated with the time stamps of the event data. Data from the Archivers' current buffers or from the database may be viewed with a custom graphical browser.

Most ambient data sources generally vary only within their own noise limits, so unnecessary data volume is avoided by limiting storage of data to movements outside a dead-band (set separately for each channel) or across an alarm threshold.

14.5.4. Integration with the accelerator

Close integration between the *BABAR* detector and the PEP-II accelerator is essential to maintain the health of the detector components and the efficiency of data collection. Data from the accelerator control system are transported via EPICS channel access to *BABAR* for display and storage, managed by a dedicated CP. In turn, background signals from the detector are made available to PEP-II to aid in injection and tuning, minimizing backgrounds and optimizing integrated luminosity. An important component of this communication is the "injection request" handshake. When the PEP-II operator requests a significant change in the beam conditions, such as injection or dumping of the beams, the request can only procede following confirmation from *BABAR*. This complements the safety interlocks that *BABAR* controls to enforce protection from radiation damage.

14.5.5. Operational experience

The ODC subsystem has been operational since the initial cosmic-ray commissioning of the detector and the beginning of data-taking with colliding beams. The core EPICS infrastructure has proven to be very robust. The large size of the subsystem, with its 15 IOCs and 10^5 data channels, produces heavy but manageable traffic on the experiment's network.

[**Need some data volume and rate numbers from ODC folks here.** Walt Innes is looking to this but tentatively reported a value of 4 MB/day]

14.6. Run Control

The run control subsystem is implemented as an application of SMI++, a toolkit for designing distributed control systems [23]. Using this software, the *BABAR* experiment is modeled as a collection of objects behaving as finite state machines. These objects represent both real entities, such as the ODF subsystem or the drift chamber high voltage controller, and abstract subsystems such as the "calibrator," a supervisor for the coordination of online components for calibration runs. The behavior of the objects are described in a specialized language (SML) which is interpreted by a generic logic engine to implement the control system.

The SML descriptions of the objects which make up the experiment simply specify their own states and transitions as well as the connections between the states of different objects. Objects perform actions on state transitions, which may include explicitly commanding transitions in other objects; objects may also be programmed to monitor and automatically respond to changes of state in other objects. Such connections express the control hierarchy of the experiment. Anticipated error conditions in components of the online system are reflected in their state models, allowing many errors to be handled automatically by the system. To reduce complexity, logically related objects are grouped together into a hierarchy of cooperating domains.

The system is highly automated; user input is generally required only to initialize the system, start and stop runs, and handle unusual errors. The user communicates with Run Control via a configurable Motif-based GUI included in SMI++.

The states and behavior of Run Control objects representing external systems are provided by a special class of intermediate software processes called *proxies*. A proxy monitors its system, provides an abstraction of it to Run Control, and receives state transition commands. These commands are interpreted and applied to the underlying hardware or software components, implementing the transitions' actions. The control level of an ODF partition is an example of such a proxy.

Communication between the various proxies and the Run Control engines is provided by DIM [24], a fault tolerant "publish and subscribe" communications package based on TCP/IP sockets, allowing Run Control to be distributed transparently over a network.

Essential to the operation of the online system is the notion of the *Runnable* status of its various Detector Control and data acquisition components, indicating that they are in a state suitable for production-quality data-taking. The Run Control logic interlocks data-taking to the logical AND of all components' *Runnable* status. While this condition is not satisfied, data-taking may not start and any existing run will be paused with an alert sent to the operator.

14.7. Common software infrastructure

14.7.1. Databases

Five major databases are used by the online system:

1. *Configuration Database:* This database, implemented using the commercial object-oriented database management system Objectivity [25], allows the creation of hierarchical associations of system-specific configuration data with a single numeric *configuration key*. This key is distributed to all online components, which can then use it to retrieve from the database all the configuration information they require. Convenient mnemonics are associated with the keys for currently relevant configurations, and may be selected for use via the Run Control GUI [26].

2. and 3. *Ambient and Conditions Databases:* These databases, also implemented using Objectivity, are based on the notion of time histories of various data items associated with the experiment. The history for each item is divided into intervals over which a specific value of the item is valid. The Ambient Database is used principally by the Detector Control subsystem to record detector parameters and environmental data at the time they are measured [26].

The Conditions Database is used to record calibration constants and alignments, a digested subset of the Ambient information, and the configuration keys in force during data taking runs. It has the additional feature, compared to the Ambient Database, that the data for a given time interval may be updated as they are refined in the course of improved understanding of the apparatus [27].

The Configuration and Conditions Databases are exported for reconstruction and physics analysis.

4. *Occurrence (Error) Log:* Informational and error messages generated in the online system are routed through the CMLOG system [28] to a central database, from which they are available for operators' real-time viewing or historical browsing, using a graphical tool, as well as for subscription by online components which may require notification of certain occurrences.

5. *Electronic Logbook:* An Oracle-based [29] logbook is used to maintain the history of the experiment's data-taking, organized by runs normally of approximately two hours' duration. It contains information on beam parameters — instantaneous and integrated luminosity, currents, and energies — as well as records of data acquisition parameters such as trigger rates, data volumes, and deadtimes, and the detector configuration used for a run. The logbook also contains text comments and graphics added by the operations staff.

A number of other databases are used in the online system for various tasks such as indexing logged data files, the repair history of online hardware and spares, and tracking software problem reports.

14.7.2. Software Release Control and Configuration Management

All of the online software is maintained in the common *BABAR* code repository, based on the freely available Concurrent Versions System software, CVS [30].

The online's Unix and VxWorks applications are built and maintained with an extension of the standard *BABAR* software release tools [31]. At the start of every data-taking run, the identities of the current production software release and any installed patches are recorded; thus it is possible at a later date to reconstruct the versions of online software used to acquire data.

14.8. Summary

The online system has exceeded its data acquisition performance goals. It is capable of acquiring colliding beam events, with an average size of 28 kB, at a $\mathcal{O}$(2500 Hz) Level 1 trigger rate and reducing this rate in Level 3 to the required $\mathcal{O}$(120 Hz) limit. This provides comfortable margins, for under normal beam conditions the Level 1 trigger rate is $\mathcal{O}$(800–1000) Hz.

The system is capable of logging data from Level 3 at much higher rates when required for diagnostics; the nominal 120 Hz figure represents a compromise between data volume and its consequential load on downstream processing and archival storage, and trigger efficiency for low multiplicity final states.

During normal data-taking the online system routinely achieves an efficiency of over 98%, taking both data acquisition livetime and the system's overall reliability into account.

14.8.1. Upgrade Path

There are several hardware options open for scaling the existing ODF subsystem. Currently most ROMs receive more than one fiber from the FEE. These fibers could be distributed over more ROMs to add processing power. There are also commercial upgrade paths for the ROMs' MVME2306 SBC boards available from Motorola. Crates can be split (up to a maximum total of 32) to create more VME event building bandwidth, as well as more fragment level CPU power and network bandwidth. Gigabit Ethernet connections could also be installed to improve the network event builder's bandwidth.

Various software upgrade options are being investigated, including optimizing the VxWorks network drivers and grouping sets of events together in order to reduce the impact of per-event overhead.

Current background projections indicate that fragment level CPU, segment level memory bus bandwidth and network event building bandwidth are the most likely bottlenecks we will encounter in the future.

Increases in the Level 1 trigger rate or in the background occupancy and complexity of events are expected to necessitate providing additional capacity for OEP, principally for Level 3 triggering. The online farm machines could be replaced with faster models. More machines could also be used, at the expense of increases coherent loading on various servers and of additional management complexity.

No requirement for significant capacity upgrades to the data logging subsystem or to Detector Control is foreseen at this time.

REFERENCES

1. Sun Ultra 5, with single 333 MHz UltraSPARC-IIi CPUs and 512 MB of RAM, Sun Microsystems, Inc., Palo Alto, California, USA.
2. Cisco model 6500, Cisco Systems, Inc., San Jose, California, USA.
3. For the period of data taking covered by this writing, a single Sun Microsystems Ultra Enterprise 450, with four ? MHz CPUs, 2 GB of RAM, and 720 GB of RAID-3 disk, acted as the data logging server as well as the ODF boot server. Beginning in 2001 all its responsibilities other than data logging were moved to a new Sun Ultra 220R machine with dual 450 MHz UltraSPARC-II CPUs, 1 GB of RAM, and an additional 200 GB RAID array.
4. These file and database servers (presently five) are primarily additional Sun Ultra 220Rs, each with about 200 GB of RAID disk. The ten application servers are a mix of Sun Ultra 60s, with dual 360 MHz

UltraSPARC-II CPUs and 1 GB of RAM, and Ultra 5s; the console servers are all Ultra 5s.

5. R. Claus et al., "Development of a Data Acquisition System for the *BABAR* CP Violation Experiment", *Proceedings of the 11th IEEE NPSS RealTime Conference*, 14-18 June 1999, Santa Fe, New Mexico, USA. (http://strider.lansce.lanl.gov/rt99/index11.html)
6. P. Grosso, et al., "The *BABAR* Fast Control System", *Proceedings of the International Conference on Computing in High-Energy Physics, 1998*, 31 Aug-4 Sep 1998, Chicago, Illinois, USA.
7. Motorola model MVME2306 boards, each with a 300 MHz PowerPC 604 CPU, 32 MB of RAM, 5 MB of non-volatile flash memory, two PCI mezzanine card slots (only one of which is currently used), a 100 Mbps Ethernet interface, and a Tundra Universe II VME interface, manufactured by the Motorola Computer Group division of Motorola Inc., Tempe, AZ 85282.
8. These use the Hewlett-Packard GLINK protocol. [need a more detailed reference here from Gunther, perhaps]
9. The VxWorks RTOS and Tornado Development interface are products of Wind River Systems, Inc., Alameda, CA 94501-1153.
10. J. Postel, "User Datagram Protocol", RFC-0768, 28 Aug 1980. Internet documentation including all RFCs (Requests for Comment) are available online from the Internet Engineering Task Force, http://www.ietf.org.
11. T. J. Pavel, et al., "Network Performance Testing for the *BABAR* Event Builder", *Proceedings of the CHEP 1998 Conference*.
12. S. Deering, "Host Extensions for IP Multicasting", RFC-1112, August 1989.
13. G. P. Dubois-Felsmann, E. Chen, Yu. Kolomensky, *et al., Flexible Processing Framework for Online Event Data and Software Triggering,* IEEE Transactions on Nuclear Science 47, 353 (2000).
14. DHP reference needed – Scott Metzler, CHEP conference or *BABAR* note.....
15. CORBA (Common Object Request Broker Architecture) Standards, Object Management Group, `http://www.corba.org/`.
16. ACE/TAO CORBA implementation, Distributed Object Computing Group, Washington University, St. Louis, Missouri, USA and University of California, Irvine, California, USA, `http://www.cs.wustl.edu/~schmidt/TAO.html`.
17. JAS reference
18. High Performance Storage System (HPSS), International Business Machines, Inc., Armonk, New York, USA, `http://www4.clearlake.ibm.com/hpss/index.jsp`.
19. L. Dalesio *et al.*, "The Experimental Physics and Industrial Control System Architecture: past, present and future", Nuclear Instruments and Methods in Physics Research A, vol. 352, pp.179, 1994.
20. The ODC IOCs are model MVME177 boards with MC68060 CPUs from Motorola Computer Group, *op. cit.*.
21. ISO standards 11519, "Road vehicles — Low-speed serial data communication" (1994), and 11898, "Road vehicles — Interchange of digital information — Controller area network (CAN) for high-speed communication" (1993). See `http://www.iso.ch/` .
22. T. Meyer and R. McKay, "The *BABAR* General Monitoring Board", *BABAR* Note 366, May 1998.
23. B. Franek and C. Gaspar, "SMI++ Object Oriented Framework for Designing and Implementing Distributed Control Systems", IEEE Trans. Nucl. Sci. 45 (1998) 1946. SMI++ was adopted, with some upgrades, from a system used in the DELPHI experiment at CERN.
24. C. Gaspar and M. Donszelman, "DIM - A Distibuted Information Management System for the DELPHI experiment at CERN", Proceedings of the IEEE Eight Conference REAL TIME '93 on Computer Applications in Nuclear, Particle and Plasma Physics; Vancouver, Canada.
25. A product of Objectivity Inc., Mountain View, CA 94041; `http://www.objectivity.com/`.
26. G. Zioulas *et al.*, "Ambient and Configuration databases for the *BABAR* online system",

Proceedings of the Real Time 99 Conference, June 14-18, 1999, Santa Fe, New Mexico.
27. [reference to be supplied, perhaps from CHEP 98].
28. The products `cdev` and `CMLOG` are software facilities produced by the Accelerator Controls group of the Thomas Jefferson National Accelerator Facility. Documentation is available from their Web site, `http://www.cebaf.gov/`.
29. A product of Oracle Corporation, Redwood Shores, CA USA. `http://www.oracle.com/`.
30. The Concurrent Versions System is an open-source distributed version control system. `http://www.cvshome.org/`.
31. Reference to Software Release Tools

15. Summary