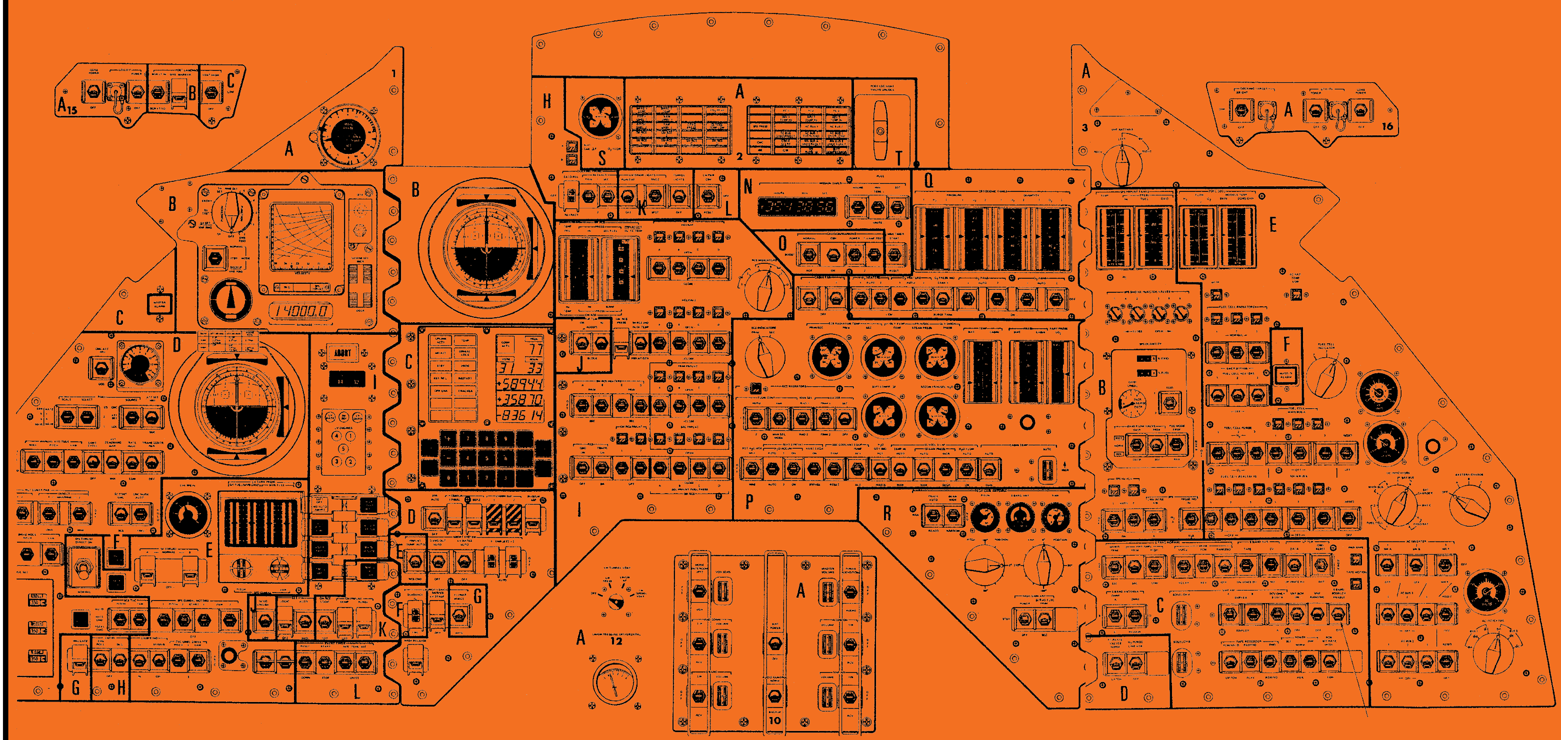
CERN/LHCC 2014-016
LHCb TDR 16
21 May 2014

UPGRADE

LHCb Trigger and Online

TDR



Technical Design Report

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

CERN/LHCC 2014-016
LHCb TDR 16
21st May 2014

LHCb Trigger and Online Upgrade Technical Design Report

The LHCb collaboration

Abstract

The LHCb experiment will be upgraded between 2018 and 2019 in order to reach unprecedented precision on the main observable of the b and c-quark sectors. This Technical Design Report addresses the *trigger-less readout system* and the *full software trigger* features.

LHCb collaboration

I. Bediaga, J.M. De Miranda, F. Ferreira Rodrigues, A. Gomes[a], A. Hicheur, A. Massafferri, A.C. dos Reis, A.B. Rodrigues
[1]*Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil*

S. Amato, K. Carvalho Akiba, L. De Paula, O. Francisco, M. Gandelman, J.H. Lopes, D. Martins Tostes, I. Nasteva, J.M. Otalora Goicochea, E. Polycarpo, C. Potterat, M.S. Rangel, V. Salustino Guimaraes, B. Souza De Paula, D. Szilard, D. Vieira
[2]*Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil*

M. Cruz Torres, C. Göbel, J. Molina Rodriguez
[3]*Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to* [2]

L. An, Y. Gao, F. Jing, Y. Li, H. Lu, S. Wu, Z. Yang, X. Yuan, Y. Zhang, L. Zhong
[4]*Center for High Energy Physics, Tsinghua University, Beijing, China*

Y. Xie
[5]*Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China, associated to* [4]

L. Beaucourt, S. Cap, D. Decamp, N. Déléage, C. Drancourt, Ph. Ghez, J.-P. Lees, J.F. Marchand, M.-N. Minard, B. Pietrzyk, W. Qian, S. T'Jampens, V. Tisserand, E. Tournefier, G. Vouters
[6]*LAPP, Université de Savoie, CNRS/IN2P3, Annecy-Le-Vieux, France*

Z. Ajaltouni, M. Baalouch, H. Chanal, E. Cogneras, O. Deschamps, I. El Rifai, M. Grabalosa Gándara, P. Henrard, M. Hoballah, R. Lefèvre, J. Maratas, S. Monteil, V. Niess, P. Perret, D.A. Roa Romero
[7]*Clermont Université, Université Blaise Pascal, CNRS/IN2P3, LPC, Clermont-Ferrand, France*

S. Akar, K. Arnaud, E. Aslanides, J.-P. Cachemiche, J. Cogan, P.-Y. Duval, F. Hachon, M. Jevaud, W. Kanso, R. Le Gac, O. Leroy, G. Mancinelli, A. Mordà, M. Perrin-Terrin, F. Réthoré, J. Serrano, A. Tsaregorodtsev
[8]*CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France*

Y. Amhis, S. Barsuk, C. Beigbeder, M. Borsato, D. Charlet, O. Duarte, O. Kochebina, J. Lefrançois, F. Machefert, A. Martín Sánchez, M. Nicol, P. Robbe, M.-H. Schune, M. Teklishyn, A. Vallier, B. Viaud, G. Wormser
[9]*LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France*

E. Ben-Haim, M. Charles, S. Coquereau, P. David, L. Del Buono, L. Henry, A. Martens, D.A. Milanes, F. Polci
[10]*LPNHE, Université Pierre et Marie Curie, Université Paris Diderot, CNRS/IN2P3, Paris, France*

J. Albrecht, T. Brambach, Ch. Cauet, M. Deckenhoff, K. Dungs, U. Eitschberger, R. Ekelhof, L. Gavardi, M. Kaballo, P. Klauke, F. Kruse, F. Meier, R. Niet, C.J. Parkinson[52], M. Schlupp, A. Shires, B. Spaan, S. Swientek, J. Wishahi
[11]*Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany*

O. Aquines Gutierrez, J. Blouw, M. Britsch, M. Fontana, D. Popov, M. Schmelling, D. Volyanskyy, H. Voss, M. Zavertyaev[b]
[12]*Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany*

S. Bachmann, A. Bien, A. Comerma-Montells, M. De Cian, F. Dordei, S. Esen, C. Färber, E. Gersabeck, L. Grillo, X. Han, S. Hansmann-Menzemer, A. Jaeger, M. Kolpin, K. Kreplin, G. Krocker, B. Leverington, J. Marks, M. Meissner, M. Neuner, T. Nikodem, P. Seyfert, M. Stahl, S. Stahl, U. Uwer, M. Vesterinen, S. Wandernoth, D. Wiedner, A. Zhelezov
[13]*Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany*

O. Grünberg, M. Heß, C. Voß, R. Waldi
[14]*Institut für Physik, Universität Rostock, Rostock, Germany, associated to* [13]

R. McNulty, R. Wallace, W.C. Zhang
[15]*School of Physics, University College Dublin, Dublin, Ireland*

A. Palano[c]
[16]*Sezione INFN di Bari, Bari, Italy*

A. Carbone[d], A. Falabella, D. Galli[d], U. Marconi, N. Moggi, M. Mussini, S. Perazzini[d], V. Vagnoni, G. Valenti, M. Zangoli
[17]*Sezione INFN di Bologna, Bologna, Italy*

W. Bonivento[43], S. Cadeddu, A. Cardini, A. Contu, A. Lai, B. Liu, G. Manca[e], R. Oldeman[e], B. Saitta[e]
[18]*Sezione INFN di Cagliari, Cagliari, Italy*

M. Andreotti[f], W. Baldini, C. Bozzi, R. Calabrese[f], M. Corvo[f], M. Fiore[f], M. Fiorini[f], E. Luppi[f], A. Mazurov[f], M. Savrie[f], I. Shapoval[50,f], G. Tellarini[f], L. Tomassetti[f], S. Vecchi
[19]*Sezione INFN di Ferrara, Ferrara, Italy*

L. Anderlini[g], A. Bizzeti[i], M. Frosini[43,g], G. Graziani, G. Passaleva, M. Veltri[h]
[20]*Sezione INFN di Firenze, Firenze, Italy*

G. Bencivenni, P. Campana[43], P. De Simone, G. Lanfranchi, M. Palutan, M. Rama, A. Sarti[m], B. Sciascia, R. Vazquez Gomez
[21]*Laboratori Nazionali dell'INFN di Frascati, Frascati, Italy*

R. Cardinale[43,j], F. Fontanelli[j], S. Gambetta[j], C. Patrignani[j], A. Petrolini[j], A. Pistone
[22]*Sezione INFN di Genova, Genova, Italy*

M. Calvi[k], L. Cassina, C. Gotti, B. Khanji, M. Kucharczyk[29,43,k], C. Matteuzzi
[23]*Sezione INFN di Milano Bicocca, Milano, Italy*

J. Fu[43], N. Neri, F. Palombo[u]
[24]*Sezione INFN di Milano, Milano, Italy*

S. Amerio, G. Busetto[r], S. Gallorini, A. Gianelle, D. Lucchesi[r], A. Lupato, M. Morandin, M. Rotondo, L. Sestini, G. Simi, R. Stroili
[25]*Sezione INFN di Padova, Padova, Italy*

F. Bedeschi, S. Leo, P. Marino[t], M.J. Morello[t], G. Punzi[s], F. Ruffini, S. Stracka[43], J. Walsh
[26]*Sezione INFN di Pisa, Pisa, Italy*

G. Carboni[l], E. Furfaro[l], E. Santovetti[l], A. Satta
[27]*Sezione INFN di Roma Tor Vergata, Roma, Italy*

A.A. Alves Jr[43], G. Auriemma[n], V. Bocci, G. Martellotti, D. Pinci, R. Santacesaria, C. Satriano[n], A. Sciubba
[28]*Sezione INFN di Roma La Sapienza, Roma, Italy*

A. Dziurda, W. Kucewicz[o], T. Lesiak, B. Rachwal, M. Witek
[29] *Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland*

M. Firlej, T. Fiutowski, M. Idzik, P. Morawski, J. Moron, A. Oblakowska-Mucha, K. Swientek, T. Szumlak
[30] *AGH - University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland*

V. Batozskaya, K. Klimaszewski, K. Kurek, M. Szczekowski, A. Ukleja, W. Wislicki
[31] *National Center for Nuclear Research (NCBJ), Warsaw, Poland*

L. Giubega, A. Grecu, F. Maciuc, M. Orlandea, B. Popovici, S. Stoica, M. Straticiuc
[32] *Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania*

G. Alkhazov, N. Bondar[43], A. Dzyuba, G. Gavrilov, O. Maev, N. Sagidova, Y. Shcheglov, A. Vorobyev
[33] *Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia*

V. Balagura, S. Belogurov, I. Belyaev, V. Egorychev, D. Golubkov, T. Kvaratskheliya, I.V. Machikhiliyan, D. Savrina[35], A. Semennikov, A. Zhokhov
[34] *Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia*

A. Berezhnoy, M. Korolev, A. Leflat, N. Nikitin
[35] *Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia*

S. Filippov, E. Gushchin, L. Kravchuk
[36] *Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia*

T. Likhomanenko, V. Shevchenko, A. Ustyuzhanin
[37] *National Research Centre Kurchatov Institute, Moscow, Russia, associated to* [34]

A. Bondar, S. Eidelman, P. Krokovny, V. Kudryavtsev, L. Shekhtman, V. Vorobyev
[38] *Budker Institute of Nuclear Physics (SB RAS) and Novosibirsk State University, Novosibirsk, Russia*

A. Artamonov, K. Belous, R. Dzhelyadin, Yu. Guz[43], A. Novoselov, V. Obraztsov, A. Popov, V. Romanovsky, M. Shapkin, O. Stenyakin, O. Yushchenko
[39] *Institute for High Energy Physics (IHEP), Protvino, Russia*

A. Badalov, M. Calvo Gomez[p], L. Garrido, R. Graciani Diaz, E. Graugés, C. Marin Benito, E. Picatoste Olloqui, V. Rives Molina, H. Ruiz, X. Vilasis-Cardona[p]
[40] *Universitat de Barcelona, Barcelona, Spain*

B. Adeva, P. Alvarez Cartelle, A. Dosil Suárez, V. Fernandez Albor, A. Gallas Torreira, J.A. Hernando Morata, G.P. Julian, A. Pazos Alvarez, E. Perez Trigo, M. Plo Casasus, A. Romero Vidal, J.J. Saborido Silva, B. Sanmartin Sedes, C. Santamarina Rios, M. Seco, P. Vazquez Regueiro, C. Vázquez Sierra, M. Vieites Diaz
[41] *Universidad de Santiago de Compostela, Santiago de Compostela, Spain*

F. Martinez Vidal, A. Oyanguren, P. Ruiz Valls, C. Sanchez Mayordomo
[42] *Instituto de Fisica Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain, associated to* [40]

F. Alessio, F. Archilli, C. Barschel, S. Benson, J. Buytaert, D. Campora Perez, L. Castillo Garcia, M. Cattaneo, Ph. Charpentier, K. Ciba, X. Cid Vidal, M. Clemencic, J. Closier, V. Coco, P. Collins, G. Corti, B. Couturier, C. D'Ambrosio, F. Dettori, A. Di Canto, H. Dijkstra, P. Durante, M. Ferro-Luzzi, C. Fitzpatrick, R. Forty, M. Frank, C. Frei, C. Gaspar, V.V. Gligorov, H. Gordon, L.A. Granado Cardoso, T. Gys, C. Haen, J. He, T. Head, E. van Herwijnen, R. Jacobsson, C. Joram, B. Jost, M. Karacson, T.M. Karbach, D. Lacarrere, B. Langhans, R. Lindner, C. Linn, G. Liu, S. Lohn,

R. Matev, Z. Mathe, S. Neubert, N. Neufeld, J. Panman, M. Pepe Altarelli, N. Rauschmayr, M. Rihl, S. Roiser, L. Roy, T. Ruf, H. Schindler, B. Schmidt, A. Schopper, R. Schwemmer, S. Sridharan, F. Stagni, V.K. Subbiah, F. Teubert, E. Thomas, D. Tonelli, M. Ubeda Garcia, J. Wicht, K. Wyllie, A. Zvyagin
[43]*European Organization for Nuclear Research (CERN), Geneva, Switzerland*

V. Battista, A. Bay, F. Blanc, J. Bressieux, M. Dorigo, F. Dupertuis, S. Giani', G. Haefeli, P. Jaton, C. Khurewathanakul, I. Komarov, V.N. La Thi, N. Lopez-March, R. Märki, B. Muster, T. Nakada, A.D. Nguyen, T.D. Nguyen, C. Nguyen-Mau[q], J. Prisciandaro, A. Puig Navarro, B. Rakotomiaramanana, J. Rouvinet, O. Schneider, F. Soomro, P. Szczypka[43], M. Tobin, S. Tourneur, M.T. Tran, G. Veneziano, Z. Xu
[44]*Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*

J. Anderson, R. Bernet, E. Bowen, A. Bursche, N. Chiapolini, M. Chrzaszcz[29], Ch. Elsasser, F. Lionetto, P. Lowdon, K. Müller, N. Serra, O. Steinkamp, B. Storaci, U. Straumann, M. Tresch, A. Vollhardt
[45]*Physik-Institut, Universität Zürich, Zürich, Switzerland*

R. Aaij, S. Ali, M. van Beuzekom, P.N.Y. David, K. De Bruyn, C. Farinelli, V. Heijne, W. Hulsbergen, E. Jans, P. Koppenburg[43], A. Kozlinskiy, J. van Leerdam, M. Martinelli, M. Merk, S. Oggero, A. Pellegrino, H. Snoek, J. van Tilburg, P. Tsopelas, N. Tuning, J.A. de Vries
[46]*Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands*

J. van den Brand, T. Ketel, R.F. Koopman, R.W. Lambert, D. Martinez Santos, G. Raven, M. Schiller, V. Syropoulos, S. Tolk
[47]*Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, The Netherlands*

G. Onderwater
[48]*KVI - University of Groningen, Groningen, The Netherlands, associated to* [46]

E. Pesen
[49]*Celal Bayar University, Manisa, Turkey, associated to* [43]

A. Dovbnya, S. Kandybei, I. Raniuk
[50]*NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine*

O. Okhrimenko, V. Pugatch
[51]*Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine*

S. Bifani, N. Farley, P. Griffith, I.R. Kenyon, C. Lazzeroni, J. McCarthy, L. Pescatore, N.K. Watson, M.P. Williams
[52]*University of Birmingham, Birmingham, United Kingdom*

M. Adinolfi, J. Benton, N.H. Brook, A. Cook, M. Coombes, J. Dalseno, T. Hampson, S.T. Harnew, P. Naik, E. Price, C. Prouve, J.H. Rademacker, S. Richards, D.M. Saunders, N. Skidmore, D. Souza, J.J. Velthuis, D. Voong
[53]*H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom*

W. Barter, M.-O. Bettler, H.V. Cliff, H.-M. Evans, J. Garra Tico, V. Gibson, S. Gregson, S.C. Haines, C.R. Jones, M. Sirendi, J. Smith, D.R. Ward, S.A. Wotton, S. Wright
[54]*Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom*

J.J. Back, T. Blake, D.C. Craik, D. Dossett, T. Gershon, M. Kreps, C. Langenbruch, T. Latham, D.P. O'Hanlon, T. Pilař, A. Poluektov[38], M.M. Reid, R. Silva Coutinho, C. Wallace, M. Whitehead
[55]*Department of Physics, University of Warwick, Coventry, United Kingdom*

S. Easo[43], R. Nandakumar, A. Papanestis[43], S. Ricciardi, F.F. Wilson
[56]*STFC Rutherford Appleton Laboratory, Didcot, United Kingdom*

L. Carson, P.E.L. Clarke, G.A. Cowan, R. Currie, S. Eisenhardt, D. Ferguson, D. Lambert, H. Luo, A.-B. Morris, F. Muheim, M. Needham, S. Playfer, A. Sparkes
[57]*School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom*

M. Alexander, J. Beddow, L. Eklund[43], D. Hynds, s Karodia, I. Longstaff, S. Ogilvy, M. Pappagallo, P. Sail, F.J.P. Soler, P. Spradlin
[58]*School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom*

A. Affolder, T.J.V. Bowcock, H. Brown, G. Casse, S. Donleavy, K. Dreimanis, S. Farry, K. Hennessy, D. Hutchcroft, M. Liles, B. McSkelly, G.D. Patel, A. Pritchard, K. Rinnert, T. Shears, N.A. Smith
[59]*Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom*

G. Ciezarek, S. Cunliffe, U. Egede, A. Golutvin[34,43], S. Hall, M. McCann, P. Owen, M. Patel, K. Petridis, I. Sepp, E. Smith, W. Sutcliffe, D. Websdale
[60]*Imperial College London, London, United Kingdom*

R.B. Appleby, R.J. Barlow, T. Bird, P.M. Bjørnstad, S. Borghi, D. Brett, J. Brodzicka, S. Chen, S. De Capua, G. Dujany, M. Gersabeck, J. Harrison, C. Hombach, S. Klaver, G. Lafferty, A. McNab, C. Parkes, A. Pearce, S. Reichert, E. Rodrigues, P. Rodriguez Perez, M. Smith
[61]*School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom*

S.-F. Cheung, D. Derkach, T. Evans, R. Gauld, E. Greening, N. Harnew, D. Hill, P. Hunt, N. Hussain, J. Jalocha, M. John, D. Johnson, O. Lupton, S. Malde, E. Smith[56], S. Stevenson, C. Thomas, S. Topp-Joergensen, N. Torr, G. Wilkinson
[62]*Department of Physics, University of Oxford, Oxford, United Kingdom*

I. Counts, P. Ilten, M. Williams
[63]*Massachusetts Institute of Technology, Cambridge, MA, United States*

R. Andreassen, A. Davis, W. De Silva, B. Meadows, M.D. Sokoloff, L. Sun
[64]*University of Cincinnati, Cincinnati, OH, United States*

J.E. Andrews, R. Cenci, B. Hamilton, A. Jawahery, J. Wimberley
[65]*University of Maryland, College Park, MD, United States*

M. Artuso, S. Blusk, A. Borgia, T. Britton, S. Ely, P. Gandini, J. Garofoli, B. Gui, C. Hadjivasiliou, N. Jurik, M. Kelsey, R. Mountain, B.K. Pal, T. Skwarnicki, S. Stone, J. Wang, Z. Xing, L. Zhang.
[66]*Syracuse University, Syracuse, NY, United States*

[a]*Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil*
[b]*P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia*
[c]*Università di Bari, Bari, Italy*
[d]*Università di Bologna, Bologna, Italy*
[e]*Università di Cagliari, Cagliari, Italy*
[f]*Università di Ferrara, Ferrara, Italy*
[g]*Università di Firenze, Firenze, Italy*
[h]*Università di Urbino, Urbino, Italy*
[i]*Università di Modena e Reggio Emilia, Modena, Italy*
[j]*Università di Genova, Genova, Italy*
[k]*Università di Milano Bicocca, Milano, Italy*

[l] *Università di Roma Tor Vergata, Roma, Italy*
[m] *Università di Roma La Sapienza, Roma, Italy*
[n] *Università della Basilicata, Potenza, Italy*
[o] *AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland*
[p] *LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain*
[q] *Hanoi University of Science, Hanoi, Viet Nam*
[r] *Università di Padova, Padova, Italy*
[s] *Università di Pisa, Pisa, Italy*
[t] *Scuola Normale Superiore, Pisa, Italy*
[u] *Università degli Studi di Milano, Milano, Italy*

We dedicate this document to the memory of Ioana Videau,
who passed away in March 2014.

Contents

Chapter 1

Introduction

The LHCb experiment will be upgraded during the Long Shutdown 2 (2018-2019) to facilitate recording proton-proton collision data at $\sqrt{s} = 14\,\text{TeV}$ with an instantaneous luminosity of $2 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$. A total dataset of at least $50\,\text{fb}^{-1}$ will be collected by the upgraded experiment in less than ten years. Analysis of this data will produce unprecedented precision in the b and c-quark flavour sectors [1]. This Technical Design Report (TDR) will address two key features of the upgrade: the *trigger-less readout system* and the *full software trigger*.

One of the main limitations of the current experiment is that the collision rate must be reduced to the readout rate of 1.1 MHz within a fixed latency. This reduction is achieved using the basic signatures available to the Level-0 hardware trigger. The largest inefficiencies in the entire trigger chain, especially for purely hadronic decays, occur at the Level-0 decision. Therefore, one of the main objectives of the LHCb upgrade is to remove this bottleneck by implementing a trigger-less readout system. The readout system will be composed of the event builder, the Timing and Fast Control (TFC) distribution, the Experiment Control System (ECS) and the Event Filter Farm (EFF). Such a system will allow the full inelastic collision rate of 30 MHz to be processed by the full software trigger.

The full software trigger will run on the LHCb EFF. The selections applied must be as similar as possible to those applied in offline analyses to maximize trigger efficiencies and to minimize systematic uncertainties. Both aspects are required to measure b and c-quark observables with high precision. Sophisticated algorithms will be employed to achieve this increasing the hadronic event yields by about a factor ten with respect to Run 1.

The requirements for the readout system and the full software trigger are presented in Chapter 2. The implementation of the readout system is described in detail in Chapter 3. Chapter 4 discusses the implementation of the full software trigger and its expected performance as a function of the output bandwidth. The readout system and trigger have evolved considerably since the Framework TDR [2]. These evolutions are described in the next two sections.

1.1 Evolution of the readout system

The main challenge for the trigger-less readout is to build a cost-effective system that can handle the sizable bandwidth of 4 TBytes/s.

The event builder described in the Framework TDR is similar to the one used during Run 1 but with a much larger bandwidth. Its design featured two main components: readout boards and a large local area network. The readout boards would have been an ATCA[1] mother board equipped with four AMC[2] mezzanine cards. In this design, each AMC interfaces the front-end electronics to the computer network and processes 108 Gbit/s. The input and output are serial optical links running the 4.8 Gbit/s GBT [3] and the 10 Gigabit Ethernet protocols, respectively. The network is composed of several core routers equipped with a large quantity of memory to handle traffic congestion. These core routers connect all of the readout boards to all of the PC servers at the head of each sub-farm. Each event fragment belonging to the same collision is pushed to one PC server in which they are assembled and then distributed to one node of the sub-farm.

Since the Framework TDR, a new approach has been developed. A new building at the surface will house the core routers and the EFF. This permits utilizing a cost-effective solution in which the readout board, the router and the EFF are located in close proximity. Long distance optical links of 300 meters will be required between the front-end electronics located underground and the readout boards located at the surface. This new design makes it possible to use high-bandwidth cost-effective data-centre link-technology for the event-building.

The central part of the event builder in such an architecture consists of dedicated PC servers. These servers interface the front-end electronics via a readout unit embedded in each PC server. Therefore, the form factor of the readout unit is PCI Express instead of AMC. The input is realised via serial optical links running the GBT protocol, while the output is directed to the PC motherboard using the PCI Express Gen3 protocol. The readout unit can process a maximum of 100 Gbits/s. The PC server is also interfaced to the computer network. All of the PC servers involved in the event building are connected by a large-scale network running 100 Gbit/s bidirectional links. This allows the exchange of event fragments between PC servers, with one of the servers collecting the fragments that belong to the same collision. The PC server is also connected to a sub-farm that runs the trigger algorithms. After an extensive R&D program, this architecture was chosen by the collaboration in March 2014.

1.2 Evolution of the trigger

The trigger presented in the Framework TDR was designed to run at an instantaneous luminosity of $1 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Its architecture, which is similar to that used in Run 1, is composed of two main blocks: the *Low Level Trigger* (LLT) and the *High Level Trigger*

[1] Advanced Telecommunications Computing Architecture
[2] Advanced Mezzanine Card

(HLT). The LLT is essentially the Run 1 Level-0 hardware trigger modified to run within the new readout architecture. The LLT selects events containing clusters with high transverse energy in the calorimeters or tracks with high transverse momentum in the muon detector. The HLT is divided into two parts, *HLT1* and *HLT2*, which are executed in sequence. HLT1 runs a partial reconstruction and HLT2 runs inclusive and exclusive selections. The estimated processing time was 20 ms per event and the output bandwidth was fixed to $20\,\mathrm{kHz} \times 100\,\mathrm{kBytes} = 2\,\mathrm{GBytes/s}$.

After the publication of the Framework TDR, it was decided that the operational luminosity will be $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. This luminosity will be kept constant during the fill using the *luminosity levelling scheme* that was successfully operated in Run 1. An additional difficulty in designing the upgrade trigger was that several options were proposed for the tracking system, each with different characteristics. At the end of 2013, the collaboration selected the following tracking technologies for use in the LHCb upgrade: the VELO Pixel [4]; the Upstream Tracker (UT); and the Scintillating Fiber Tracker (SciFi) [5]. The VELO Pixel and UT reduce the time of the tracking sequence by a factor three [5], while the SciFi permits performing the complete tracking sequence in the HLT in less than 10 ms. The maximum processing time allowed for each event in the EFF running in 2020 has been estimated to be 13 ms. Therefore, the complete tracking sequence can be run in the HLT with time remaining for selections [6].

An extensive R&D program was carried out to establish the feasibility of a trigger running at $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Three designs were studied: the full software trigger based on the HLT; the previous option assisted by a co-processor running in FPGA called the *Tracking Processing Unit*; and the implementation of the LLT in the new readout architecture followed by an HLT. In March 2014, following a detailed review, the full software trigger became the baseline design since it provides the maximum flexibility and is robust against fast obsolescence of technological products. The main conclusion from this review was that the upgrade trigger is feasible at $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The Tracking Processing Unit was studied by the Milan and Pisa groups. It performs the *upstream tracking* by looking for tracks between the VELO and the UT using the so-called *artificial retina* algorithm [7]. It is implemented in the FPGA of the common readout board allowing massively parallel computing. Detailed simulations of the hardware in the LHCb framework show that the performance are similar to the offline upstream tracking and that tracks are found in less than 500 ns. The description of the tracking processing unit and of its performance can be found in Ref. [8]. This approach is interesting as it reduces the time spent in the HLT tracking sequence by 25% and confirms that an efficient trigger is doable in the upgrade condition. However, it also adds complexity to the trigger system and its offline simulation without reducing significantly the size of the farm required. After a detailed comparison with the baseline option this alternative was not retained.

The LLT reduces the efficiency for hadronic channels which is the reason it was not chosen as part of the baseline design. However, the LLT can act as a safety net to protect the event building and to regulate the rate at the input of the EFF. Furthermore, the LLT would be used at the start of Run 3 if the EFF is not fully in place at the start of data taking. The implementation of the LLT in the new readout architecture was studied and it was shown that the LLT can be implemented in the readout board for any form factor [9]. The CPU power available in the event building farm permits implementing the LLT in software. This solution represents the best compromise between cost, flexibility and added security.

Chapter 2

Requirements

The requirements for the trigger-less readout system are summarised in Table 2.1.

Table 2.1: Boundary conditions for the online system.

Event rate	40 MHz
Mean nominal event size	100 kBytes
Readout board bandwidth	up to 100 Gbits/s
CPU nodes	up to 4000

The maximum bandwidth for the readout board is fixed by the 16-lanes PCIe Gen3 protocol. It is theoretically limited to 128 Gbit/s. The quoted number is below this limit to keep some safety margin and to match the bandwidth required by 24 GBT links when they are fully loaded with a data bandwidth of 4.5 Gbit/s.

The maximum number of CPU nodes comes from the power, cooling and space constraints of the new data-centre.

The aim of the full software trigger is to select beauty and charm particles decaying into a large variety of final states with the highest efficiency and purity, minimising systematic uncertainties. The requirements for the full software trigger are summarised in Table 2.2.

Table 2.2: Requirements for the full software trigger.

Instantaneous luminosity	$2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$
Pile-up	7.6
Input rate	30 MHz
Maximum processing time per event	13 ms
Output bandwidth	$20\,\mathrm{kHz} \times 100\,\mathrm{kB} = 2\,\mathrm{GByte/s}$

The input rate is defined by the maximum number of bunches of protons allowed in the machine per beam (2808) due to the gaps associated with the rise time of the PS/SPS/LHC injection kickers and the LHC dump kicker.

The maximum processing time is determined by the number of CPU nodes in the EFF and by the number of HLT processes running in each node. The given number has been obtained by scaling the farm of Run 1 assuming a factor 16 from Moore's law. It will discussed in detail in Sect. 3.6.

The maximum output bandwidth is the one given in the Framework TDR, which can be optimised in the future to enhance the physics output of the experiment. The trigger performance as a function of the output bandwidth will be reviewed in the Sect. 4.6.7.

Chapter 3

Online

The Online system provides the infrastructure for the operation of the LHCb experiment, in particular the detectors and the trigger. It consists of the common readout-system, the data acquisition, the experiment control system, the timing and fast control system, the event-filter farm and general online infrastructure.

In many respects the system represented here is a natural evolution of the current LHCb online system described in Ref. [10], where simply more modern technologies replace some of the current ones.

The guiding principles of the online system are simplicity of design, use of standards and common solutions wherever possible and an emphasis of architecture over specific technology.

3.1 System design

The upgraded LHCb experiment will be running at a constant instantaneous luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The yield in events useful for physics will be maximised by switching to a fully synchronous readout of each bunch-crossing. Consequently no more trigger decision is sent to the front-end electronics, making the upgraded LHCb readout completely trigger-free. This requires a change of all front-end electronics of all the detectors. Also several detectors will be replaced or upgraded. The details can be found in the Framework TDR [2] and in various sub-system TDRs [4, 5, 11].

A new readout system is required to accommodate all these changes and to make best use of technologies which have become available since the original design of LHCb, which is described in Ref. [12].

The common detector link (GOL) will be replaced by the radiation-hard Versatile Link [13]. For cost reasons it has been decided in the very beginning that all detectors must perform zero-suppression on the front-end, before sending data to the DAQ. The Versatile Link can be operated as bi-directional and as simplex link. Because of the large number of links required for the data acquisition it has further been decided to separate data from control and timing information. Data will be pushed over simplex links, while control and timing information use bi-directional links. On all ECS/TFC and the majority of the DAQ links the GBT protocol [3] will be used.

An important aspect of the system is that the same generic hardware will be used to implement the data acquisition, fast control and slow control elements outside the detector, namely the PCIe40 board, described in detail in the Sect. 3.3. The different functionalities will be selected by firmware.

The event-builder connects the readout-boards to the filter-farm nodes, where the HLT will be running. The cost of the event-builder is minimised by using cost-effective data-centre technology in the network and ensuring short distances between components. Data-centre technologies in the network require the use of PCs as end-points. The most compact system on the other hand is achieved by concentrating all DAQ and TFC and most ECS hardware in the data-centre on the surface. This in turn requires to operate the detector Versatile Links over a relatively long distance and is discussed extensively in the Sect. 3.2.

The overall readout architecture is illustrated in Fig. 3.1. The role of the ECS is largely

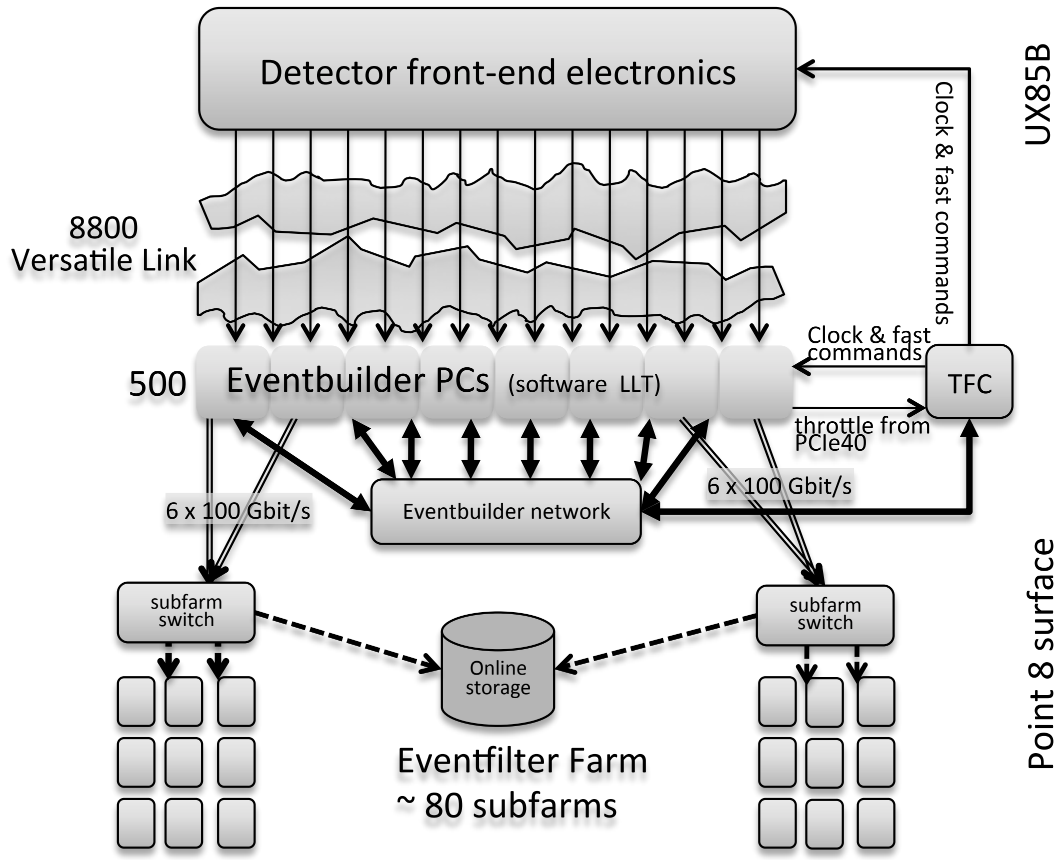


Figure 3.1: The architecture of the upgraded LHCb readout-system. All elements shown in the diagram are connected to and controlled by the ECS.

unchanged with respect to the original system [10]. Partitioning facilitates debugging and commissioning. The use of standard frameworks (Joint COntrols Project JCOP [14]) ensures a coherent, easy-to-operate system, with a high degree of automation and intelligent auto-recovery to optimise efficiency.

Apart from the distribution of the critical timing signals and fast commands to the front-end the TFC implements as before a central, robust flow-control, the so-called *throttle*. Backpressure in the readout-system, from the PCIe40 boards onwards, will eventually make one of the PCIe40 trigger the *throttle* and ensure that synchronously the influx of new events is stopped until the back-pressure has stopped.

Finally the event-filter farm and the storage system need above all to be scalable and cost-effective. The farm is also designed to be open to the use of upcoming technologies, while the baseline system has traditional dual-socket x86 servers in mind.

3.1.1 Size of the system

The size of the system is given in Table 3.1. Some of the numbers, such as the number of readout-boards are rounded up limits useful for system design and budgeting. The exact numbers will be most likely be somewhat lower and they will be determined once all the front-end designs are frozen. In any case the system design scales well and the exact numbers do not impact on the design. Most of the Versatile Links for the DAQ

Table 3.1: Summary of the readout-system

Versatile Links for DAQ	8800
Mean nominal total event-size	100 kB
PCIe40 boards for DAQ	500
Versatile Links / readout board (DAQ)	up to 48
Event builder-PCs	500
PCIe40 boards for ECS and TFC (SOL40)	77
Core switch ports (100 Gbit/s)	500
Event-filter nodes	up to 4000
Output rate	20 kHz
Nominal instantaneous output rate	2 GB/s

will use the so-called wide mode of the GBT, corresponding to an effective bandwidth of 4.5 Gbit/s. It is assumed that a single PCIe40 card is used in each event-builder PC. The *limitation* of 4000 event-filter nodes comes from the power, cooling and space constraints of the data-centre, assuming a node to need 1U rack-space and about 400 W of power.

Nominal event-size

The nominal event-size is estimated from the number of Versatile Links as reported in the various detector TDRs [4, 5, 11]. Rounding up this gives 8800 links with a usage factor of about 80%. Therefore, assuming 30 MHz of non-empty bunch-crossings this leads to a nominal event size of 100 kB as shown in Table 3.1. This number agrees within 15% with an estimation from the detector occupancies. Ongoing work on better compactification and compression, for instance by suppression of multiple identical headers in the PCIe40 FPGAs has not been taken into account. Moreover the system scales well and the cost for the event-builder and the storage-system described in this TDR vary essentially linearly with the event-size.

Output bandwidth and Online storage

Online storage serves solely to bridge outages in the connection to permanent tape storage in the IT department off-site. One week is deemed sufficient. In the past there never have been lengthy interruptions. The output bandwidth *to* the Online storage has to match the set output rate from the HLT. Only 20 disks are required to handle the nominal output rate of 2 GB/s since one disk can sustain 100 MB/s. The number of disks increase to 100 when the output rate is 10 GB/s[1].

The output *from* the Online storage to the permanent storage is determined by the capacity of the LAN connection between the LHCb online system and the Tier0 in Meyrin. Assuming the use of 10 Gigabit technology one pair of fibres can transport 1 GB/s[2]. More than 20 pairs are available. It is clear that technically also the 10 GB/s case poses no problem.

3.2 Long distance cabling

As has been argued in Sect. 3.1, that optimal density and minimal distances are achieved by bringing the readout-boards to the surface. This however requires that the Versatile Links from the detector operate over the entire distance of approximately 300 m between the underground areas, UX85B, and the location of the planned data-centre at the surface.

The Versatile Link was originally conceived with a distance of about 60 to 80 m in mind, sufficient to connect the detector front-end to the readout-electronics behind the shielding wall. In the following the feasibility of running it over 300 m will be demonstrated.

3.2.1 Implementation at Point 8

The preferred path of the optical links is through the PM shaft. There is an alternative way traversing the shielding wall and going up through the PZ shaft, but it is less convenient. The length of both paths is however almost the same.

The total number of fibres for both DAQ and ECS/TFC has been estimated to be 17000 as an upper limit, including spares. Each fibre has three break-points: one on the

[1] In practice at least 3 times as many disks are required to ensure reliability and give enough bandwidth for two reads

[2] This is the cheapest way of doing this, it is of course possible to use multiple wavelengths (CWDM or DWDM) on the same fibre, this requires more expensive optics on both sides.

surface, one in one of the distribution locations close to the detector and one at or very close to the actual detector front-end electronics.

Two options are considered for the fibre-installation: pre-connectorized cables and blown loose fibres which are pre-terminated on one end. Blown fibres need to be terminated with a pigtail, which is spliced onto the loose end of the fibre. The splice introduces an additional attenuation of at most 0.3 dB. Apart from that the optical budget does not depend on the installation of the fibres, and all of the following is applicable to both installation options. Cost will determine the final choice.

3.2.2 Measurements

Several performance measurements have been done on various fibres to determine the feasibility of a 300 m readout on a small set of optical components. The components used are a dual Versatile Link transmitter, Versatile Link transceiver and an Avago MiniPod receiver and transmitter. The latter has been chosen for the PCIe40 readout board.

The laser diode used in the Versatile Link [13] is based on a commercial 10 Gbit/s transmitter. These transmitters are designed to drive a 10 Gbit/s signal over a distance of up to 300 m of OM4 fibre. The actual link speed for the Versatile Link is 4.8 Gbit/s and it is reasonable to assume that at this speed the 300 m transport should be no problem over OM4. Since the link is running only at 4.8 Gbit/s we also evaluated the link performance with low grade OM3 fibres. OM3 fibres have the same attenuation per distance unit but have weaker constraints on modal dispersion and are usually used for shorter distances.

At the 5 GHz signalling rate, which is used by the Versatile Link, the mode dispersion is much less pronounced than at 10 GHz, which is believed to be the main reason why OM4 fibres, which differ only in their reduced mode dispersion from OM3, do not show any advantage in our measurements. Hence their significantly higher cost is not justified in our application.

The usual approach for determining the feasibility of a fibre installation is to calculate the optical power budget for the proposed system. This budget calculation is done by subtracting the receiver sensitivity from the optical launch power and comparing the obtained power margin with all the sources of optical signal loss in the system. Loss sources are:

- Attenuation inside the fibre through scattering
- Attenuation through connectors and splices
- Signal degradation through dispersion
- Transmitter and Receiver ageing
- Radiation damage

Most of these items can be obtained by looking at the specifications of the components involved. Unfortunately there are no commercial links running at 4.8 Gbit/s and so there is no specification for the dispersion value.

For the radiation damage the values defined by the Versatile Link project [15] have been used. They correspond to an integrated dose of 10 kGy and a 1 MeV neutron equivalent fluence of $5 \times 10^{14}\,\mathrm{n/cm^2}$. This is significantly higher than the worst case estimated for LHCb which is $10^{13}\,\mathrm{n/cm^2}$.

Another factor is the optical receiver. Since it is a commercial component, made for 10 Gbit/s operation we assume that it will perform better at 4.8 Gbit/s than specified. To quantify this effect we also determined the sensitivity of a commercial receiver at 10 and 4.8 Gbit/s.

The following subsections summarize the optical dispersion and receiver sensitivity values determined for a 4.8 Gbit/s link [16] Since we did not know the exact distance at the time of the measurements, the conservative distance of 400 m is used.

3.2.3 Optical Dispersion of OM3 and OM4

We measured both OM3 and OM4 fibres of the major fibre manufacturers to establish their usability in the proposed readout scheme. At the speed relevant for the Versatile Link no difference between OM3 and OM4 could be found [16]. Given the price-difference, OM4 will only be considered if the long-term tests in 2015 will indicate any unexpected advantage of OM4 over OM3.

3.2.4 Optical receiver sensitivity at 4.8 Gbit/s

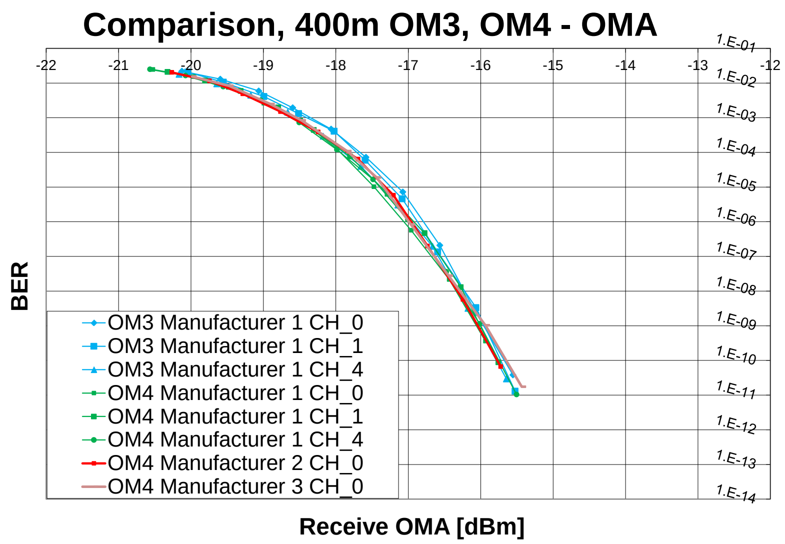


Figure 3.2: Measurement of the bit error rate as a function of the receives optical modulation amplitude for different OM3 and OM4 fibres.

The 400 m fibre introduces only marginal dispersion and so the receiver sensitivity can be directly obtained from Fig. 3.2. It shows the bit error rate as a function of the received optical modulation amplitude. The receiver sensitivity is obtained by extrapolating the curve to the bit error rate of 10^{-13} which is the standard value for 10 Gbit/s Ethernet as well as the Versatile Link. The receiver sensitivity is approximately -15 dBm at 4.8 Gbit/s instead of the -11.1 dBm specified by the manufacturer for 10 Gbit/s operation. For reasons that are detailed in [16] a more conservative value of -14.2 dBm is assumed for the sensitivity of the receiver.

3.2.5 Long distance cabling feasibility

Table 3.2: Revised optical power budget calculation after measuring the distortion penalties for various fibres at 4.8 Gbit/s. The value for fibre loss is calculated for a range of 400 m to be consistent with the measurements we did. The components used for these measurements are a dual Versatile Link transmitter (VTTx), Versatile Link transceiver (VTRx) and an Avago MiniPod (MP) receiver and transmitter.

Description	Unit	DAQ VTTx to MP		Control MP to VRx	
		spec.	meas.	spec.	meas.
Transmitter OMA	dBm	-5.2	NM	-3.2	NM
Receiver sensitivity	dBm	-11.1	-14.2	-13.1	NM
Power budget	dB	5.9	9.0	9.9	9.9
Fibre loss (2.3 dB/km)	dB	0.9	NM	0.9	NM
Connectors (0.5 dB/pair)	dB	1.5	NM	1.5	NM
Disp. (400 m, 4.8 Gbit/s)	dB	2.4	0.5	2.4	0.5
TX Radiation penalty	dB	0	NM	-	-
RX Radiation penalty	dB	-	-	2.5	NM
Fibre Radiation penalty	dB	0.1	NM	0.1	NM
Margin	dB	1.0	6.0	2.5	4.4

Table 3.2 summarises the power budget calculation for the 4.8 Gbit/s readout link. For each direction, DAQ and Controls, the values as specified by the manufacturer (spec.) and the one measured (meas.) are given. The specified value is used when it could not be measured (NM).

The margins are given for a bit error rate of 10^{-13}, which is the standard for 10 Gbit/s Ethernet and also the Versatile Link. The link is valid if the margin is at least 0 dB. Best practice is to have 3 dB to accommodate regular ageing of the installation. For a link running at 4.8 Gbit/s, the optical margin is 6.0 dB for the DAQ direction and 4.4 dB for the controls direction. These margins make the link definitely workable.

The tests performed on the long distance optical link are encouraging but need to be completed with further tests under realistic conditions. A couple of optical fibre ribbons will be installed between the underground area and the surface in the second part of 2014. They will allow to test the deployment procedure and to run long-term tests in-sit with a significant number of links. In addition, a larger number of optical devices will be measured.

3.3 Readout board

The readout board is a generic component which has been designed for the data acquisition of all detectors, the distribution of the timing and fast commands and the slow control [17].

Several prototypes have been developed between 2011 and 2013. Their role was both to check the feasibility of mapping the readout system over an ATCA architecture and to validate critical technical points like high density design, signal integrity issues in using high speed serial links or DDR3 RAMs, power consumption and cooling as well as the use of complex communication standards.

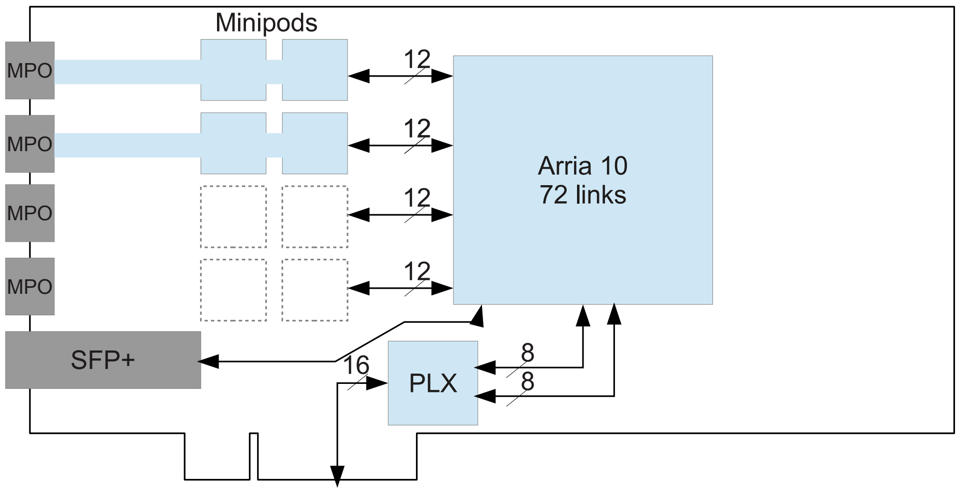


Figure 3.3: Schematic of the readout board when it is implemented using the PCI Express standard.

Detailed studies of the evolution of the network technologies and global optimization of the readout system have shown that a cost effective implementation can be achieved when the readout board is embedded in a PC server. Therefore, the collaboration decided to move to the PCI Express (PCIe) standard for the readout board in March 2014.

The schematic of the PCIe card is shown in Fig. 3.3. The board will present 48 bi-directional optical links for interfacing the FE electronics and one bidirectional optical link for interfacing the TFC. All of them are connected to a large-size Arria 10 FPGA[3]. The latter is also interfaced to the CPU through two 8-lanes PCIe Gen3 buses connected to a PLX PCIe switch to form a 16-lanes PCIe Gen3 bus. Although each input link can carry up to 10 Gbit/s, the maximum data transfer rate is fixed by the PCIe Gen3 output

[3] The production of the Arria 10 is expected to start in 2015.

to about 110 Gbits/s which corresponds to 24 input links, fully loaded, running the GBT protocol with an effective bandwidth of 4.5 Gbit/s.

The performance obtained with the ATCA prototypes allow us to conclude that the feasibility of implementing the readout board using the PCIe standard is assured [18]. An additional cooling and mechanical study is however needed but we can be helped by the numerous cooling solutions available on the market for graphics cards.

A prototype of the PCIe board is in preparation and should be ready end 2014 if no difficulties appear with the implementation of the PCIe standard and in the migration to the new FPGA family.

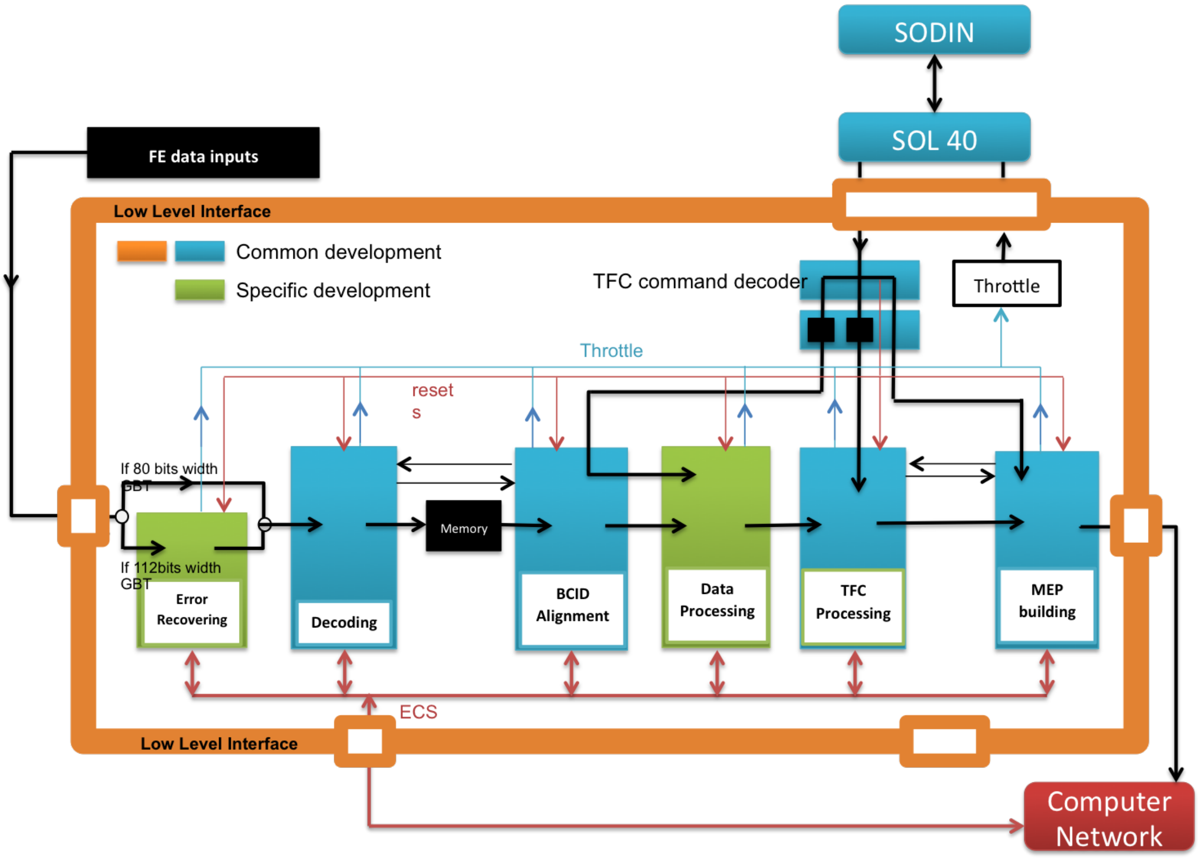


Figure 3.4: The main block of the readout board firmware when it is programmed for data acquisition.

Several functionalities can be obtained with the generic board by programming differently the FPGA. Different flavours will be prepared for the data acquisition of the detectors, TFC and ECS. The firmware of the board will contain an interface layer code, *Low Level Interface*, common to all the flavours as shown in Fig. 3.4. Its aim is to interface the hardware with the user code firmware using common blocks like a GBT decoder or PCIe IP core. The environment to develop the firmware for each flavour of the boards will be common across the entire LHCb experiment, with only the user code being exclusive.

The same approach has also been put in place for a global simulation framework. This collaborative method has been proven to be very effective in reducing nonconformities and to enforce compatibility with specifications. It also reduces considerably the number of components to be developed and saves developer effort in firmware design.

The user code dedicated to the functionality of the readout of events from the trigger-less FE faces considerable challenges. Events will arrive from the FE to the PCIe boards asynchronously across all input links due to the variable latency in compression/zero-suppression mechanisms so the code has to be able to handle a big spread in latency between fragments of the same event. The readout code of the board must decode the frames from the FE and realign them according to their bunch crossing identifier. It then builds an multi-event packet (MEP) and sends it to the DAQ network. Common effort is on-going to find the best technological solutions to this challenge.

3.4 Timing and fast control

The TFC [19] is responsible for controlling and distributing clock, timing and trigger information, synchronous, and asynchronous commands to the entire readout system as described in the global LHCb readout architecture [20]. The system must maintain synchronization across the readout architecture, provide the mechanisms for special monitoring triggers and manage the dispatching of the events to the EFF. It regulates the transmission of events through the entire readout chain taking into account throttles from the readout boards, the LHC filling scheme, back-pressure from the readout network and physics decisions if any. The specifications, functionalities and the full details of the system are published in Ref. [21].

Generally, the information generated and propagated by the TFC system to the entire readout system are:

- the LHC reference clock at 40 MHz, that is the Master clock of all the electronics synchronized to the Master clock of the LHC accelerator;
- commands to synchronously control the processing of events in the readout board;
- commands to synchronously control the processing of events at the Front-End (FE) electronics;
- calibration commands for the detector electronics;
- destination of the Multi-Events Packets and their load balancing.

In addition, FE electronics configuration is generated by the ECS and relayed by the TFC system to the FE boards.

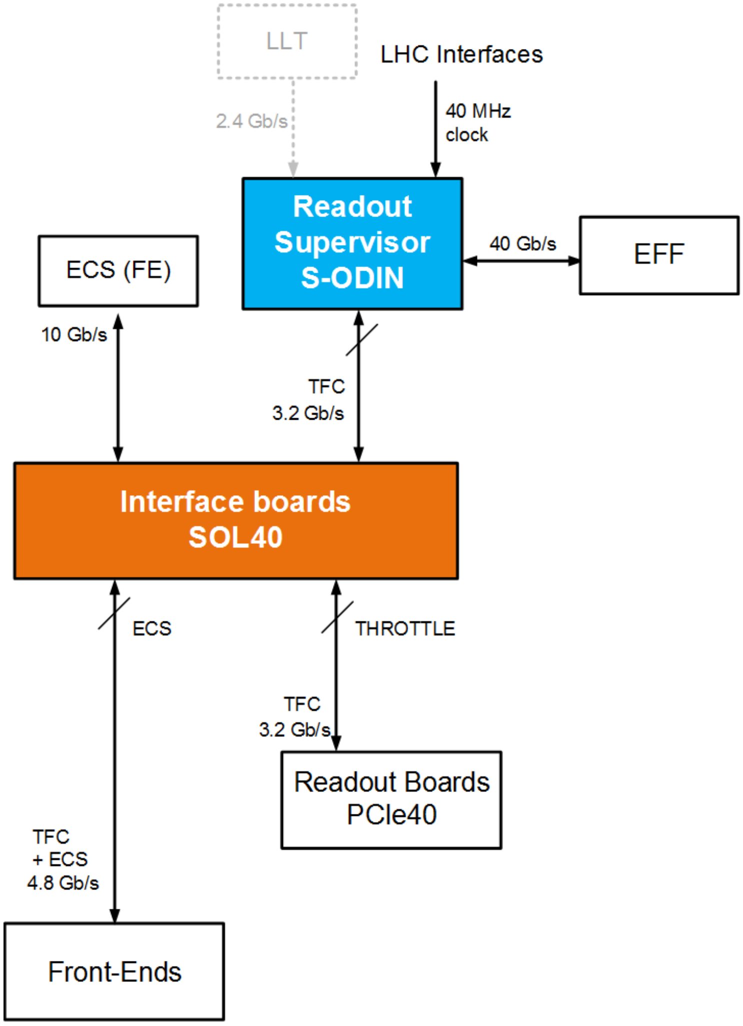


Figure 3.5: Logical architecture of the new TFC system.

3.4.1 TFC architecture, timing and control distribution

The logical scheme of the upgraded TFC architecture and the data flow is represented in Fig. 3.5.

The readout supervisor *S-ODIN* is the TFC Master, being responsible for generating the necessary information and commands.

The sub-detector readout electronics comprises FE and the Back-End (BE) boards. Both are connected to the S-ODIN via a set of 3.2 Gbit/s high-speed bi-directional optical links via multiple interface board, *SOL40*. These connections define the partition granularity. The topology of the connections, defining a partition, is controlled by the TFC to run any ensemble of sub-detectors simultaneously.

In the LHCb upgrade, the FE electronics is trigger-less, *i.e.* no triggers are sent downstream towards the detector, contrary to the current LHCb system. Therefore the

TFC system ensure that the whole readout system is synchronous across the generation, transmission and processing of events. It includes *throttle* mechanism to absorb possible back-pressure from data congestion at the BE and from high usage of the processing farm.

Architecturally, the new TFC system heavily profits from FPGA technologies and the bi-directional capability of the GBT transceiver [3] which carries simultaneously detector data, timing and readout control information, as well as ECS information. The communication in the non-radiation area is also based on serial transmission protocols implemented with the commercial high-speed transceivers available in modern FPGAs. Thus, each element in the TFC system can be seen as a separate FPGA equipped with commercial high-speed transceivers.

The SOL40 boards serves three main purposes:

- Interface all the readout boards to the S-ODIN by fanning-out the synchronous timing and trigger information and fan-in throttle information.
- Interface all the FE electronics to the S-ODIN by relaying the clock, timing and commands information onto fibres towards the FE electronics [22].
- Relay the ECS information [23].

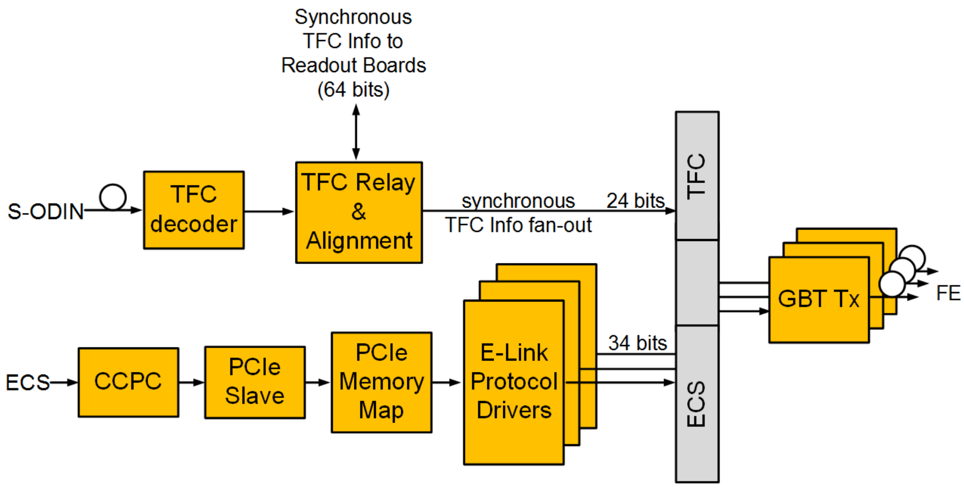


Figure 3.6: Schematic view of the packing mechanism to merge TFC and ECS information on the same GBT links towards the FE electronics.

The TFC and ECS information are merged in the SOL40 boards and transmitted to the FE via GBT links. The logical scheme of the merging is shown in Fig. 3.6. The TFC information is packed into the GBT word at 40 MHz, while the ECS information is packed on best effort to fill up the remaining available bits in the GBT protocol.

The SOL40 boards may be cascaded and configured differently to support different requirements in terms of number of links and bandwidth.

3.4.2 Functionalities of the TFC system

The main functionalities of the TFC are:

- *Readout control*: control of the entire readout system is done by one of the TFC Masters in the pool. The control of the readout implies controlling the trigger rate, balancing the load of events at the processing farm, balancing the load of buffers in the electronics. The TFC system auto-generates internal triggers for calibration and monitoring purposes in a programmable way as well as a full set of commands in order to keep the system synchronous. The details of the specifications for the FE and BE are described in detail in Ref. [22].

- *Event description*: a data bank containing information about the identity of an event as well as the trigger source is transmitted by the central TFC Master to the farm for each event as part of the event data.

- *Event Management*: control of the availability of processing nodes and assignment of the destination for each event based on a credit-scheme mechanism.

- *Partitioning*: this is achieved by instantiating a set of independent TFC Masters in the same FPGA, each of which may be invoked for local sub-detector activities or used to run the whole of LHCb in a global data taking. An internal programmable switch fabric allows routing the information to the desired destination.

- *Coarse and fine time alignment*: the clock reception and control system [24] provides means of aligning the global timing of the experiment. The TFC distribution network transmits a clock to the readout electronics with a known and stable phase at the level of about 50 ps and very low jitter (< 5ps). The latency of the distributed information is fully controlled and maintained constant. Local alignment at the FE and the BE of the individual TFC links is required to assure synchronization of the experiment. It relies on the synchronous reset commands together with Bunch Identifiers and Event Identifiers checks.

- *Luminosity monitoring*: a combination of physics event types is selected by the TFC system in order to allow real-time and absolute luminosity measurements.

- *Run statistics*: information about the trigger rates, run dead-time, number of events accepted, types of events accepted, bunch currents, luminosity and load of buffers are stored in a database to allow retrieving run statistics and information per run or per LHC fill.

3.4.3 Hardware implementation

The functionality and tasks of the TFC system can be achieved by profiting from the same technology backbone of the entire readout system, namely PCIe40 card. The details of TFC aspects in such technology are discussed in [25].

3.5 Event building

Event-building requires to bring the data from all readout-boards into a single CPU node. A local area network (LAN) is used for this. Several LAN technologies are or will be available, however at the time of writing there are only two which have a certain market-share and are known outside very specialized contexts: Ethernet (IEEE 802.3) and InfiniBand [26]. Ethernet exists today in 10 Gbit/s and 40 Gbit/s versions (10G and 40G) and FDR InfiniBand offers effectively about 50 Gbit/s. Measurements and costing are based on these two technologies.

In both cases a variant with 100 Gbit/s speed will be available at the time of the upgrade which will be cheaper and simply reduce the number of necessary links by a factor two. Several architectures are possible [27], only the most cost effective one is presented in the next sections.

3.5.1 Bidirectional event-building

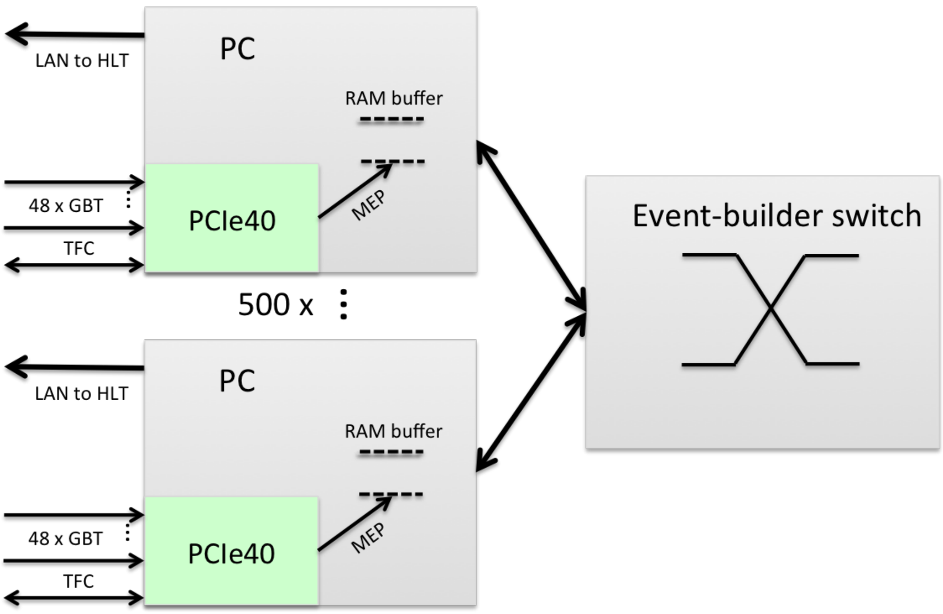


Figure 3.7: The PCIe based readout system. The PCIe40 readout boards are directly connected to the event-builder PCs through 16-lane PCIe edge-connector.

A simplified view of the bidirectional event-building is shown in Fig. 3.7. The main steps of the event building are the following:

- Up to 48 of versatile links are connected to a PCIe40 card. Each card is hosted by a dedicated PC.

- Data are pushed by the PCIe40 FPGA into the main-memory of the hosting PC. Data from several bunch-crossings are coalesced into a multi-event fragment packet (MEP) to reduce the message rate and ensure efficient link-usage.

- Event-building is then done by combining all MEP containing data from the same bunch-crossings in a single PC.

- For each MEP one PC is elected to be the *event-builder* PC. All non-elected PCs will send their MEP to this PC. They will use the same link for this purpose, which they also use to receive the MEPs when they are themselves the elected event-builder. In this way the link is used in both directions and the number of ports in the high-speed event-building network is only as large as the number of event-builder PCs.

The PCs can do some processing of the events as discussed in Sect. 3.5.2. The remaining events will be sent to a sub-farm of compute nodes, where the high-level trigger will process them. The event-builder PC has a dedicated network-link to a simple distribution switch, where the compute units are also connected. This can, but need not be, the same link-technology as used for the event-building. The ultimate choice will be determined by cost.

The detailed view of the bidirectional event-building is shown in Fig. 3.8.

3.5.2 PC-based event-builder performance

The maximum bandwidth of the PCIe40 card is fixed to about 100 Gbit/s by the 16 lanes PCIe Gen3 protocol. Therefore, the load on the event-builder server is quite high. It is at the level of 200 Gbit/s full-duplex when there is no data-reduction before events are sent to the farm-nodes. Such a system has become possible since the advent of the so-called *SandyBridge*[4] micro-architecture which is the first CPU handling the PCIe Gen3 protocol.

We have built a realistic test-system to measure performance, stability and resource-usage. Figure 3.9 shows the data-flow in one event-builder server. The other event-builder servers and the farm-nodes have been emulated by four different servers. The amount of transferred data from one server is the same as it will be in the final system. Since a PCIe40 was not available, it has been emulated using a GPU card from Nvidia. This generator produces the same data-pattern as the FPGA firmware, with all associated protocol overheads. It is 100% compatible with the FPGA version and it can send data over 16-lanes PCIe Gen3 at maximum speed.

The prototype event-builder is using InfiniBand FDR dual-port cards with 16 PCIe Gen3 lanes. This allows event-building and the sending of completed events at 100 Gbit/s over two bundled ports of ~ 54 Gbit/s for each connection. The final system will look the same, except that the link-bundle will be replaced by a single 100 Gbit/s link.

The event-building protocol briefly looks as follows: the event-manager, implemented by the readout-supervisor, elects one of the event-builder PCs for each MEP. All non-elected

[4] `http://en.wikipedia.org/wiki/Sandybridge`

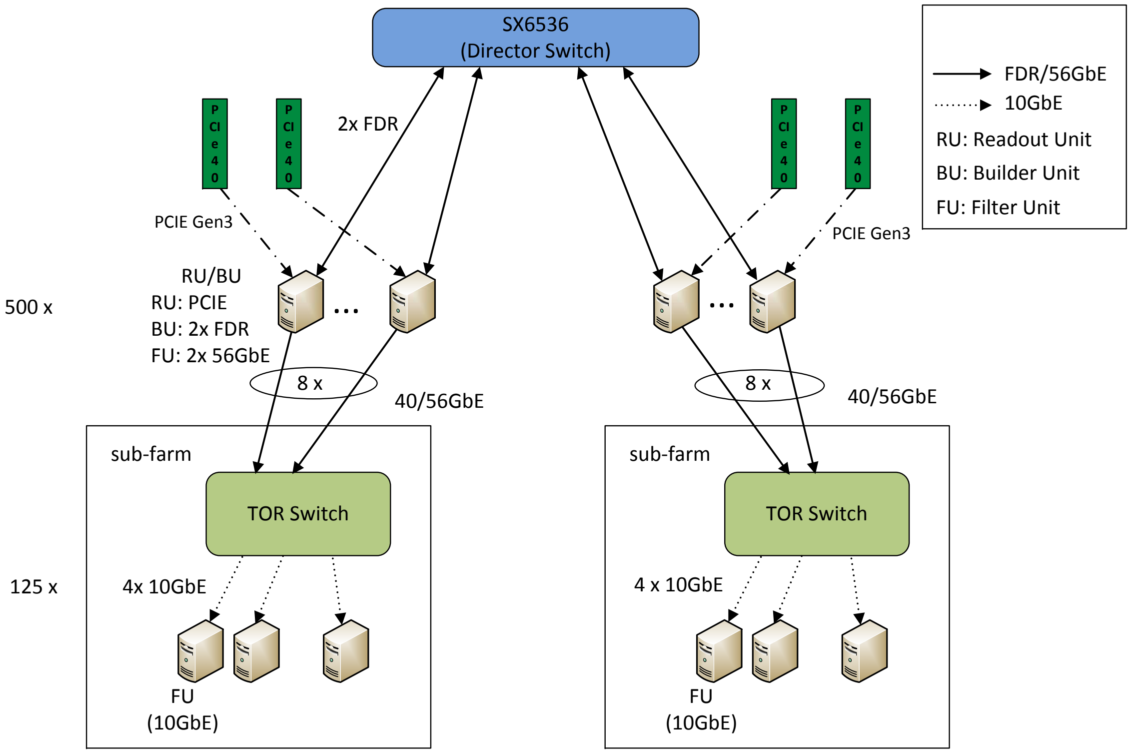


Figure 3.8: Bidirectional event-building using FDR InfiniBand for event-building and Ethernet for the event distribution. The PCIe40 cards are contained in the PCs labelled RU/BU.

PCs will send the MEP, which they got pushed by their PCIe40 card into their memory, to this elected PC. The event-builder PCs are elected in a round-robin fashion. Load balancing is achieved using a simple credit scheme. Normally however it is expected that every event-builder PC sees approximately the same amount of data. The event-building is zero-copy in the sense that the only copy operation is the one from the DMA engine (either of the FPGA or the network interface card) into the memory of the receiving PC. The transfer can be initiated by the senders (*push*) or by the elected event-builder PC (*pull*), which makes in practice little difference. Figure 3.10 shows a long-term test of the event-building performance. Consistently about 90% of the theoretical maximum link-speed (about 104 Gbit/s) is achieved. The server is sustaining four times this I/O as required. Best practice OS tuning for high-performance I/O has been applied[5].

[5]The CPUs used in the test are Intel E5-2670 v2 with a C610 chipset. The servers are equipped with 1866 MHz DDR3 memory in optimal configuration. Hyper-threading has been enabled.

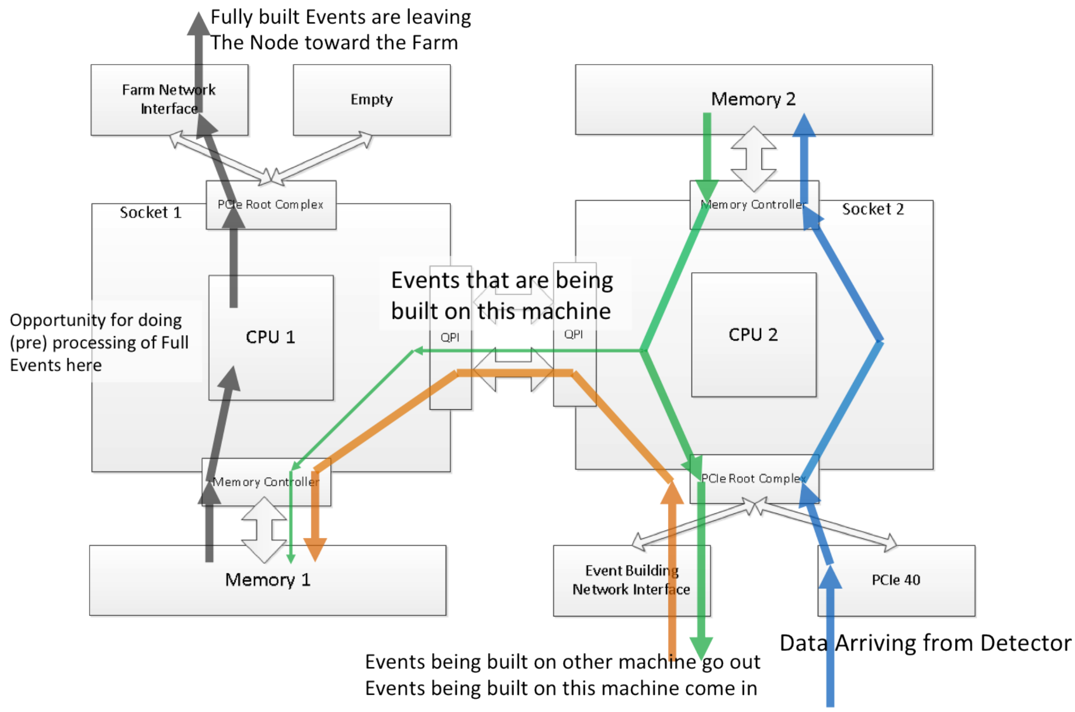


Figure 3.9: Data-flow in the event-builder server

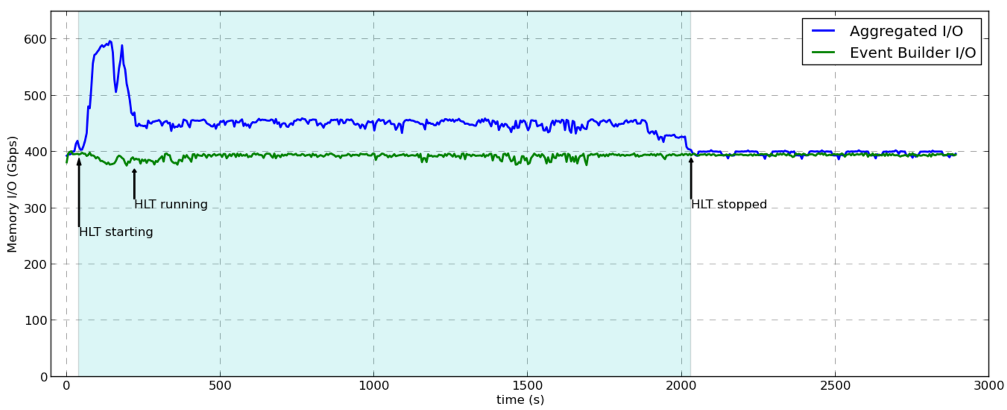


Figure 3.10: The performance of the event-building expressed as memory bandwidth (Event Builder I/O) as a function of time. The Aggregate I/O shows the additional memory bandwidth due to running *parasitic* High-Level-Trigger (HLT) jobs as described in the text.

3.5.3 Residual resources in event-builder machines

As can be expected from a purely zero-copy event-building the CPU-load is rather modest. At about 400 Gbit/s more than 80% of the CPU resources are free. The CPU-needs for book-keeping are constant, with growing server-performance their relative weight will drop.

In the current architecture the memory pressure is more important than the CPU and it is currently the limiting factor for opportunistic usage of the event-builders. We have run, in parallel to the event-building, the LHCb trigger application *Moore*, in off-line mode where data come from a file in parallel to the event-building.

On the test-machine we can launch 18 instances of *Moore* without negatively influencing the event-building application as seen in Fig. 3.10. This corresponds roughly to half of the capacity of the machine, if it were only used for triggering. In fact the limitation does not come from the CPU needs of the event-building, which is rather small (about 15%), but from the total available memory-bandwidth in the server.

The available memory bandwidth will increase in future server architectures[6] while the bandwidth-needs of the event-builder remain constant at 200 Gbit/s per PCIe40 card.

Very conservatively we therefore estimate that at least 80% of the event-building server will be available for opportunistic use by the high-level trigger or a software version of the low-level trigger.

3.6 Event filter farm

The event-filter-farm for the upgraded LHCb experiment, referred to as *farm* in the rest of this section, will be responsible for reducing the event-rate from the 30 MHz of colliding bunches to the accepted output rate to storage.

The farm will be installed on the surface area in a containerized data-centre. This data-centre will be bought, possibly in several stages starting in 2018. It is also assumed that the data acquisition (event-building) system is located in the same place.

We base all numbers on a standard model of a dual-socket Intel Xeon based server, which have been using successfully in LHCb in the past five years. The architecture of the upgraded LHCb event-builder is such that it can connect to any type of compute unit as long as such a unit can be attached to the network.

3.6.1 CPU performance estimate

Based on the dual Intel server processor model, we have tried to estimate the CPU power available in 2020. The results are shown in Figure 3.11.

In Ref. [28] it has been estimated that the growth-rate of server performance at equal cost is about 1.25 per year. This is shown in the lower-most (red) curve. This is significantly lower than what we have seen in the acquisitions we have conducted. The growth-rates we have seen are between 1.48 and 1.74. The difference is due to the high specialization of the event-filter farm for a single high-throughput application, while the number in Ref. [28] is for general purpose data-centres. Taking the mean value between our lowest measured growth-rate and 1.25 gives a growth rate of 1.37 which is what we have assumed here.

[6] In fact it will already go up by almost 50% in the generation following the one on which the present tests have been performed.

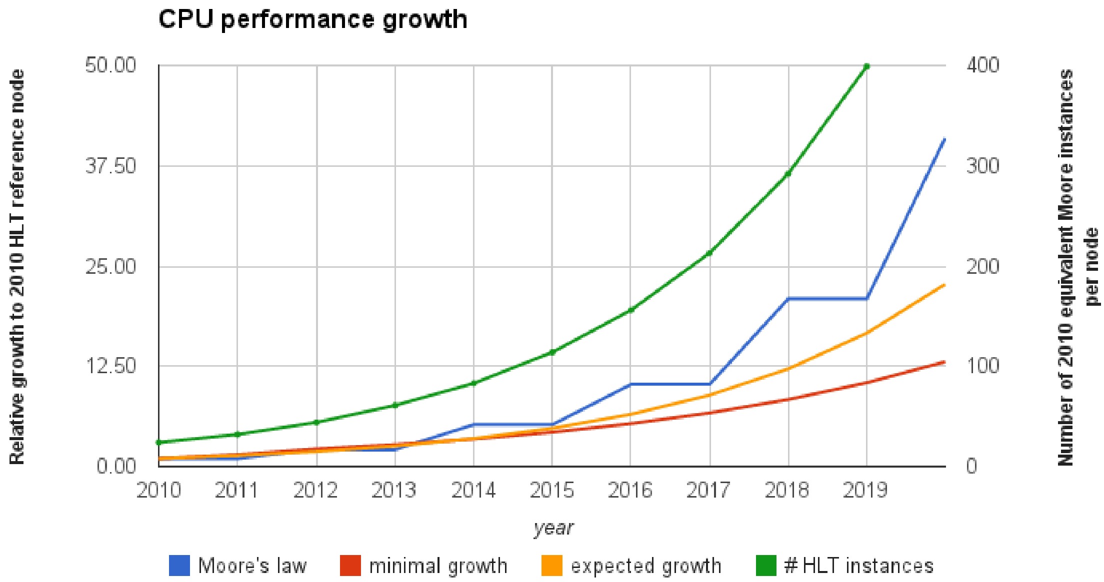


Figure 3.11: Expected CPU performance growth relative to the reference HLT-node of 2010 for various models described in the text (left axis). The curves are normalised to the performance of the 2010 reference node. The right axis and top-most (green) curve indicate the performance in terms of the number of Moore instances on the 2010 reference node.

This is shown in the second to lower-most (yellow) curve. Moore's law is indicated in the upper (blue) curve, where the growth is calculated from the increase in number of transistors per unit area. This is probably too optimistic. For an acquisition in 2019 we arrive at cumulated growth-factor with respect to the reference node[7] of about 16.

One can also express this growth in terms of the number of instances of a Moore application with the performance as measured on the reference node. In these units, shown on the right-hand axis in Fig. 3.11, one can see that in 2019 we expect to be able to run 400 instances of the Moore application on a server. Therefore, the CPU time budget for each Moore application is 13 ms assuming a farm of 1000 servers, and an input rate of 30 MHz.

It should be noted that our estimation is only about 30% above the most pessimistic scenario. No improvement in the software on current architectures nor any improvements from R&D on many-core architectures has been taken into account. Significant efforts will be devoted to exploiting these technologies, see also Sect. 3.9.5.

The above extrapolation assumes that the memory bandwidth grows such that the individual instances of the *Moore* application do not influence each others performance. While techniques such as *forking* help with reducing memory contention, intense R&D

[7] The reference node is a dual-socket Intel X5650 (*Westmere*) machine. Each processor has 6 physical cores and two virtual processing units (*hyper-threads*) and is clocked at 2.67 GHz. The machines have 24 GB of RAM total.

will be devoted to reducing the growth in memory bandwidth needs.

3.6.2 Implementation

The data-centre at Point 8 is dimensioned to be able to house at least 4200 rack-units (Us) of equipment. Two MW of power and cooling will be available.

The rackspace is needed for servers, network equipment and patch-panels. The typical server for the HLT is very compact, probably one half U. However we keep open the option for less dense technologies should they be more cost-effective.

If the base-line HLT is done with 1000 servers each server will require 32 Gbit/s network bandwidth. This fits very well with announced future chipsets, which will integrate 4×10G Ethernet or alternatively 40G Ethernet on the main-board. If it is more cost-effective an InfiniBand or other high-speed card can be easily added to the servers, since these half-U servers provide space for one add-in communication card.

The data-centre has also to house the 500 event-building PCs. We conservatively estimate that these PCs need two Us, even if it is very likely that 1 U will be sufficient.

A typical rack-layout could be composed of the following *sub-farm* unit: four event-builder PCs, one top-of-the rack switch with 4×100G Ethernet uplinks and 10 worker-nodes each connected with 4×10G Ethernet. The event-builder PCs would be connected via an optical direct attach cable to the central event-builder switch(es). Two such sub-farms can be easily fit into one 42U rack, which would make the total data-centre requiring about 70 racks. An example is shown in Fig. 3.12.

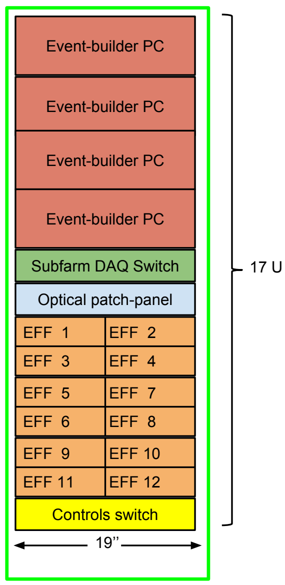


Figure 3.12: A possible rack layout combining event-builder and event-filter farm servers. Two such arrangements would fit easily into a 42U standard rack.

It should be noted that this is only one possible implementation - other layouts are possible. The final decision will be taken when the containerized data-centre is ordered in 2018.

3.7 Experiment control system

The ECS is in charge of the configuration, monitoring and control of all areas of the experiment: the Detector Control System (DCS), the Data Acquisition System and the HLT. It provides and homogeneous and coherent interface between the operators and all experimental equipment, as shown in Fig. 3.13.

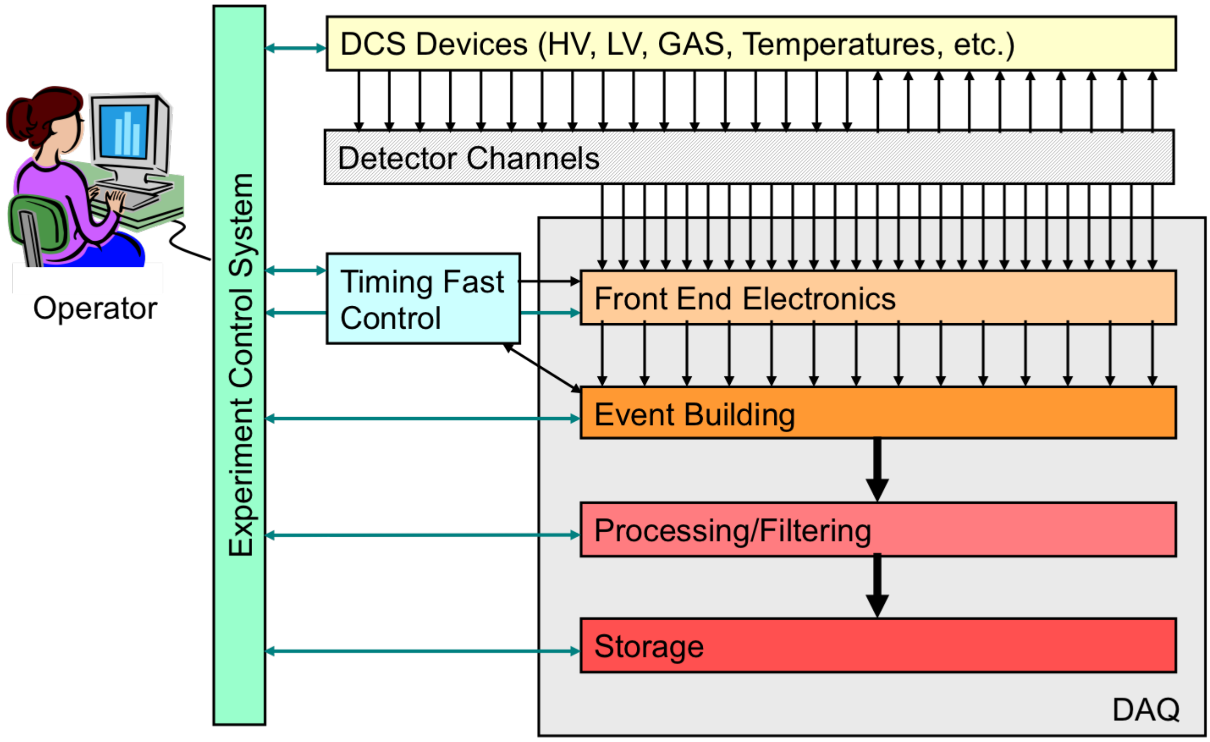


Figure 3.13: Scope of the Experiment Control System.

The ECS for the upgraded detector will be an evolution of the current system, described in the ECS chapter of the original LHCb Online System TDR [10]. It will continue to be developed in the context of the Joint Control Project (JCOP) [14], a common project between the four LHC experiments and a central Controls group at CERN. It defined a common architecture and a framework to be used by the experiments in order to build their detector control systems.

3.7.1 Architecture

JCOP adopted a hierarchical, highly distributed, tree-like, structure to represent the structure of sub-detectors, sub-systems and hardware components. This hierarchy allows a high degree of independence between components, for concurrent use during integration, test or calibration phases. It also allows integrated control, both automated and user-driven, during physics data-taking. LHCb adopted this architecture and extended it to cover all areas of the experiment.

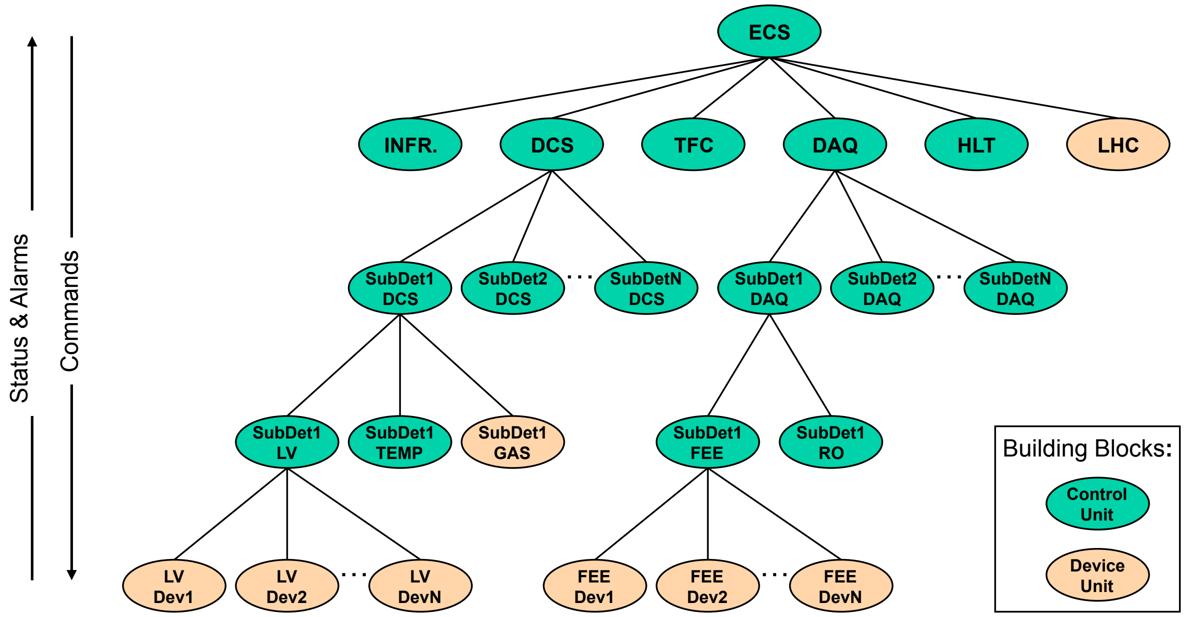


Figure 3.14: Simplified ECS Architecture.

Figure 3.14 shows a simplified version of LHCb's control system architecture. The building blocks of this tree can be of two types: *Device Units*, the tree leaves, which are capable of *driving* the equipment to which they correspond and *Control Units* (CUs) which correspond to logical sub-systems and can monitor and control the sub-tree below them.

3.7.2 Framework

The JCOP Framework provides for the integration of the various components in a coherent and uniform manner. It was implemented based on a Supervisory Control And Data Acquisition system called PVSSII, now WinCC-OA[8].

While WinCC-OA offers most of the needed features to implement a large control system, the Control Units described above are abstract objects and are better implemented using a modelling tool. For this purpose SMI++ [29] was integrated into the framework. SMI++ is a toolkit for designing and implementing distributed control systems, its

[8]Siemens ETM homepage

methodology combines three concepts: object orientation, Finite State Machines (FSM) and rule-based reasoning. The JCOP Framework was also complemented with LHCb specific components, providing for the control and monitoring of LHCb equipment, for example, DAQ electronics boards, DCS power supplies or HLT algorithms.

3.7.3 DAQ & Electronics Control

The new upgraded electronics will be integrated into the control system following the same philosophy. Standard LHCb components will be developed which will allow users to configure, monitor and interact with their electronics. The upgrade electronics specifications document [20] contains requirements and guidelines for electronics developers, so that common software can be implemented.

As described in the TFC section, the ECS interface to the FE electronics will be implemented via SOL40 interface boards, using the GBT system. This bi-directional link allows the writing and reading of configuration and monitoring data. The GBT-SCA chip provides an interface between the GBT and standard protocols such as I2C, SPI or JTAG and can be mounted on the FE modules, as shown in Fig. 3.15.

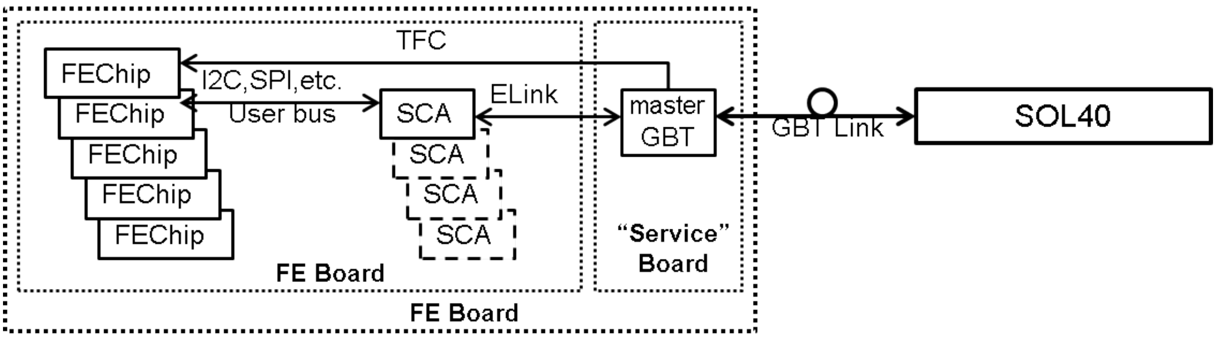


Figure 3.15: FE Electronics ECS Interface

In the baseline implementation, the SOL40 boards are PCIe cards inside a PC. A generic server running inside the PC will be developed centrally and will provide the interface between the FE electronics and WinCC-OA. Similarly to the FE electronics, the software for the configuration and monitoring of BE boards, like readout boards, S-ODIN will be provided centrally in the form of JCOP components providing for the high level description and access of all electronics components.

3.7.4 Guidelines & Templates

Configurable framework components will be distributed to the sub-detector and sub-system teams in order to build their specific control systems. In order to assure the coherence and homogeneity of the system also quite detailed guidelines, specifying naming conventions and colour codes will be prepared and distributed. Whenever possible also *templates* will be centrally provided, *i.e.* the code necessary to implement the guidelines and conventions or the code to implement the finite state machine behaviour specified for the different LHCb domains.

3.7.5 Operations & Automation

As in the current system all standard procedures and, whenever possible, any error recovery procedures will be automated using the Framework FSM tools [30]. The experiments operation, in terms of user interfaces, will be again based on the JCOP Framework and WinCC-OA, providing a Run Control, a DCS Control panel, Alarm screens similar to the current ones. As an example the current Run Control shown in Fig. 3.16.

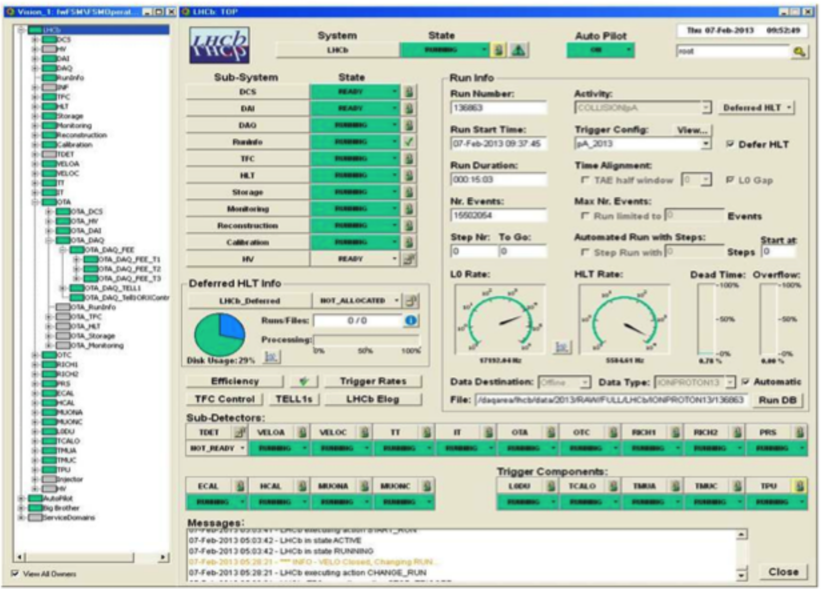


Figure 3.16: LHCb's current run-control.

3.8 Infrastructure

The Online system needs a lot of infrastructure which will be very briefly described in this section.

The readout-system will be located in a containerized data-centre at the surface of Point 8 [31] whose first part will be put in place in 2018. This is a modular system and more capacity can be deployed quickly and will be added as needed.

For cost reasons these containers will contain no or only minimal battery-backed-up power. Sensitive equipment, in particular the storage systems and the servers running the virtualized infrastructure for the ECS will be located in the existing surface server-room, which has a fully redundant, battery-backed-up power distribution.

3.8.1 ECS network and storage

As in the existing LHCb system the ECS will use a dedicated network infrastructure, separated entirely from the DAQ network. The ECS network will be a traditional Ethernet LAN with most devices connecting via Gigabit Ethernet. The core ECS will be deployed on virtual machines running all background SCADA services and of physical machines connected to hardware (GBT, CAN, etc. . .). These infrastructures will be fully redundant and be designed for high availability [32].

Common shared file-systems will be available to all computers, irrespective of the operating system.

A high-performance storage system capable of storing at least 5 GB/s[9] will be available with a capacity to cover at least seven days of LHC running.

3.8.2 Usage of Online facilities outside LHC operations

The event-filter farm and event-builder PCs represent a substantial amount of CPU power. They will be made available for off-line processing during periods outside LHC operations. The exact implementation of the software infrastructure will be determined before Run 3 to ensure maximum compatibility and interoperability with the LHCb grid-sites. The goal is to achieve an overall usage-factor of the facilities throughout the year in excess of 80%.

3.9 Project organisation

This section is devoted to project organisation for the online system and the readout board as well as their costs and schedules.

3.9.1 Readout board project

The institutes participating in the readout board project are listed in Table 3.3 and the division of responsibilities in the Table 3.4. They do not include the contribution of each sub-detector to develop their own firmwares.

Table 3.3: List of institutes participating in the readout board project.

Country	Institute(s)
France	Centre de Physique des Particules de Marseille (CPPM)
	Laboratoire d'Annecy-le-vieux de Physique des Particules (LAPP)
Italy	Sezione INFN di Bologna (Bologna)
Switzerland	CERN

[9] 5 GB/s is more than the nominally needed 2 GB/s at 20 kHz. The extra capacity ensures that in case of outages the backlog of data can be transferred in parallel to normal operation.

Table 3.4: Division of responsibilities in the readout board project.

Tasks	Institute(s)
Conception, design, pre-series, low level interface	CPPM
Production	CPPM, Bologna
Coordination of the firmware developments and generic firmware	LAPP
WinCC-OA supervision	CERN
Commissioning	All institutes

Concerning the schedule, the prototype of the PCIe board is in preparation and should be ready end 2014. Six months will be required to produce a pre-series of about 50 boards and 18 months to produce the final batch of about 500 boards.

The cost of the board depends on the number of optical drivers. It is given for different configurations in Table 3.5 when the board is equipped with an Arria 10 FPGA from Altera and when the optical transmitters are the MiniPod from Avago. The Table also includes the cost of the pre-series corresponding to 10% of the final production.

Table 3.5: Cost of the PCI Express readout board. The numbers in parenthesis defined the number of optical links in input followed by the number of optical links in output.

	Cost [kCHF]
Unit price PCie40 (24/0)	5.8
Unit price PCie40 (48/0)	6.6
Unit price PCie40 (36/36)	7.9
Unit price PCie40 (48/48)	8.8
Pre-series of 50 boards	380

3.9.2 Online project

The institutes participating in the online project, where *online* here means everything except the readout-board, are listed in Table 3.6 and the division of responsibilities in Table 3.7.

Table 3.6: List of institutes participating in the DAQ, ECS and TFC projects

Country	Institute
Italy	Sezione INFN di Bologna (Bologna)
Switzerland	CERN

3.9.3 Schedule for the Online project

R&D and technology tracking on network technologies will go on until 2017 when a decision on the network technology for the event-building will be taken. Tendering will

Table 3.7: Division of responsibilities in the Online system.

Tasks	Institute(s)
Eventbuilding network	CERN
Eventbuilding PCs	CERN, Bologna
Eventbuilding firmware	CERN, Bologna
Event filter farm	CERN
Experiment Control System	CERN
Online Infrastructure	CERN
Timing and Fast Control	CERN
Commissioning	All institutes

be followed by acquisition of the event-building PCs and event-building network in 2018 to be ready for detector commissioning in 2019. The Online infrastructure for the farm (containerized data-centre) will be selected in 2017 and the first part will be bought in 2018. Deployment of the ECS and TFC equipment will be done in early 2018 to be ready well in advance before sub-detector commissioning. The event-filter farm will be bought as late as possible to get the best performance. Only a minimal subset will be bought before 2019 for farm-commissioning. Experience in LHCb shows that event-filter farm nodes can be added smoothly even during data taking.

3.9.4 Cost of the Online project

The cost of the Online system is based on past experience from the many call for tenders for the current system. For the event-builder quotes based on InfiniBand equipment from Mellanox have been used. The individual components are costed in Table 3.8 and would be funded through the Common Projects [33]. The cost for the TFC is given for reference since it belongs to the *general electronics* item of the Common Projects.

Table 3.8: Cost of the Online System

	Cost [kCHF]
Event builder (network and PCs)	3600
Optical Fibres	1700
Controls network	905
Controls system (ECS)	930
Event-filter farm	2800
Infrastructure	775
Timing and Fast Control (TFC)	500

With the foreseen funding, the LHCb upgrade would be equipped with a trigger-less readout and an event filter farm equipped with $\mathcal{O}(1000)$ nodes in 2020.

3.9.5 R&D on many-core computing

The main aim of the R&D on the many-core computing is to optimize the cost / performance ratio for the EFF [34]. It would also help to mitigate the risk related to the number of trigger processes per CPU node which might not scale as the Moore's law in the coming years. The R&D would study the relative performance of the trigger algorithms on different computing platforms like the Intel Xeon/Phi and GPGPUs, and the related issues of code portability.

In the spirit of R&D no fixed responsibilities are attributed however a coordination will be put in place, which will ensure a healthy balance between the various technologies and approaches.

The institutes participating in the many-core R&D are listed in Table 3.9.

Table 3.9: List of institutes participating in the many-core R&D.

Country	Institute(s)
Germany	TU Dortmund
Italy	University and INFN Padova
Netherlands	NIKHEF, University of Groningen
Spain	University of Barcelona (with technical associate La Salle, Universitat Ramon Llull)

Chapter 4

Trigger

The trigger system of the upgraded LHCb detector is developed based on experience gained during Run 1. Between 2009 and 2018, the hardware trigger, *Level-0* (L0), reduces the rate to the 1.1 MHz at which the whole detector can be read out. Events passing L0 are reconstructed by the HLT, a software application which runs on every processor of the EFF. In 2012, 0.35 GByte/s of events are written to permanent storage for further offline analysis.

The key challenges of a full software trigger are the limitations due to CPU budget, defined by the size of the EFF, discussed in Sect. 3.6 and the limited output bandwidth, which is constrained by offline computing resources. In Sect. 4.5.2 it is be shown that the track reconstruction can be performed at close to offline quality with the full input rate. This allows both constraints to be factorised as the full track sample can be reconstructed without intermediate selections. Note that this is a fundamental difference to the HLT used in Run 1 which starts with a partial reconstruction step that requires a rate reduction due to CPU constraints for the reconstruction and thus tightly couples rate and CPU constraints.

The all-software trigger offers unprecedented flexibility in designing selections, and in particular allows efficient triggering on low-momentum signatures which would normally be out of reach at a hadron collider. For this reason, it is important to first consider the rates at which various types of events are produced at the energies and luminosities planned for the upgraded detector. We therefore begin by describing the anatomy of events produced under the planned nominal upgrade conditions in Sect. 4.1.

The trigger system runs different set of algorithms, as sketched in Fig. 4.1: a software integrated LLT, full event reconstruction, and event selection which are detailed in Sect. 4.4, 4.5 and 4.6 respectively.

The LLT is an evolution of the current L0 trigger and uses limited information from the calorimeters and the muon stations. It is shown in this document that a LLT will not be necessary, but it will be kept as backup solution. This backup could reduce the input rate to the software trigger by a factor of two with limited cost in physics sensitivity. The advantage of maintaining this LLT is that it can be rapidly deployed in the face of changing beam conditions, should the LHC choose a filling scheme different to that which

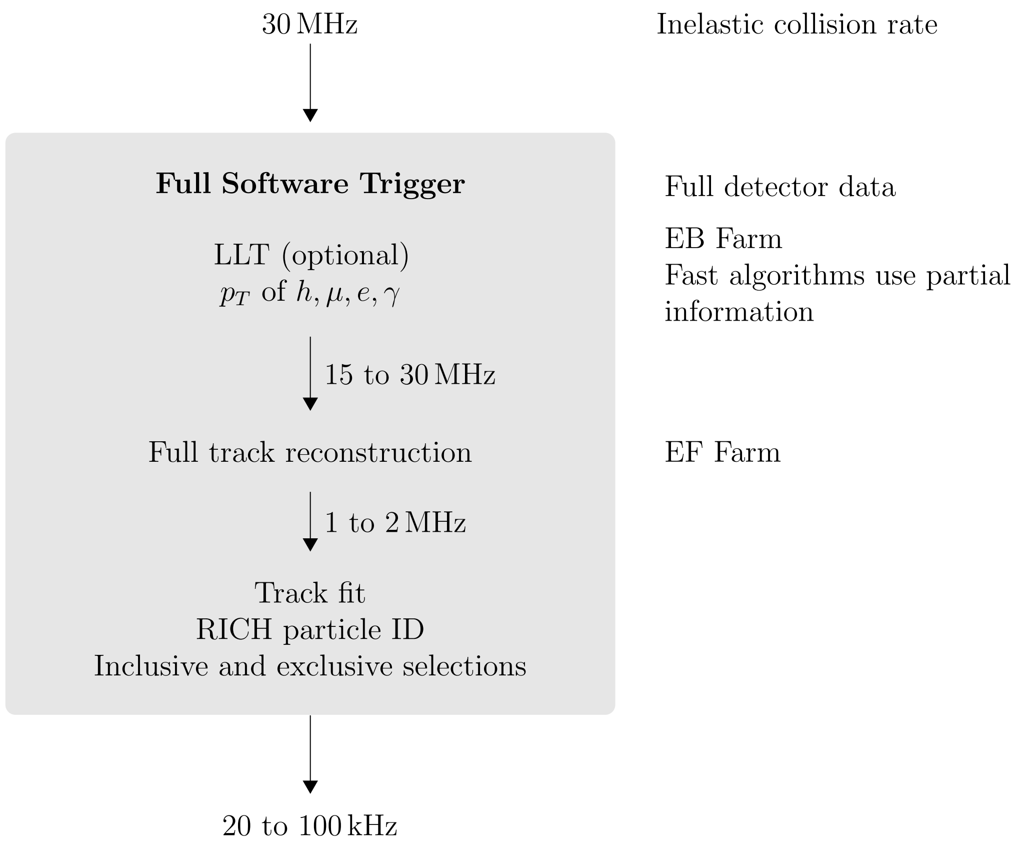


Figure 4.1: Schematic view of the full software trigger.

we presently expect.

The full event reconstruction reconstructs tracks with a precision very close to offline. Based upon this information, a trigger selection is performed that reduces the data rate by a moderate factor, at which point the kalman filter based track fit and the RICH based particle identification can be performed. The rate reduction is such that sufficient time is provided for the RICH ring finding algorithms as discussed in Sect. 4.5.3 . This particle ID information is then used to reduce the output rate to a level that can be processed by the offline computing.

One possible implementation of an inclusive beauty trigger is presented in Sect. 4.6.2 Its performance is discussed in terms of efficiency on selected signal channels, background event rates, and CPU time needed to perform the trigger selection. A proof of principle for efficient and low rate selections of exclusive beauty decays is given, in addition to a discussion of trigger selections for hadronic beauty decays, where the entire trigger chain can be performed without introducing selection criteria that bias the lifetime distribution.

Section 4.7 describes the robustness of the system. The behaviour of the track reconstruction at luminosities higher and lower than the nominal one is discussed, as well as a strategy to cope with imperfections in the incoming data. Finally, Sect. 4.8 concludes with a discussion of the costs and the people/institutes working on the proposed software HLT.

Throughout this document we use several Monte-Carlo simulated samples produced under different upgrade scenarios, in addition to samples produced under Run 1 conditions. The relevant conditions are: The average number of both elastic and inelastic proton-proton collisions per event, referred to as ν; the instantaneous luminosity, $\mathcal{L}$, and the beam energy, $\sqrt{s}$. During Run 1, spillover was not included in the simulation as the detector readout of LHCb is robust to spillover at 50 ns bunch spacing. In the upgrade the LHC will operate on a 25 ns bunch spacing, and so we simulate this in our upgrade samples. Table 4.1 describes the conditions and naming conventions of these samples.

Table 4.1: Conditions corresponding to the data-taking scenarios described in this document.

Name	$\mathcal{L}$ [$\times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$]	ν	$\sqrt{s}$ [TeV]	Spillover
Run 1	0.4	2	7-8	N
Upgrade, nominal luminosity	2	7.6	14	Y
Upgrade, reduced luminosity	1	3.8	14	Y
Upgrade, increased luminosity	3	11.4	14	Y

4.1 Event anatomy

In this section, we outline our estimate of the production rates of heavy flavour particles in the context of the upgraded LHCb detector using Monte-Carlo events [35]. The purpose of this study is to understand the output data rates of an ideal trigger, which selects all events containing interesting physics signatures reconstructible in the LHCb acceptance [6].

The conditions are described in Table 4.1. Firstly, a sample of minimum-bias events of 200 k events consistent with Run 1 conditions at $\sqrt{s} = 7$ TeV. Secondly, a large sample of *generator level*[1] minimum bias of 10 M events consistent with the nominal upgrade conditions. Thirdly, a fully simulated sample of 100 k events in which the upgraded VELO pixel hardware has been simulated. We look for 20 types of parent particles and their charge conjugates in each event, covering most of the known ground-state beauty and charm hadrons, as well as long-lived light-quark hadrons. We make no immediate assumptions about their final-state topology or their suitability for physics analysis purposes, and we do not require that they come directly from a primary vertex. This list is not, of course, complete, and as a result some of the decay rates presented will be slight underestimates, but it is representative of almost all possible topologies of interesting events[2].

We also determine the number of candidates that have a reconstructible vertex within the VELO. A track is considered to be within the VELO acceptance if it has positive

[1] *generator level only* means that the sample has not been propagated through the simulated LHCb detector, digitised and reconstructed

[2] We ignore for now the production of strongly decaying short-lived resonances which may become an active area of research in the future, and the decays of exotic particles which are also of considerable interest.

momentum in the forward direction and traverses at least three VELO stations. A candidate is considered to have decayed in the VELO acceptance if at least two child tracks traverse at least three VELO stations each.

4.1.1 Generator-level yields

Table 4.2 presents the per-event yields for b, c, and light, long-lived hadrons, for both Run 1 and nominal luminosity upgrade datasets, as well as the percentage of these candidates which decay within the VELO and are fully contained within the LHCb acceptance. It highlights the relative complexity of proton-proton collisions pre- and post-upgrade. After the upgrade we can expect a factor of five increase in the per-event rate of charm hadrons, a factor six increase in the per-event rate for beauty hadrons and a factor four increase in light, long-lived particles all of which leave a secondary vertex in the VELO and are fully contained within the LHCb acceptance.

Table 4.2: Per event yields and efficiencies for generator-level Monte-Carlo. ϵ(VELO) is the efficiency for candidates having at least two tracks traversing at least three modules in the VELO. ϵ(LHCb) is the efficiency for candidates having all child tracks contained in the LHCb acceptance.

	Run 1, Original VELO geometry		
Category	Yield in 4π	ϵ(VELO)	ϵ(VELO) $\times$ ϵ(LHCb)
b-hadrons	0.0258 ± 0.0004	$30.5 \pm 0.6\%$	$11.1 \pm 0.4\%$
c-hadrons	0.297 ± 0.001	$21.9 \pm 0.2\%$	$14.2 \pm 0.1\%$
light, long-lived hadrons	8.04 ± 0.01	$6.67 \pm 0.02\%$	$6.35 \pm 0.02\%$
	Upgrade, nominal luminosity, VELO pixel geometry		
Category	Yield in 4π	ϵ(VELO)	ϵ(VELO) $\times$ ϵ(LHCb)
b-hadrons	0.1572 ± 0.0004	$34.9 \pm 0.1\%$	$11.9 \pm 0.1\%$
c-hadrons	1.422 ± 0.001	$24.73 \pm 0.04\%$	$15.12 \pm 0.03\%$
light, long-lived hadrons	33.291 ± 0.006	$7.022 \pm 0.004\%$	$6.257 \pm 0.004\%$

The increase in per-event yield for light, long-lived particles has consequences for the design of the trigger. Any trigger which looks for displaced vertices in the VELO in the Run 1 dataset would have a yield of $8.04 \times 6.67\% = 0.53$, or approximately one candidate every two events. In the upgrade this yield increases to 2.1. *Every* event contains two such light hadron decays, saturating any trigger of this type unless further information is available to make a decision. In addition, the event rate in the upgrade will be double that of Run 1, due to the 25 ns bunch spacing. In the following subsection we will indicate the estimated input bandwidth to the trigger for each of these signal categories.

4.1.2 Reconstructed yields

Table 4.3 presents the per-event yields and efficiencies for truth-matched candidates[3] that have been fully simulated and partially reconstructed within the LHCb simulation framework. Current analysis experience shows that candidates with a parent p_T above 2 GeV/c and a lifetime above 0.2 ps have the potential to be usable in offline analyses. For this reason, we show the efficiency of these two selection criteria when applied to the vertex formed by the partially reconstructed final state. Any partially reconstructible candidates passing these selection criteria are considered to be potentially interesting for further offline analysis and hence define the sample which an ideal inclusive trigger would select.

Table 4.3: Per-event yields determined from minimum-bias events after partial offline reconstruction. The first row indicates the number of candidates which had at least two tracks from which a vertex could be produced. The last rows show the output rate of a trigger selecting such events with perfect efficiency, assuming an input rate of 15 MHz from the LHC and an event size of 50 kB, as during Run 1 in the first instance, and 30 MHz at 100 kB for the upgrade in the second.

offline-reconstructed, Run 1			
	b-hadrons	c-hadrons	light, long-lived hadrons
Reconstructed yield	$(4.0 \pm 0.1) \times 10^{-3}$	0.0196 ± 0.0003	0.0792 ± 0.0006
$\epsilon(p_T > 2\,\mathrm{GeV}/c)$	$83 \pm 1\%$	$47.2 \pm 0.8\%$	$2.0 \pm 0.1\%$
$\epsilon(\tau > 0.2\ \mathrm{ps})$	$89 \pm 1\%$	$64.2 \pm 0.7\%$	$99.53 \pm 0.05\%$
$\epsilon(p_T) \times \epsilon(\tau) \times \epsilon(\mathrm{LHCb})$	$29 \pm 1\%$	$22.3 \pm 0.6\%$	$1.9 \pm 0.1\%$
Output rate (kHz)	17.3	66.9	22.8
Output rate ($\mathrm{GB\,s^{-1}}$)	0.9	3.3	1.1
offline-reconstructed, Upgrade, nominal luminosity			
	b-hadrons	c-hadrons	light, long-lived hadrons
Reconstructed yield	0.0317 ± 0.0006	0.118 ± 0.001	0.406 ± 0.002
$\epsilon(p_T > 2\,\mathrm{GeV}/c)$	$85.6 \pm 0.6\%$	$51.8 \pm 0.5\%$	$2.34 \pm 0.08\%$
$\epsilon(\tau > 0.2\ \mathrm{ps})$	$88.1 \pm 0.6\%$	$63.1 \pm 0.5\%$	$99.46 \pm 0.03\%$
$\epsilon(p_T) \times \epsilon(\tau) \times \epsilon(\mathrm{LHCb})$	$27.9 \pm 0.3\%$	$22.6 \pm 0.3\%$	$2.17 \pm 0.07\%$
Output rate (kHz)	270	800	264
Output rate ($\mathrm{GB\,s^{-1}}$)	27	80	26

The signal rates facing the upgraded detector are very large. We could select 27 GByte/s of $b\bar{b}$ events alone using an inclusive trigger, and three times as many $c\bar{c}$ events. Within these constraints it is clear that any inclusive trigger strategy must have a poor efficiency on at least some signal modes, because it will need to be downscaled to cope with large signal rates regardless of selection efficiency. Fig. 4.2 shows the evolution of the rate as a

[3] Candidates are partially reconstructed by forming a vertex from two charged tracks which are truth-matched to genuine pions, kaons, protons, muons, or electrons. The vertex is then truth-matched to a composite particle, and no additional selection criteria are applied.

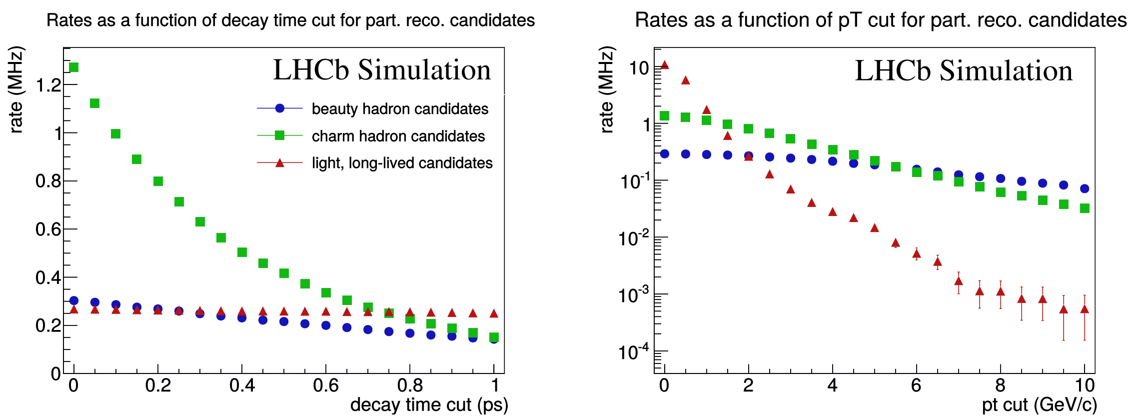


Figure 4.2: The rate of secondary vertices associated to partially reconstructible decays: (left) as a function of decay time for candidates with $p_T > 2$ GeV/c and (right) as a function of transverse momentum selection criteria for candidates with $\tau > 0.2$ ps.

function of the lifetime and p_T thresholds. Once the signal purities are high this approach amounts to downscaling signal.

The conclusion to be drawn from this section is that the allowable output rate of signal is the greatest challenge facing the upgraded trigger system. The problem is no longer the classical "find one Higgs event among 10 billion background events", but the more complex "discriminate in a minimally biasing way between various topologically similar signals". While this scenario is an inherent design feature of B factories where large production cross-sections and low event multiplicities mean that every event is of interest, it represents a new paradigm for hadron collider experiments. In certain cases there will be signal modes in the upgrade where even a 100% pure trigger must be downscaled to fit into the output bandwidth allocated. In these circumstances, the trigger algorithms should be as similar as possible to the offline event selections, since the trigger will frequently *be* the offline event selection. If it is possible to cut harder offline then it makes more sense to implement this cut in the trigger and reduce the prescale. This is the logic behind the trigger design described in the remainder of this document.

4.2 Trigger sequence

The full software trigger is a sequence of algorithms which are run sequentially on the event-builder farm as well as on the EFF.

The available processing time is different between the two farms, not only because of the different numbers of nodes, but also because in the event-builder nodes a part of the CPU and memory-bandwidth resources are required for the event-building and hence not available for trigger algorithms. It is at the level of approximately 2 ms when the event-builder consists of 500 servers and at the level of 13 ms for the EFF containing 1000 nodes.

The organization of the sequence between the two farms depends upon the running conditions and the available CPU power. In the early stage of the Run 3 data, when the luminosity is low and the full EFF is not yet available, we can run the LLT algorithms and partial reconstruction in the event-builder farm. As CPU power becomes available we can then move to full tracking in the EFF.

4.3 Global event cuts

All trigger systems are designed to maximally exploit the available computing resources. Whenever spare computing power is available, it is used to bring the event reconstruction and selection closer to what would be done in the ideal offline case. Since more complex events take longer to process, it is necessary to ask whether the physics content of the most complicated and expensive events is commensurate with the physics interest. In the case of LHCb, events with the largest multiplicities typically have the worst signal purities. Removing the most complicated events using *Global Event Cuts* (GECs) provides an overall increase in performance by allowing reconstruction criteria to be brought closer to that of the offline algorithms for the simpler events which remain.

There are different ways to measure the event multiplicity: one can count the number of PVs, the number of tracks reconstructed, or simply the hit multiplicity of a subdetector. All these measures are well correlated. The final choice will depend on the performance of the relevant subdetector as installed in 2018. For the studies presented here, the selection is made based on the sum of the multiplicities of the ECAL and HCAL, GEC = $N_{\mathrm{ECAL}} + N_{\mathrm{HCAL}}$. This variable is well correlated with the other possible measures of the event multiplicity, as shown in Fig. 4.3.

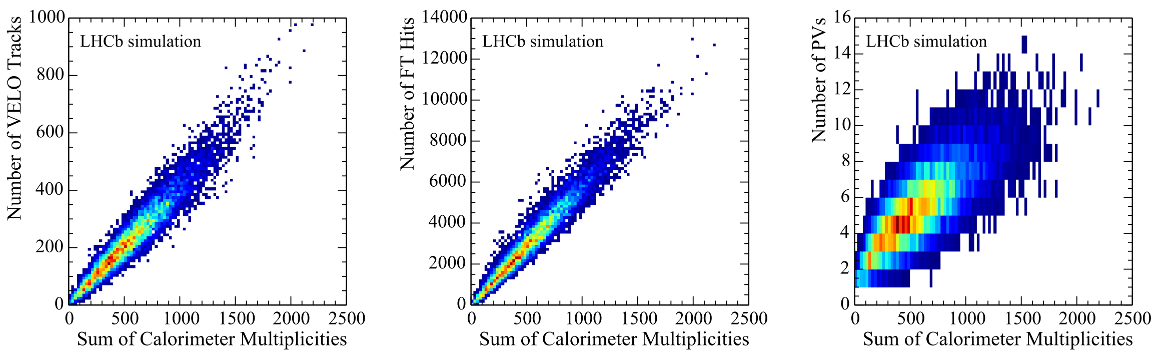
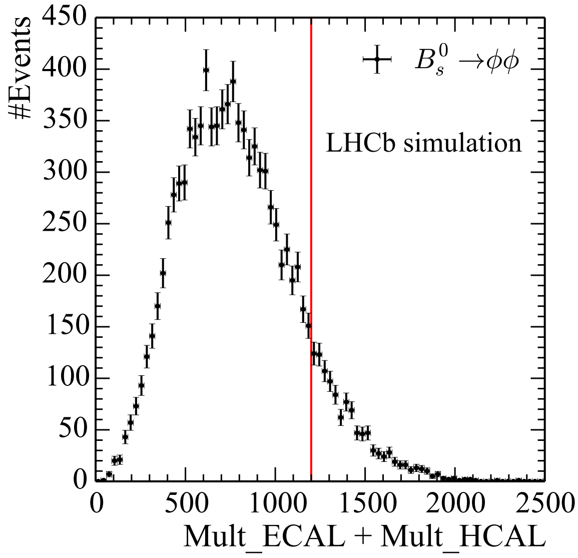
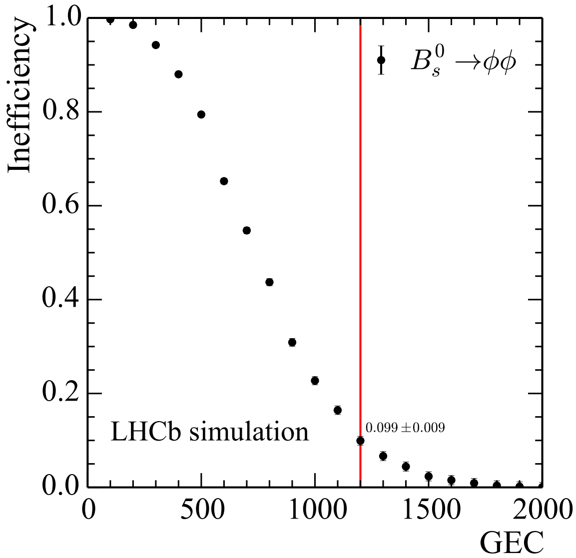


Figure 4.3: Correlation of the sum of the calorimeter multiplicities (GEC) with other global event variables: (left) Number of Velo tracks; (middle) number of FT hits and (right) number of reconstructed primary vertices.

The efficiencies of these GECs need to be evaluated on a signal sample. We use $B_s^0 \to \phi\phi$ events simulated with nominal upgrade conditions. The distribution of the calorimeter multiplicity is shown in Fig. 4.4a, where the tail of events towards higher multiplicities can be seen. The integrated inefficiency for a number of GEC requirements



(a) Event distribution for $B_s^0 \to \phi\phi$.



(b) GEC inefficiency on $B_s^0 \to \phi\phi$.

Figure 4.4: Nominal upgrade conditions: (a) Distribution of calorimeter multiplicities in signal events. (b) The inefficiency introduced by GECs. The red vertical lines represent the nominal GEC of 1200.

are presented in Fig. 4.4b. While a thorough optimisation of the GEC will be performed prior to data taking, for now we choose a GEC of 1200 which removes the tail of events with highest multiplicities while maintaining a 90% signal efficiency as shown in Fig. 4.4a and Fig. 4.4b.

The choice of GEC applied in these studies can be compared with the optimal working point in Run 1 in which the hadronic triggers selected only events with SPD multiplicities below 600, which translates into an inefficiency of approximately 15% at a luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. In the remainder of the document the algorithm timing for both reconstruction and selection algorithms will be measured as a function of applied GECs.

4.4 Low Level Trigger algorithms

The goals of the algorithms implementing the selection for the LLT are to identify electron, hadron and muon candidates in the electromagnetic (ECAL) and hadronic (HCAL) calorimeters and in the muon detector. The algorithms selects the candidates of each type which have the highest transverse energy (E_{T}, for the electron and hadron candidates) and the highest transverse momentum (p_{T}, for the muon candidates). The E_{T} or p_{T} of these candidates are compared to thresholds to decide if the event is transferred to the next level of the trigger sequence. These algorithms are executed in the event building farm, as explained in Sec. 3.5.3.

4.4.1 Calorimeter

Electron and hadron candidates are defined as clusters of 2×2 cells in the ECAL and the HCAL, respectively. Their associated E_{T} is the sum of the energies measured in each cell of the cluster. With the upgraded LHCb detector, no distinction is possible between an electron cluster and a photon cluster when using only the calorimeter information. This is why in the context of the LLT, *electron* candidates is a term that refers both to electrons and photons.

In addition to the E_{T} of the most energetic hadron and electron candidates, the calorimeter LLT algorithms compute the total energy deposited over the entire ECAL and HCAL and the ECAL and HCAL multiplicities. The latter are defined as the number of cells with an energy deposit larger than a given threshold. These quantities may be used for the global event cuts described in Sec. 4.3.

The first steps of the computations needed to obtain the electron and hadron candidates in the LLT are realised in custom electronics (Front-End boards). This hardware architecture is described in detail in Ref. [11]. In summary, the hardware-level processing consists in a rough calibration of the energy deposited in the calorimeter cells and in the computation of the E_{T} of the 2×2 clusters in each Front-End board. These clusters are written to the raw data in order to be further processed in the event building farm by the algorithm implementing the final calorimeter LLT selection or to be possibly used as electron or photon seeds in the first stage of the HLT sequence.

This algorithm extracts and decodes the raw data, selects the highest E_{T} electron and hadron candidates amongst the ones received from the Front-End boards, and computes the total multiplicity and energy summing over all clusters recorded in the event. The processing time is equal to $10\,\mu\mathrm{s}$ of CPU time per event [9] and has been evaluated from fully simulated signal events, generated in conditions corresponding to an instantaneous luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The algorithm is designed so as its processing time is independent of the instantaneous luminosity of the LHC. The cluster pre-processing in the Front-End boards is necessary to reduce the processing time. Without it, this time would be equal to 3 ms and would not fit inside the timing budget.

4.4.2 Muon

The algorithm implementing the search for the muon candidates for the LLT starts by retrieving the digitised information for the four muon stations M2–M5 from the muon detector raw data. Each muon station is divided into several sectors containing logical channels, either pads or strips. The logical channels are used to build the maps of the logical pads which have a hit inside them. The Cartesian coordinates of the logical pads are then calculated using a realistic detector geometry.

A muon track is defined as four aligned hits in the muon stations. The aligned hits are searched for iteratively, starting from the muon station M5, and containing down to M2. At each iteration, an extrapolated point in station M_i is obtained as the intersection of the station's plane with a straight line linking a hit found in station M_{i+1} to the LHCb *pp* interaction point.

For each muon candidate, the transverse momentum is estimated from the coordinates of the hits in M2 and M3, and written in the raw event to be possibly used in the HLT. The p_{T} calculation is done in the thin lens approximation of the dipole magnetic field, without further approximation on small angles.

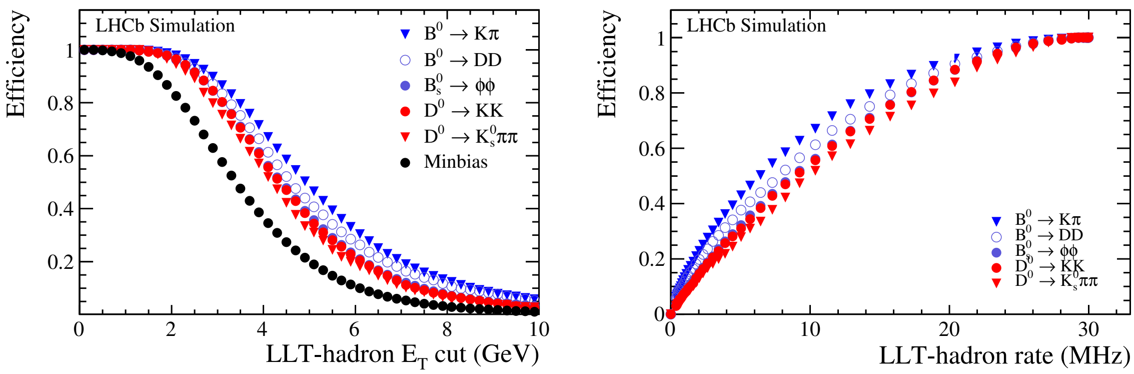
The processing time of this algorithm is on average 0.7 ms of CPU time per event [9]. It has been estimated in a similar way as the calorimeter algorithm processing time, from simulated events corresponding to a luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

4.4.3 Performances

The performances of the algorithms described above, in selecting, at the LLT stage, decay channels representative of the LHCb physics program of the upgrade [1] are reported here. The LLT efficiency for these channels and the minimum bias retention rates are estimated from full Monte-Carlo simulation generated in the upgrade conditions, without applying any GEC.

The performances of the calorimeter algorithms are computed for the decay modes $B^0 \to K^+\pi^-$, $B^0 \to D^+(K\pi\pi)D^-(K\pi\pi)$, $B^0_s \to \phi(KK)\phi(KK)$, $D^0 \to K^0_{\mathrm{S}}\pi^+\pi^-$ and $D^0 \to K^+K^-$, taking only the *hadron* candidates into account for the event selection, and similarly for the measurement of the minimum bias retention rate.

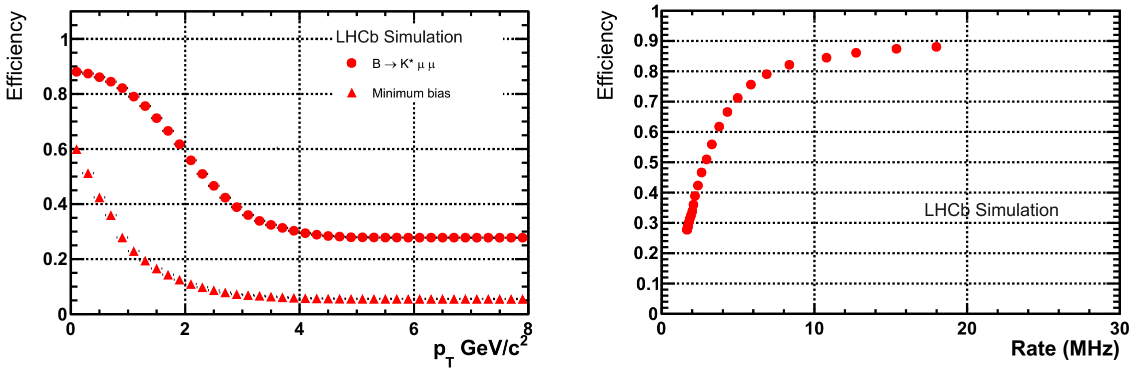
Figure 4.5a shows the efficiency that an event containing the signal decay is selected by the calorimeter algorithm, as a function of the value of the threshold placed on the E_{T} of the hadron candidates. Figure 4.5b shows the same quantity as a function of the minimum bias retention rate.



(a) (b)

Figure 4.5: LLT efficiencies as a function (a) of the hadron E_{T} threshold and (b) of the minimum bias retention rate, considering only the selection based on hadron candidates.

The performances of the muon algorithm are evaluated using the $B^0 \to K^*\mu\mu$ decay mode. The efficiency of the LLT muon selection is defined as the fraction of events for which at least one of the signal muon has a p_{T} above a given threshold. It is presented as



(a) (b)

Figure 4.6: LLT efficiencies as a function of (a) the muon p_{T} cut and (b) the minimum bias retention rate, considering only the selection based on muon candidates.

a function of this threshold on Fig. 4.6a together with the retention fraction for minimum bias data defined as the fraction of minimum bias event giving a candidate above the threshold. Figure 4.6b shows the efficiency represented as a function of the minimum bias retention rate achieved considering the selection obtained with the muon algorithm alone.

These studies show that a retention factor of 2 can be achieved with the LLT, having efficiencies on hadronic signal B and D decays between 65 and 80%, and for channels with muons of about 85%.

4.5 Track reconstruction and particle identification

Several different track reconstruction algorithms exist in LHCb. Some consider only one tracking detector, while others combine information from several sub-detectors. A full discussion of the upgraded tracking system and all available reconstruction algorithms is given in Ref. [5]. The track reconstruction sequence optimised for the trigger is discussed in detail in Ref. [36]. The reconstruction of tracks in the VELO is known as *VELO tracking*, extending VELO tracks with information from the UT is performed by the *VELO-UT algorithm*, and the *Forward tracking* is responsible for adding hits in the SciFi to either VELO or VELO-UT tracks.

There is no magnetic field in the region of the VELO. As a result VELO only tracks have no momentum information. Once a track has been extended to the UT the momentum can be measured with a resolution of 15%. Tracks with measurements both in the UT and SciFi have a momentum resolution of $\approx 0.5\%$.

In the present offline reconstruction, every algorithm is executed and the results are combined. While there is a large overlap between the tracks found, the combination of all algorithms outperforms any single track reconstruction sequence. The trigger system shares the track reconstruction algorithms with the offline, but the present constraints of the trigger system mandate a dedicated sequencing and configuration of these same algorithms. The priority is to reconstruct the most valuable tracks first, with more specialised track reconstruction algorithms only being used later in the decision making process. Figure 4.7 shows a diagram of the track reconstruction sequence used in the trigger, as well as the main offline reconstruction sequence. Track reconstruction in the trigger begins with

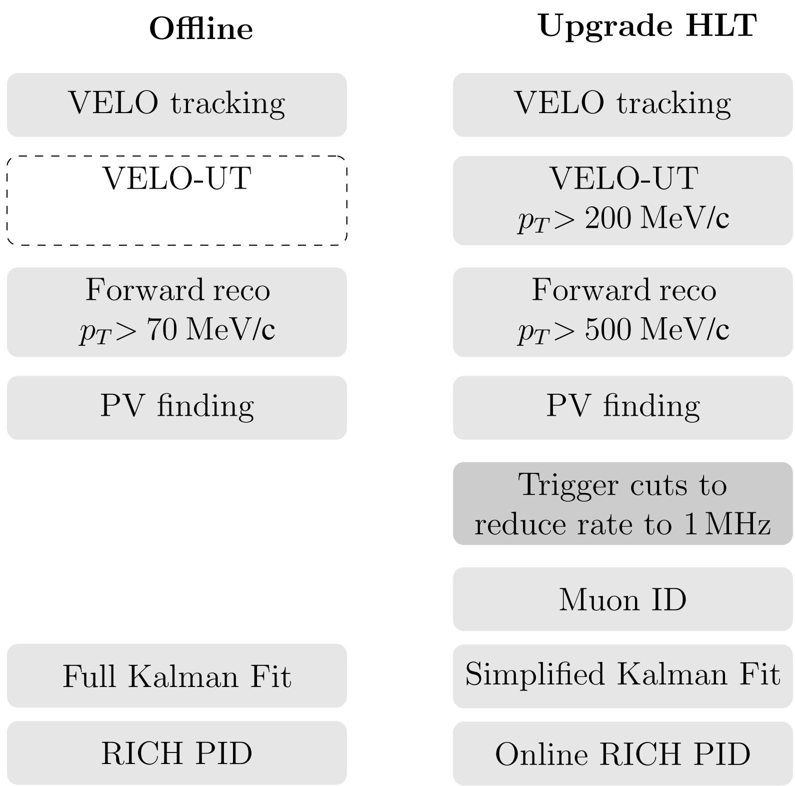


Figure 4.7: Track reconstruction sequences used (left) in the offline and (right) in the trigger reconstruction. The offline reconstruction considers all VELO tracks for extension in the SciFi, whereas in the trigger information from the UT sub-detector is used to determine the charge and remove low p_T tracks before the Forward tracking. The use of the UT significantly reduces the execution time of the Forward tracking.

execution of the full VELO tracking. Information from the UT sub-detector is then used to extend every VELO track which is consistent with a transverse momentum of at least $0.2\,\mathrm{GeV}/c$. For the subset of tracks which were successfully extended, the charge and momentum is estimated. These tracks are then extended further by searching for hits consistent with $p_T > 0.5\,\mathrm{GeV}/c$ in the SciFi sub-detector. The size of the search regions used to extend tracks in the SciFi are reduced by taking into account the charge and momentum measured in the UT. The execution time is further improved by rejecting tracks with $p_T < 0.4\,\mathrm{GeV}/c$.

The main offline track reconstruction sequence for long tracks uses the same VELO tracking as the trigger. However, instead of first adding information from the UT sub-detector, all VELO tracks are extended by adding hits from the SciFi sub-detector. This sequence of algorithms provides most of the tracks used in LHCb physics analyses. Throughout this document this configuration is referred to as the *offline reconstruction*, and it is the configuration against which the trigger tracking performance is compared.

4.5.1 Track reconstruction efficiencies

Tracking efficiencies are measured on a sample of simulated $B_s^0 \to \phi\phi$ decays with $\nu = 7.6$. All efficiencies discussed in this section are absolute efficiencies measured relative to the standard LHCb definition of reconstructible tracks, defined in Ref. [5]. Detector acceptance effects are not included in the overall reconstruction efficiency, since they are already taken into account in the definition of *reconstructible*, while sub-detector hit inefficiencies are accounted for.

Table 4.4: The reconstruction efficiency in per cent achieved by the HLT tracking sequence for different categories of tracks. The efficiency is measured with respect to particles which are reconstructible as long tracks. The first two columns give the efficiency without and with GEC, while the third shows the reconstruction efficiency achieved relative to the offline track reconstruction.

	no GEC	GEC< 1200	relative
Ghost rate	10.9%	5.9%	-
long	42.7%	42.9%	50.4%
long, from B	72.5%	72.8%	80.3%
long, $p_T > 0.5\,\mathrm{GeV}/c$	86.9%	87.4%	97.2%
long, from B, $p_T > 0.5\,\mathrm{GeV}/c$	92.3%	92.5%	98.7%

Table 4.4 summarizes the track finding efficiency for the HLT sequence. The reduced efficiencies for the first two categories are due to tracks with $p_T < 0.5\,\mathrm{GeV}/c$ being included in the denominator of the efficiency. For tracks that originate from beauty decays, leaving hits in all tracking detectors, and satisfying a p_T requirement of $500\,\mathrm{MeV}/c$, the efficiency in the entire tracking sequence is 92.3%, without applying any GEC. Requiring GEC< 1200 increases the efficiency only slightly, to 92.5%. This shows the excellent stability of the track finding sequence at high detector occupancies.

The algorithm used to perform the VELO track reconstruction is exactly the same as used offline. In the offline case, all VELO tracks are processed by the Forward tracking without requiring a UT hit. The final column in Table 4.4 gives the efficiency of the track reconstruction in the trigger relative to the efficiency of the offline track reconstruction. The reconstruction in the trigger achieves efficiencies close to those of the offline reconstruction by design, as it re-uses the same algorithms. The relative track finding efficiency of the HLT tracking sequence compared to the offline sequence is 98.7% for tracks with $p_T > 500\,\mathrm{MeV}/c$.

Requiring a track to be in the acceptance of the UT sub-detector reduces the efficiency. This loss in efficiency is, however, expected to be largely recoverable by reconstructing tracks outside the UT acceptance but inside the acceptance of the SciFi as a special class. These tracks are directly passed from the VELO to the Forward track reconstruction, without any UT requirements.

The ghost rate is additionally reduced by a factor of four in the trigger sequence, compared to the offline reconstruction. The requirement of additional hits in the UT sub-detector is an efficient method to suppress ghost tracks. It has to be noted, however, that the offline ghost rate as seen by physics analyses is reduced by requirements on the quality of the Kalman Filter based track fit, which are not applied here.

The uniformity of the HLT reconstruction efficiency as a function of η, ϕ, p_T, p, number of PVs, and the distance of closest approach to the beam line are shown in Fig. 4.8. The ratios of offline and trigger reconstruction efficiencies are shown as well. The dependence of the efficiency on these variables is not significantly affected by the VELO-UT reconstruction algorithm.

In summary, the presented track finding sequence for the HLT reconstructs 98.7% of tracks with a $p_T > 500\,\mathrm{MeV}/c$ relative to the offline reconstruction and with a factor four reduction in ghost rate.

4.5.2 CPU cost of track reconstruction

This section discusses the CPU time needed to perform the track reconstruction. All algorithm timings are measured on the Run 1 EFF using a set of four reserved farm nodes (HLTe0901-4), which are the same nodes used in the estimation of the available farm budget given in Sect. 3.6. All timing measurements are performed by running a single instance of the trigger algorithms and by averaging the time of 10 runs over 10000 events.

The sum of all reconstruction algorithms is shown for a range of GEC values in Fig. 4.9a. It can be seen that the reconstruction time does not critically depend on the existence of GECs to remove high multiplicity tails. However, removing the busiest events, which anyway have worse signal purity offline, leads to a speedup of 20% even using the presently unoptimised choice of GEC$<$ 1200.

The dependence of the forward track reconstruction time on the internal requirement on the minimal track p_T is shown in Fig. 4.9b. In case additional ressources would be available, a loosening of the $500\,\mathrm{MeV}/c$ requirement to $200\,\mathrm{MeV}/c$ would cost an additional 1.2 ms and would allow to significantly increase the trigger efficiency, *e.g.* for charm or strage meson decays. Similarly, if ressources need to be saved, the p_T requirement can be tightened and about 1 ms can be gained with moderate looses for beauty triggers.

The CPU timing for each algorithms is given in Table 4.5, both for the scenario without GEC and for the default GEC requirement of 1200. The total time is evaluated to be 5.4 ms with or 6.6 ms without the use of GECs. Compared to the total timing budget for the upgrade farm, which is estimated to be 13 ms (see Sect. 3.6). Both cases fit comfortably within the budget. The default scenario, running the full software trigger with a GEC requirement of 1200, consumes less than 40% of the available CPU resources and

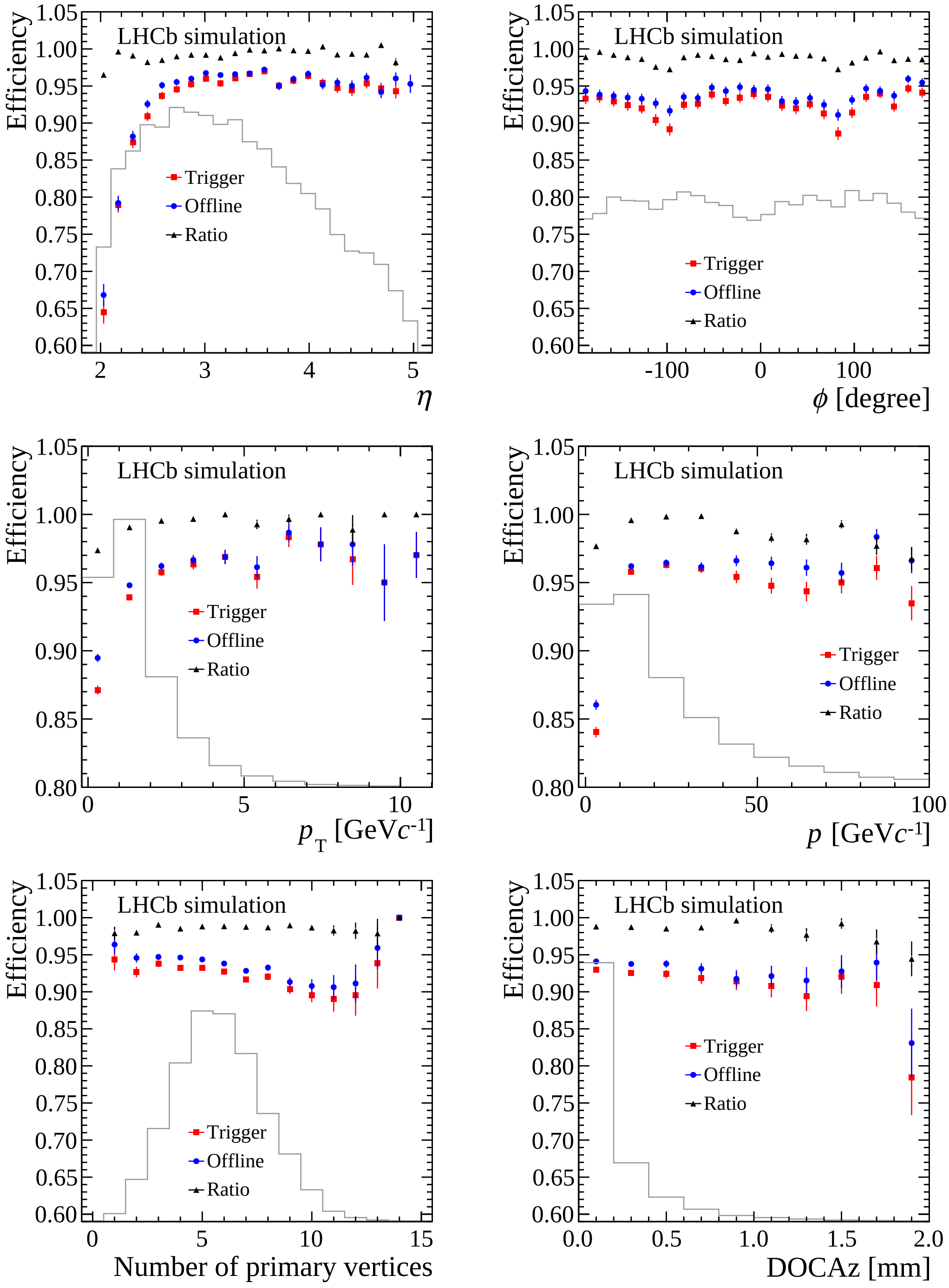
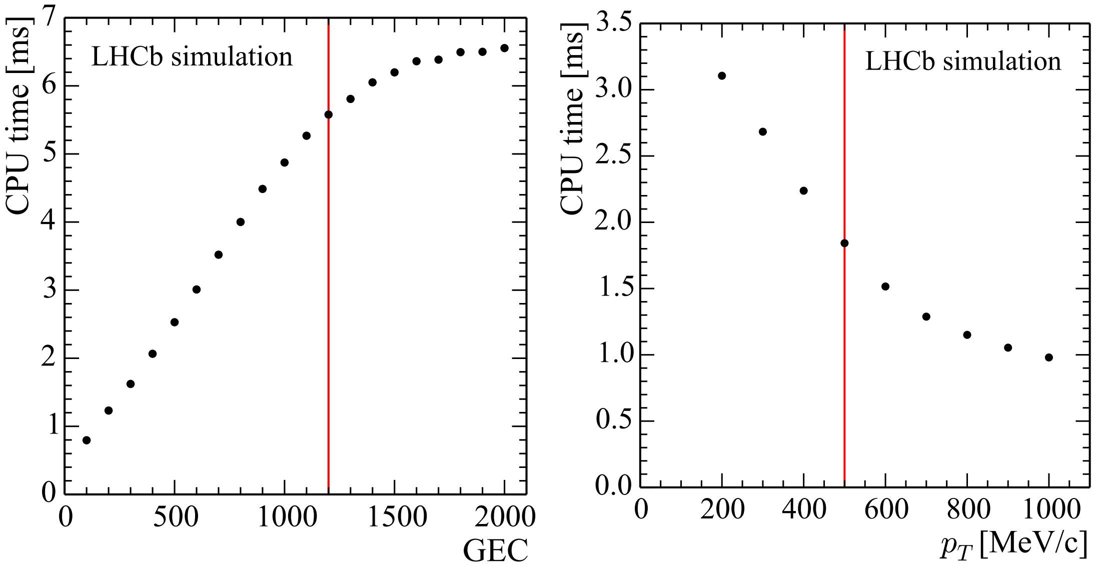


Figure 4.8: Offline and HLT tracking sequence efficiencies for long tracks from b-hadrons with $p_T > 0.5\,\mathrm{GeV}/c$, GEC< 1200. The HLT sequence is shown in red squares, the offline sequence in blue circles, and the ration between both in black triangles. The solid grey histogram shows the distribution of reconstructible particles.



(a) Trigger time versus GEC

(b) Forward tracking time versus p_T

Figure 4.9: (a) Total time spent in trigger reconstruction as a function of GEC cut applied. (b) Forward tracking CPU time as a function of the internal p_T requirement.

Table 4.5: Timing measurements on minimum bias events produced under nominal upgrade conditions. The total is the sum of the preceding rows. For the GEC$<$ 1200 timing, the output rate is scaled from 29 MHz to 30 MHz in the last row to provide a direct comparison.

	CPU time[ms]	
Tracking Algorithm	No GEC	GEC = 1200
VELO tracking	2.3	2.0
VELO-UT tracking	1.4	1.3
Forward tracking	2.5	1.9
PV finding	0.40	0.38
Total @29 MHz		5.6
Total	6.6	5.4

still provides almost all tracks with $p_T > 500\,\text{MeV}/c$ without any intermediate selection requirements. It has to be underlined again that the absolute reconstruction timing numbers, measured on the same CPUs, are around a factor three faster than for the current LHCb detector even though the instantaneous luminosity is a factor five higher. The design of the tracking system of the upgraded LHCb detector will therefore make possible to fully reconstruct all events at a 30 MHz input rate, for the first time at a hadron collider.

4.5.3 RICH particle identification

During Run 1, there was tentative use of the RICH particle identification (PID) in the HLT, but it required many sacrifices in order to make the calculation quicker and did not include up to date calibrations. Starting in Run 2, the detector will be calibrated between the HLT1 and HLT2 levels. This will allow the use of the RICH particle identification in HLT2 processing. It has been shown that information with the same quality as the offline reconstruction can be achieved in an affordable CPU time [37].

The use of RICH PID in the trigger will become more and more necessary due to constraints on the output bandwidth. This will benefit prompt charm decays in particular.

The time taken to perform the calculations, limited to the kaon and pion hypotheses, have been measured on a minimum bias sample simulated in the upgraded conditions. It takes 74 ms for events satisfying the GEC at 1200. It should be noted that the calculation will only be performed at a reduced rate of 1-2 MHz (see Fig. 4.1). Therefore, the effective CPU cost of the RICH particle identification gets reduced to $2.5-5$ ms and thus constitutes between 20 and 40% of the CPU budget.

4.6 Trigger selections and efficiencies

In order to maximise the physics output of the experiment, we plan to use a combination of both inclusive and exclusive trigger selections. Approximately one half of the bandwidth will be allocated to the inclusive b-hadron trigger which is expected to be used by the majority of analyses involving beauty hadrons. Studies involving charm hadrons, however, are mostly selected in the trigger using exclusive selections. A valuable feature of the full event reconstruction in the LHCb upgrade trigger is that many unique selections can be employed.

In this section, we perform a proof-of-principle study of the LHCb upgrade trigger strategy. The results presented in this section demonstrate that this strategy will work in the upgrade running conditions at LHCb. For the time being, the output bandwidth is fixed to 2 GB/s which can be optimised in the future to enhance the physics output of the experiment. Therefore, we consider three output scenarios in this section: 2 GB/s (20 kHz); 5 GB/s (50 kHz) and 10 GB/s (100 kHz). More details on the studies presented in this section can be found in Ref. [38].

4.6.1 Benchmark channels

A small set of decay modes has been chosen for detailed study:

- The decay mode $B^0 \to K^{*0}\mu^+\mu^-$ is an important channel for the LHCb upgrade. It serves as an ideal b hadron to many-body decay, having a trigger efficiency in Run 1 similar to several other important channels of similar topology.

- The decay $B^0_s \to \phi\phi$ is interesting due to its unique signature: four kaons from a secondary vertex where each $M(K^+K^-)$ falls into a narrow mass window around

Table 4.6: Offline selections for benchmark decay modes. Loose track PID requirements and resonance mass criteria are also applied but not shown. No B_s^0 p_T criteria is applied for $B_s^0 \to \phi\phi$. However, $p_T(\phi_1) \times p_T(\phi_2) > 2(\,\mathrm{GeV}/c)^2$ is required. For the D^0 mode, the *slow* pion from the $D^* \to D^0\pi^+$ decay is also selected and required to have $p_T > 100\,\mathrm{MeV}/c$. For the D^+ channel there are additional requirements on minimum displacement of the tracks from the primary vertex (MIP > 0.05 mm), maximum distance of the closest approach between any pair of D^+ child tracks (max DOCA< 0.5 mm), a cosine of the angle between the D^+ momentum direction and the direction between the D^+ primary vertex and decay vertex (DIRA> 0.9). A scalar sum of the D^+ daughter transverse momenta is required to be larger than 1 GeV/c. Wide mass range for the D^+ is meant to cover also the D_s^+ region.

	$B^0 \to K^{*0}\mu^+\mu^-$	$B_s^0 \to \phi\phi$	$D^0 \to K_S^0\pi^+\pi^-$	$D^+ \to K^-K^+\pi^+$
track p_T	$> 250\,\mathrm{MeV}/c$	$> 400\,\mathrm{MeV}/c$	$> 250\,\mathrm{MeV}/c$	$> 200\,\mathrm{MeV}/c$
track χ^2_{IP}	> 4	> 4	> 4	> 4
$M - M(\mathrm{PDG})$	$\pm 250\,\mathrm{MeV}/c^2$	$\pm 250\,\mathrm{MeV}/c^2$	$\pm 70\,\mathrm{MeV}/c^2$	$^{+630}_{-370}\,\mathrm{MeV}/c^2$
τ	> 0.5 ps	> 0.3 ps	–	–
candidate p_T	$> 3.5\,\mathrm{GeV}/c$	–	$> 2\,\mathrm{GeV}/c$	$> 2.5\,\mathrm{GeV}/c$
$\chi^2_{\mathrm{vertex}}/n_{\mathrm{dof}}$	< 15	< 15	< 15	–

$M(\phi)$. The software upgrade trigger will offer the unique opportunity to perform highly efficient exclusive selections for decays of this type. The decay $B_s^0 \to \phi\phi$ is an ideal benchmark for the demonstration of this capability.

- The decay modes $D^0 \to K_S^0\pi^+\pi^-$ and $D^+ \to K^-K^+\pi^+$ are used to study exclusive charm selections.
- The decays $B^0, D^0 \to h^+h^-$ are used to study lifetime-unbiased trigger selections.

Offline selections for each of the exclusive modes motivated by those used in Run 1 are listed in Table 4.6.

The performance of the upgrade trigger is studied for a number of additional b-hadron decay modes as well. However, for all other modes no offline selection has been made available at the time of this study. Instead, we filter (to distinguish from selecting) these modes by making the following requirements: all charged decay products are required to have been reconstructed with $p_T > 250\,\mathrm{MeV}/c$, $\chi^2_{\mathrm{IP}} > 4$; and the generated b-hadron is required to satisfy $p_T > 5\,\mathrm{GeV}/c$ and $\tau > 1\,\mathrm{ps}$. This filter restricts the generated sample to the subset of b-hadron decays that would typically be selected offline. We note that the trigger efficiency for offline-*selected* versus offline-*filtered* $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\phi$ are consistent to within a few percent. The results shown here for the inclusive b-hadron trigger are benchmarked using offline-filtered signal samples.

4.6.2 Topological selection

The LHCb topological trigger (TOPO) is described in detail in Ref. [39]. This trigger inclusively selects $b\bar{b}$ events using a subset of the tracks originating from the decay of a b-hadron. The strategy is to find displaced vertices made from 2, 3 or 4 tracks that do not emanate directly from a PV. A loose selection is applied when forming these vertices. The purity (rate) is increased (decreased) using a boosted decision tree (BDT). This algorithm was first designed for LHCb running in 2011 and is described in detail in Ref. [40]. The BDT-based inclusive trigger has provided LHCb with highly-pure $b\bar{b}$ samples in Run 1.

Implementation

To demonstrate the feasibility of the TOPO in upgrade running conditions we use the same basic strategy as that of Run 1, with an updated preselection and BDT trained on 14 TeV Monte Carlo. For this study, we fix the relative bandwidth division between the 2, 3 and 4-body lines to be the same as what was used in Run 1. The Run 1 TOPO made a looser BDT requirement on candidates that contained muons. This same strategy is employed here. We also assume that some form of ghost killing algorithm will exist in the upgrade with performance equivalent to the current ghost probability at LHCb. This assumption has very little impact on the performance as is discussed below.

In addition to the updated preselection and BDT training, several improvements to the timing performance have been made by requiring that the TOPO only uses as input tracks that do not point back to a PV in addition to selection criteria on the scalar and vector p_T sum of the tracks. With these criteria the timing of the TOPO is found to be well below 0.1 ms, making it impossible to measure it reliably.

Performance

Applying the upgrade TOPO algorithm to minimum bias events, simulated with the upgrade conditions, yields a $b\bar{b}$ purity of 100% for a 10 kHz TOPO output rate. The purity obtained is excellent for the output rates considered in this study. The number of background candidates from non-$b\bar{b}$ categories that pass the TOPO is small. No $c\bar{c}$ or pile-up[4] events pass until the output rate is increased to about 25 kHz. The pile-up background should be removable with the inclusion of a dedicated algorithm using additional information from the VELO.

The performance on a variety of offline-filtered b-hadron decay modes is given in Table 4.7 and compared to the Run 1 trigger efficiencies. For a subset of the decays studied, the efficiency *vs* output rate is shown in Fig. 4.10 (all such plots are shown in Ref. [38]). For the 10 kHz output, the upgrade TOPO efficiencies are roughly the same as in Run 1. At 25 kHz, the upgrade efficiencies are about the same as Run 1 for decays that contain multiple muons, about 50% larger for semileptonic decays, and 2-4 times larger for fully hadronic decay modes. Even larger gains are obtained for hadronic decay modes

[4] We define pile-up backgrounds as candidates that contain at least one track from a PV. Most of these candidates are due to poorly reconstructed or non-reconstructed PVs.

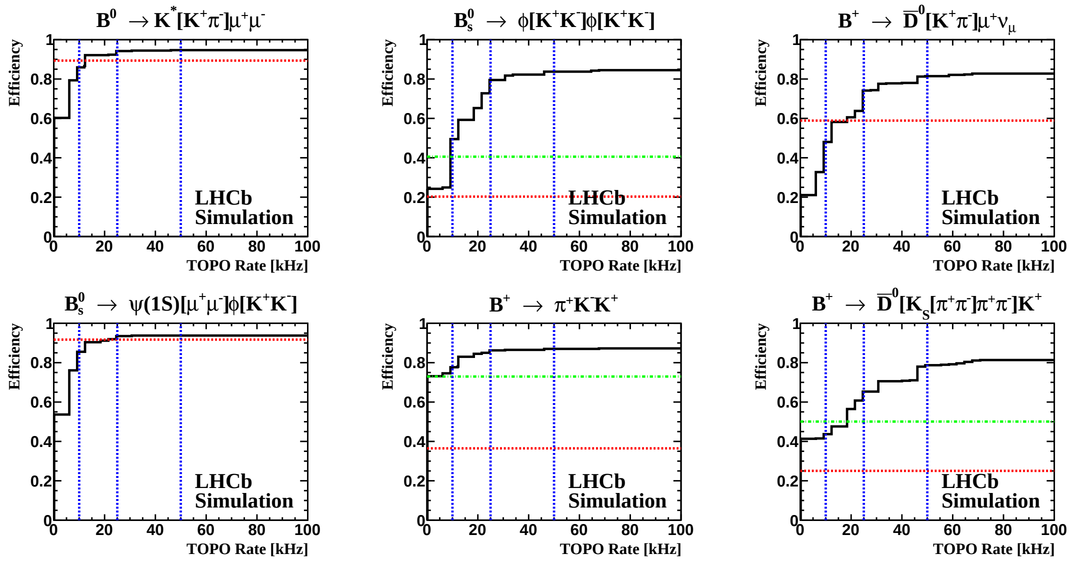


Figure 4.10: Efficiency on offline-filtered signal events *vs* TOPO output rate for a subset of the decays studied. The red dotted line shows the Run 1 trigger efficiency, while the dot-dashed green line shows twice the Run 1 efficiency for hadronic final states. The vertical dotted lines show the three output-rate scenarios considered in this study.

by going to an output rate of 50 kHz. The benefits of moving to a fully software trigger are clearly displayed in these results.

4.6.3 Lifetime unbiased hadronic triggers

The availability of all high-p_T tracks, irrespective of their displacement from PVs, at the first trigger stage makes it possible to select hadronic decay modes in a lifetime unbiased manner. This will be the first time that such triggers can be deployed at full input rate at a hadron collider. In this context, *lifetime unbiased* means that there are no selection criteria on quantities which are correlated with the signal particle's decay-time, apart from an explicit lower cutoff on the decay-time itself. Thus, what is unbiased is the shape of the decay-time distribution. A downscaled sample of events at small decay-times will be kept in order to study decay-time resolution in a data-driven manner. The benefits of this approach are that one removes any need to control decay-time resolution or acceptance functions which reduces the systematic uncertainties of a lifetime-based measurement.

Implementation

A complete description of the implementation is given in Ref. [38]. The challenges of this approach are to control the time taken to form all possible track combinations and the output rate. Of these the timing is the more critical issue, since it affects the general feasibility of the method, while the output rate needs to be tuned for each decay mode

Table 4.7: TOPO performance on offline-filtered (see text for definition) signal samples. The Run 1 efficiency is the full L0xHLT1xHLT2, where HLT2 is assumed to be only the Run 1 TOPO, for each mode. *N.b.*, the Run 1 trigger achieves this efficiency while running at a 5 times lower luminosity. Naive scaling of output rates by luminosity leads to a best comparison between the full software trigger and Run 1 to be the 25 kHz upgrade output rate efficiency values.

		Upgrade TOPO Output Rate		
Decay	Run 1	10 kHz	25 kHz	50 kHz
$b \to s$ penguins				
$B^0 \to K^*[K^+\pi^-]\mu^+\mu^-$	89%	85%	94%	94%
$B^0 \to K^*[K^+\pi^-]e^+e^-$	43%	38%	79%	85%
$B_s^0 \to \phi[K^+K^-]\phi[K^+K^-]$	20%	49%	79%	83%
semi-leptonic decays				
$B^0 \to D^{*-}[\pi^- D^0[K^+\pi^-]]\mu^+\nu_\mu$	63%	58%	81%	90%
$B^0 \to D^-[K^+\pi^-\pi^-]\mu^+\nu_\mu$	40%	27%	61%	74%
$B^+ \to \bar{D}^0[K^+\pi^-]\mu^+\nu_\mu$	58%	48%	74%	81%
$B^+ \to \bar{D}^*[\bar{D}^0[K^+K^-\pi^0]\pi^0]\mu^+\nu_\mu$	39%	25%	64%	72%
$B^+ \to \bar{D}^0[K_S^0[\pi^+\pi^-]\pi^+\pi^-]\mu^+\nu_\mu$	32%	17%	58%	69%
$B_s^0 \to K^-\mu^+\nu_\mu$	59%	52%	67%	71%
$B_s^0 \to D_s^-[K^+K^-\pi^-]\mu^+\nu_\mu$	47%	29%	71%	79%
$\Lambda_b^0 \to p^+\mu^-\bar{\nu}_\mu$	54%	44%	59%	60%
charmless decays				
$B^+ \to \pi^+K^-K^+$	36%	77%	86%	87%
$B_s^0 \to K^-K^+\pi^0$	21%	32%	47%	60%
decays with charmonium				
$B_s^0 \to \psi(1S)[\mu^+\mu^-]\phi[K^+K^-]$	91%	85%	93%	93%
$B_s^0 \to \psi(2S)[\mu^+\mu^-]\phi[K^+K^-]$	93%	86%	93%	94%
$B_s^0 \to \psi(1S)[\mu^+\mu^-]K^+K^-\pi^+\pi^-$	91%	79%	95%	95%
hadronic open charm decays				
$\Lambda_b^0 \to \Lambda_c^+[p^+K^+\pi^-]\pi^-$	33%	67%	87%	90%
$B^+ \to \bar{D}^0[K_S^0[\pi^+\pi^-]\pi^+\pi^-]K^+$	25%	43%	65%	78%
$B^+ \to \bar{D}^0[K^+\pi^-]K^+\pi^-\pi^+$	26%	30%	83%	93%
$B^0 \to D^+[K^-\pi^+\pi^+]D^-[K^+\pi^-\pi^-]$	18%	7%	56%	80%
hadronic τ lepton mode				
$B^0 \to D^{*-}[\pi^-\bar{D}^0[K^+\pi^-]\tau^+[\pi^+\pi^+\pi^-\bar{\nu}_\tau]\nu_\tau$	17%	1%	64%	90%

based on the physics priorities of the experiment. For two-body decays like $B^0, D^0 \to h^+h^-$, all tracks with momenta above the RICH kaon threshold ($p > 9.3\,\mathrm{GeV}/c$) are combined and a vertex is fit. The combination is required to have a scalar sum $p_T > 2.5$ GeV/c and to fall in an appropriate mass window. These selection criteria can be applied prior to the time-consuming vertex fit and are therefore crucial for controlling the combinatorics timing. Other selection criteria are used to reduce the output rate but do not affect the timing. For the quasi-two-body decay $B_s^0 \to \phi\phi$, first a single track above the RICH pion threshold is combined with an oppositely-charged track. If the invariant mass is near that of the ϕ meson, a vertex fit is run. If such a ϕ meson candidate can be made, then a second ϕ candidate is searched for using all tracks with $p_T > 500\,\mathrm{MeV}/c$.

For n-body decays that cannot be factored into a quasi-2-body combination, a lifetime-unbiased selection will not fit into the CPU budget. It is possible, however, to perform a minimally-lifetime-biased n-body selection. All tracks above the RICH pion treshold are first associated to a PV using IP information. If the mininum IP is above some cutoff, then the track is left unassociated to any PV. When combining n tracks, only tracks that are either associated to the same PV or are unassociated are considered. Thus, the only combinations not considered are those that contain tracks associated to different PVs. This approach is found to cause only a few percent inefficiency at small lifetimes and no observable inefficiency above about 1 ps.

Performance

The timing of these selections is measured using the same configuration and minimum bias samples as used for the reconstruction timing measurements. The two-body and $B_s^0 \to \phi\phi$ timings are 0.1 ms/event. There is enough CPU time available for adding more such selections to the trigger within the available budget. In the case of the generic selection, the timing is 0.2 ms/event. We conclude that the generic n-body minimally-lifetime-biasing timing is under control and that these selections fit comfortably into the trigger timing budget as well.

For $B^0 \to h^+h^-$ decays we measure the rate separately for $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ decays, the two main time dependent analyses in this family. In both cases PID requirements are applied. The measured rate is about 1 kHz for $B^0 \to \pi^+\pi^-$ and 100 Hz for $B_s^0 \to K^+K^-$ decays, while the efficiency on offline-filtered $B^0 \to h^+h^-$ candidates is about 60%. For $D^0 \to h^+h^-$ decays we measure the rate separately for the four combinations : $\pi^+\pi^-$, K^+K^-, and the Cabibbo-favoured and doubly Cabibbo-suppressed $K^\pm\pi^\mp$ combinations. In all cases PID requirements are applied. For $D^0 \to K^+K^-$ we measure a rate of 2 kHz of which only 500 Hz contain a misidentified pion. For the Cabibbo-favoured $D^0 \to K^-\pi^+$ decay, we measure a rate of 20 kHz with high signal purity (see Ref. [38]). The Cabibbo-favoured mode can be downscaled by a factor of 10 without any losses in physics performance as analyses which use these modes are not statistically limited by their Cabibbo-favoured sample. For $D^0 \to \pi^+\pi^-$ decays and the Cabibbo-suppressed modes we measure a total output rate of about 40 kHz. Downscaling these modes would directly translate into losses in physics performance so these selections will need to be tightened. The effects of the selection criteria applied can be studied using the unbiased K^+K^- and downscaled Cabibbo-favoured decay. Further study is required to determine how to maximize the physics output from these modes while satisfying the output-bandwidth constraints.

For the lifetime unbiased $B_s \to \phi\phi$ selection, the rate without applying PID requirements is 1.8 kHz. Applying PID criteria that are 90% efficient on $B_s \to \phi\phi$ decays reduces the selection output rate down to a negligible 12 Hz. At this rate a large number of such selections can be added. The efficiency with respect to the lifetime biasing offline selection given in Table 4.6 is 89%.

4.6.4 Exclusive charm and beauty triggers

For a number of decay modes, a higher efficiency is obtained using an exclusive selection. For the decay $B_s^0 \to \phi\phi$ the offline selection given in Table 4.6, biasing the lifetime, can be implemented in the trigger by increasing the p_T cut to 500 MeV and by requiring $\sum |p_T| > 4$ GeV/c. These cuts are 95% efficient on offline-selected candidates. The output rate without applying PID requirements is about 700 Hz. The output rate is sufficiently low to permit running PID algorithms. Applying nearly 100% efficient PID requirements results in an output rate of less than 10 Hz. Given the present bandwidth requirements, a large number of such selections could be included without difficulty. As only displaced tracks are used in these selections, the timing requirements are negligible. Exclusive selections of b-hadron decays can be performed in the LHCb upgrade trigger.

Charm decays such as $D^0 \to K_S^0\pi^+\pi^-$ also fall into the same category where an exclusive selection is more efficient than an inclusive one. The offline selection for $D^0 \to K_S^0\pi^+\pi^-$ is given in Table 4.6. Increasing the track p_T requirement to 500 MeV/c is about 60% efficient on offline-selected candidates. The slow pion that originates from the $D^* \to D^0\pi^+$ decay does not have sufficient momentum to be efficiently selected using long tracks. Therefore, the slow pion is selected using a VELO-UT track with a requirement of $p_T > 200$ MeV/c. The D^* is cleanly selected by requiring $\Delta m = M(D^0\pi) - M(D^0) < 160\,\mathrm{MeV}/c^2$ even with the lower momentum resolution of the VELO-UT pion. This trigger selection is about 50% efficient on offline-selected candidates but produces an output rate of about 9 kHz. Further tightening the selection, by requiring D^0 mass to be within 60 MeV/c^2 of the nominal value, $\Delta m < 155$ MeV/c^2 and the $\tau(D^0) > 0.2$ ps, reduces the output rate to 1.3 kHz and has signal efficiency of 40%.

Multibody decays of D^+ and D_s^+ are expected to produce large trigger output rates which cannot be reduced with the D^* tagging. Therefore quite a tight trigger selection is applied to the $D^+ \to K^-K^+\pi^+$ decays which are reconstructed using only long tracks significantly displaced from the PV ($\chi^2_{\mathrm{IP}} > 6$). We accept only the D^+ candidates having a good quality decay vertex, significant flight distance from the PV, $p_T > 3$ GeV/c, and invariant mass within $\pm 60\,\mathrm{MeV}/c^2$ of the nominal mass (the full trigger selection is given in Ref. [38]). Selecting long tracks has efficiency of only about 12% on offline-selected candidates but reduces the output rate by a factor of two and significantly decreases the multiplicity. All the other requirements reduce the trigger efficiency to 9% and give the output rate of 56 kHz. Requiring $\tau(D^+) > 0.5$ ps and applying PID cuts on both final state kaons results in an efficiency of 6% and an output rate of 33 kHz. These results show that for the $D^+_{(s)}$ decays we may need to perform a multivariate analysis to optimise the selection and suppress the output rate more efficiently. Furthermore, the track p_T requirement has non-uniform efficiency across the Dalitz plot. Thus, if the trigger lines need to be prescaled for these decays, such acceptance effects can be taken into account by prescaling according to Dalitz-plot location.

The charm production cross section is so large that efficiently selecting charm decays, even exclusively, while producing output rates $\mathcal{O}(1\,\mathrm{kHz})$ is difficult and not feasible for many decay modes. Therefore it is difficult to make any estimates of the minimum bandwidth required for efficient charm hadron selections. One can naively scale the charm lines from Run 1 taking into account an increase of the $c\bar{c}$ cross section, luminosity and impact of removing the L0 hardware trigger, which gives an estimated output rate of 70 kHz as compared to 2 kHz in Run 1. Therefore, to improve the trigger performance with respect to Run 1 one needs to come up with a new strategy for charm trigger lines. Tagging all the D^0 decays with $D^* \to D^0\pi^+$ would reduce the related trigger rate by almost one order of magnitude without affecting the physics potential. Increasing signal purity is very helpful also for studies of dynamics of the D^0 multibody decays and branching ratio measurements. It would still be necessary to keep some of the $D^0 \to K^-\pi^+$ decays untagged for charm spectroscopy measurements as well as calibration studies and estimation of various systematic effects. Similarly to D^0 decays, one could select Λ_c^+ from $\Sigma_c(1455)^{++,0}$ and $\Sigma_c(1520)^{++,0}$ decays to the $\Lambda_c^+\pi^{+,-}$ final states, since these Σ_c states are as copiously produced as Λ_c baryons. In addition to the tagged Λ_c lines there should also be an untagged $\Lambda_c^+ \to pK^-\pi^+$ line for charm baryon spectroscopy and all the calibrations. Finally, the general inclusive selection from Run 1 is not feasible in the upgrade trigger, but a semi-inclusive line for the tagged D^0 and/or the tagged Λ_c should be feasible.

4.6.5 Inclusive and exclusive di-muon selections

Not all decay modes with two or more muons are efficiently selected by the trigger selections, *e.g.* the decay $\tau \to 3\mu$. For such modes it is possible to perform an exclusive selection that is about 60–70% efficient relative to the offline selection that outputs a rate of about 50-100 Hz. For modes like $B_s^0 \to \mu^+\mu^-$, an exclusive selection that is close to 100% efficient on offline-selected candidates produces an output rate of about 100 Hz.

Inclusive selection of detached di-muons is also possible in the upgrade trigger using only displaced tracks that are identified as muons. Selection criteria are applied on the quality and flight distance of the di-muon vertex, and on the mass and p_T of the di-muon system. This selection outputs a rate of 1–3 kHz, depending on the muon identification criteria applied. Its efficiency on $B^0 \to K^*\mu^+\mu^-$ candidates relative to the offline selection is about 75–80%, while for $\tau \to 3\mu$ candidates the efficiency is about 50–70%.

The efficiency for the decay $B^0 \to K^*\mu^+\mu^-$ is lower than what is obtained in the topological trigger. However, for related decay modes, such as $B^0 \to K_\mathrm{S}^0\mu^+\mu^-$, where the hadrons are unlikely to contribute to the topological trigger efficiency, the inclusive di-muon selection will help improve the total trigger efficiency. Furthermore, for various low-mass di-muons, such as light dark matter searches, this inclusive selection provides efficiency where no other trigger selection applies.

4.6.6 Electroweak and high-p_T selections

LHCb currently has vibrant programs in electroweak and high-p_T physics. While these programs are expected to continue to be important in Run 3, their trigger selections require minimal computing resources and output bandwidth. Therefore, they are omitted from direct study in this document.

4.6.7 Output-bandwidth-scenarii

In this section, three possible output-bandwidth scenarios are considered: 2 GB/s (20 kHz); 5 GB/s (50 kHz) and 10 GB/s (100 kHz). The objective here is to show a realistic bandwidth division for each option and to identify the physics gains by increasing the output rate. The bandwidth divisions are shown in Table 4.8.

Table 4.8: Possible output-bandwidth scenarios for the upgrade trigger, along with plausible bandwidth divisions for each (ϵ denotes small).

Selection	Output Rate (kHz)		
Topological	10	20	50
Lifetime unbiased	1	4	5
Exclusive beauty	ϵ	1	3
Inclusive di-muon	–	–	2
Charm	9	20	40
Total	20	50	100
Bandwidth [GBs^{-1}]	2	5	10

The topological trigger efficiencies for each scenario are given in Table 4.7 for b-hadrons decays. For a total output rate of 20 kHz, corresponding to a topological output rate of 10 kHz, the topological trigger efficiency is roughly the same as the Run 1 trigger. At a total output rate of 50 kHz, the efficiencies are about the same as Run 1 for decays that contain multiple muons, about 50% larger for semileptonic decays, and 2–4 times larger for fully hadronic decay modes. Even larger gains are obtained for hadronic decay modes by going to an upgrade topological output rate of 50 kHz. Of the core physics goals, making a precise measurement of the CKM angle[5] γ gains the most by increasing the output rate of the topological trigger.

For lifetime-unbiased selections, at a total output rate of 20 kHz there is only room for a few tight b-hadron selections. At 50 kHz total output rate there is sufficient bandwidth available to run efficient lifetime-unbiased b-hadron selections and a few tight lifetime-unbiased charm selections. At 100 kHz, many more charm selections could be added to the lifetime-unbiased list.

Exclusive-beauty selections can be summarised as follows: (20 kHz) only very low-rate lines like $B_s^0 \to \phi\phi$ and lines for golden modes like $B_s^0 \to \mu^+\mu^-$ may be run; (50 kHz) a handful of important exclusives may be added; and (100 kHz) about 10–20 more exclusives can be added. There is only sufficient bandwidth to run inclusive-di-muon selections in the highest output-rate scenario considered here.

[5]This is the *CP*-violating phase in the SM. It is measured at tree level using hadronic open charm decays of the form $B \to DX$.

In the lowest output-rate scenario considered, there is only room for a few golden charm modes to be selected exclusively. At a total output rate of 50 kHz, some tight inclusive-charm selections can be added. At the highest output rate considered here, there is sufficient bandwidth for the following: efficient exclusive-charm selections; efficient inclusive-charm selections; and several lifetime-unbiased-charm selections as well. The charm production rates are large enough that it simply is not possible to efficiently select charm decays and write out at a low rate.

In summary, at a total output rate of 20 kHz the physics program at LHCb will need to be restricted. At 50 kHz a diverse beauty program will be possible, while a charm program of similar scope to that of Run 1 can be carried out. At 100 kHz the beauty program reaches its full potential, while the charm program records the *legacy* dataset of charm physics.

This study strongly motivates writing data out at a high rate. The limit on what can be written will be determined by the offline-computing resources available. One way to increase the physics output without increasing the need for additional offline-computing resources is to decrease the event size. This may be possible for certain types of events, *e.g.*, those selected by charm triggers. If recording a subset of the event information is sufficient for offline analysis, then the output rate can be increased without increasing the offline-computing resources required to analyse the data. Another option would be to put certain types of data onto tape and delay analysing them until the offline-computing resources required become available. These approaches will be exercise during the Run 2.

4.7 Robustness

This section addresses the robustness of the proposed trigger system and discusses the interplay between the proposed trigger system and physics analysis potential.

4.7.1 Data-simulation differences

All studies performed in this document are made using the latest and most complete available simulation of the upgraded LHCb detector. The discussed performance depend on the simulated signal and inelastic pp collision kinematics, the detector responses, and on the event multiplicities. While the first two are well simulated in the current detector, the hit and track multiplicities are found to be underestimated by simulation in Run 1 running conditions. Studies on Run 1 data and the corresponding Monte-Carlo simulation have shown that the CPU timing of the trigger sequence in the simulation can be underestimated by up to 50%.

We counter-balance the potential optimism of our performance numbers by making very conservative assumptions when estimating the available computing budget. Specifically, the CPU budget of 13 ms as discussed in Sect. 3.6 does not take into account the buffering of events to the local disks known as deferred processing[6]. As shown in Fig. 4.1, the HLT

[6] This local buffering of events has been used very successfully in the second half of the Run 1.

application will be split into two phases, the first of which will treat all events and store the accepted ones locally. These locally buffered events will be processed in the inter-fill gaps, which is expected to gain a factor of two in effective CPU resources.

An additional safety margin is provided by the LLT, described in Sect. 4.4. The LLT can reduce the rate by a factor of two with limited performance losses in beauty signals, however at significant cost in charm efficiency. The LLT is used as temporary safety net and as such is not considered in any CPU timing budget made in this document.

4.7.2 Partial reconstruction for the upgrade trigger

The full software trigger performs a complete event reconstruction upfront. This is a much more challenging and advanced trigger system than the one used during Run 1 [39, 41]. The latter applies hard selection criteria on the reconstructed VELO tracks and then upgrades only a small fraction of them to long tracks. The intention is to inclusively reconstruct the highest p_T track originating from a b-hadron decay [42].

The strategy of partially reconstructing only a small subset of the tracks limits the flexibility of the trigger system and reduces the efficiency for many signals, especially for non-beauty or multi-body decays. It is highly desirable, therefore, to keep the full event reconstruction, as discussed in Sect. 4.5. However, the partial reconstruction approach is a backup solution which preserves most of the absolute trigger efficiency for the core b-hadron physics which LHCb will study.

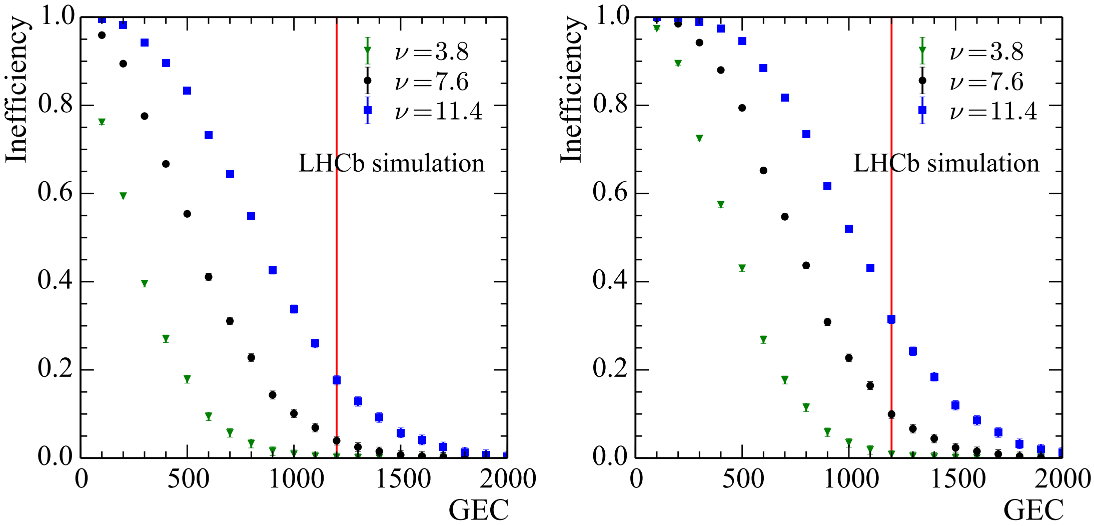
Running a partial reconstruction based trigger, the timings improve relative to Table 4.5. The total time needed by the VELO-UT and the VELO algorithms is reduced by 25% and the forward tracking by 52%. The total time is 3.2 ms instead of 5.4 ms when the GEC is equal to 1200. Additionally, due to the tight search windows, the timing cost of this approach is largely insensitive to the total event multiplicity.

4.7.3 Performance at modified luminosities

After Long Shutdown 2, we do not expect to start data taking immediately at the nominal upgrade luminosity of 2×10^{33} cm^{-2} s^{-1}. Rather, we expect to increase luminosity gradually until we reach nominal, or if performance enhancements continue throughout preparations for the upgrade we may even wish to operate at rates higher than that which we have planned for. We therefore study the performance of the trigger system using both the reduced and increased luminosity samples described in Table 4.1. This additionally provides an estimate of the robustness of the performance at multiplicities above those we expect. These working points are the same as the ones used in the recent tracking TDR [5], they are discussed in further detail in Ref. [36].

The performance of the track reconstruction is tested for two scenarios: without GECs and with a GEC of 1200.

The inefficiency introduced by GECs for both signal and minimum bias events is presented in Fig. 4.11 for the three luminosities studied as a function of their ν : 3.8 for the reduced luminosity sample, 7.6 for the nominal sample and 11.4 for the increased



(a) Minimum bias

(b) $B_s^0 \to \phi\phi$

Figure 4.11: Minimum bias- and signal inefficiency as a function of the applied GEC. Inefficiencies on signal MC at GEC $= 1200$ are $(0.8 \pm 1.0)\,\%$ at $\nu = 3.8$, $(9.9 \pm 0.9)\,\%$ at $\nu = 7.6$, and $(31.4 \pm 0.9)\,\%$ at $\nu = 11.4$.

luminosity. The GEC of 1200 is fully efficient at the reduced luminosity, 10% inefficient at nominal luminosity, and 31% inefficient at the increased luminosity.

Performance at $1 \times 10^{33}\ \mathrm{cm^{-2}\,s^{-1}}$

Table 4.9: CPU timing for different luminosity running scenarios. The total is the sum of the preceding rows. For the GEC< 1200 timing, the output rate is scaled from 29 MHz (nominal luminosity, 25 MHz for increased luminosity) to 30 MHz in the last row to provide a direct comparison.

	CPU time[ms]					
	reduced luminosity		$2 \times 10^{33}\ \mathrm{cm^{-2}\,s^{-1}}$		increased luminosity	
Algorithm	No GEC	GEC = 1200	No GEC	GEC = 1200	No GEC	GEC = 1200
VELO	0.85	0.84	2.3	2.0	4.4	3.2
VELO-UT	0.69	0.68	1.4	1.3	2.3	1.7
Forward	0.85	0.83	2.5	1.9	6.3	3.2
PV finding	0.18	0.18	0.40	0.38	0.69	0.54
Sum @ 25 MHz	-	-	-	-	-	8.7
Sum @ 29 MHz	-	-	-	5.6	-	-
Sum @ 30 MHz	2.6	2.5	6.6	5.4	14	7.2

The performance at a reduced luminosity is summarised in Table 4.9. The overall time of the reconstruction sequence is reduced to 2.5 ms instead of 5.4 ms when the GEC is equal to 1200. Therefore, less than 18% of the planned CPU capacity is required for the full event reconstruction. The remaining CPU time might be used to lower the p_T thresholds in the Forward and VELO-UT tracking algorithms.

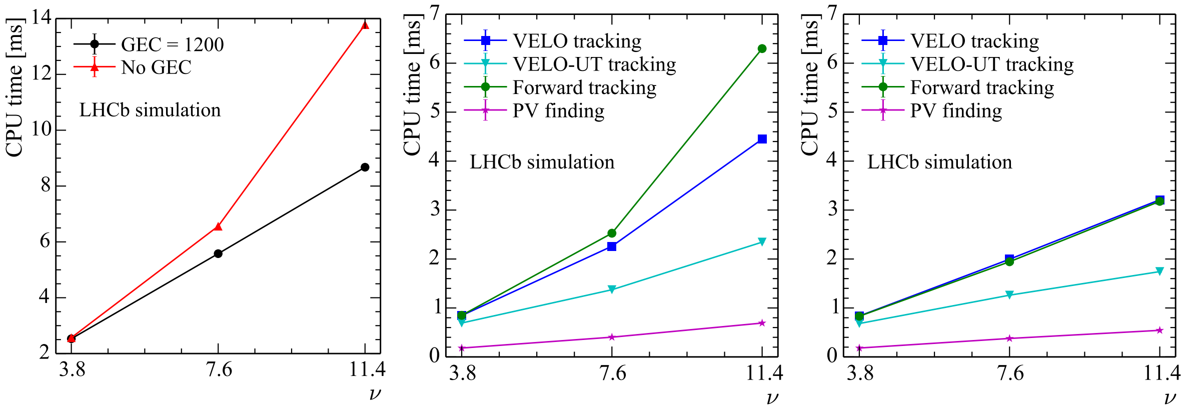
The GEC cut of 1200 affects the CPU time of the tracking sequence only marginally, while for lower cut values, the behaviour is fully analogous to the nominal conditions discussed previously.

Performance at 3×10^{33} cm^{-2} s^{-1}

The performance of the trigger reconstruction sequence at a the increased luminosity working point is given in Table 4.9. At this luminosity a GEC requirement of 1200 introduces a large inefficiency, measured of about 30% on b-signal.

It takes 7.2 ms instead of 5.4 ms when the GEC is equal to 1200. However, for more efficient running, the initial farm will not be sufficient since the current implementation of the algorithms would need about the whole budget without GECs.

In that exercice, neither the tracking algorithms nor the trigger sequence were tuned for increased luminosity. During the preparation of this document both the execution time and efficiency of all tracking algorithms were improved dramatically. It is plausible that such improvements will continue over the next decade.



(a) Total timing (b) Individual, no GEC (c) Individual, GEC

Figure 4.12: The time cost of the total track reconstruction sequence as a function of ν (left), and the individual timings of all tracking algorithms with respect to ν without (middle) and with (right) a GEC requirement of 1200.

Performance of individual tracking algorithms as a function of luminosity

The CPU time required by the individual tracking algorithms is shown in Fig. 4.12b and 4.12c as a function of the GEC requirement. The VELO, VELO-UT and PV finding algorithms scale linearly with ν. The Forward tracking algorithm scales linearly with luminosity when GEC selection criteria are applied. Without a GEC requirement it scales faster than linearly. This behaviour of the Forward tracking algorithm is well known [43] and possible improvements of it are currently under study.

4.7.4 Performance with degraded single hit resolution

The simulation of the SciFi tracker assumes that the single hit resolution is 42 μm. Corresponding to test measurements of short fibre tracker modules in a cosmic ray experiment [44]. A 2-bit read-out scheme will be used for the SciFi tracker which will result in a resolution of about 60 μm. Additional misalignment or noise could further worsen the single hit resolution.

Table 4.10: Track reconstruction efficiencies for different single hit resolutions.

	Efficiency [%]	
Resolution	long, $p_T > 0.5\,\mathrm{GeV}/c$	long, from B, $p_T > 0.5\,\mathrm{GeV}/c$
42 μm	87.4	92.5
62 μm	86.7	92.1
82 μm	86.5	92.9
100 μm	86.0	91.4

The track reconstruction efficiency is compared over a range of single hit resolutions from 42 μm to 100 μm. Table 4.10 shows that the reconstruction efficiency decreases by 1% as the single hit resolution worsens by a factor two. The track reconstruction is robust against changes of the single hit resolution in the SciFi tracker. The loss in tracking efficiency can be recovered by retuning the tracking algorithms for each particular hit resolution scenario.

4.8 Project organisation

The trigger upgrade is managed by the trigger upgrade coordinator, who is also a deputy project leader of the High Level Trigger project. The trigger upgrade coordinator also works in close collaboration with the L0 trigger project.

The institutes currently working on the trigger system are listed in Table 4.11. To fully exploit the physics potential of the upgraded trigger system it is essential that the project has sufficient people and that a team of core software, reconstruction and trigger selection experts is assembled that will be fully dedicated to the development and optimisation of the system. It is clear that the potential of the proposed full software trigger can only be exploited with a dedicated effort in software optimisation, performed in collaboration of physicists and software engineers.

The migration from the current hardware L0 implementation to the software LLT implementation will be the responsibility of the institutes already involved in the maintenance of the L0, as detailed in Table 4.11. The software LLT algorithms will be fully integrated into the global trigger software. The small remaining hardware parts dedicated specifically to the LLT are included in the FE electronics developed by the calorimeter and muon projects.

Table 4.11: Breakdown of the contributions to the trigger system.

System	Institutes
LLT:	
Calo	LAL Orsay/LAPP Annecy
Muon	CPPM Marseille
Decision	LPC Clermont-Ferrand
HLT:	
Code control / releases	CERN / VU Amsterdam
Software framework development	CERN / VU Amsterdam / Nikhef / EPFL Lausanne
Bandwidth division	CERN / Rio
Interface with online	CERN / Nikhef / VU Amsterdam
Online monitoring	CERN
Online adaptation of reconstruction	CERN / TU Dortmund / U. Heidelberg / Pisa
Online calibration	Oxford / Manchester / Cincinnati Bristol / Cambridge / Nikhef
Inclusive beauty trigger	MIT
Charm triggers	Cincinnati / Glasgow / U. Heidelberg / UFRJ/Rio / Padova
Muon triggers	CERN / TU Dortmund / Nikhef
Calorimeter objects	Barcelona / LPC Clermont-Ferrand
PV reconstruction	Krakow / CERN
RICH implementation	CERN / Cambridge / Oxford / Edinburgh
Trigger simulation	CERN

4.8.1 Trigger project schedule

Development of the algorithms used in the full software trigger will continue until the start of data taking in 2020. Specifically, the following points are emphasised:

Using Run 2 to validate possible strategies

Several proposed technical improvements are currently being implemented and will be tested in Run 2:

Split between HLT1 and HLT2: The software trigger has been split into two standalone applications. HLT1 will preprocess every event, while HLT2 will be run asynchronously, which allows sufficient time for detector alignment and calibration before the execution of HLT2. This splitting of the HLT is presently being commissioned.

Data handling: Several proposals for streaming the data have been implemented for Run 2 and will be commissioned for the upgrade. A large data set will be recorded, but not reconstructed and analysed immediately. Another data set will be saved in a reduced-size format and the trigger output analysed directly.

Physics priorities

The upgraded LHCb experiment will have the unique opportunity to collect an unprecedented sample of beauty and charm decays. The full software trigger permits selecting pure samples of such decays for offline processing. However, due to the limit on the output bandwidth, decisions will need to be made on the focus of the physics programme. These decisions will be guided by the following:

Run 2 physics outcome: An advantage of a full software trigger is that it can be quickly adapted to changes in the priorities of our physics programme, for example, in response to new discoveries during Run 2 of the LHC or elsewhere.

Theoretical understanding: Currently, theoretical understanding of observables sensitive to New Physics in b-meson decays is quite advanced [1,45]. In charm or b-baryons, however, Standard Model contributions to similar observables are less well understood [46] and hence unambiguous identification of the presence of New Physics is more difficult for these decays. Theoretical progress made in the coming years will determine the focus, in particular of the charm physics programme and thus, the charm trigger selections.

Continuous benchmarking & optimisation

The technology of microprocessors, both x86 and alternative architectures, will be monitored continuously making it possible to choose the most cost efficient option for the EFF. The trigger software will undergo optimisation, both within the trigger group and as part of future collaboration-wide optimisation activities. We foresee an optimisation programme that will adapt the experimental software to optimally exploit modern hardware, both in the general design of the software framework and in individual algorithms.

Output bandwidth

A decision on the best use of the available offline computing resources, in particular the total output bandwidth, will be made in 2018, based on the points discussed above. It should be noted that the final physics output of the experiment is constrained by the quantity of RAW data recorded. Analyses can be postponed until sufficient resources become available to process this data in a strategy commonly referred to as *data parking*.

Acknowledgements

We would like to thank our external reviewer Pere Mato (CERN, PH/SFT) for his careful reading and many pertinent remarks and suggestions.

We gratefully acknowledge the contributions of the GS/SE, EN/MEF and EN/EL groups from CERN for their invaluable help in the preparation and realization of the long-distance optical fibre cabling prototype installation.

Bibliography

[1] LHCb Collaboration, *Updated sensitivity projections for the LHCb Upgrade*, LHCb-PUB-2013-015. Sep, 2013.

[2] LHCb collaboration, I. Bediaga *et al.*, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007. Apr, 2012.

[3] P. Moreira *et al.*, *The GBT, a proposed architecture for multi-Gbps data transmission in high energy physics*, in *Proceedings of the Topical Workshop on Electronics for Particle Physics*, p. 332, November, 2007. CERN-2007-007.

[4] LHCb Collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021. November, 2013.

[5] LHCb Collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001. February, 2014.

[6] J. Albrecht, V. Gligorov, and G. Raven, *Review Document: Full Software Trigger*, LHCb-PUB-2014-036. Mar, 2014, On behalf of the the HLT software group.

[7] L. Ristori, *An artificial retina for fast track finding*, vol. 453, pp. 425 – 429, 2000. Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics, doi: http://dx.doi.org/10.1016/S0168-9002(00)00676-8.

[8] A. Abba *et al.*, *A specialized track processor for the LHCb upgrade*, LHCb-PUB-2014-026. Mar, 2014.

[9] S. Akar *et al.*, *Review document: Low Level Trigger (LLT)*, LHCb-PUB-2014-037. Mar, 2014.

[10] LHCb Collaboration, *LHCb online system, data acquisition and experiment control: Technical Design Report*, CERN-LHCC-2001-040. December, 2011.

[11] LHCb Collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022. November, 2013.

[12] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[13] L. Amaral *et al.*, *The versatile link, a common project for super-LHC*, Journal of Instrumentation **4** (2009), no. 12 P12003.

[14] D. R. Myers *et al.*, *The LHC experiments Joint COntrols Project, JCOP*, in *Proc. Int. Conf. on Accelerator and Large Experimental Physics Control Systems (Trieste) Italy, 1999*, 1999.

[15] A. Xiang *et al.*, *Versatile Link Technical Specification, part 1*, CERN-EDMS-1140664. April, 2011.

[16] R. Schwemmer *et al.*, *Evaluation of 400 m, 4.8 Gbit/s Versatile Link lengths over OM3 and OM4 fibres for the LHCb upgrade*, JINST (2013), no. JINST_109P_1113.

[17] J.-P. Cachemiche *et al.*, *Readout board specifications for the LHCb upgrade*, CERN-EDMS-1251709. May, 2011.

[18] C. Abellan *et al.*, *Feasibility study of implementation of the LHCb readout board in several form factors*, CERN-EDMS-1357162. February, 2014.

[19] F. Alessio and R. Jacobsson, *Timing and Fast Control for the upgraded readout architecture of the LHCb experiment at CERN*, IEEE TNS **60** (2013) 5.

[20] K. Wyllie *et al.*, *Electronics Architecture of the LHCb Upgrade*, LHCb-PUB-2011-011. Jun, 2013.

[21] F. Alessio and R. Jacobsson, *System-level Specifications of the Timing and Fast Control system for the LHCb Upgrade*, LHCb-PUB-2012-001. Feb, 2014.

[22] F. Alessio and R. Jacobsson, *Readout Control Specifications for the Front-End and Back-End of the LHCb Upgrade*, LHCb-PUB-2012-017. Feb, 2014.

[23] F. Alessio and R. Jacobsson, *A new readout control system for the LHCb upgrade at CERN*, JINST **7** (2012) C11010.

[24] F. Alessio, Z. Guzik, and R. Jacobsson, *LHCb Global Timing and Monitoring of the LHC Filling Scheme*, LHCb-PUB-2011-004. Jan, 2011.

[25] F. Alessio, *Timing and Fast Control aspects in a fully trigger-less PCIe-based readout architecture for the LHCb upgrade*, CERN-EDMS-1376144. Feb, 2014.

[26] InfiniBand Trade Association, *IBTA website*, `http://www.infinibandta.org/`.

[27] G. Liu and N. Neufeld, *DAQ Architecture for the LHCb Upgrade*, in *Proc. CHEP2013*, 2014.

[28] *ECFA High Luminosity LHC Experiments Workshop: Physics and Technology Challenges. 94th Plenary ECFA meeting*, .

[29] B. Franek and C. Gaspar, *SMI++ - an object oriented Framework for designing distributed control systems*, IEEE Trans. Nucl. Sci. **45** (1998) 1946.

[30] C. Gaspar, *The LHCb Experiment Control System: On the path to full automation*, in *Conf. Proc. ICALEPCS2011, Grenoble, France*, 2011.

[31] L. Brarda *et al.*, *A new data-centre for the LHCb experiment*, Journal of Physics: Conference Series **396** (2012), no. 1 012009.

[32] E. Bonaccorsi, N. Neufeld, and F. Sborzacchi, *Performance evaluation and capacity planning for a scalable and highly available virtulization infrastructure for the LHCb experiment*, Journal of Physics: Conference Series (2013), no. 1.

[33] LHCb collaboration, *Addendum to to the Memorandum of Understanding for Collaboration in the Construction of the LHCb Detector: Common Project Items*, CERN/RRB 2012-119A. Apr, 2014.

[34] A. Badalov *et al.*, *GPGPU opportunities at the LHCb trigger*, LHCb-PUB-2014-034. CERN-LHCb-PUB-2014-034. May, 2014.

[35] C. Fitzpatrick and V. V. Gligorov, *Anatomy of an upgrade event*, LHCb-PUB-2014-027. Mar, 2014.

[36] T. Head, J. Albrecht, and K. Dungs, *Track finding scenarios for the upgrade trigger* , LHCb-PUB-2014-028. Mar, 2014.

[37] S. Benson and O. Lupton, *PID Performance in the LHCb High Level Trigger*, LHCb-PUB-2014-038. May, 2014.

[38] M. Williams *et al.*, *Trigger selections for the LHCb upgrade*, LHCb-PUB-2014-031. Mar, 2014.

[39] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[40] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, `arXiv:1210.6861`.

[41] J. Albrecht, V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, `arXiv:1310.8544`.

[42] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003.

[43] Y. Amhis, M. De Cian, T. Nikodem, and F. Polci, *Description and performance studies of the Forward Tracking for a scintilating fibre detector at LHCb*, LHCb-PUB-2014-001. January, 2014.

[44] B. Beischer *et al.*, *A high-resolution scintillating fiber tracker with silicon photomultiplier array readout*, Nucl. Instrum. Meth. **A622** (2010) 542, `arXiv:1011.0226`.

[45] LHCb Collaboration, R. Aaij *et al.*, *Implications of LHCb measurements and future prospects*, Eur. Phys. J. C **73** (2012) 2373. 178 p.

[46] A. Lenz and T. Rauh, *D-meson lifetimes within the heavy quark expansion*, Phys. Rev. **D88** (2013) 034004, `arXiv:1305.3588`.