Belle II Computing

Fabrizio Bianchi

INFN and University of Torino

Incontro con Referees, 11/09/2014

Disclaimer

- Alla commissione I di luglio ho fatto una presentazione dettagliata dello stato computing di Belle II:
- https://agenda.infn.it/getFile.py/access?contribId=21&resId=0&materialId=slides&confId=8247

- Di seguito riassumero' i punti salienti e mi concentrero' sulle richieste 2015

Belle II Computing Model (1)

- Outline of central production:
  - Raw data permanently stored on tape
  - Calibrated and processed to produce mDST
  - MC events processed in the same way
  - Detector and MC data skimmed to create subset of events (skims) targeting a set of physical analysis
  - Data challenges to test the system before the data taking
- Tiered structure:
  - Raw Data Centers (KEK, PNNL, GridKa, CNAF, KISTI, etc.): where the raw data are stored and processed. "Raw Data Centers" also serve as "Regional Data Centers" and "MC Production Sites".
  - Regional Data Centers (DESY, INFN Tier2 Federation, etc.): large computing centers where a (partial) copy of the mDST data is stored. "Regional Data Centers" also serves as "MC Production Sites".
  - MC Production Sites: where MC production/reconstruction and physics analysis is performed

Belle II Computing Model (2)

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

Assumptions to Estimate Resources

- Luminosity
- Cross section: 7.5 nb (5.0 nb) for data (MC)
- Sizes of event in raw data from DAQ people: 300 KB
- Size of event in miniDST from estimate/measurement: 45 KB
- CPU for Reco (MC generation and reconstruction) from current code: 45 (90) HEPSpec x sec /evt for data (MC)
- Storage and CPU for Analysis from scaled Belle experience
  - Storage as a mDST copy
  - CPU power as half of a raw data processing
- 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

Luminosity Profile

Year	2016	2017	2018	2019	2020	2021	2022
Lum/year (ab^{-1})	0.16	0.41	3.94	9.42	11.70	11.90	12.47
Total Lum (ab^{-1})	0.16	0.57	4.51	13.93	25.63	37.53	50.00

- Severe cut on machine operation budget at KEK for Japanese Fiscal Year 2014
  - Not possible to start the commissioning of SuperKEKB as planned
  - The collaboration and the Lab are working together to develop a plan that will minimize the impact of the cut

Proposed Belle II Computing Infrastructure in Italy

- Our request: resources at CNAF and in a Tier2 Federation (Napoli - ReCaS, Pisa and Torino)
- Roles according to Belle II computing model:
  - Raw Data Center: CNAF
  - Regional Data Center: CNAF + Tier2 Federation
  - MC Production Sites: CNAF + Tier2 Federation + opportunistic use of available resources
- At CNAF:
  - 40% of resources for MC production and analysis (Table at page 9)
  - Starting in 2019: INFN share of raw data (including CPU to process them)
- In the Tier2 Federation:
  - 60% of resources for MC production and analysis shared between Napoli, Pisa and Torino
  - Resources allocate to Belle II in different Tier2 sites should be comparable
  - Add resources only in a site each year and rotate among sites.

Costing of Resources in Italy (1) (w/o second copy of raw data) Resources at CNAF

Year		2015	2016	2017	2018	2019	2020	2021	2022
Disco(PB) da comprare (rimpiazzo dopo 5 anni)		0.10	0.06	0.00	0.52	1.08	1.50	1.42	1.48
Disco: Costo per PB in KE (-20% anno)		250.00	200.00	160.00	128.00	102.40	81.92	65.54	52.43
Disco costo totale (KE)	507.69	25.00	12.00	0.00	66.56	110.59	122.88	93.06	77.59
CPU(KHEPSpec) da comprare (rimpiazzi dopo 4 anni)		4.00	0.00	5.20	0.00	9.20	13.20	21.20	13.60
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)	285.98	48.00	0.00	39.94	0.00	45.22	51.90	66.69	34.23
Gran totale al CNAF	793.66	73.00	12.00	39.94	66.56	155.81	174.78	159.75	111.82

Costing of Resources in Italy (1) (w/o second copy of raw data) Resources at Tier2 Federation

Year		2015	2016	2017	2018	2019	2020	2021	2022
Disco(PB) da comprare (rimpiazzo dopo 5 anni)		0.30	0.09	0.00	0.78	1.62	2.25	2.13	2.22
Disco: Costo per PB in KE (-20% anno)		220.00	176.00	140.80	112.64	90.11	72.09	57.67	46.14
Disco costo totale (KE)	703.15	66.00	15.84	0.00	87.86	145.98	162.20	122.84	102.42
CPU(KHEPSpec) da comprare (rimpiazzi dopo 4 anni)		12.00	0.00	7.80	0.00	13.80	19.80	31.80	20.40
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)	500.96	144.00	0.00	59.90	0.00	67.83	77.86	100.03	51.34
Gran totaleal nei TIER2	1204.11	210.00	15.84	59.90	87.86	213.81	240.06	222.87	153.76

S. Pardi

Network Estimates in Italy (Gbps)

	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Napoli	0,20	0,20	0,20	0,20	0,50	1,50	2,00	3,00	3,50	4,00
Torino	0,20	0,20	0,20	0,20	0,50	1,50	2,00	3,00	3,50	4,00
Pisa	0,20	0,20	0,20	0,20	0,50	1,50	2,00	3,00	3,50	4,00
CNAF	0,50	0,70	0,70	0,70	1,00	2,00	4,50	6,00	6,50	7,50

Include:

- mDST (Data & MC) transfer
- Raw Data import at CNAF (from 2019)
- Raw Data Movement between CNAF and Tier2

Riunione del 30/7/14

- Scopo: chiarire alcuni punti sollevati nella riunione della Commissione I di luglio.
  - Presenti: Merola, Zoccoli, Lucchesi, Russo, De Nardo, Carlino, Forti
- PON Recas creata da tre enti: INFN, Univ Na, Univ BA. Sedi su NA, BA, CS, CT.
  - Atto d'obbligo di sostenere per 5 anni l'infrastruttura creata. Responsabilità solidale di tutti e tre gli enti. Quindi l'infrastruttura deve essere sostenuta collettivamente dalle tre entità.
  - Necessario una qualche forma associativa (prob. comitato di gestione) in essere entro fine anno, quando finisce il PON-Recas.
  - L'accordo dovrebbe essere tra INFN-NA-BA, mentre CS e CT si aggiungo con rapporti di collaborazione.
- Per cio' che riguarda Belle II:
  - Il comitato di gestione di PON-RECAS (collettivamente) ha deciso di rendere disponibili all'esperimento Belle-II (in cui ci sono INFN e UniNA) risorse di computing che (ma e' un accidente) sono state acquisite dall'Universita' di Napoli.
  - Nel MoU dovremo scrivere che ci sono risorse anche dell'universita' di Napoli, ma con la firma unica dell'INFN.
  - Precisamente: l'INFN entra come federazione di Tier2, costituita al momento da Torino, Pisa e PON-Recas Napoli, che contiene risorse INFN e Universita' di Napoli.

Yearly Spreadsheet Computing

SUBSYTEM	ITEM	CAT	TOTAL	Unit	2013	2014	2015	2016	2017	2018
COMP	*Disco (TB)*		*3600*	*TB*		*50*	*450*	*600*	*600*	*1900*
COMP	*Costo unitario disco (-20% ogni anno)*		---	*kE/TB*	*0.37*	*0.37*	*0.24*	*0.19*	*0.15*	*0.12*
COMP	Disco(KE)		**309**	kE		19	106	28	0	156
COMP	*CPU (KHEPSpec)*		*86.7*	*kHEPSpec*		*0.7*	*15*	*15*	*28*	*28*
	Costo unitario CPU (-30% ogni anno)			*kE/ kHEPSpec*	*14*	*14*	*12.00*	*9.60*	*7.70*	*6.10*
COMP	CPU(KE)		**290**	kE		10	180	0	100	0
COMP	Missioni coordinamento		**30**	kE	5	5	5	5	5	5
COMP	**TOTALE COMP**		**628.3**	**kE**	**5**	**33**	**291**	**33**	**105**	**161**
	VALORE CONTRIBUTO PON-RECAS		**206**		0	0	206	0	0	0
	COSTO a CARICO INFN		**422.3**		**5**	**33**	**85**	**33**	**105**	**161**

Necessita' 2015: Hardware

- Disco:
  - MC Production Campaign: 400 TB
    - Total MC to be produced: 4 PB, O(10^8 eventi)
    - Italian share: 10%
  - Detector Study: 25 TB
  - Software & Physics Tools: 25 TB
  - Partecipazione a test networking e data transfer

- CPU:
  - MC Production Campaign: 13 KHEPSpec
  - Detector Study: 1 KHEPSpec
  - Software & Physics Tools: 1 KHEPSpec

Richieste Hardware 2015

- Servono 450 TB + 15 KHEPSpec
- CNAF: Abbiamo 50 TB + 0.7 KHEPSpec. Vogliamo arrivare a 150 TB + 5 KHEPSpec. Servono:
  - 100 TB -> 25 KE
  - 4.3 KHEPSpec -> 52 KE
- Napoli (fornite da RECAS)
  - 300 TB
  - 10 KHEPSpec
  - N.B: 100 TB e 5 KHEPSpec li stiamo gia' usando (generosita' di ReCaS)

Richieste Missioni 2015

- Coordinamento (To): 5 KE
- Partecipazione Workshop Computing
  - 1 workshop 5 gg ad Hawaii (3K) + 1 workshop 2 gg a KEK (0.5 KE) = 3.5 KE/persona
  - Partecipanti:
    - Napoli: Del Prete, De Nardo, Pardi, Russo
    - Padova: La Caprara, Stroili
    - Pisa: Ciampi, Mazzoni
    - Torino: Bianchi, Lusso
  - Quote richieste: Na (3), Pd(2), Pi (2), To(1) = 28 KE
    - DeNardo (Na) usera' fondi di coordinamento
    - Bianchi (To) usera' fondi di coordinamento
- Gran totale: 33 KE

Backup

Global Resource Estimate

Year	2015	2016	2017	2018	2019	2020	2021	2022
Tape II copy raw data [PB]	0.00	0.00	1.17	9.23	28.51	52.45	76.80	102.32
Tape @ KEK [PB]	0.00	0.33	1.17	9.23	28.51	52.45	76.80	102.32
Tape [PB]	0.00	0.33	2.34	18.46	57.01	104.90	153.60	204.64
Disk DST [PB]	0.00	0.01	0.02	0.18	0.57	1.05	1.54	2.05
Disk mDSTx4[PB]	0.00	0.24	0.87	6.89	21.28	39.16	57.34	76.40
Disk DST+mDSTx4	0.00	0.25	0.89	7.08	21.85	40.21	58.88	78.44
Disk MCx3 [PB]	0.00	0.70	2.49	11.09	21.37	34.14	47.13	60.74
Disk Analysis[PB]	0.00	0.03	0.10	0.82	2.53	4.66	6.83	9.09
Disk Challenge[PB]	4.00	5.00	2.00	0.00	0.00	0.00	0.00	0.00
Disk [PB]	4.00	5.98	5.49	18.98	45.76	79.01	112.83	148.27
CPU processing [kHS]	0.00	1.87	4.79	46.00	109.98	136.60	138.93	145.59
CPU reprocessing [kHS]	0.00	6.85	24.40	96.53	74.54	137.15	200.82	267.55
CPU MC [kHS]	0.00	9.85	25.25	121.30	145.01	180.11	183.18	191.96
CPU MC repr. [kHS]	0.00	27.09	96.52	143.37	138.15	220.69	304.65	392.63
CPU Analysis [kHS]	0.00	0.93	3.33	26.33	81.32	149.61	219.08	291.87
CPU Challenge [kHS]	133.25	111.04	161.80	0.00	0.00	0.00	0.00	0.00
CPU [kHepSPEC]	133.25	157.64	316.08	433.53	548.99	824.16	1,046.67	1,289.61

Second Copy of Raw Data

- After 2019, second copy distributed: 30% PNNL, 20% each Italy and Germany, 30% Asia (?)
  - Other countries (Canada) could host a significant fraction
  - Work is in progress to have a plan
- INFN share can be hosted at CNAF buying:
  - Tapes
  - Tape drives
  - CPU for (re)processing
- Costing depend very much on the changes of tape technology between now and 2019

What We Have Done So Far

- Web site: https://web2.infn.it/Belle-II/
- Belle VO active at CNAF, Frascati, Legnaro-Padova, Napoli, Pisa, Roma3, Torino.
- Contributing to the MC production and to network bandwidth test (S. Pardi, Napoli)