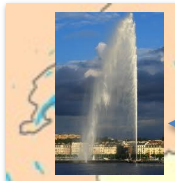
Calcolo Scientifico: appunti di viaggio

Concezio Bozzi

Seminario CNAF

12 Maggio 2021

About myself



2015

99-09

1999

Ferrara

Pisa

1987

Chieti

*1968

I am senior staff researcher at INFN

My research activity:

SLB decays
Vertex DETector

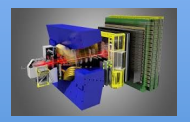
R&D on silicon strip tracker

Silicon Vertex Tracker, B mixing, Vub,
Distributed computing, simulation production

Vub averages

TCAD&G4 simulations

PHYSICS



semileptonic analyses: B^0 mixing, hadronic R(D*)
Measurement of χ_b production and spectroscopy

COMPUTING

Computing resource manager
R&D on software and computing for LHCb upgrade
Editor of SW & computing upgrade TDRs
Computing Project Leader

I have served in various committees and review boards

CERN

INFN Istituto Nazionale di Fisica Nucleare

MINISTERO DELL'ISTRUZIONE, DELL'UNIVERSITÀ E DELLA RICERCA

But what I really like is:

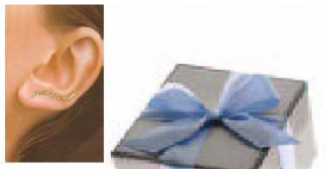
A lightning summary of flavour physics

Search for physics beyond the Standard Model, by measuring decays of heavy hadrons

ATLAS / CMS:
open the Box

real particles

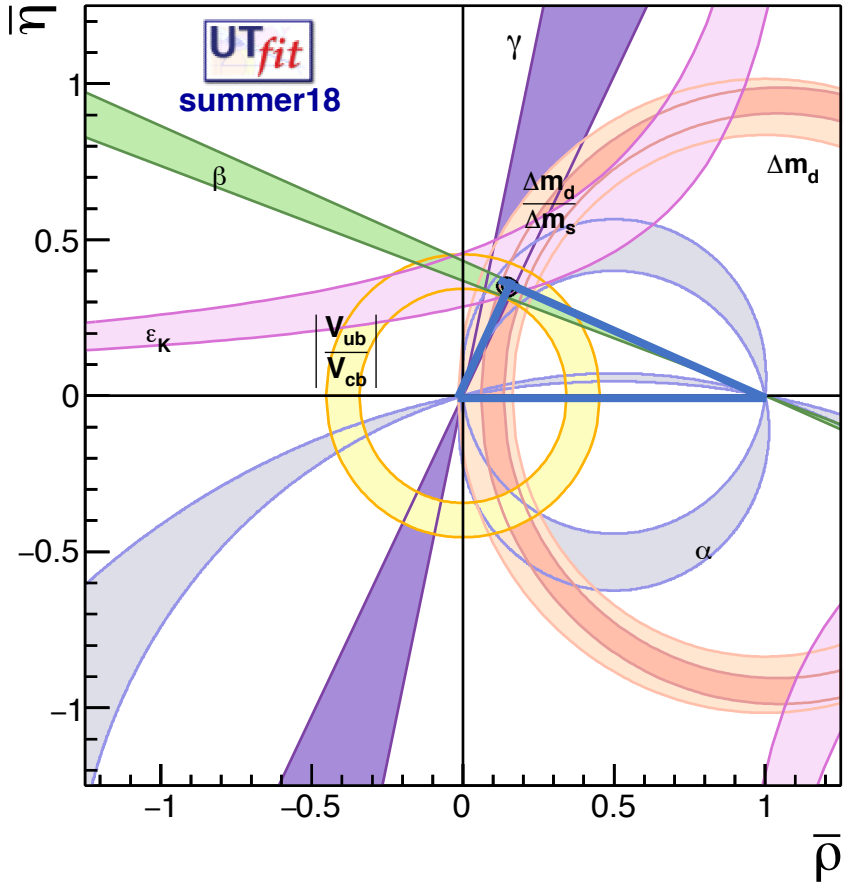
Direct observation



Flavour physics:
shake the Box, listen

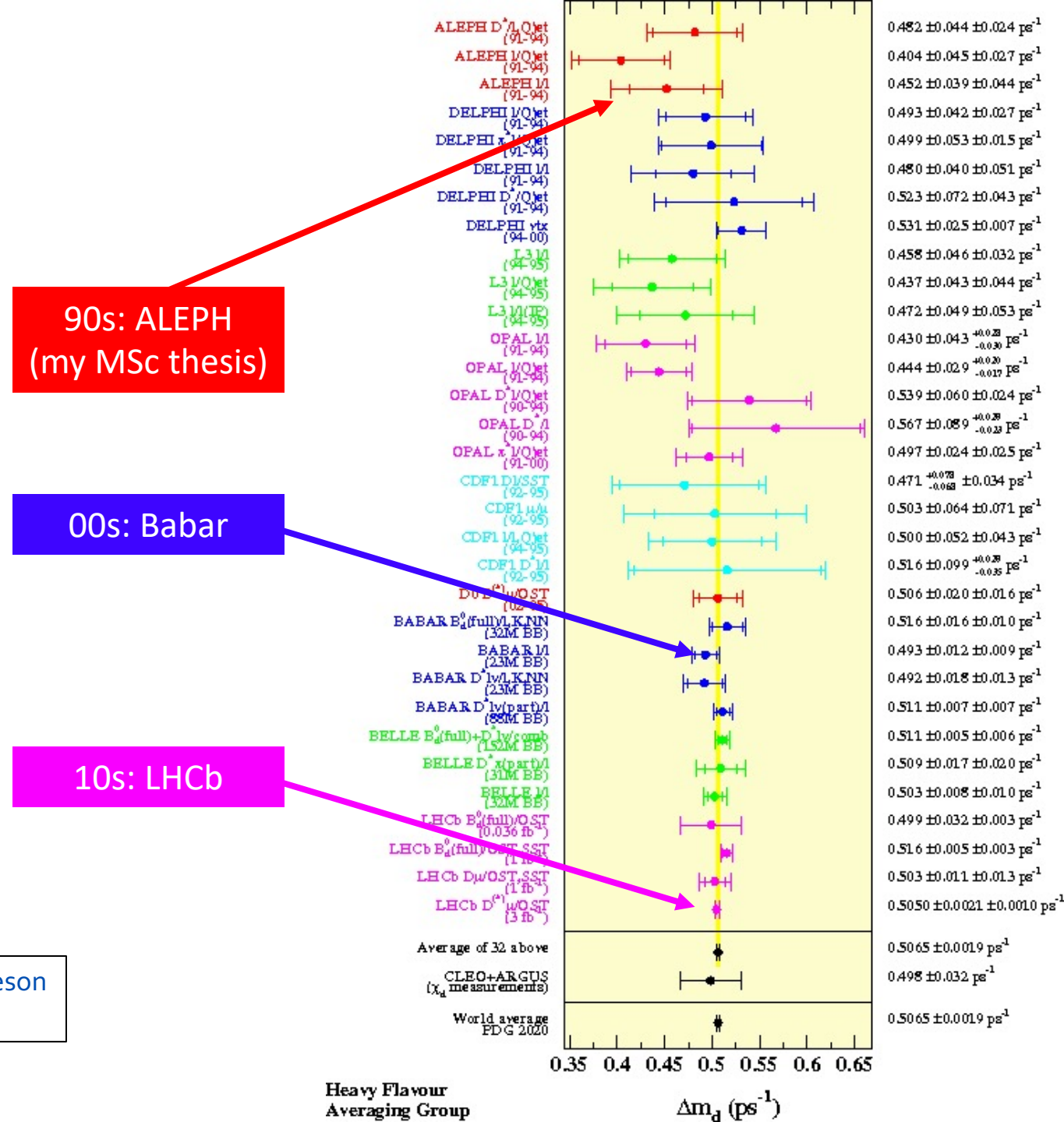
virtual particles

Indirect observation



Precision

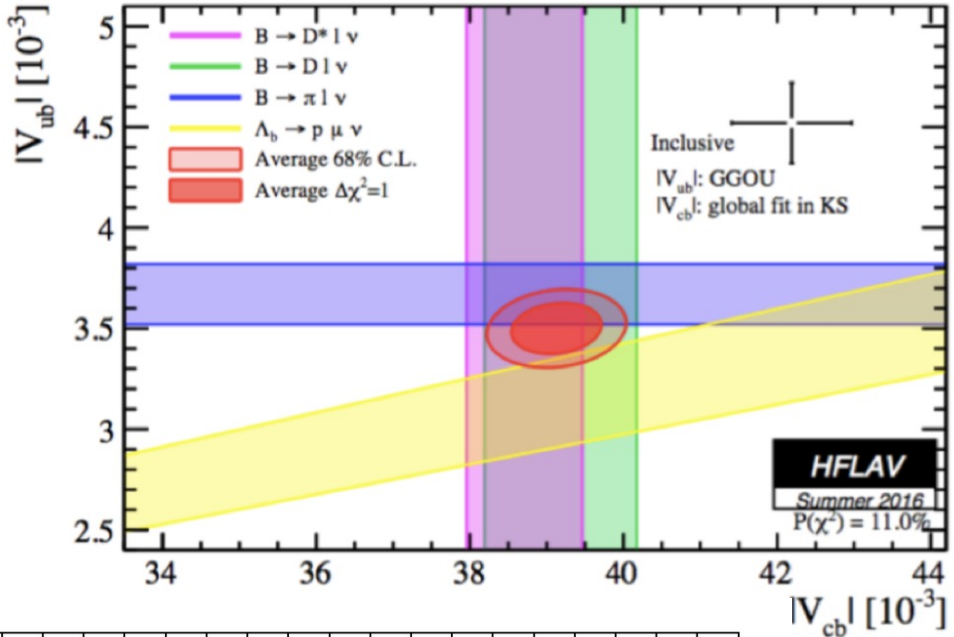
- Measurements of the neutral B meson oscillation frequency
- Performed by using very different technologies
  - “snailmail” grid in the 90s
  - “human” grid in the 00s
  - “orthodox” grid in the 10s

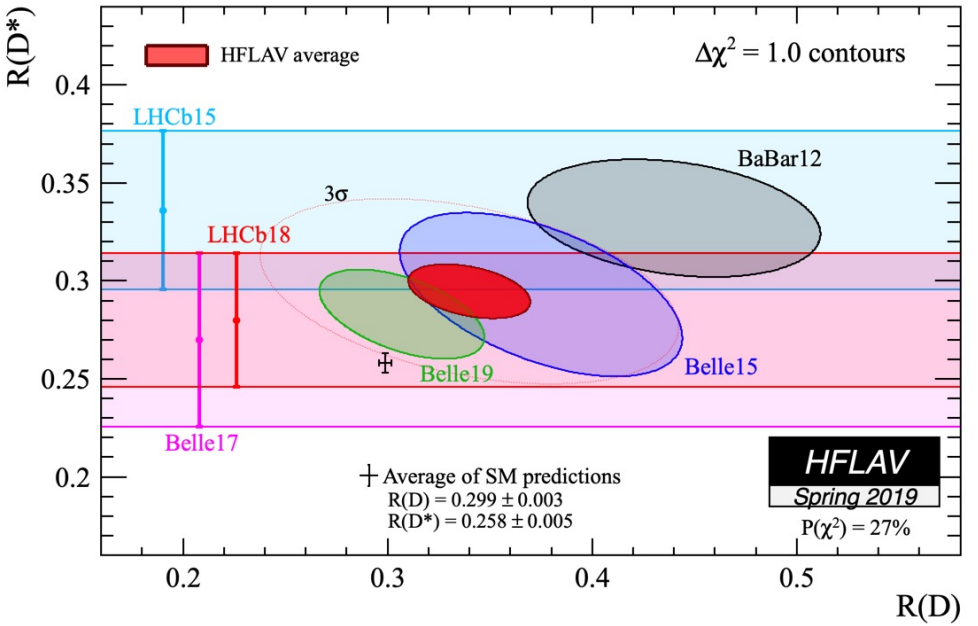


R. Aaij et al (LHCb Collaboration), A precise measurement of the B_0 meson oscillation frequency, Eur.Phys.J.C 76 (2016) 7, 412

Tensions with SM

- Measurements of CKM matrix element $|V_{ub}|$ with Babar data
- Tests of lepton flavour universality with LHCb data
  - From (*impossible*) to (*impossible*)^2





$$R(D^{(*)}) = \mathcal{B}(B \to D^{(*)}\tau\nu) \,/\, \mathcal{B}(B \to D^{(*)}\mu\nu)$$

J. P. Lees et al (Babar Collaboration), Study of B--> X_u ln decays in BB(bar) events tagged by a fully reconstructed B-meson decay and determination of $|V_{ub}|$, Phys.Rev.D 86 (2012), 032004

R. Aaij et al (LHCb Collaboration), Test of Lepton Flavor Universality by the measurement of the B_0 -> D^{*-} t^+ n_t branching fraction using three-prong tau decays, Phys.Rev.D 97 (2018) 7, 072013

R. Aaij et al (LHCb Collaboration), Measurement of the ratio of the B_0 -> D^{*-} t^+ n_t and B_0 -> D^{*-} μ^+ $n_\mu$$ branching fractions using three-prong tau-lepton decays, Phys.Rev.Lett. 120 (2018) 17, 171802

Simulation production: a “human” grid…

IEEE Nuclear Science Symposium – Rome October 2004

https://doi.org/10.1109/NSSMIC.2004.1462659

BABAR SIMULATION PRODUCTION – A MILLENNIUM OF WORK IN UNDER A YEAR

D. A. Smith, *SLAC, Menlo Park, CA, USA*
F. Blanc, *Univ. of Colorado, Boulder, CO, USA*
C. Bozzi, D. Andreotti, *INFN, Ferrara, Italy*

Abstract

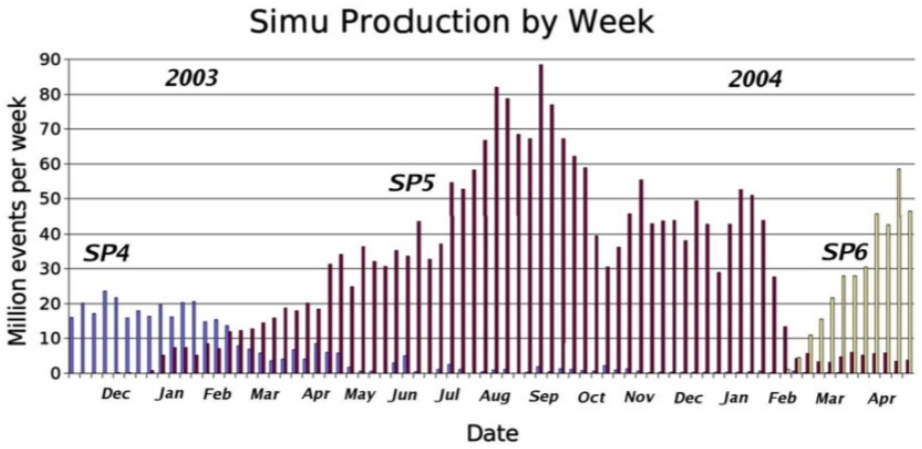
The BaBar experiment requires simulated events beyond the ability of a single computing site to provide. This paper describes the evolution of simulation and job management methods to meet the physics community requirements and how production became distributed to use resources beyond any one computing center. The evolution of BaBar simulation along with the development of the distribution of the computing effort is described.

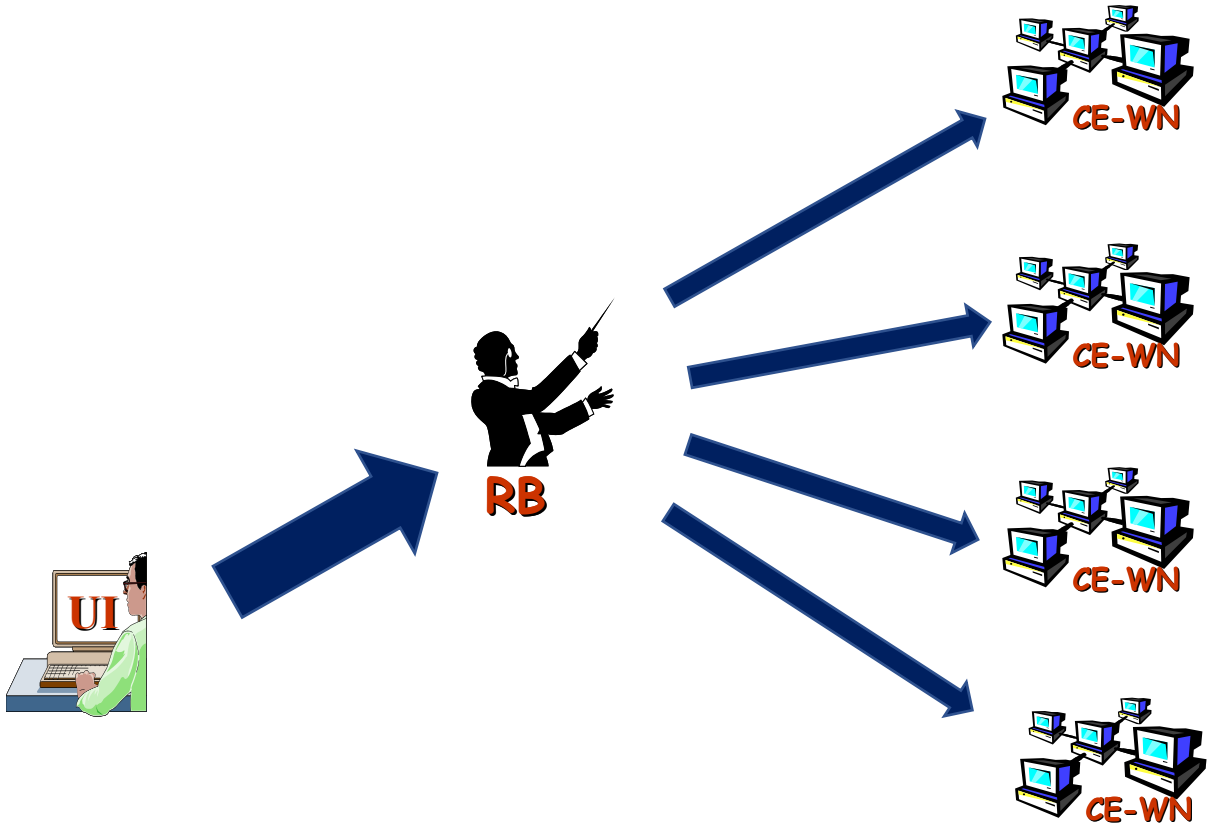
As the computing effort is distributed to more sites there is a need to simplify production so the effort does not multiply with number of production centers. Proper tools are created to be flexible in handling errors and failures that happen in the system and respond accordingly, to reduce failure rates and production effort.

billion events. SP5 in 2003 would produce events for run cycles 1-3, and need 1.6 billion events. For SP6 it was recognised that the new reconstruction code would not produce significantly different events than what was produced in SP5, so SP6 would only produce events for run cycle 4, and SP5 could be used for analysis of run cycles 1-3. This change resulted in SP6 only needing 1 billion events to match the request.

This resulted in the fact that SP5 would be the largest requested production cycle in BaBar, and would need a greater amount of distribution of the computing effort to get done on time. This effort was performed and finished earlier this year, and I will concentrate on this effort as a description of a complete large scale computing effort.

RESOURCES NEEDED





Simulation production: a “human” grid…

IEEE Nuclear Science Symposium – Rome October 2004

https://doi.org/10.1109/NSSMIC.2004.1462659

BABAR SIMULATION PRODUCTION – A MILLENNIUM OF WORK IN UNDER A YEAR

D. A. Smith, *SLAC, Menlo Park, CA, USA*
F. Blanc, *Univ. of Colorado, Boulder, CO, USA*
C. Bozzi, D. Andreotti, *INFN, Ferrara, Italy*

Abstract

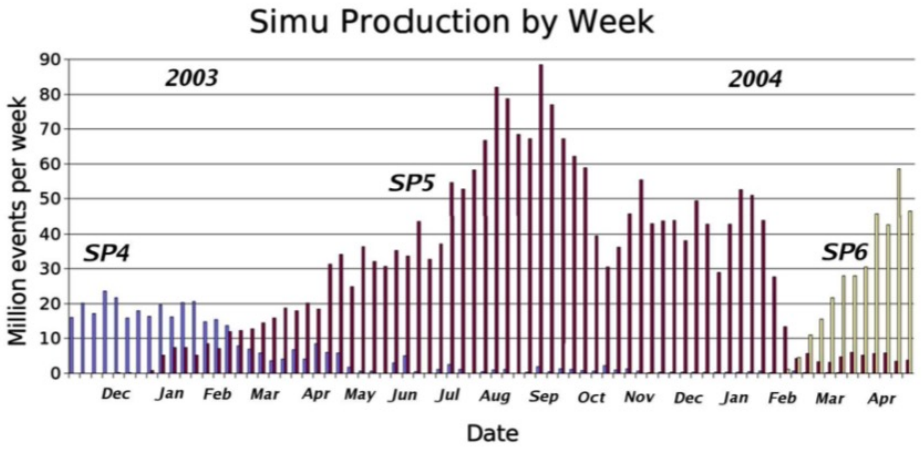
The BaBar experiment requires simulated events beyond the ability of a single computing site to provide. This paper describes the evolution of simulation and job management methods to meet the physics community requirements and how production became distributed to use resources beyond any one computing center. The evolution of BaBar simulation along with the development of the distribution of the computing effort is described.

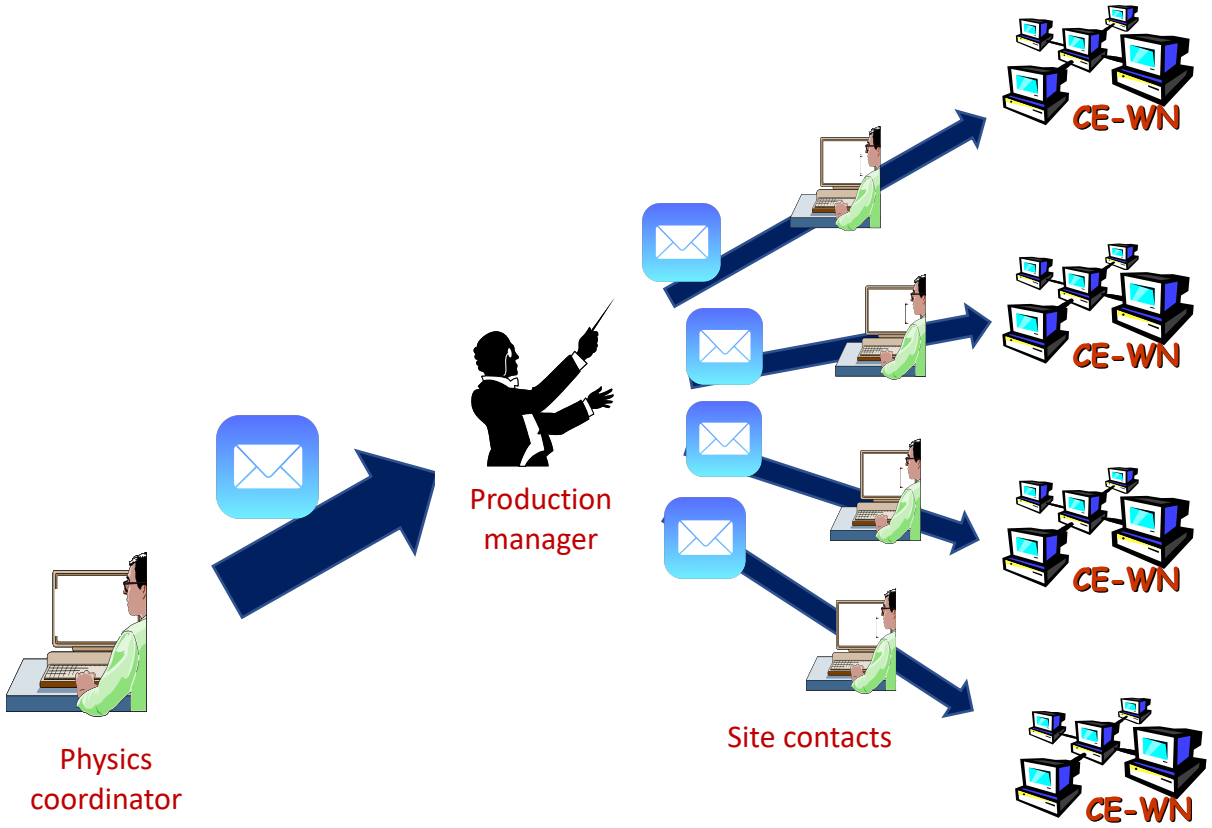
As the computing effort is distributed to more sites there is a need to simplify production so the effort does not multiply with number of production centers. Proper tools are created to be flexible in handling errors and failures that happen in the system and respond accordingly, to reduce failure rates and production effort.

billion events. SP5 in 2003 would produce events for run cycles 1-3, and need 1.6 billion events. For SP6 it was recognised that the new reconstruction code would not produce significantly different events than what was produced in SP5, so SP6 would only produce events for run cycle 4, and SP5 could be used for analysis of run cycles 1-3. This change resulted in SP6 only needing 1 billion events to match the request.

This resulted in the fact that SP5 would be the largest requested production cycle in BaBar, and would need a greater amount of distribution of the computing effort to get done on time. This effort was performed and finished earlier this year, and I will concentrate on this effort as a description of a complete large scale computing effort.

RESOURCES NEEDED





Simulation production: a “human” grid…

IEEE Nuclear Science Symposium – Rome October 2004

https://doi.org/10.1109/NSSMIC.2004.1462659

BABAR SIMULATION PRODUCTION – A MILLENNIUM OF WORK IN UNDER A YEAR

D. A. Smith, *SLAC, Menlo Park, CA, USA*
F. Blanc, *Univ. of Colorado, Boulder, CO, USA*
C. Bozzi, D. Andreotti, *INFN, Ferrara, Italy*

Abstract

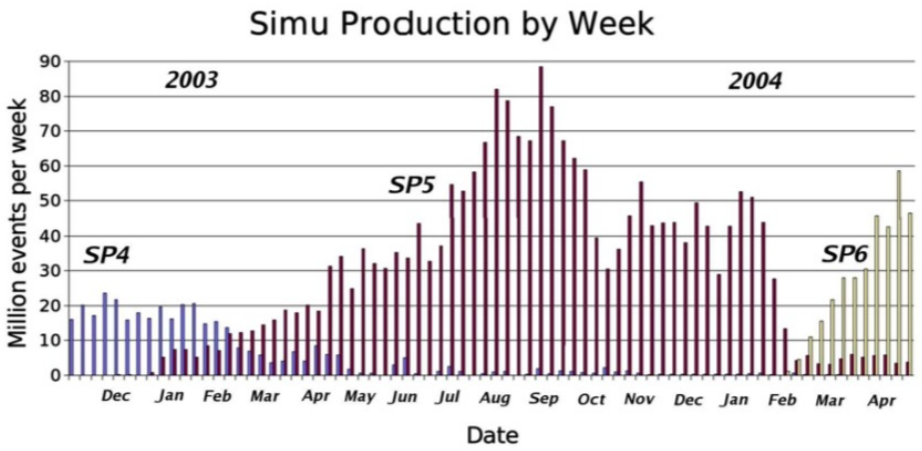
The BaBar experiment requires simulated events beyond the ability of a single computing site to provide. This paper describes the evolution of simulation and job management methods to meet the physics community requirements and how production became distributed to use resources beyond any one computing center. The evolution of BaBar simulation along with the development of the distribution of the computing effort is described.

As the computing effort is distributed to more sites there is a need to simplify production so the effort does not multiply with number of production centers. Proper tools are created to be flexible in handling errors and failures that happen in the system and respond accordingly, to reduce failure rates and production effort.

billion events. SP5 in 2003 would produce events for run cycles 1-3, and need 1.6 billion events. For SP6 it was recognised that the new reconstruction code would not produce significantly different events than what was produced in SP5, so SP6 would only produce events for run cycle 4, and SP5 could be used for analysis of run cycles 1-3. This change resulted in SP6 only needing 1 billion events to match the request.

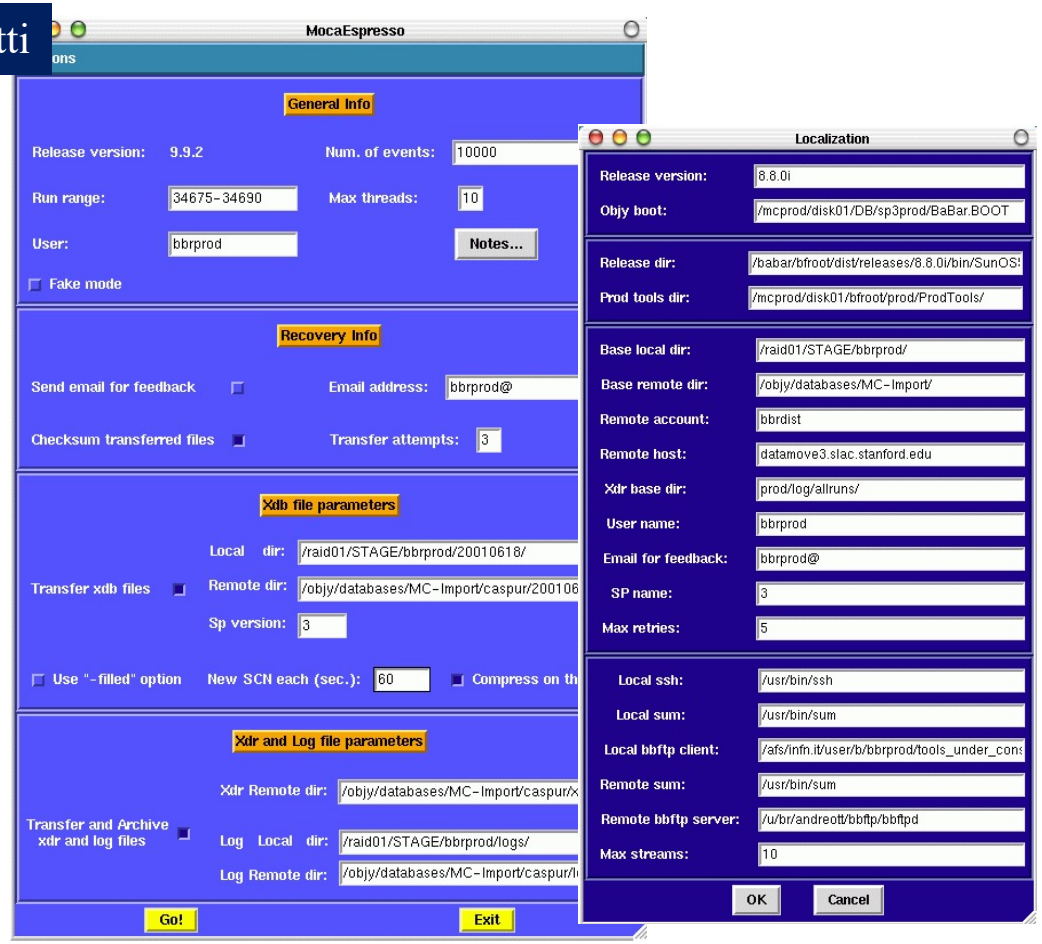
This resulted in the fact that SP5 would be the largest requested production cycle in BaBar, and would need a greater amount of distribution of the computing effort to get done on time. This effort was performed and finished earlier this year, and I will concentrate on this effort as a description of a complete large scale computing effort.

RESOURCES NEEDED



SLAC **MocaEspresso** BABAR

D. Andreotti



…but also a real (prototype) grid

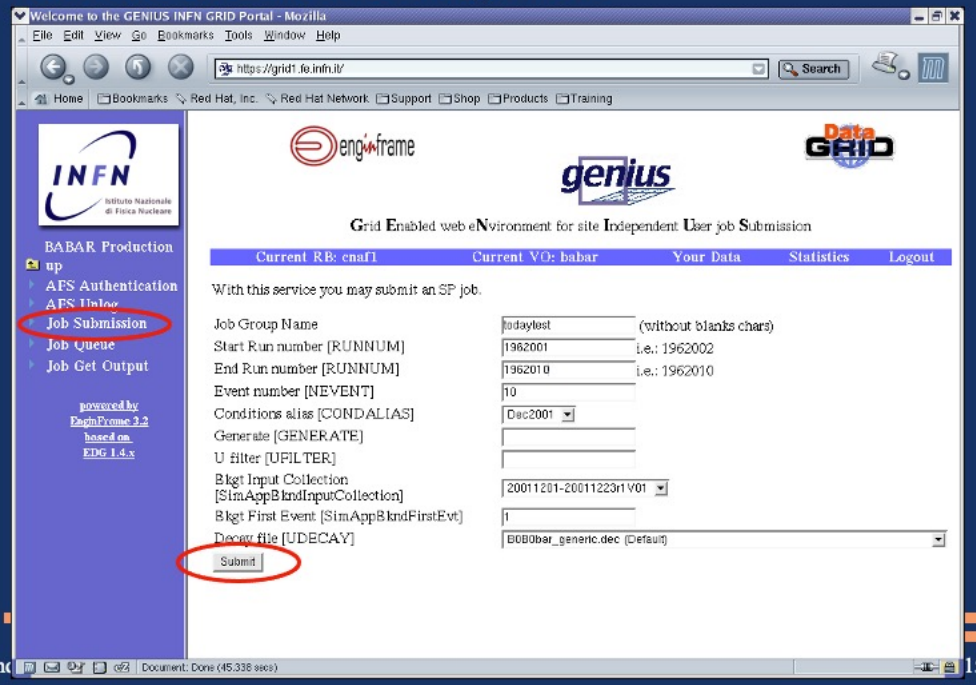
Using Grid for the Babar Experiment

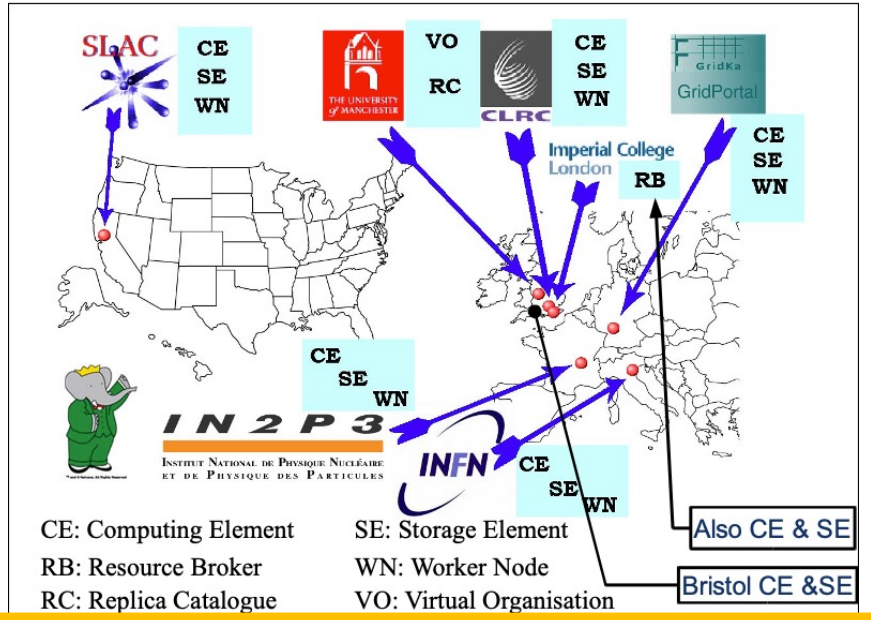
Concezio Bozzi
INFN Ferrara
Italy

uction on LCG

zio Bozzi

Create and submit jobs

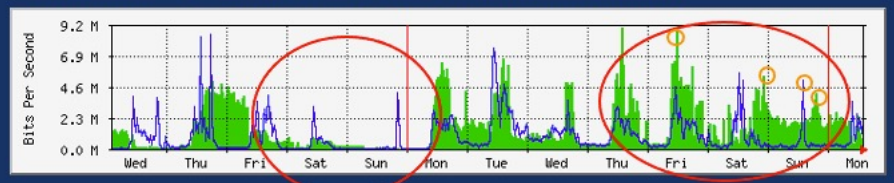




First results

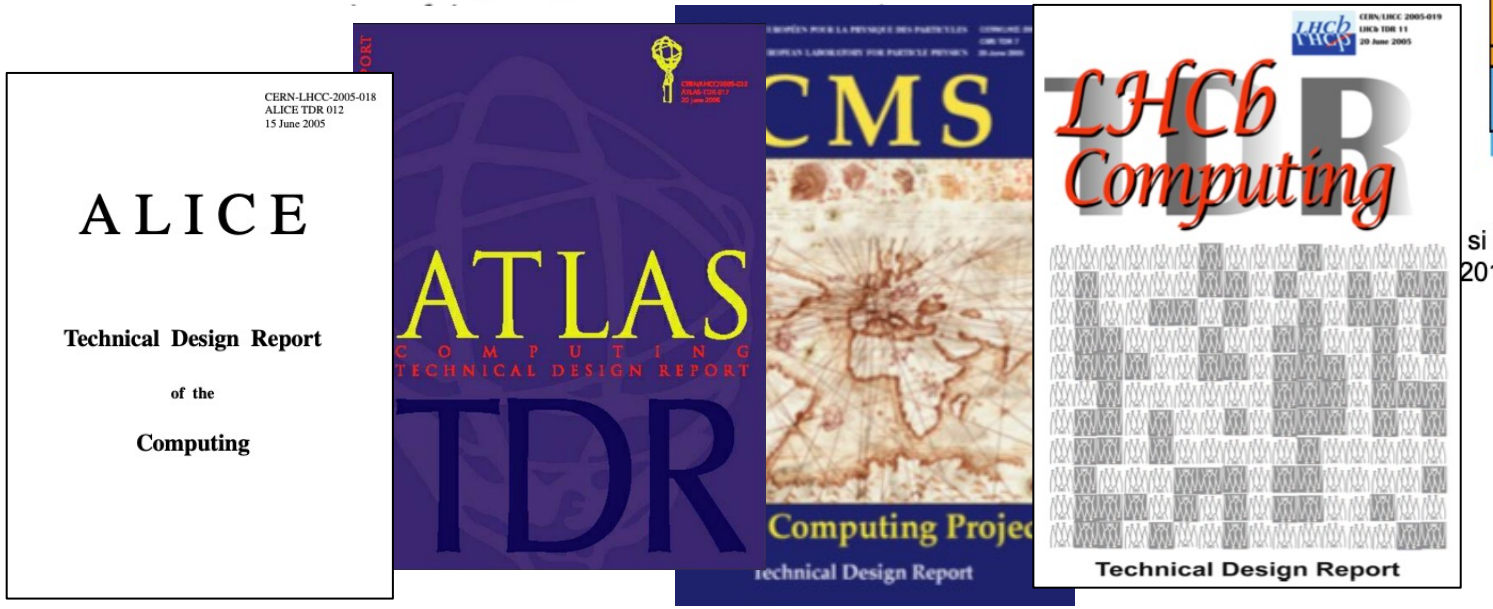
- Some hundreds of jobs submitted to all italian sites (10) in different times of the day for ~1 week.
- Jobs submitted to Ferrara had ~95% of success.
- Jobs submitted to other italian resources had ~60% of success.
- Main failure causes have been
  - Problems of the Resource Broker (due to Globus services: ~7%)
  - Problems in remote access to our Objy DB (due to simultaneous accesses and network overload: ~33%)
  - No showstoppers

Network load during the tests (Ferrara bandwidth: 12Mbit/s)



Reviewing computing models and resources at CNAF, INFN, CERN, and DoE

2007 – 2012	Italian representative, appointed by the INFN Executive Board, in the CERN WLCG Computing Resources Scrutiny Group. Referee of ATLAS and CMS computing
2010	member of the DoE committee reviewing computing and operations of the US LHC collaborations, Argonne Nat. Lab. IL (US)
2005 – 2011	member of the INFN committee reviewing scientific computing of the LHC experiments in Italy, the Italian Tier1 at CNAF, the INFN Grid and other EU-funded projects. Committee chair in 2007-2011
2005	member of CERN committee in charge of the review of the Computing Technical Design Reports of the LHC experiments

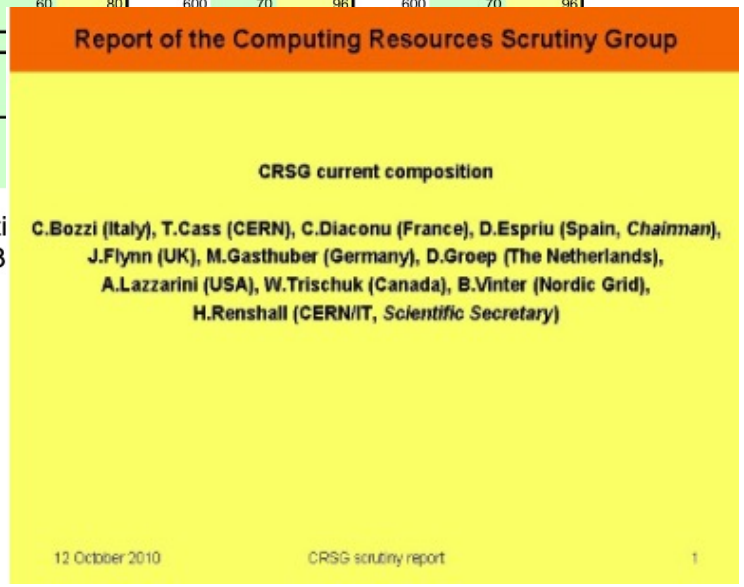


Il piano per il Tier1

CNAF PLAN SEPTEMBER 2011

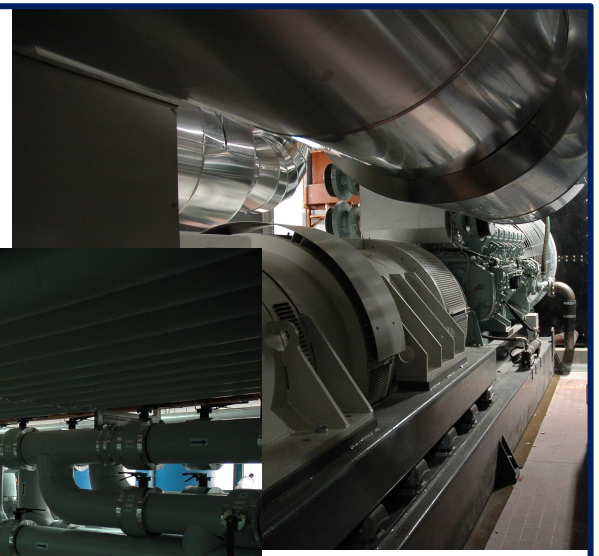
Experiment	2011 CPU HS06	2011 DISK TB-N	2011 TAPE TB	2012 CPU HS06	2012 DISK TB-N	2012 TAPE TB	2013 CPU HS06	2013 DISK TB-N	2013 TAPE TB
ALICE	22200	1501	2400	25890	1749	3098	29830	1653	5377
ATLAS	22600	2480	3000	25900	2700	3600	27300	3000	4000
CMS	18300	2400	6500	18850	2860	6630	18850	3510	7670
LHCB	9750	525	520	16950	1425	930	16500	1665	1200
Total LHC TIER1	72850	6906	12420	87590	8734	14258	92480	9828	18247
BaBar	2360	350	0	2360	350	0	2360	350	0
SuperB (dal 2011)	2500	100	0	2500	200	0	2500	200	0
CDF	7000	300	15	8000	467	15	8000	467	15
KLOE				0	33	625	0	33	625
LHCB TIER2	5400	0	0	7200	0	0	7200	0	0
TOTALE GRUPPO I	17260	750	15	20060	1050	640	20060	1050	640
AMS2	2457	143	50	5400	384	220	5400	384	220
ARGO	800	160	752	1200	224	986	1200	224	986
AUGER	1200	110	0	1600	160	0	1600	160	0
FERMI/GLAST	1400	60	40	1400	60	40	1400	60	40
MAGIC	450	30	50	500	45	70	500	45	70
PAMELA	600	60	80	600	70	96	600	70	96
Virgo	7500								
TOTALE GRUPPO II	14407								
All experiments	104517								
All w/ overlap factor	87098								
CNAF TOTAL (PLAN)	87098								
Effective overlap									
CNAF to be procured	21171								
Fresh resources									

si posticipa al 2013 la sosti
2012 è in backup. Nel 2013



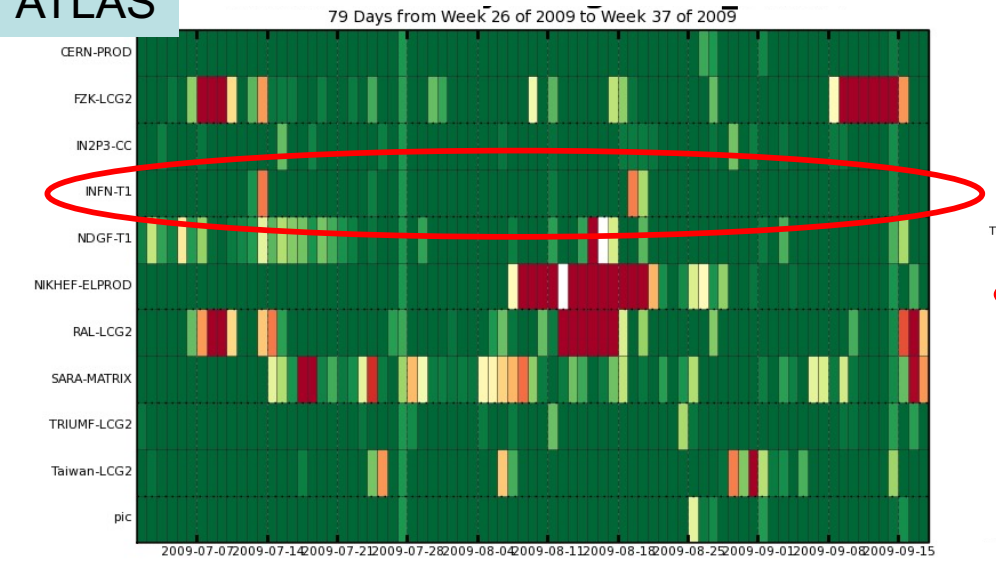
Buone notizie

- Lavori infrastrutturali completati
- Inaugurazione ufficiale a inizio giugno
  - http://www.cnaf.infn.it/main/index.php/Chi_Siamo/Video_Inaugurazione_Tier_One
- Il centro funziona bene e fornisce adeguato supporto agli esperimenti

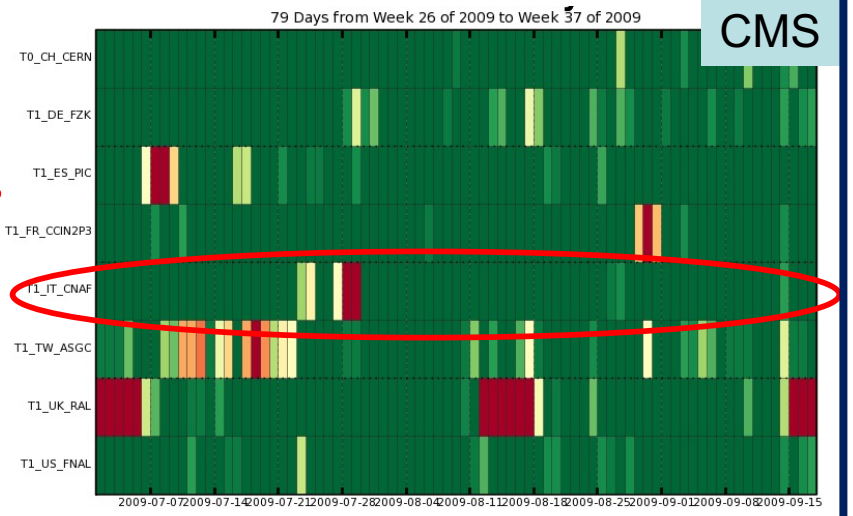


Affidabilità del Tier1 CNAF

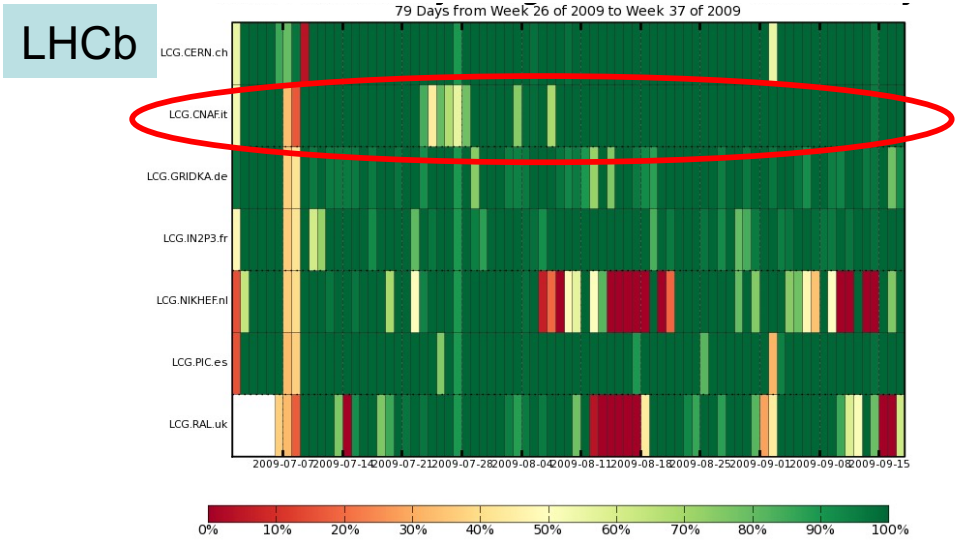
ATLAS



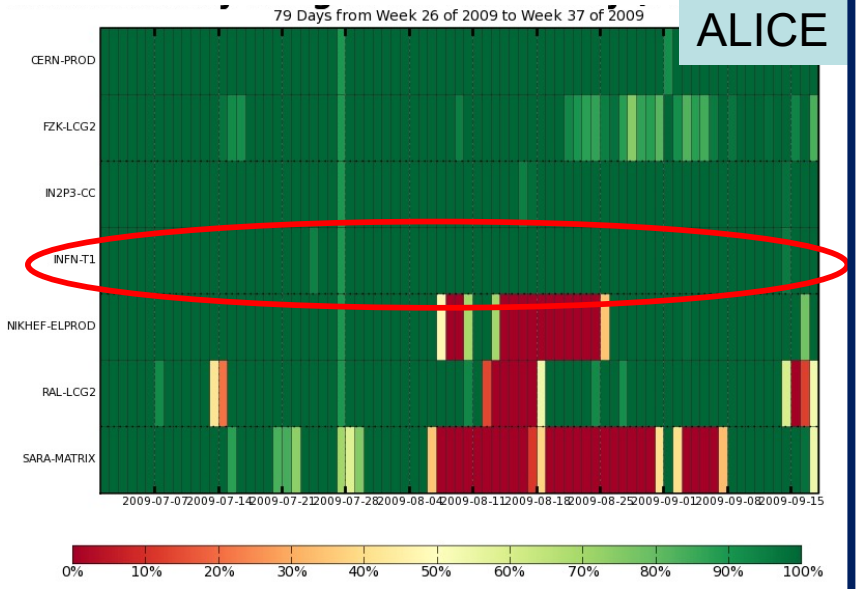
CMS



LHCb



ALICE



Managing LHCb resources since 2013

LHCb Computing Resources: 2022 requests

LHCb Public Note

Issue: 0
Revision: 0

Reference: LHCb-PUB-2021-002
Created: 1st February 2021
Last modified: 15th February 2021

Prepared By: LHCb Computing Project
C. Bozzi/Editor

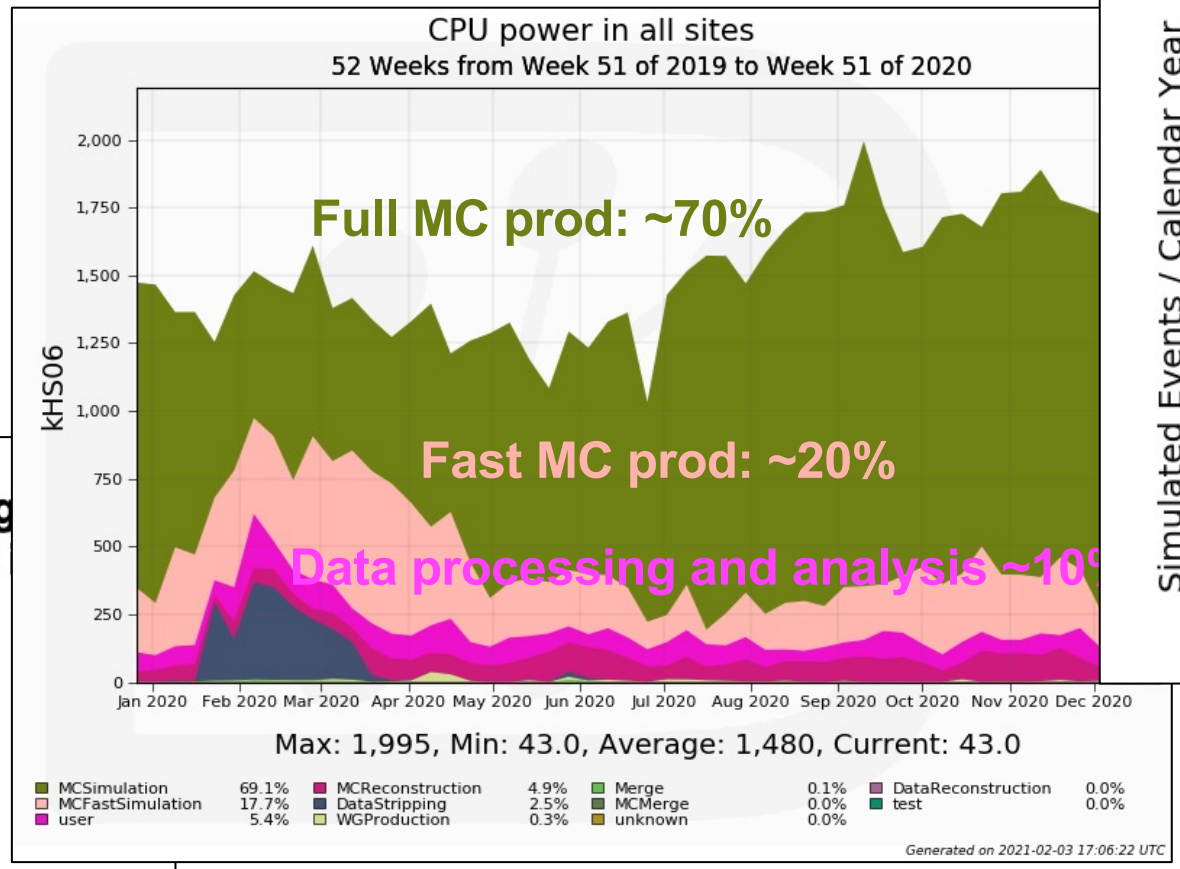
LHCb		LHCb-PUB-2020-001 2021 Request	LHCb-PUB-2020-001 2021 2021 req. / 2020 CRSG	LHCb-PUB-2020-005 2022 Request	LHCb-PUB-2020-005 2022 2022 req. / 2021 CRSG	THIS DOCUMENT 2022 Request	THIS DOCUMENT 2022 2022 req. / 2021 CRSG
WLCG CPU	Tier-0	175	179%	235	134%	189	108%
	Tier-1	574	195%	770	134%	622	108%
	Tier-2	321	166%	430	134%	345	107%
	HLT	50	500%	50	100%	50	100%
	Sum	1120	188%	1485	133%	1206	108%
Others		50	n/a	50	n/a	50	100%
Total		**1,170**	**193%**	**1,535**	**131%**	**1,256**	**107%**
Disk	Tier-0	18.8	109%	33.3	177%	26.5	141%
	Tier-1	37.6	119%	66.6	177%	52.9	141%
	Tier-2	7.2	168%	12.8	177%	10.2	141%
	Total	**63.7**	**120%**	**112.7**	**177%**	**89.6**	**141%**
Tape	Tier-0	44	121%	81	184%	81	184%
	Tier-1	76	135%	139	184%	139	184%
	Total	**119.7**	**130%**	**219.9**	**184%**	**219.9**	**184%**

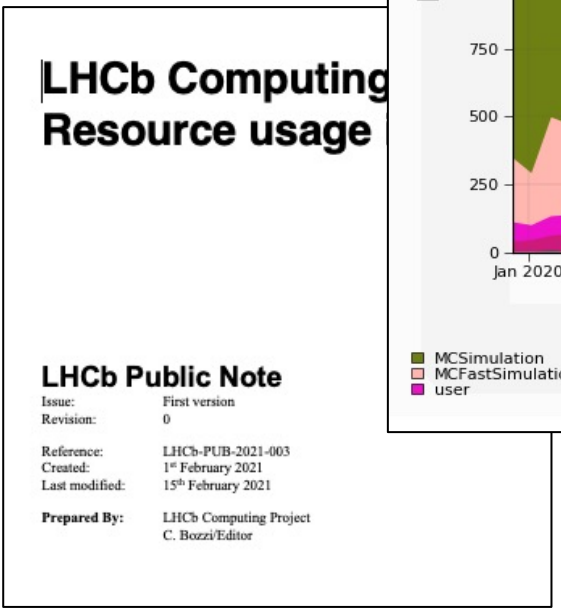
CPU Work in WLCG year (kHS06.years)	2021	2022 LHCB-PUB-2020-005	2022 THIS DOCUMENT
First pass sprucing	70	160	80
End-of-year sprucing	70	160	80
Simulation	760	870	870
Core and distributed computing infrastructure	10	10	10
User Analysis and working group productions	260	335	220
Total Work (kHS06.years)	**1170**	**1535**	**1260**
LHCb-TDR-018 (2021 pledge)	**860 (934)**	**1580**	**1580**

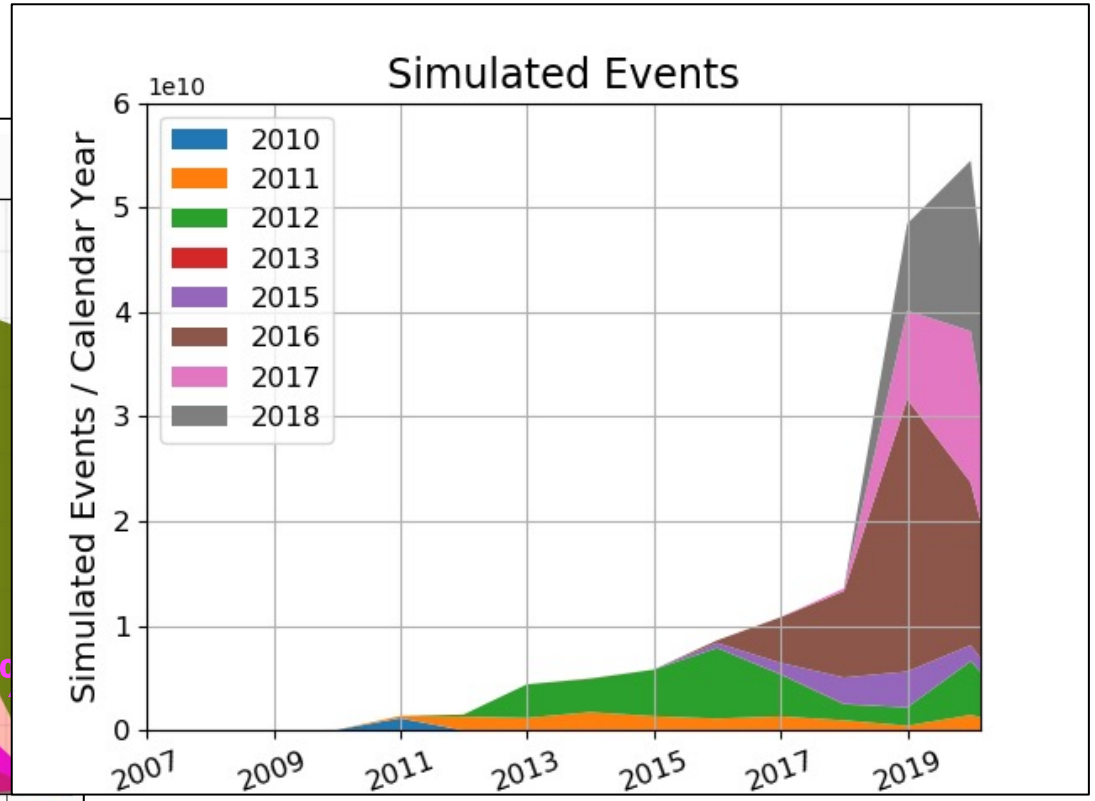
Disk storage usage forecast (PB)		2021	2022 LHCb-PUB-2020-005		2022 This document	
Real data	Run1+Run2 pp data	37.8	17.9	73.7	10.2	65.9
	Run1+Run2 PbPb + SMOG					
	Run3: FULL		13.7		13.7	
	Run3: TURBO		30.3		30.3	
	Run3: TURCAL		3.7		3.7	
	Run3: Minimum bias		2.4		2.4	
	Run3: PbPb + SMOG2		5.6		5.6	
Simulated data	Run1+Run2 Simulated Data	10.0	8.7	10.9	8.7	10.0
	Run3 simulated data		2.2		2.2	
Other	User data	15.9	8.5	28.2	1.8	12.8
	Buffers		19.7		11.0	
Total		**63.7**	**112.7**		**89.6**	
LHCb-TDR-018 (2021 pledge)		**66.0 (58.7)**	**111.0**		**111.0**	

Tape storage usage forecast (PB)		2021	2022	
Run1 + Run2	RAW data (pp+HI+fixed target)	81.1	38.4	82.1
	RDST data (pp+HI+fixed target)		13.7	
	ARCHIVE		30.0	
Run3	pp data (FULL+TURBO+TURCAL)	38.6	120.1	137.8
	minimum bias / no-bias		0.6	
	Heavy Ion Data + fixed target		5.6	
	ARCHIVE (data+MC)		11.5	
Total		**119.7**	**219.9**	
LHCb-TDR-018 (2021 pledge)		**142.0 (108.7)**	**243.0**	

Managing LHCb resources since 2013

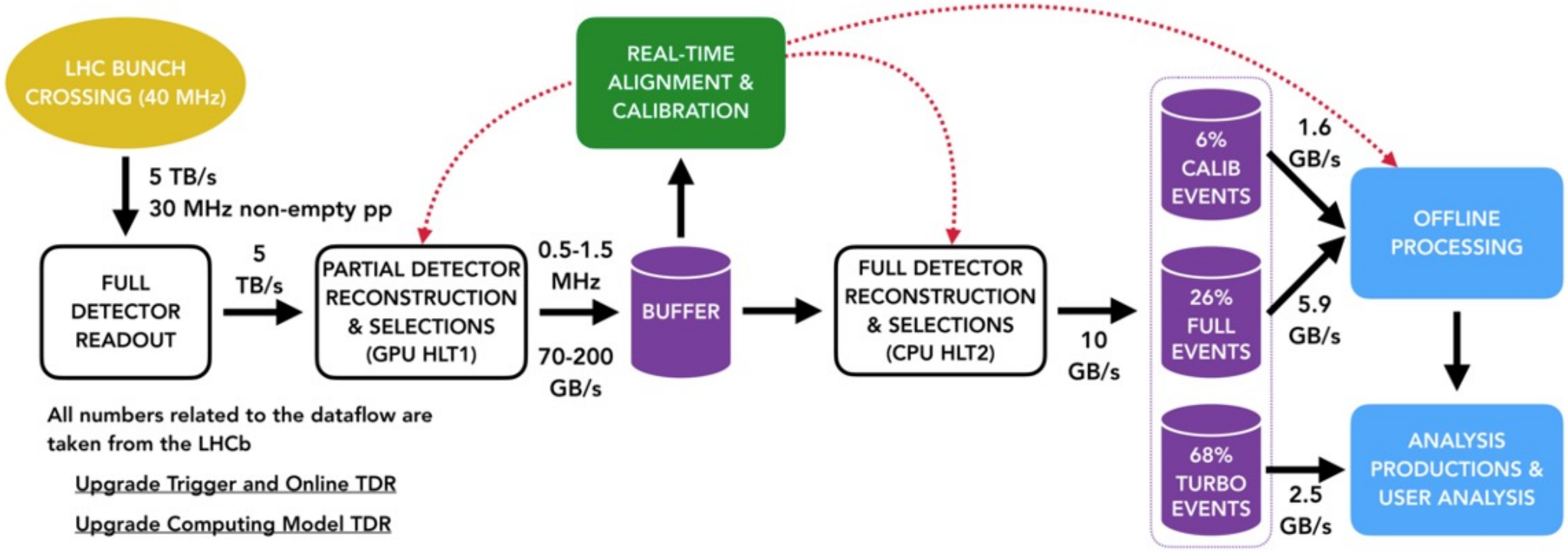






Year	Simulated events (10^9)	Stored events (10^9)	Ratio	CPU work kHS06.y	CPU per event kHS06.s	LFS TB
2017	10.3	4.2	40.3%	817	2.50	640
2018	12.0	3.0	25.3%	1009	2.65	550
2019	45.0	6.9	15.2%	1290	0.90	1110
2020	53.0	16.8	31.7%	1357	0.81	2010

Preparing SW and computing for the LHCb Upgrade



All numbers related to the dataflow are taken from the LHCb

Upgrade Trigger and Online TDR

Upgrade Computing Model TDR

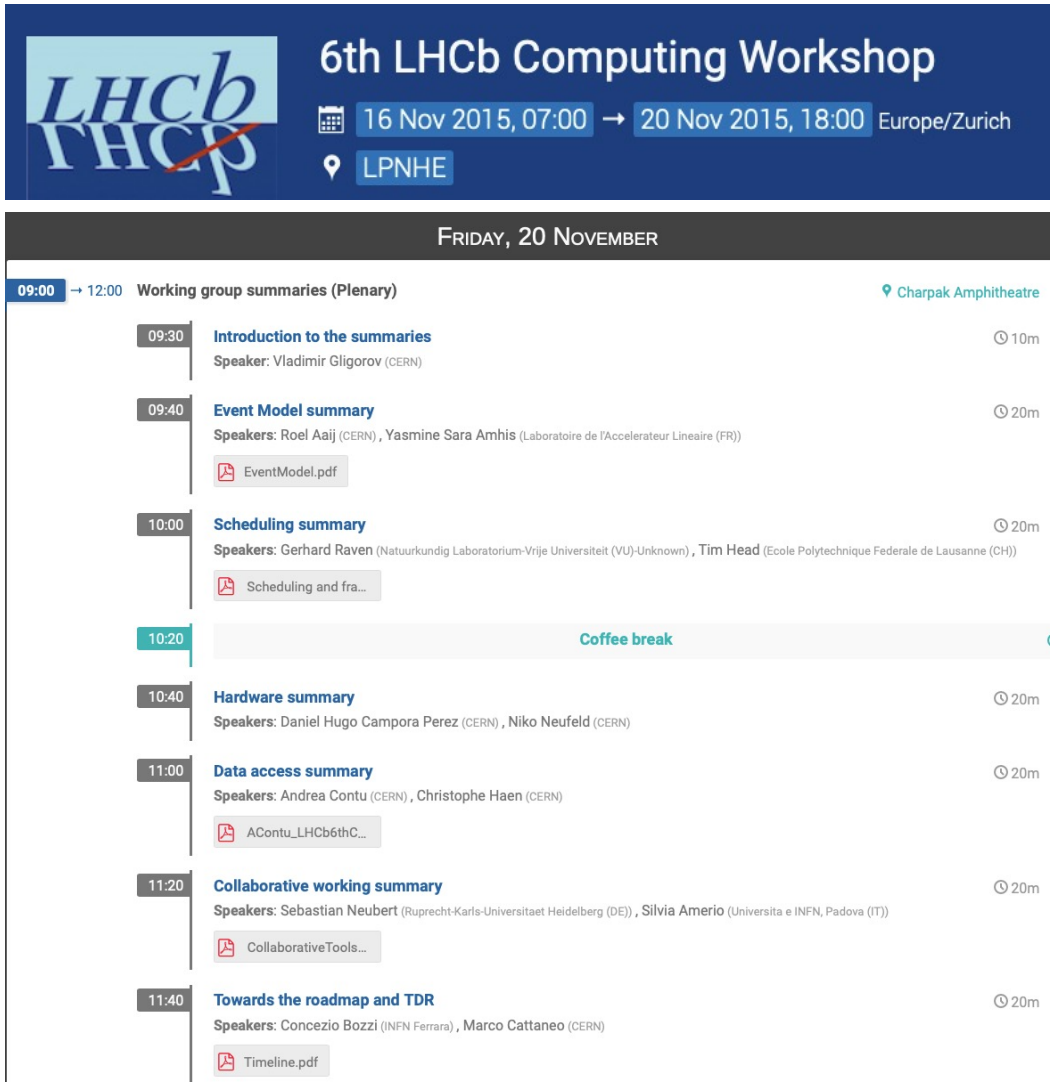
30x increase in throughput from the upgraded detector, Without corresponding jump in offline computing resources

Preparing SW and computing for the LHCb Upgrade

- Software performance: much to gain!
  - Better utilization of current multi-processor CPU architectures
  - Enable code vectorization
  - Modernize data structures
  - Reduce memory usage
  - Optimize cache performance
  - Remove dead code
  - Replace outdated technologies
  - Enable algorithmic optimization

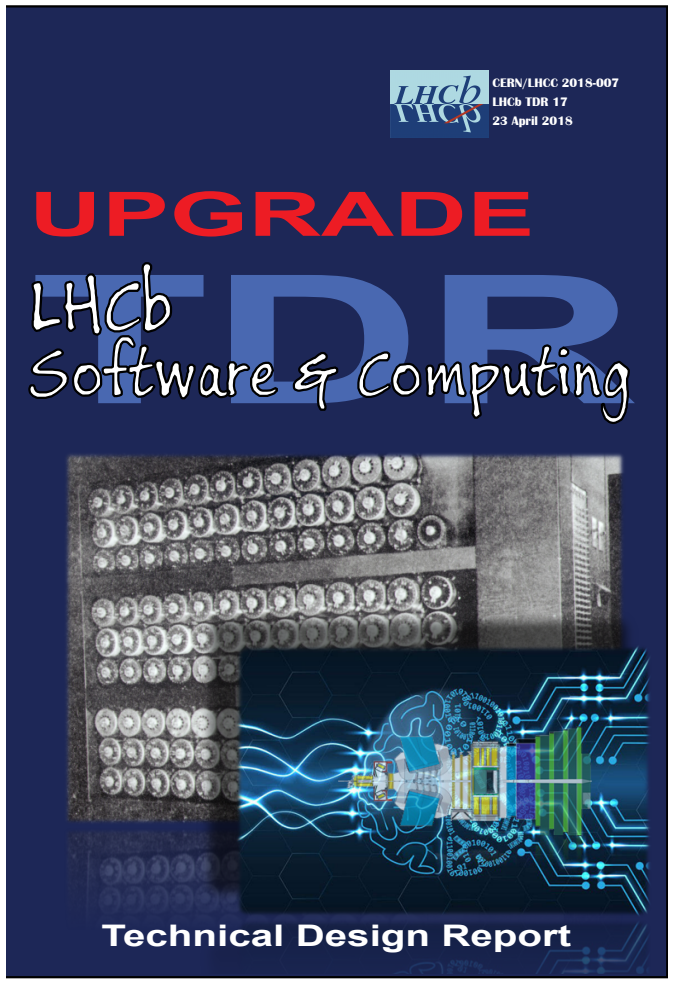
S. Roiser and C. Bozzi, The LHCb Software and Computing Upgrade towards LHC Run 3, J.Phys.Conf.Ser. 1085 (2018) 3, 032049

C. Bozzi and S. Roiser, The LHCb software and computing upgrade for Run 3: opportunities and challenges, J.Phys.Conf.Ser. 898 (2017) 11, 112002



Preparing SW and computing for the LHCb Upgrade

- Software performance: much to gain!
  - Better utilization of current multi-processor CPU architectures
  - Enable code vectorization
  - Modernize data structures
  - Reduce memory usage
  - Optimize cache performance
  - Remove dead code
  - Replace outdated technologies
  - Enable algorithmic optimization



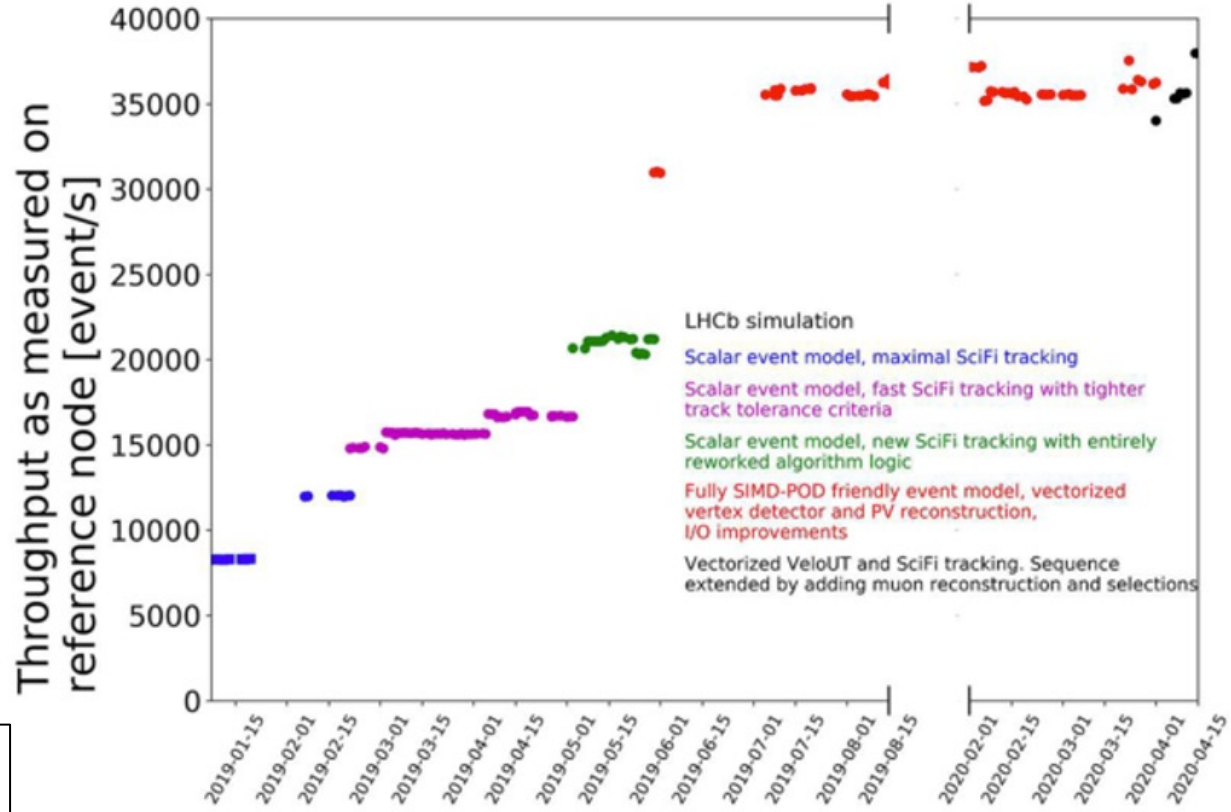
LHCb-TDR-017
Editor: C. Bozzi

S. Roiser and C. Bozzi, The LHCb Software and Computing Upgrade towards LHC Run 3, J.Phys.Conf.Ser. 1085 (2018) 3, 032049

C. Bozzi and S. Roiser, The LHCb software and computing upgrade for Run 3: opportunities and challenges, J.Phys.Conf.Ser. 898 (2017) 11, 112002

Preparing SW and computing for the LHCb Upgrade

- Software performance: much to gain!
  - Better utilization of current multi-processor CPU architectures
  - Enable code vectorization
  - Modernize data structures
  - Reduce memory usage
  - Optimize cache performance
  - Remove dead code
  - Replace outdated technologies
  - Enable algorithmic optimization



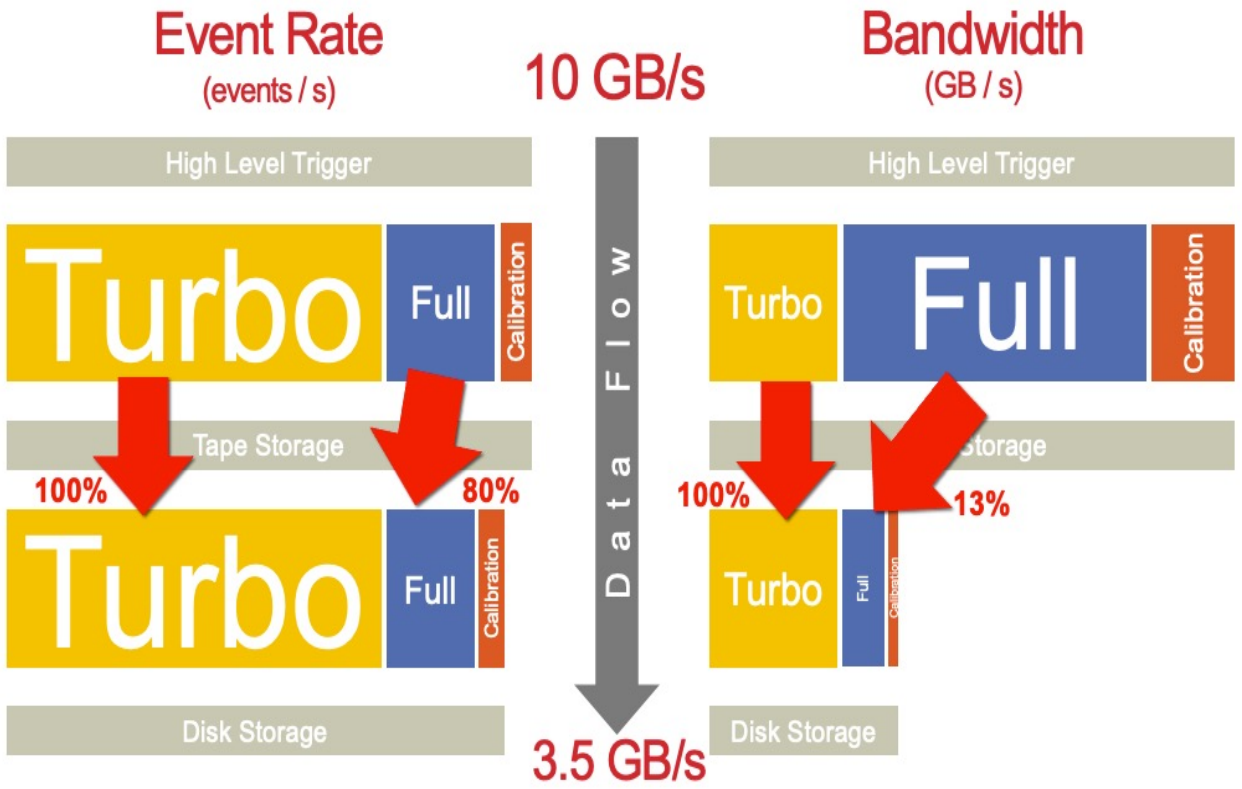
S. Roiser and C. Bozzi, The LHCb Software and Computing Upgrade towards LHC Run 3, J.Phys.Conf.Ser. 1085 (2018) 3, 032049

C. Bozzi and S. Roiser, The LHCb software and computing upgrade for Run 3: opportunities and challenges, J.Phys.Conf.Ser. 898 (2017) 11, 112002

Defining the computing model for LHCb Upgrade

- Concepts developed and implemented during Run 2 to become predominant
  - Split HLT → real-time alignment and calibration
  - TURBO stream for majority of physics program → RAW events discarded
  - FULL and CALIBRATION streams to insure flexibility → filter & slim offline
- Offline CPU computing needs dominated by simulation
  - Number of events to be simulated scales with luminosity
  - Simulation time per event scales with pileup

  → CPU simulation explodes → need for faster simulations
- Offline storage driven by trigger output bandwidth
  - MC saved in μDST, so little impact on storage



C. Bozzi and S. Roiser, Towards a computing model for the LHCb Upgrade, EPJ Web Conf. 214 (2019), 03045

C. Biscarat et al, System performance and cost modelling in LHC computing, EPJ Web Conf. 214 (2019), 03019

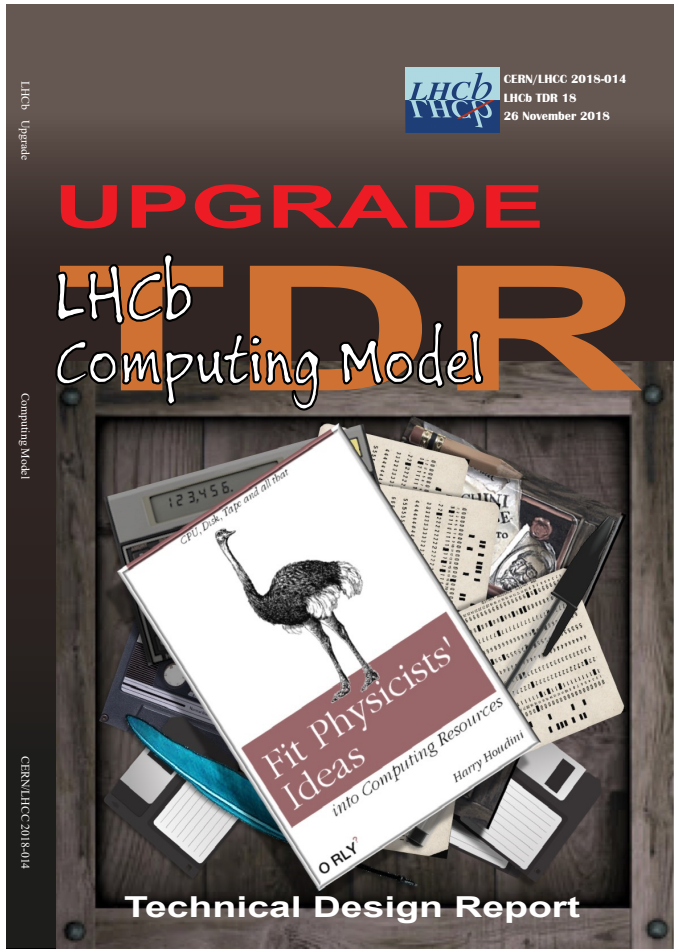
Defining the computing model for LHCb Upgrade

- Concepts developed and implemented during Run 2 to become predominant
  - Split HLT → real-time alignment and calibration
  - TURBO stream for majority of physics program → RAW events discarded
  - FULL and CALIBRATION streams to insure flexibility → filter & slim offline
- Offline CPU computing needs dominated by simulation
  - Number of events to be simulated scales with luminosity
  - Simulation time per event scales with pileup

  → CPU simulation explodes → need for faster simulations
- Offline storage driven by trigger output bandwidth
  - MC saved in μDST, so little impact on storage

C. Bozzi and S. Roiser, Towards a computing model for the LHCb Upgrade, EPJ Web Conf. 214 (2019), 03045

C. Biscarat et al, System performance and cost modelling in LHC computing, EPJ Web Conf. 214 (2019), 03019



LHCb-TDR-018

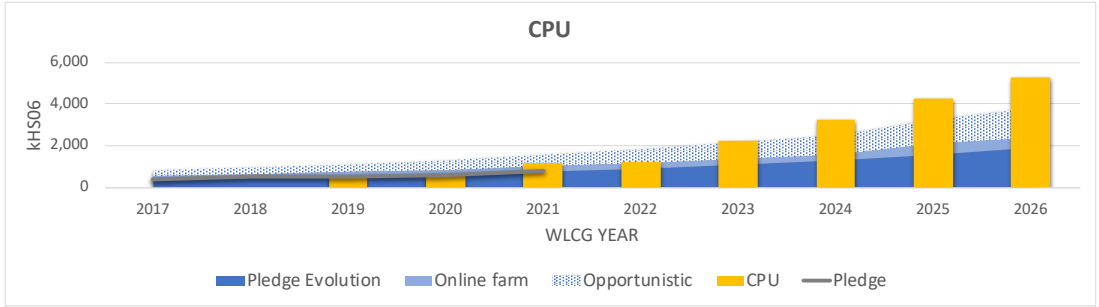
Editors: C. Bozzi, S. Roiser

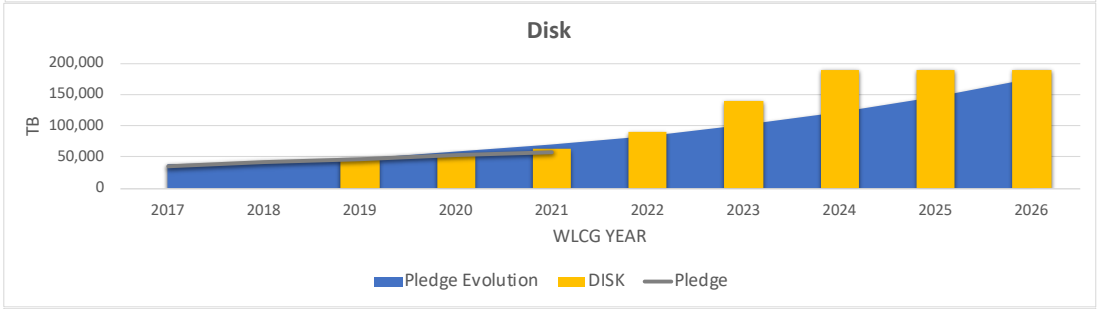
Preparing SW and computing for the LHCb Upgrade

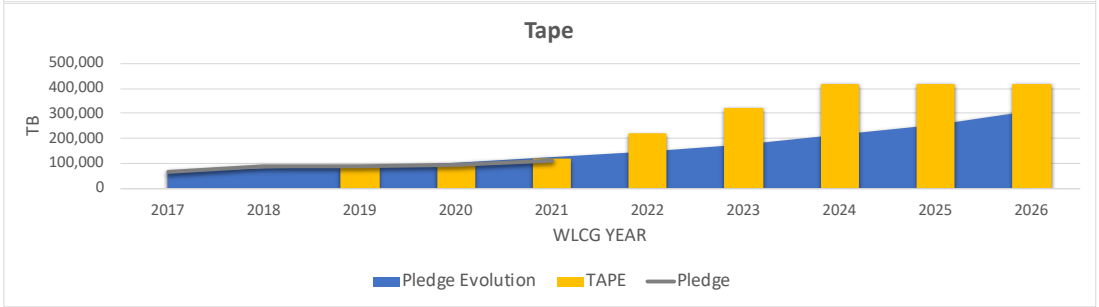
Run3 Computing model

- Concepts developed and implemented during Run 2 to become predominant
  - Split HLT → real-time alignment and calibration
  - TURBO stream for majority of physics program → RAW events discarded
  - FULL and CALIBRATION streams to insure flexibility → filter & slim offline
- Offline CPU computing needs dominated by simulation
  - Number of events to be simulated scales with luminosity
  - Simulation time per event scales with pileup

  → CPU simulation explodes → need for faster simulations
- Offline storage driven by trigger output bandwidth
  - MC saved in μDST, so little impact on storage





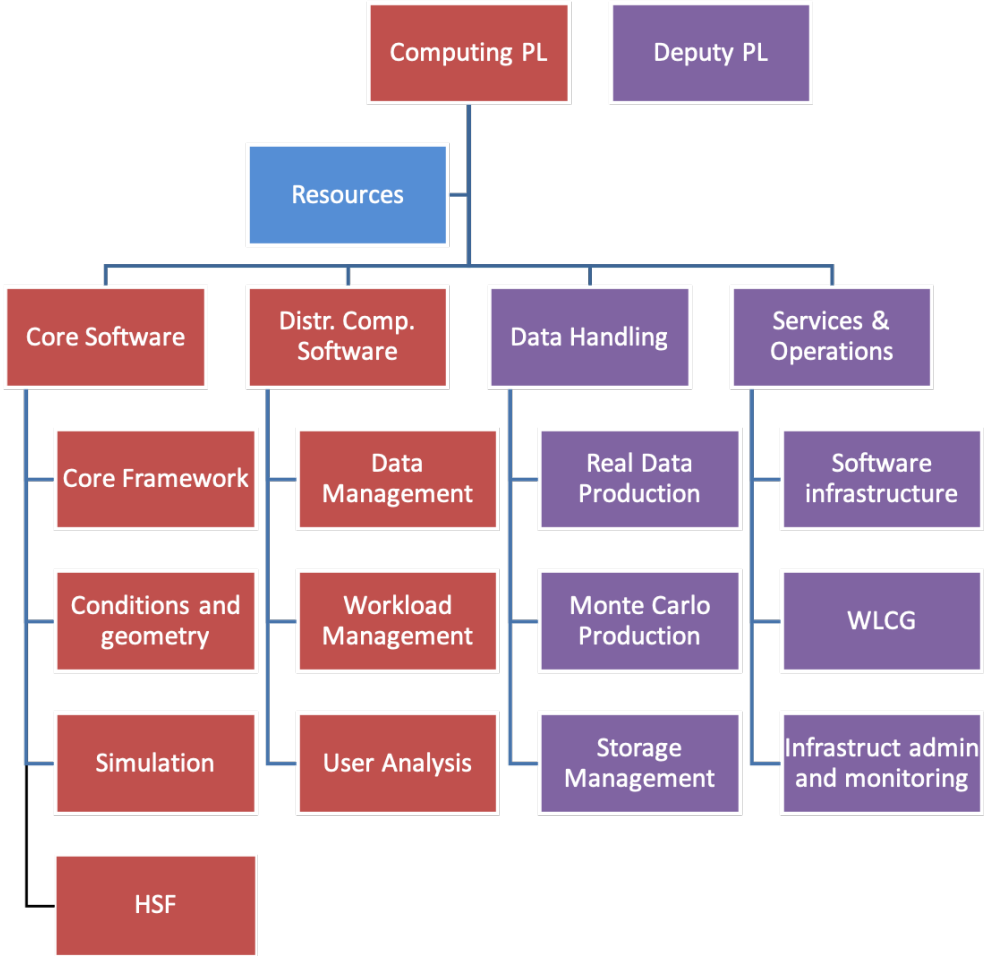


C. Bozzi and S. Roiser, Towards a computing model for the LHCb Upgrade, EPJ Web Conf. 214 (2019), 03045

C. Biscarat et al, System performance and cost modelling in LHC computing, EPJ Web Conf. 214 (2019), 03019

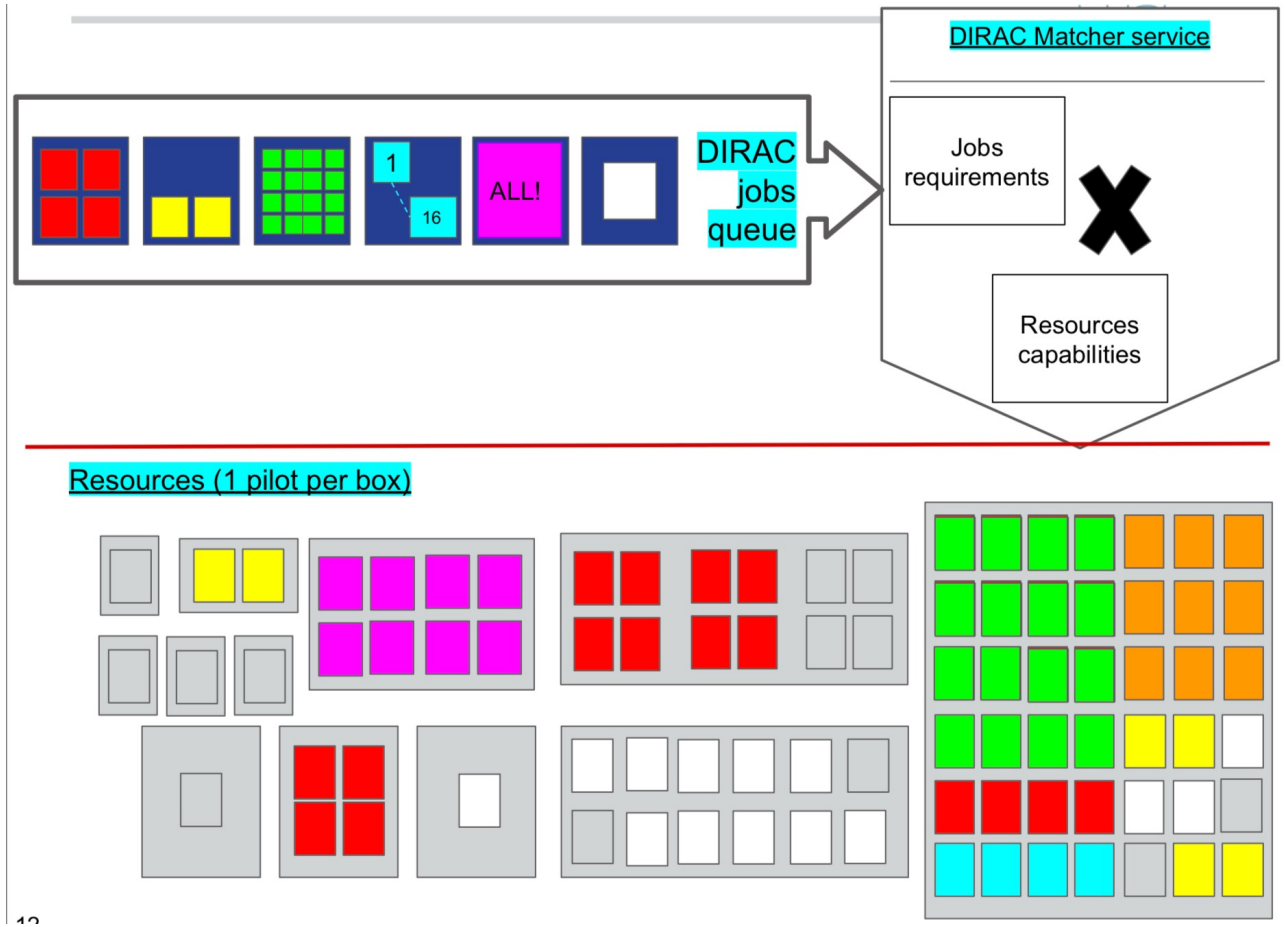
LHCb computing project leadership

- CB project leader starting January 2019
- Research and development activities in core software and distributed computing
- Operational tasks in data handling and processing, services, core software and distributed computing infrastructures
- Project interfaces to WLCG, funding agencies, HSF, other LHCb software projects (Real-Time Analysis, Data Processing and Analysis, Simulation)



Resources at CINECA

- PRACE grant (with other LHC experiments) to exploit the Marconi/A2 partition at CINECA
- Infrastructure set up to comply with experiments requirements
  - CVMFS, network connectivity
  - Collaboration with CNAF and CINECA
  - DIRAC development to exploit many-core architectures
    - Using DIRAC “pool”, an inner computing element
    - Parallel jobs matching



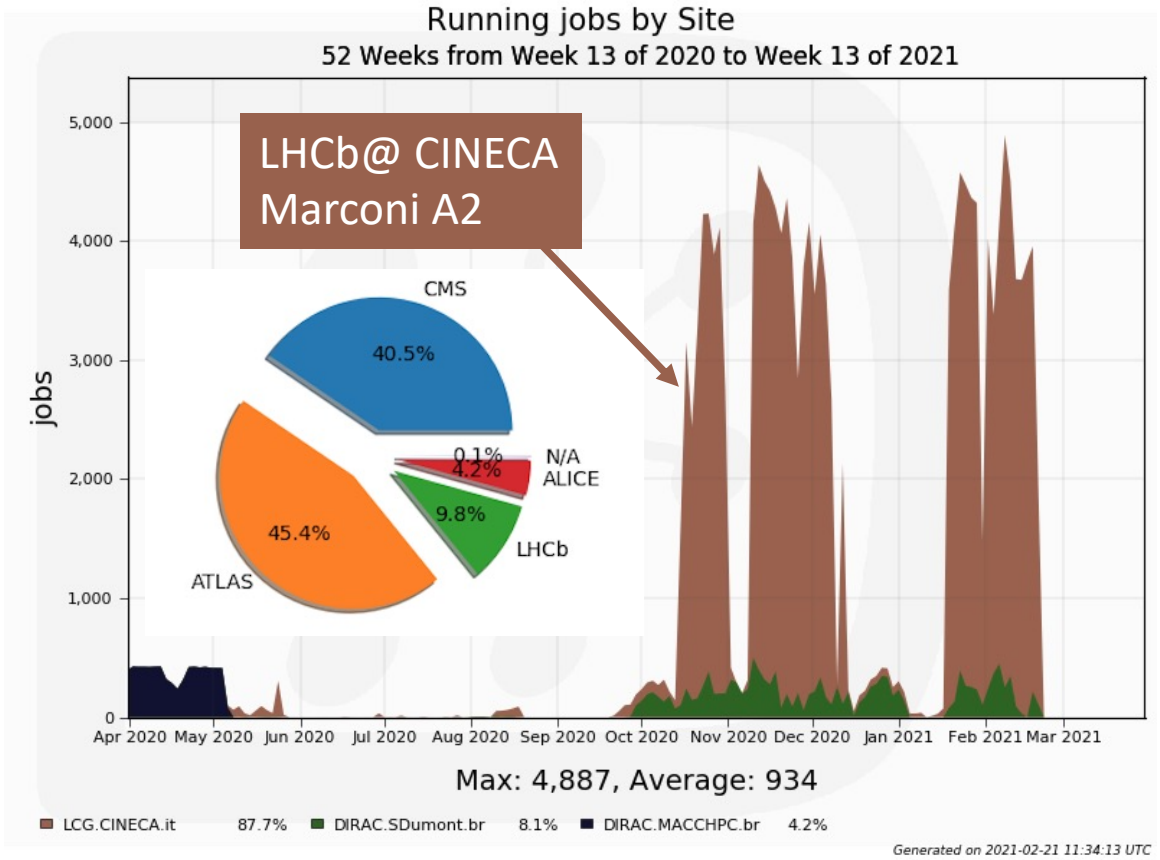
Resources at CINECA

- PRACE grant (with other LHC experiments) to exploit the Marconi/A2 partition at CINECA
- Infrastructure set up to comply with experiments requirements
  - CVMFS, network connectivity
  - Collaboration with CNAF and CINECA
  - DIRAC development to exploit many-core architectures
    - Using DIRAC “pool”, an inner computing element
    - Parallel jobs matching

EPJ Web of Conferences **245**, 09009 (2020) https://doi.org/10.1051/epjconf/202024509009
CHEP 2019

Extension of the INFN Tier-1 on a HPC system

Tommaso Boccali[1], *Stefano* Dal Pra[2], *Daniele* Spiga[3], *Diego* Ciangottini[3], *Stefano* Zani[2], *Concezio* Bozzi[4], *Alessandro* De Salvo[5], *Andrea* Valassi[6], *Francesco* Noferini[7], *Luca* dell’Agnello[2], *Federico* Stagni[6], *Alessandra* Doria[8], *Daniele* Bonacorsi[9]



Outlook

Thanks for your attention!