

Belle II Computing Model

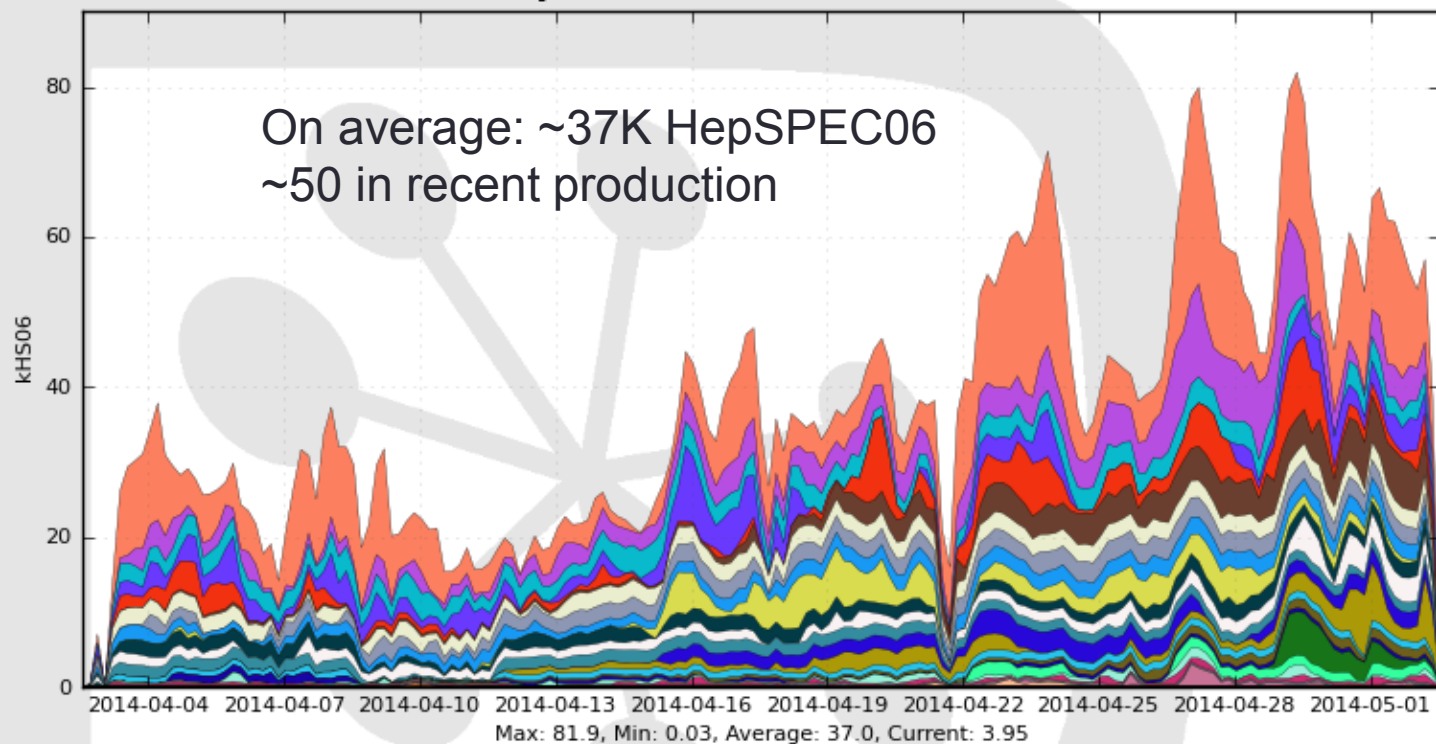
Fabrizio Bianchi

INFN and University of Torino

Spring 2014 MC Campaign

Normalized CPU usage by Site

30 Days from 2014-04-02 to 2014-05-02



Pisa 5.5%
CNAF 4.4%
Napoli 4.4%
Frascati 0.9%
Legnaro 0.8%
Torino 0.6 %

LCG.DESY.de	22.9%	LCG.Napoli.it	4.4%	LCG.Frascati.it	0.9%
LCG.KIT.de	9.2%	LCG.CYFRONET.pl	4.0%	LCG.KISTI.kr	0.9%
LCG.UA-ISMA.ua	6.6%	LCG.SIGNET.si	3.8%	LCG.Legnaro.it	0.8%
LCG.KEK2.jp	5.5%	LCG.Melbourne.au	3.5%	LCG.TORINO.it	0.6%
LCG.PISA.it	5.5%	LCG.CESNET.cz	2.8%	SSH.KMI.jp	0.4%
DIRAC.BINP.ru	5.4%	DIRAC.UVic.ca	2.8%	DIRAC.Niigata.jp	0.4%
LCG.KMI.jp	5.2%	LCG.ULAKBIM.tr	2.0%	OSG.Nebraska.us	0.3%
DIRAC.PNNL.us	5.0%	DIRAC.KrakowCloud.pl	1.2%	OSG.FNAL.us	0.2%
LCG.CNAF.it	4.4%	LCG.McGill.ca	1.0%	... plus 6 more	

Generated on 2014-05-02 14:36:33 UTC

Belle II Computing Model

- Two copies of raw data
 - First copy at KEK
 - Second copy distributed: 30% PNNL, 20% each Italy and Germany, 30% Asia (Kisti ?)
- 4 copies of miniDST for Data
 - KEK, Asia, America, Europe
- 1 copy of miniDST for MC
 - Where it was produced
- Processing: 60% at KEK, 40% distributed
- Reprocessing: were second copy of raw data is.
 - Possibility: until 2018 included, second copy and reprocessing only at PNNL
- MC Production & Analysis: everywhere according to PhD fraction

Assumptions to Estimate Resources

- Luminosity (still start of data taking in 2016)
- Sizes of event in raw data from DAQ people
- Size of event in miniDST from estimate/measurement
- CPU for Reco (MC generation and reconstruction) from current code
- CPU for Analysis from scaled Belle experience
- Two versions of mDST on disk
 - The one under production plus previous one
- Reprocessing:
 - 4 in first and second year of data taking
 - 2 in third year
 - One reprocessing every 2 years after that

Global Resource Estimate

Year	2014	2015	2016	2017	2018	2019	2020	2021	2022
Tape @ PNNL [PB]	0.00	0.00	0.00	0.00	9.23	28.51	52.45	76.80	102.32
Tape @ KEK [PB]	0.00	0.00	0.33	1.17	9.23	28.51	52.45	76.80	102.32
Tape KEK offset [PB]	2.00	2.00	1.67	0.83	0.00	0.00	0.00	0.00	0.00
Tape [PB]	2.00	2.00	2.00	2.00	18.46	57.01	104.90	153.60	204.64
Disk DST [PB]	0.00	0.00	0.01	0.02	0.18	0.57	1.05	1.54	2.05
Disk mDSTx4[PB]	0.00	0.00	0.24	0.87	6.89	21.28	39.16	57.34	76.40
Disk DST+mDSTx4	0.00	0.00	0.25	0.89	7.08	21.85	40.21	58.88	78.44
Disk MCx3 [PB]	0.00	0.00	0.70	2.49	11.09	21.37	34.14	47.13	60.74
Disk Analysis[PB]	0.00	0.00	0.03	0.10	0.82	2.53	4.66	6.83	9.09
Disk Challenge[PB]	2.00	4.00	5.00	2.00	0.00	0.00	0.00	0.00	0.00
Disk KEK offset [PB]	0.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Disk PNNL offset [PB]	0.70	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Disk [PB]	3.19	4.40	5.98	5.49	18.98	45.76	79.01	112.83	148.27
CPU Data [kHS]	0.00	0.00	1.87	4.79	46.00	109.98	136.60	138.93	145.59
raw data proc @ PNNL [kHS]	0.00	0.00	0.75	1.91	18.40	43.99	54.64	55.57	58.23
CPU mDST repr.	0.00	0.00	6.85	24.40	96.53	74.54	137.15	200.82	267.55
CPU MC [kHS]	0.00	0.00	9.85	25.25	121.30	145.01	180.11	183.18	191.96
CPU MC repr. [kHS]	0.00	0.00	27.09	96.52	143.37	138.15	220.69	304.65	392.63
CPU Analysis [kHS]	0.00	0.00	0.93	3.33	26.33	81.32	149.61	219.08	291.87
CPU Challenge [kHS]	133.25	133.25	111.04	161.80	0.00	0.00	0.00	0.00	0.00
CPU KEK offset [kHS]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CPU PNNL offset [kHS]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CPU [kHepSPEC]	133.25	133.25	157.64	316.08	433.53	548.99	824.16	1,046.67	1,289.61

Resources in Italy

O(10% of previous slide)

Site or Tier 1 fraction	Italy	This can be a site name as given on the Sites sheet or a number
Data fraction		If this is empty the data fraction will be equal to the number given on the Sites sheet or 0 for numerical Tier 1
Used Tier 1 fraction	0.10	The fraction of MC production and user analysis done at the site
Used data fraction	0.40	The fraction of mDST data stored at the site

Regional Center resources:

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Tape (raw data) [PB]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.86	8.64	13.51	18.62
Disk (data) [PB]	0.00	0.00	0.00	0.00	0.02	0.08	0.66	2.03	3.73	5.46	7.28
Disk (MC) [PB]	0.00	0.00	0.00	0.00	0.03	0.12	0.55	1.07	1.71	2.36	3.04
Disk (analysis) [PB]	0.00	0.00	0.00	0.00	0.00	0.01	0.08	0.25	0.47	0.68	0.91
Disk (challenges) [PB]	0.00	0.10	0.20	0.40	0.50	0.20	0.00	0.00	0.00	0.00	0.00
Disk [PB]	0.00	0.10	0.20	0.40	0.56	0.42	1.29	3.35	5.90	8.50	11.22
CPU (raw pro) [kHepSPEC]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.80	10.93	11.11	11.65
CPU (raw repro) [kHepSPEC]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.08	22.60	35.34	48.68
CPU (MC) [kHepSPEC]	0.00	0.00	0.00	0.00	3.61	11.97	25.46	27.11	38.58	47.26	56.86
CPU (an.) [kHepSPEC]	0.00	0.00	0.00	0.00	0.09	0.33	2.63	8.13	14.96	21.91	29.19
CPU (ch.) [kHepSPEC]	0.00	6.29	13.32	13.32	11.10	16.18	0.00	0.00	0.00	0.00	0.00
CPU [kHepSPEC]	0.00	6.29	13.32	13.32	14.81	28.48	28.09	54.12	87.07	115.62	146.38
CPU [# of cores 14.72HS/core]	0.00	427.46	905.21	905.21	1,006.09	1,934.70	1,908.20	3,676.51	5,915.15	7,854.49	9,944.14
WAN (data) [Gbit/s] (inward)	0.00	0.00	0.00	0.00	0.02	0.08	0.39	0.50	0.74	0.92	1.12
WAN (MC) [Gbit/s] (outward)	0.00	0.00	0.00	0.00	0.02	0.08	0.21	0.15	0.21	0.25	0.30
WAN (MC) [Gbit/s] (inward)	0.00	0.00	0.00	0.00	0.05	0.16	0.42	0.29	0.41	0.50	0.60
WAN (ch.) [Gbit/s](in/out)	0.00	0.16	0.32	0.64	0.64	0.48	0.00	0.00	0.00	0.00	0.00
WAN [Gbit/s] (outward)	0.00	0.16	0.32	0.64	0.66	0.56	0.21	0.15	0.21	0.25	0.30
WAN [Gbit/s] (inward)	0.00	0.16	0.32	0.64	0.71	0.72	0.81	0.79	1.15	1.42	1.72

Plus user space on disk

Costing of Resources in Italy (w/o second copy of raw data)

Cost for Analysis and MC Production										
Item	Totale 2014-2022	2014	2015	2016	2017	2018	2019	2020	2021	2022
Disco (PB) da tabella ufficiale		0.20	0.40	0.56	0.56	1.32	3.50	5.90	8.50	11.20
Disco (PB) user space (10% e poi 5% dal 2019)		0.02	0.04	0.06	0.06	0.13	0.18	0.30	0.43	0.56
Disco(PB) totale		0.22	0.44	0.62	0.62	1.45	3.68	6.20	8.93	11.76
Disco(PB) da comprare (rimpiazzo dopo 5 anni)			0.44	0.18	0.00	0.84	2.22	2.96	2.91	2.84
Disco: Costo per PB in KE (-20% anno)			265.00	212.00	169.60	135.68	108.54	86.84	69.47	55.57
Disco costo totale (KE)	1125.09	0.00	116.60	37.31	0.00	113.43	241.29	257.03	201.87	157.55
CPU(KHEPSpec) da tabella ufficiale		13.00	15.00	15.00	28.00	28.00	36.00	54.00	79.00	85.00
CPU(KHEPSpec) analisi n-tuple (10% e poi 5% dal 2019)*0.4		0.52	0.60	0.60	1.12	1.12	0.72	1.08	1.58	1.70
CPU(KHEPSpec) totale		13.52	15.60	15.60	29.12	29.12	36.72	55.08	80.58	86.70
CPU(KHEPSpec) da comprare (rimpiazzi dopo 4 anni)			15.60	0.00	13.52	0.00	23.20	18.36	39.02	6.12
CPU: Costo Unitario per KHEPSpec in KE (-20% anno)			12.00	9.60	7.68	6.14	4.92	3.93	3.15	2.52
CPU costo totale (KE)	615.41	0.00	187.20	0.00	103.83	0.00	114.03	72.19	122.75	15.40
Gran totale	1740.50	0.00	303.80	37.31	103.83	113.43	355.33	329.23	324.62	172.96

Second Copy of Raw Data

- Need to buy:
 - Tapes
 - Tape drives
 - CPU for (re)processing
- Uncertitudes:
 - When there will be a change of tape technology ?
 - Amount of tape drives needed
- Working with CNAF expert to have a costing estimate

Where to locate the resources in Italy ?

- CNAF: 40% + fraction of raw data (including CPU to process them)
- Tier2 Federation: 60% shared between Napoli, Pisa and Torino
 - Resources allocate to Belle II in different sites should be comparable
- Add resources only in a site each year and rotate among sites.

Next Steps

- Plan for resources deployment to be discussed with referees and endorsed by Commissione I and Giunta
- Preparing a document. Outline:
 - Introduction
 - Belle II Computing Model
 - Proposal for resource deployment
 - Status of sites and impact of deploying Belle II resources at CNAF, Napoli, Pisa, and Torino

Schedule (near future)

Jan. 13-14

US Belle II Computing Review

April 3, 2014

3rd MC production starts (goal: 2nd MC x 2 ↑)

May 9, 2014

Now we are here

mid. May, 2014

MC prod. 3.5 starts

gbasf2 refactoring finishes

toy production system

Aug, 2014

MC prod. 4 starts

target functionality of refactored gbasf2

toy production system

Database?

increase of CPU / SE

challenge for constant MC production

Dec, 2014

MC prod. 5 starts w/ full production

Jan. 2015

SuperKEKB accelerator commissioning starts

stable grid operation

DFC
implementation

xrootd

ANA-100G
test

KEK - PNNL
L2VPN

raw data
transfer test
online - offline

Schedule (for computing resources)

