Status of ILC Computing Model and Cost Study

Akiya Miyamoto
Norman Graf
Frank Gaede
Andre Sailer
Marcel Stanitzki
Jan Strube

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

Introduction

- Cost of computing needs for the detectors was not included in the TDR.
- For the LHC experiments, this was a very large item and (still is) a major source of headaches.
  - The Grid has been considered “The 5th experiment” at LHC.
- For ILC, I consider computing more like another sub-detector, not least because of the PFA paradigm.
  - But this needs to be proven with hard numbers

Studies in SiD / ILD

ILD (Miyamoto) presented first preliminary results of detailed studies at AWLC14 and extended studies over the summer.

SiD only had DAQ numbers at 1 TeV.

For LCWS14, SiD used ILD extrapolation of background numbers to obtain data rates at lower ILC stages.

SiD Data Rate

Table II-9.2
Overview of read-out details for the various subdetectors. Occupancies and data volumes are for a full bunch train at 1 TeV and include beam-induced background as well as charge sharing between pixels/strips. Safety factors of five and two have been applied to the rates of incoherent pairs and $\gamma\gamma \rightarrow$ hadrons respectively. BeamCal and Lumical are expected to be using the Bean chip with a buffer depth of 2820.

	cell size (mm^2)	number of channels (10^6)	av. to max. occ. (%)	approx. # bits per hit (bit)	data volume (Mbyte)
VXD barrel	0.02×0.02	408	8 - 60	32	130
VXD disks inner	0.02×0.02	295	4 - 70	32	50
VXD disks outer	0.05×0.05	980	0.5 - 20	32	20
Main tracker barrel	0.05×100	16	33 - 300	32	20
Main tracker disks	0.05×100	11	4 - 500	32	2
ECAL barrel	3.5×3.5	72	2 - 45	40	7
ECAL endcap	3.5×3.5	22	33 - 2300	40	36
HCAL barrel	10×10	30	0.07 - 200	40	0.1
HCAL endcap	10×10	5	96 - 3600	40	24
LumiCal	2.5×var.	0.061	≫100	16	340
BeamCal	2.5(5.0)×var.	0.076	≫100	16	430

1060
→ 5.3 GB/s

Assuming 2450 BX/train @ 5 Hz

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

Data Rates

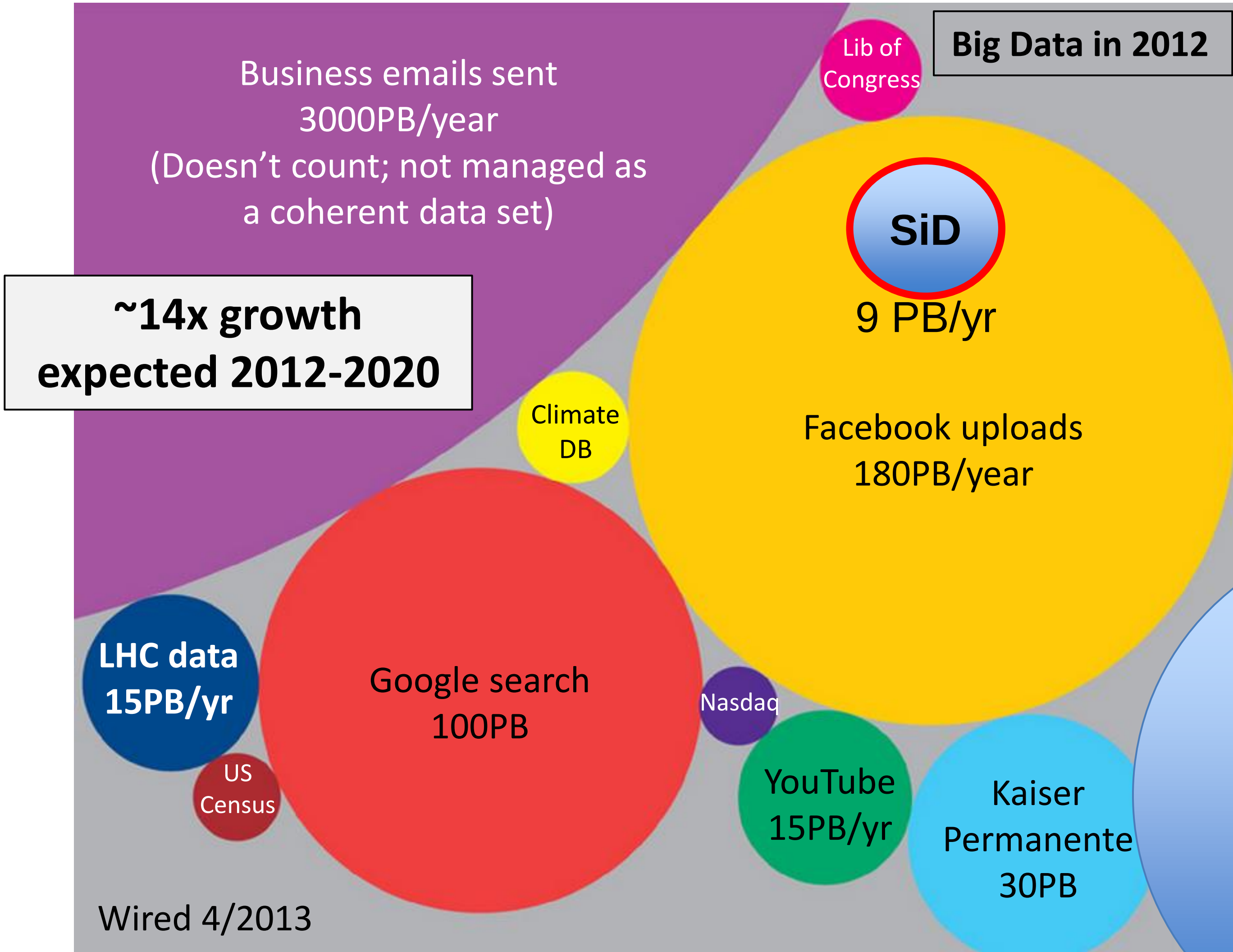
From the Detectors:
~1.1 GB raw data @ 500 GB, both for SiD/ILD

- Assuming half of a 1.6×10^7 s year:
  - ~9 PB raw data / year / detector

Slightly less at lower energies:
350 GeV SiD: ~4.9 PB, ILD ~6.3 PB
250 GeV SiD: ~3 PB, ILD ~5.5 PB
Large uncertainties from estimation of pair bg.
ILD: Detailed simulation study; SiD: extrapolation from background rates and number of BX/train

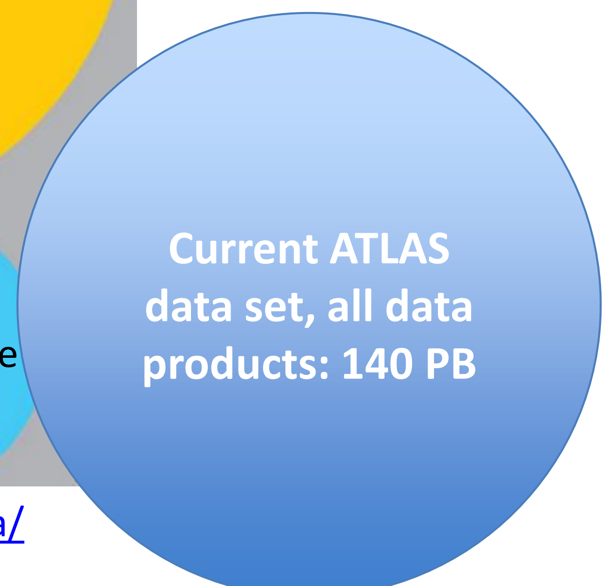
From: Torre Wenaus, CHEP 2013

Data Management
Where is LHC in Big Data Terms?



In 2012: 2800 exabytes created or replicated
1 Exabyte = 1000 PB

Computing needs and networking capabilities of society as well as science will increase in the future.



http://www.wired.com/magazine/2013/04/bigdata/

Caveats

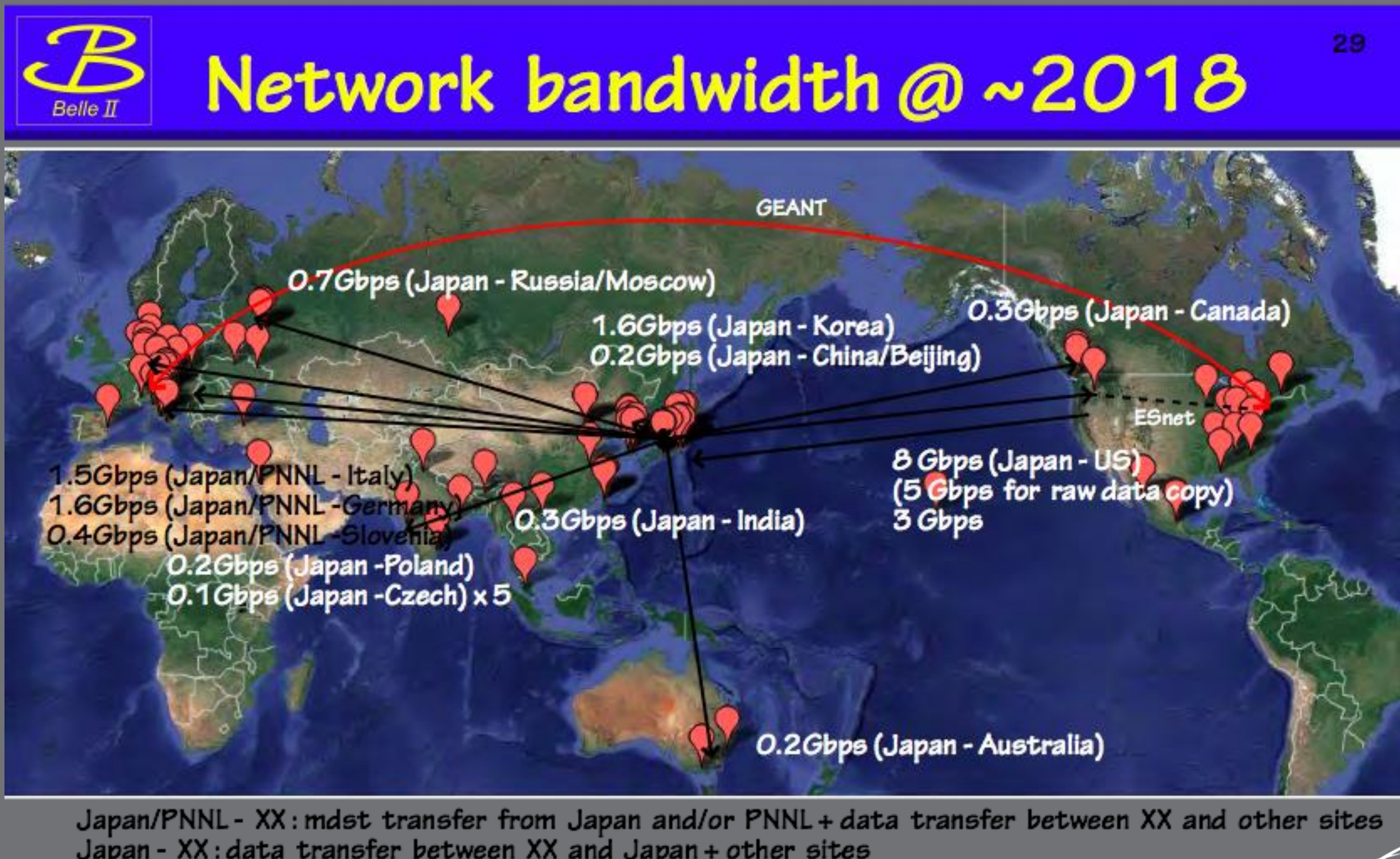
Data rate looks scary at first sight: > LHC! **But:**

- This will be in ~2030. HL-LHC will have set a new bar by then.
- There will be many experiments with data rates in the same ball park by then.
- As mentioned, large safety factors on background. Using 5, while 2 is advertized by the machine.
- Data reduction techniques being investigated. (Size is mostly from pair bg)

Given Constraints

- Want to avoid driving large investment in networking.
  - Take advantage of “natural” developments.
- Want to avoid a large computing center at the IP.
  - Size is constrained, infrastructure limited.
- Want to avoid a large investment in a Tier-0
  - Tier-0 at LHC: 45%-90% of all 11 Tier-1 combined
  - ILC campus computing should be driven by user community on site, not detector data rates.
  - But we want two copies of the raw data, one on site.

Japanese Connections

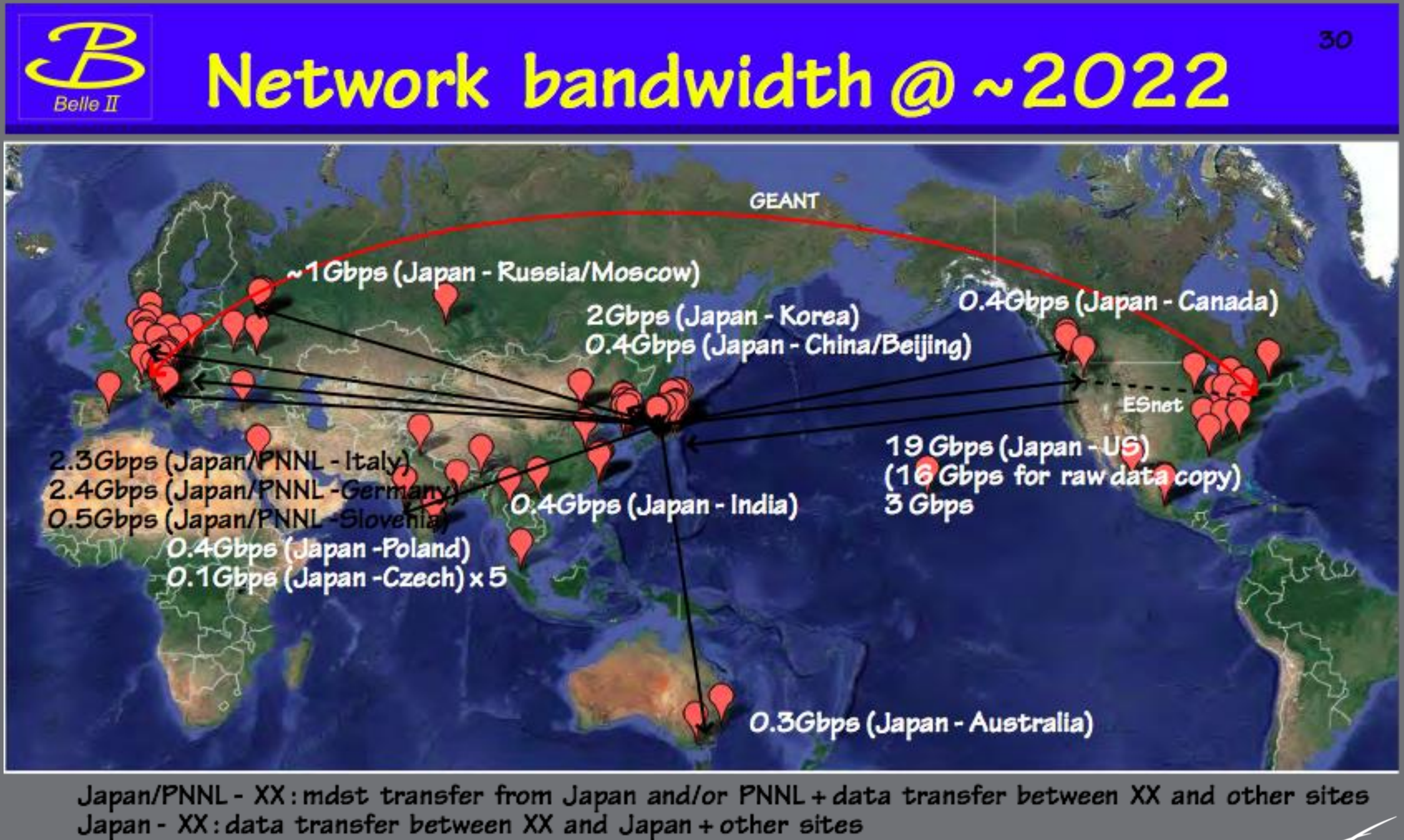


From T. Hara, ALICE WS, Mar 5, 2014

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

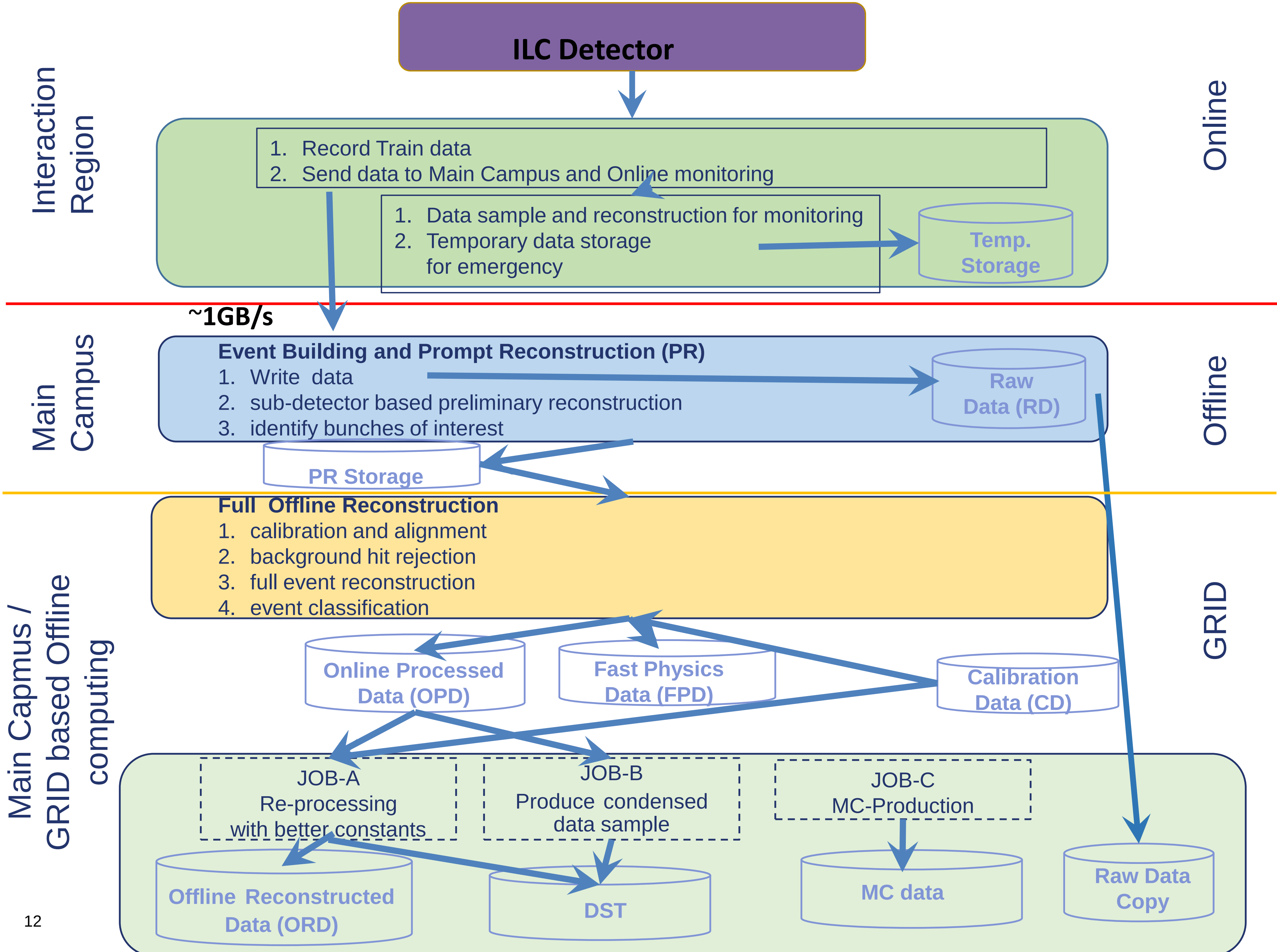
Japanese Connections



From T. Hara, ALICE WS, Mar 5, 2014

Strategy

1. Move the data out of the mountain as quickly as possible
   - Need few days buffer for network outage
2. Store a raw copy on campus
3. Process (filter) events for analysis and ship to grid.
   - CPU needs for this being investigated. Filtering could happen on the grid. Under study.
   - A complete copy of the raw data needs to be shipped for redundancy anyway.
   - Plan to use existing “Tier-1” centers.
   - Probably want certain “streams” with fast feedback.

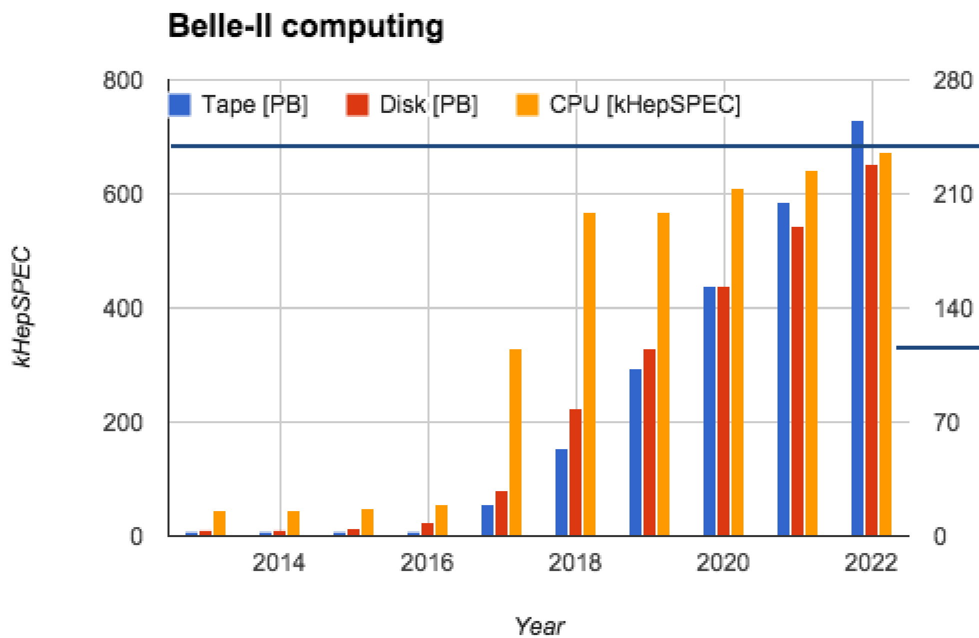


ILC Detector
Interaction Region
Online
1. Record Train data
2. Send data to Main Campus and Online monitoring
1. Data sample and reconstruction for monitoring
2. Temporary data storage for emergency
Temp. Storage
~1GB/s
Main Campus
Offline
Event Building and Prompt Reconstruction (PR)
1. Write data
2. sub-detector based preliminary reconstruction
3. identify bunches of interest
Raw Data (RD)
PR Storage
Main Capmus / GRID based Offline computing
GRID
Full Offline Reconstruction
1. calibration and alignment
2. background hit rejection
3. full event reconstruction
4. event classification
Online Processed Data (OPD)
Fast Physics Data (FPD)
Calibration Data (CD)
JOB-A Re-processing with better constants
JOB-B Produce condensed data sample
JOB-C MC-Production
Offline Reconstructed Data (ORD)
DST
MC data
Raw Data Copy

CPU estimation

Currently only have ILD estimation. SiD is unlikely to provide comparable level of detail.
Based on detailed studies of computing time for simulation and reconstruction of different physics channels at different energies (Miyamoto):
Campus computing (event filter): ~26k HepSPEC06
ILC total: 415k HepSPEC06
This is combined for both experiments.
Assumes 2x ILD computing.

Comparison with other experiments



ATLAS Tier-0+Tier-1 CPU 2015 (667 kHS)

(total: 1175 kHS)

ATLAS 2015 Disk (total): 108 PB

Tape: 98 PB

Scaling law (from LHC experiments experience): constant investment buys 20% more CPU, 15% more storage every year

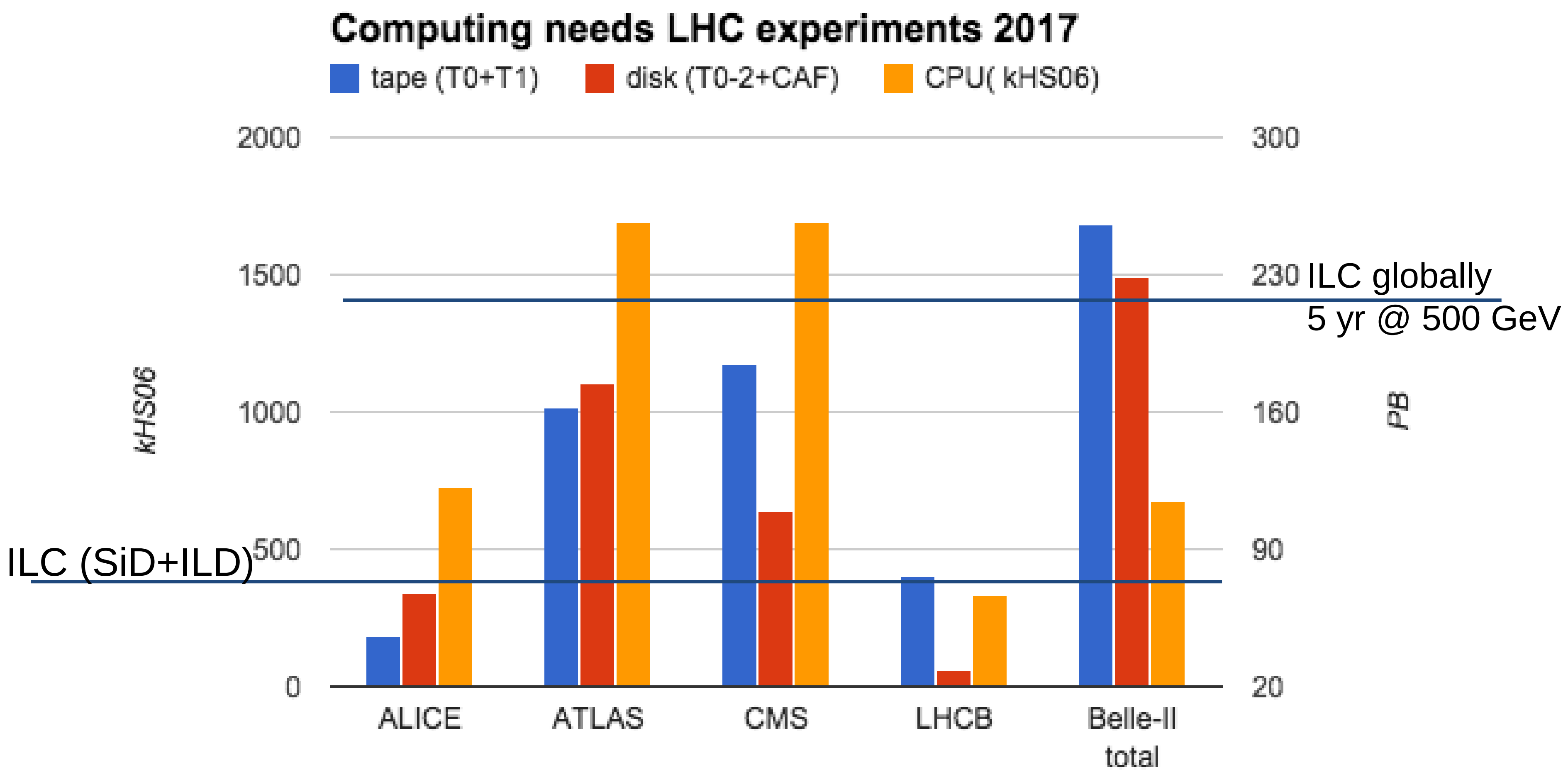
→ ~13 years to gain order of magnitude with same investment in CPU

→ ~16.5 years to gain order of magnitude in storage

Example:

Belle investment in 2016:	**Value in 2029:**
Tape: 16.44 PB	Tape: 101 PB
Disk: 19.98 PB	Disk: 123 PB
CPU: 273.31 kHS	CPU: 2924 kHS
kHS = kHepSPEC	

Pacific Northwest NATIONAL LABORATORY



For comparison: ILC 500 GeV

- 9 PB / yr / experiment raw data; 220 PB for 5 yr, ILC globally, incl. MC, analysis
- 415 kHS06 total

Manpower estimates

Task	Details	FTE
Management	Operation in general, negotiation among sites, maintain rules for computing, deliver information	6
Network	Maintenance, monitor, trouble handling develop faster network	6
Certificate	Maintenance, person ID, issue certificate, trouble handling	4
Security	Maintenance, Monitor, incident handling, prevention	4
Storage	Maintenance, trouble handling, improvements	4
CPU servers	Maintenance, trouble handling, improvements	4
Software maintenance	Maintenance, trouble handling, improvements	6
Core Software development	Design, implementation, maintenance, trouble handling, improvements	6
User contacts	User portal, monitor, dispatch problems, trouble handling	12
Misc. server management	Operation, maintenance, trouble handling of various servers such as EDMS, Indico, Web, cloud, mail, etc	4
Videoconferencing Support	Maintenance & User support	4
Total		60

Summary

- Input to MEXT process is being prepared
- Large uncertainties on data rates and CPU
    - data rates conservative, CPU unknown
- Data rates from detectors would compete with LHC today
    - > 10 years from now, our data rates will not require large additional investments
- Expected continued development of networks lets us take advantage of a distributed infrastructure
    - Allows granular contributions from FAs

Some investment needed to study data distribution strategies. Reconstruction and Analysis Software needs large investments