ALICE実験におけるWLCGと広島大学Tier-2の現状

WLCG勉強会
東大理学部1号館　508号室
2013年4月12日

杉立　徹（広島大）
for the ALICE-J GRID Team

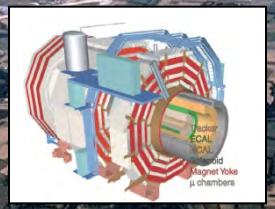
A Large Ion Collider Exp. at CERN

$\sqrt{s}$ = 8 TeV in proton + proton

$\sqrt{s_{NN}}$ = 2.76 TeV on Pb + Pb

$\sqrt{s_{NN}}$ = 5.02 TeV on p + Pb

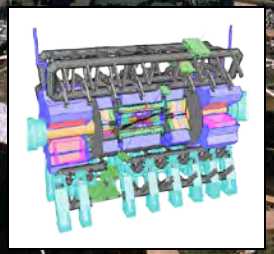
$\sqrt{s_{NN}}$ at LHC = 28 x RHIC =320 x SPS = 1000 x AGS



CMS

LHCb

ATLAS

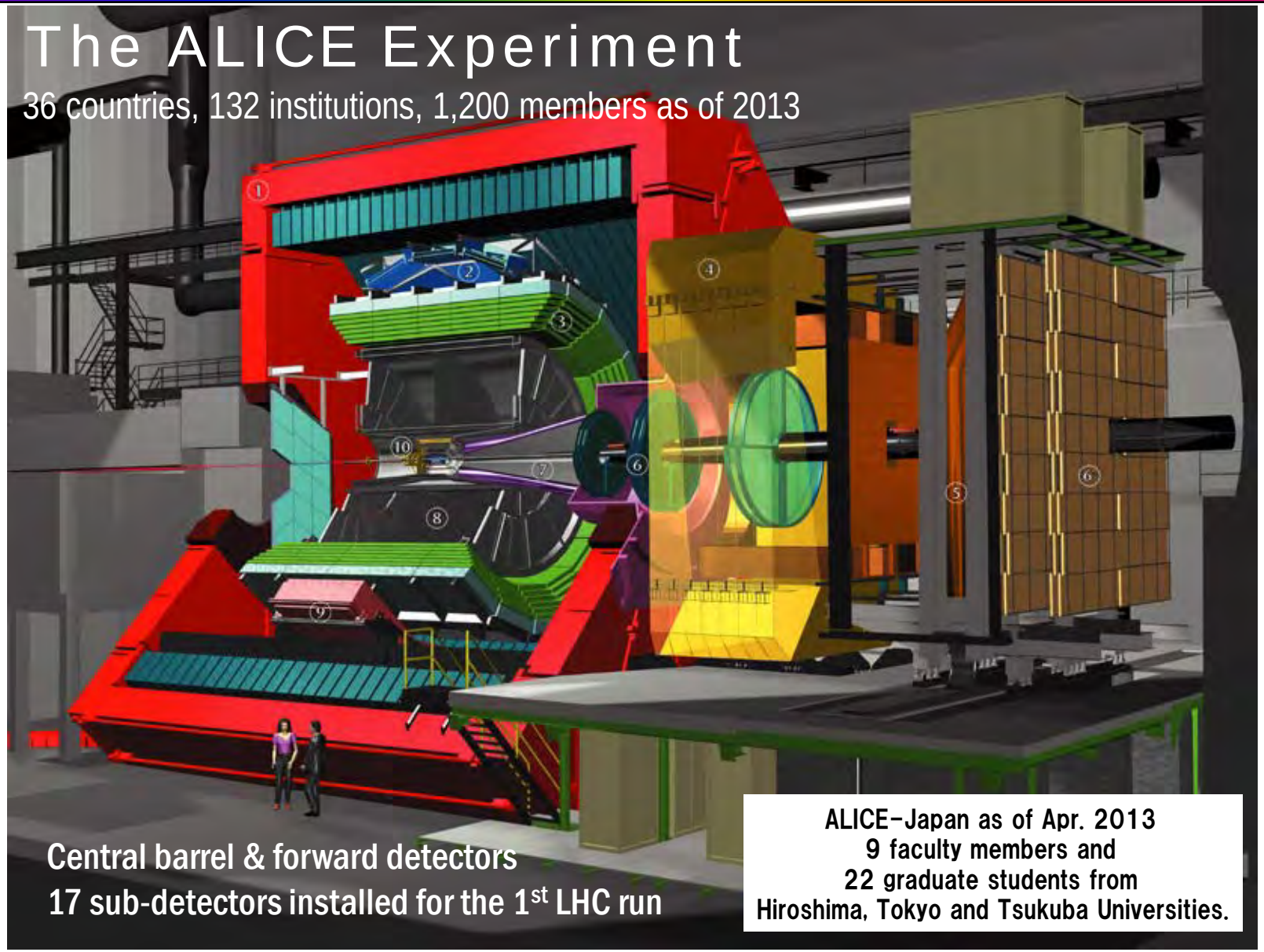


ALICE

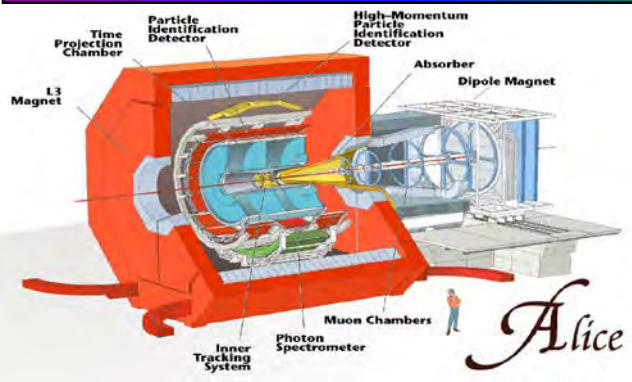
- understand *"strong" QCD*, and
- reveal dynamics at *early Universe*.

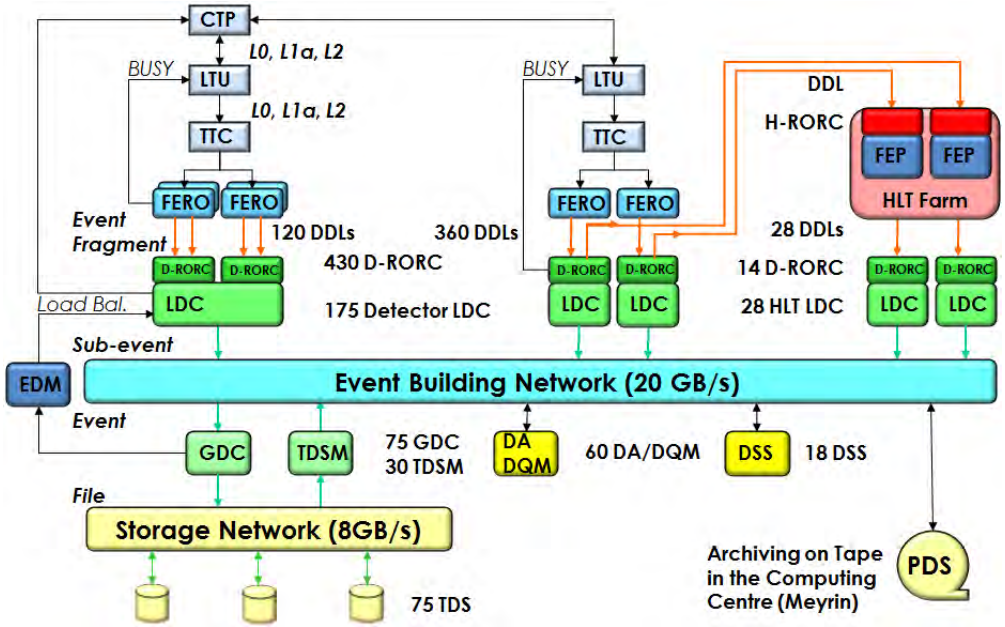
1200 members from 132 institutes in 36 countries

The ALICE Experiment at Point-2

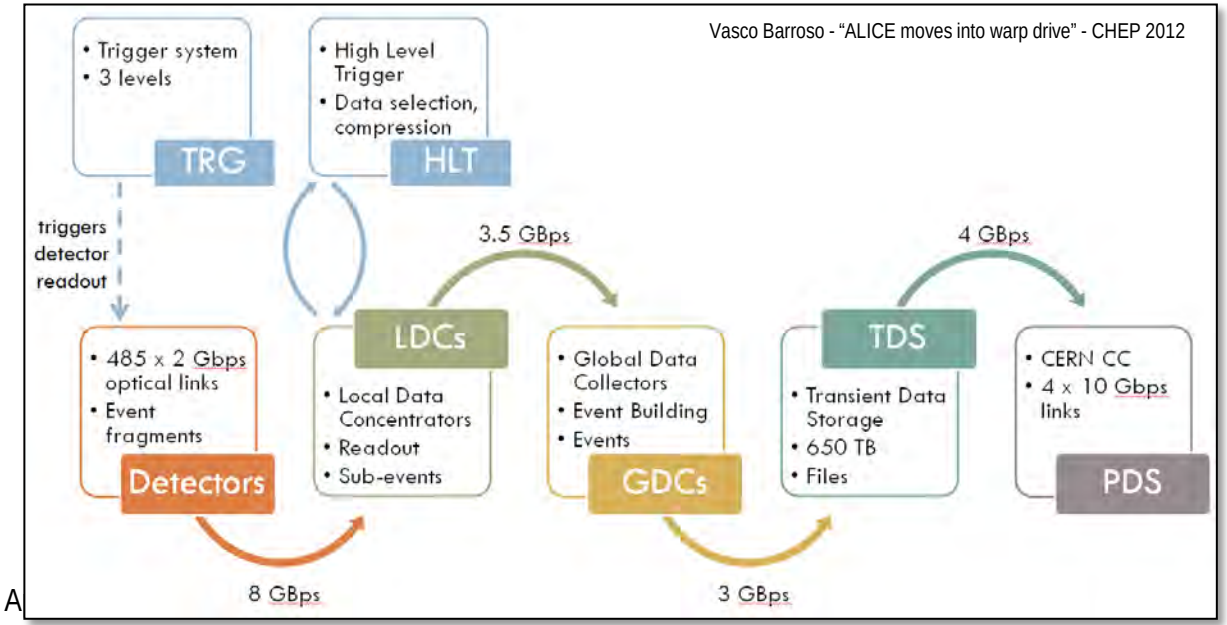


ALICE Data Stream to Tape

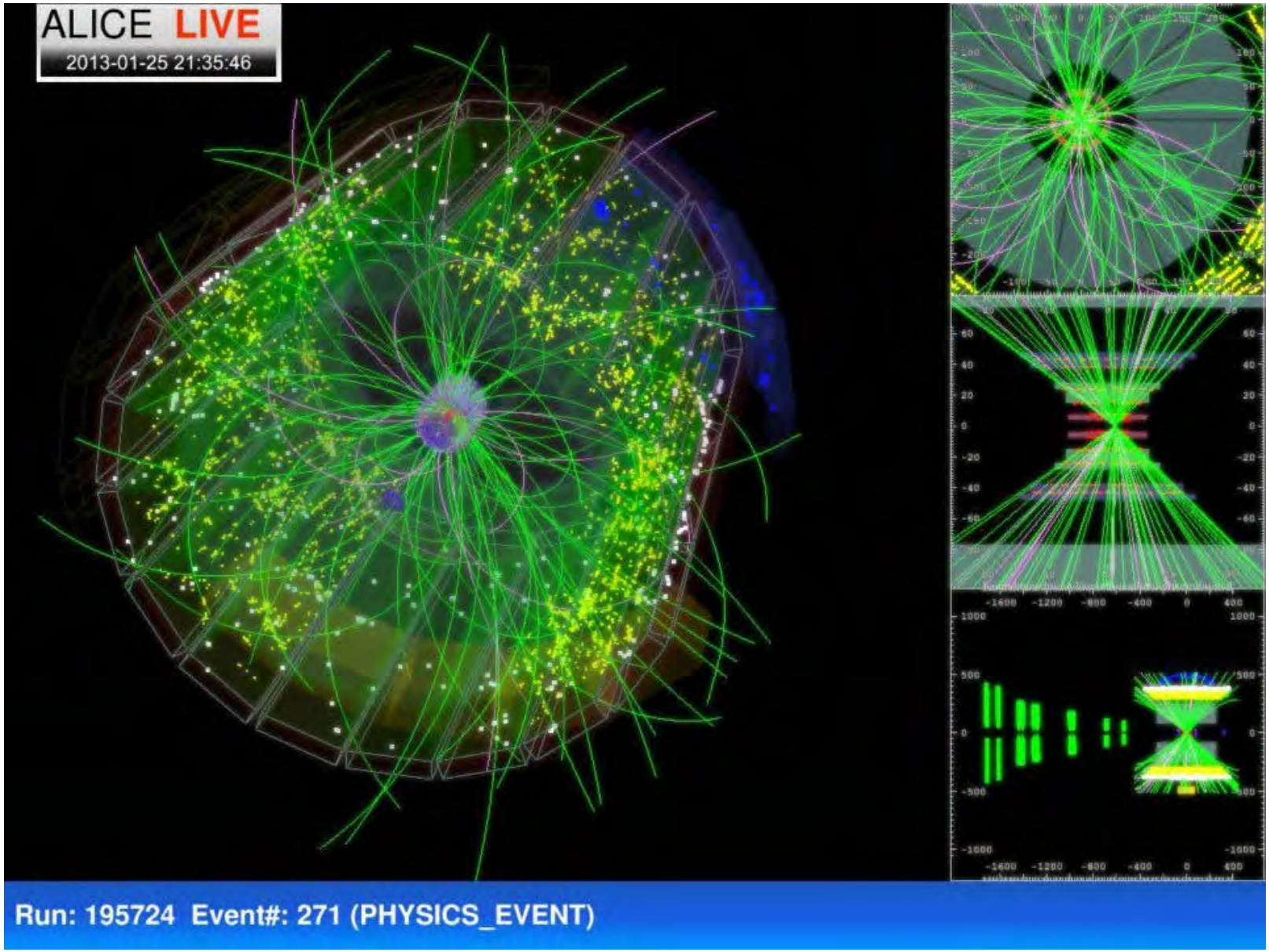




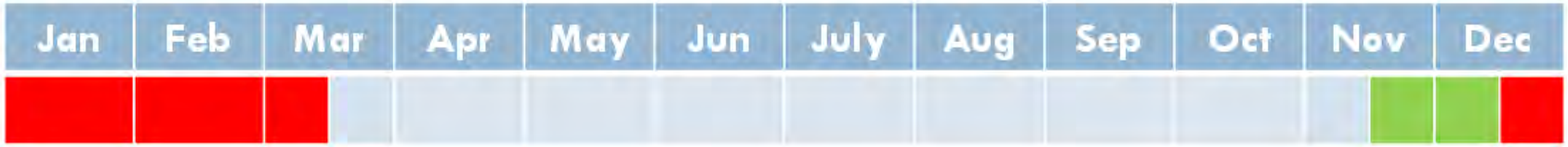
- **Run plan:**
  - **p+p in 10^7 sec (7 months)**
  - **A+A in 10^6 sec (1 months)**
- **Raw data: 2.5 PB/year**
  - **$2x10^8$ events/year for A+A**
  - **12.5 MB/event**
- **Event Summary Data**
  - **3 MB/event for A+A**
- **Analysis Object Data**
  - **300 KB/event for A+A**



ALICE Live of p+Pb Event



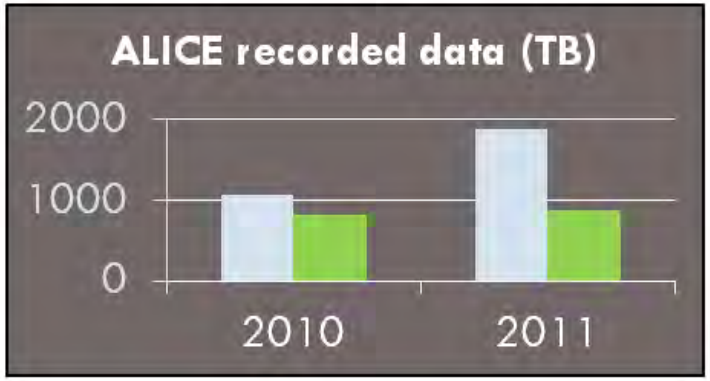
Typical LHC Year

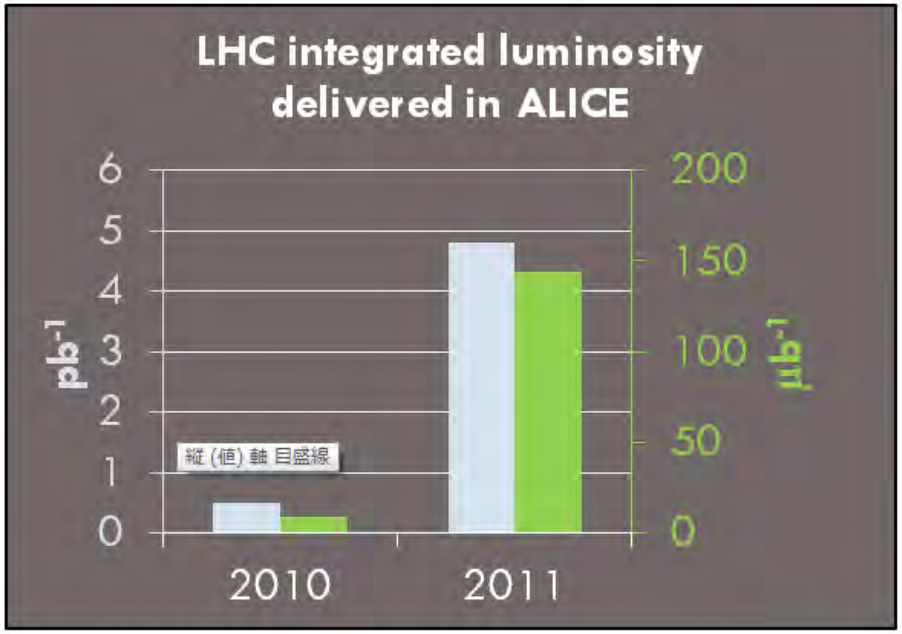


Shutdown for maintenance

proton-proton collisions

Heavy-ion collisions





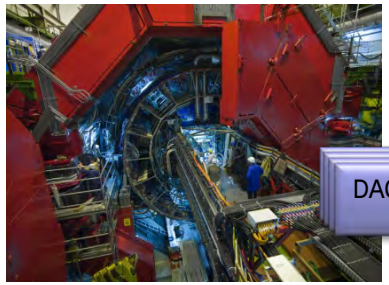
Vasco Barroso - "ALICE moves into warp drive" - CHEP 2012 21/05/2012

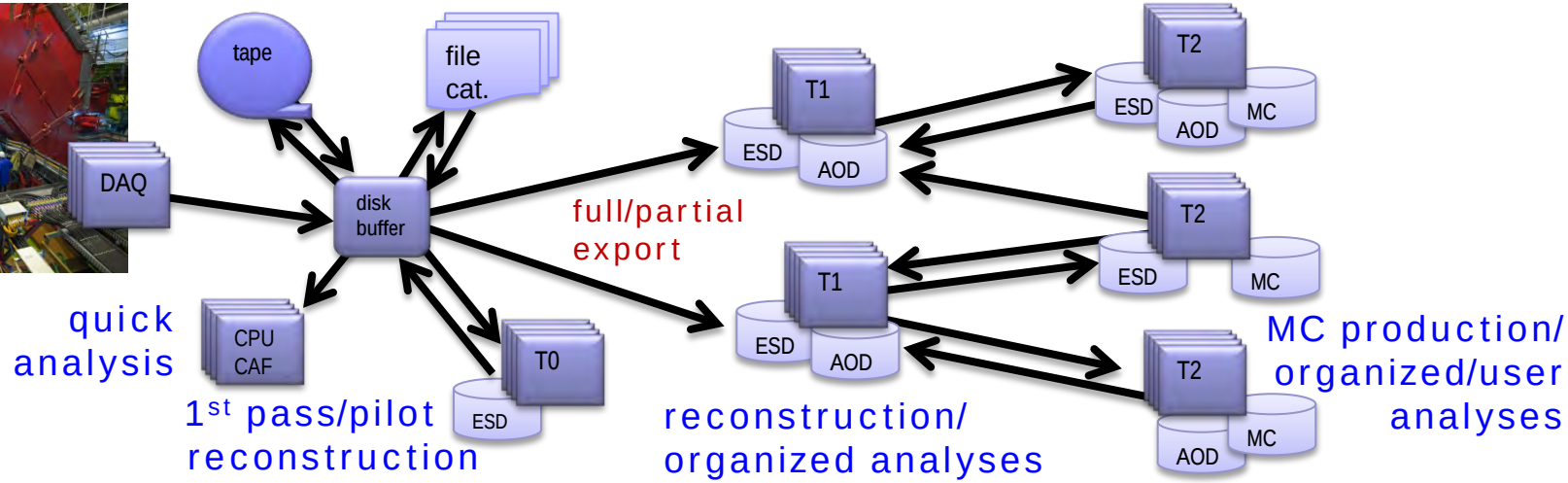
Data Samples as of Today

This is my private compilation and can be different from official numbers.

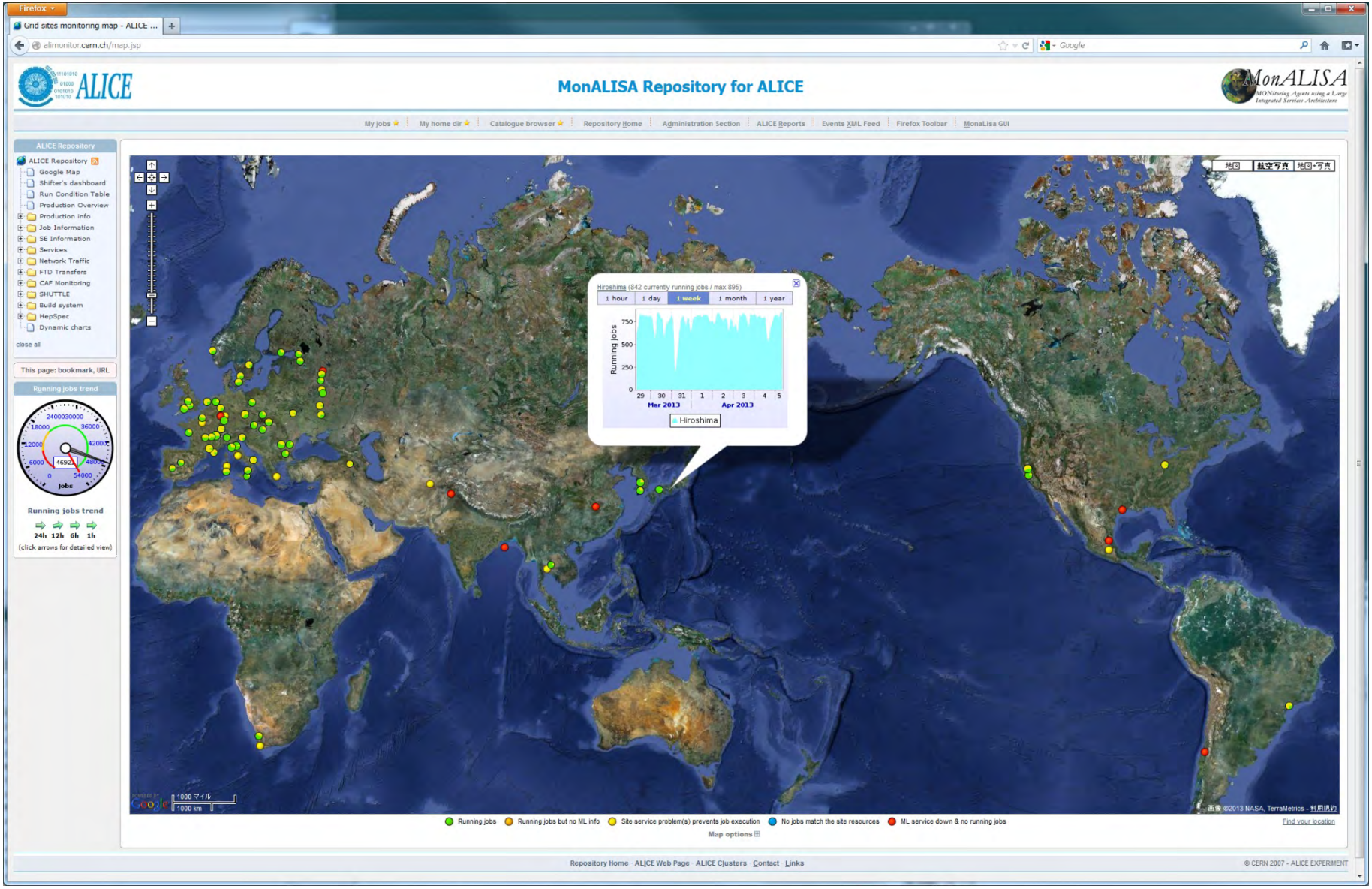
year	collisions	# of events recorded on tape
2013	p+Pb/Pb+p @5.02TeV	140M minimum bias and about 10M events with specific triggers out of L_{int}=31.94 nb^{-1} delivered in 4 weeks.
2012	p+p @8TeV	Continuing p+p data taking until X'mass shutdown (Wk 50) L_{int}=2.47 pb^{-1} recorded out of 3.49 pb^{-1} triggered as of Oct. 11.
2012	p+Pb@5.02TeV	1.8M events during the pilot run in Sept.
2011	Pb+Pb @2.76TeV/A	132M events in various triggers in 3.2PB (24MB/event) in 24days L_{int}=143.6 μb^{-1}
2011	p+p @7TeV	10^9 events in minimum bias trigger at 10kHz in 100 days L_{int}=2 pb^{-1} for rare triggers
2010	p+p @2.76TeV	74M events in minimum bias trigger in 35 hours 10M events and L_{int}=18 nb^{-1} for rare triggers
2010	Pb+Pb @2.76TeV/A	30M events in minimum bias trigger
2010	p+p @7TeV	800M events in minimum bias trigger 50M events in muon triggers 20M events in high N_{ch} triggers
2010	p+p @900GeV	8M events in minimum bias trigger
2009	p+p @2.36TeV	40k events in minimum bias trigger
2009	p+p @900 GeV	300k events in minimum bias trigger

ALICE Computing Model





ALICE Tiers

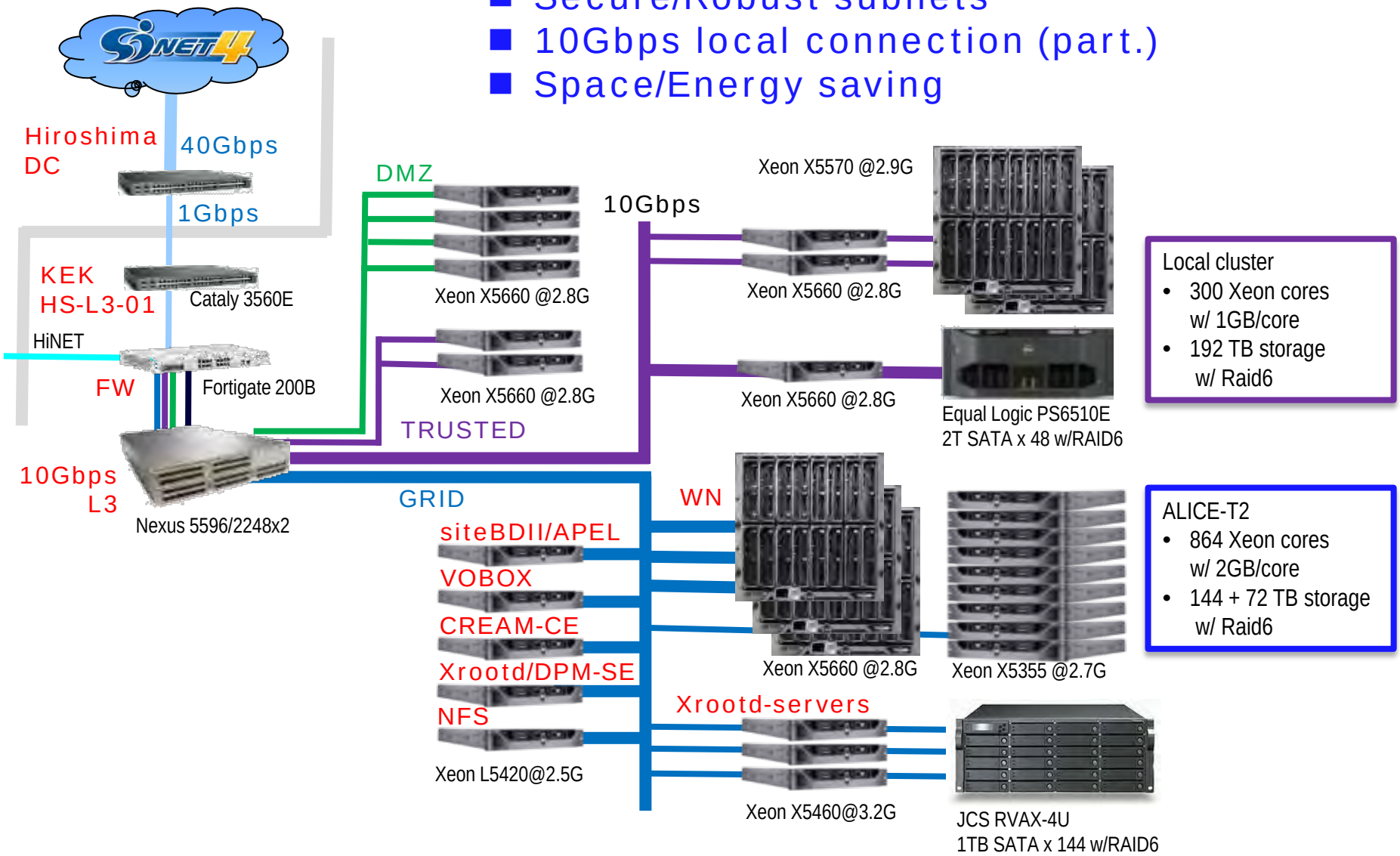

Firefox
Grid sites monitoring map - ALICE ...
alimonitor.cern.ch/map.jsp
ALICE
MonALISA Repository for ALICE
MonALISA
My jobs
My home dir
Catalogue browser
Repository Home
Administration Section
ALICE Reports
Events XML Feed
Firefox Toolbar
MonaLisa GUI
ALICE Repository
ALICE Repository
Google Map
Shifter's dashboard
Run Condition Table
Production Overview
Production info
Job Information
SE Information
Services
Network Traffic
FTD Transfers
CAF Monitoring
SHUTTLE
Build system
HepSpec
Dynamic charts
close all
This page: bookmark, URL
Running jobs trend
Jobs
Running jobs trend
24h 12h 6h 1h
(click arrows for detailed view)
Hiroshima (842 currently running jobs / max 895)
1 hour
1 day
1 week
1 month
1 year
Running jobs
Mar 2013
Apr 2013
Hiroshima
Running jobs
Running jobs but no ML info
Site service problem(s) prevents job execution
No jobs match the site resources
ML service down & no running jobs
Map options
Find your location
Repository Home · ALICE Web Page · ALICE Clusters · Contact · Links
© CERN 2007 - ALICE EXPERIMENT

ALICE Tier-2 at Hiroshima

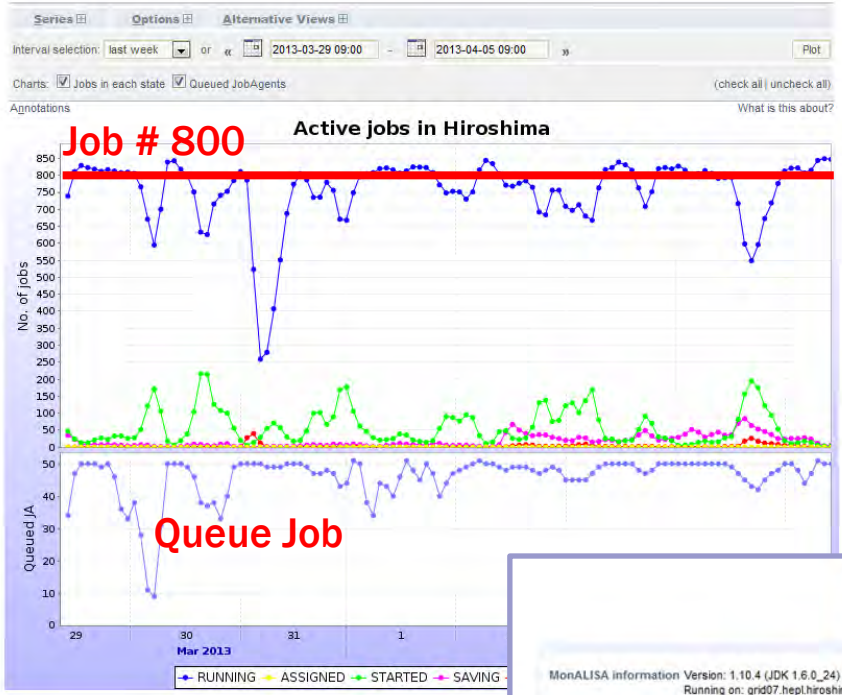
- The ALICE LCG site "JP-HIROSHIMA-WLCG" with grid middleware EMI-1 on SLC5 **until the last week**.
- LCG service; APEL, sBDII, CREAM-CE, XROOTD, DPM-SE, VOBOX... **minimum LCG service at minimum cost**
- WN resources; **1164 Xeon-cores in total**
  Xeon5355(4cores@2.6GHz) x 2cpu x 32boxes &
  Xeon5365(4cores@3.0GHz) x 2cpu x 20blades &
  Xeon5570(4cores@2.9GHz) x 2cpu x 26blades &
  Xeon5670(6cores@2.9GHz) x 2cpu x 3blades &
  Xeon5660(6cores@2.8GHz) x 2cpu x 42blades
- Storage cap; **408TB disks on 6 servers** and **no MS**
- Around **2/3 resource** deployed in the ALICE GRID
- Network B/W: 1Gbps on 40Gbps-SINET4 in Japan
- WLCG support by ASGC in Taiwan
- ~~ALICE associated Tier-1 in Lyon~~
- Responsible by Prof. Toru Sugitate
- Operated by TS and 中宮義英, and remote technical supports by a part-time SE of 創夢（株）in Tokyo

New Configuration from Feb. 2012

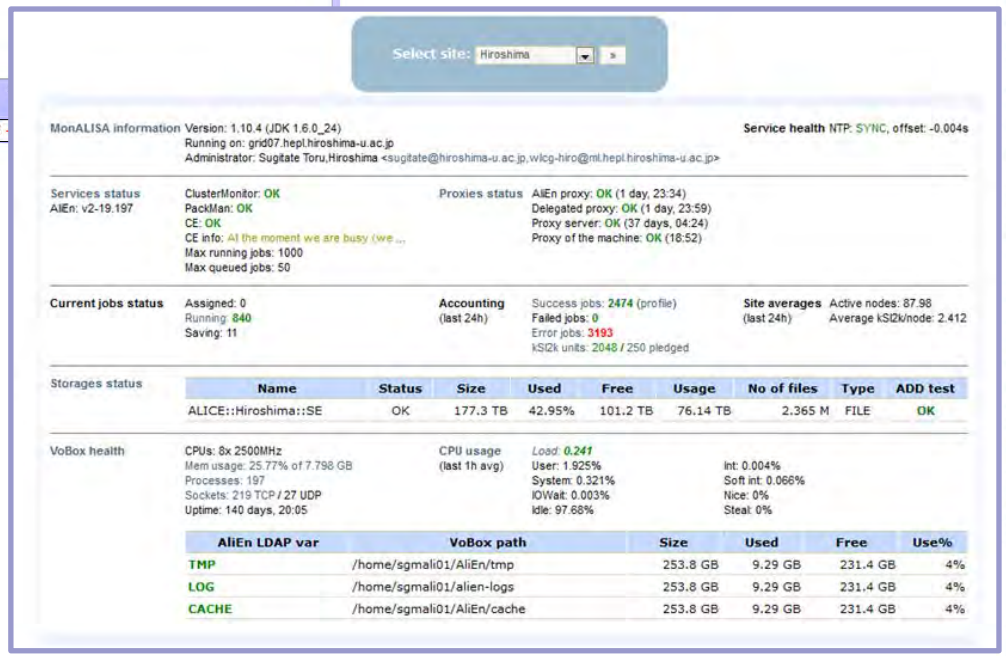
- Secure/Robust subnets
- 10Gbps local connection (part.)
- Space/Energy saving



JP-HIROSHIMA-WLCG



- 864 Xeon-cores in CRAEM-CE.
- Xrootd disk storage of 177 TB in 4 disk servers.
- Stably accepting over 800 jobs and process around 5000 jobs a day,
- corresponding to about 2% contribution to entire ALICE.



Network Speed and Traffics

Hiroshima »

Links: FDT, Kernel p

‹Hiroshima›

Chart view »

IN from

No.	ID	Site	Speed (Mbps)	Hops	RTT (ms)	Streams
1.	1386285	NERSC	302.01	18	153.77	1
2.	1253777	CCIN2P31	167.61			1
3.	1385029	IPNL	167.61			1
4.	1385003	CCIN2P3	150.85			1
5.	1385042	LUNARC	150.85			1
6.	1457825	LLNL	142.47			1
7.	1384919	NIHAM	117.45			1
8.	1384932	Legnaro	108.95			1
9.	1384942	Trujillo	100.67			1
10.	1384921	CyberSar_Cagliari	100.37			1
11.	1407787	Talca	92.28			1
12.	1384945	NECTEC	92.10			1
13.	1384924	CSC	83.89			1
14.	1384955	Subatech	83.89			1
15.	1384947	Athens	75.50			1
16.	1385013	Cyfronet	75.43			1
17.	1388160	NIPNE	75.43			1
18.	1384990	Troitsk	75.43			1
19.	778391	GSI-SCLAB	58.72			1
20.	1254150	Wuhan	58.72			1
21.	1385027	CERN-L	58.66			1
22.	1384985	KFKI	58.66			1
23.	1388123	Catania	58.61			1
24.	1384961	GSI-SGE	50.33			1
25.	1266811	PDC	50.33			1
26.	1385017	IHEP	50.18			1
27.	918809	UCT				
28.	1385001	UNA				
29.	1384952	RRC				
30.	1385010	SUT				
31.	1384993	Mad				
32.	1384929	PAK				
33.	1384964	JINR				
34.	1359842	Saol				
35.	892029	CSIF				
36.	1385025	Kolk				
37.	1388095	Tori				
38.	818602	KNU				
39.	804893	WIT				
40.	1385046	GRIF				
41.	1384996	MEP				
42.	1385019	ISM				
43.	1384981	OSC				
44.	1384998	RAL	8.38			1

OUT to

No.	ID	Site	Speed (Mbps)	Hops	RTT (ms)	Str
1.	1266070	PDC	150.85	16	313.72	
2.	491471	Houston	34.89	15	171.57	
3.	878783	KPI	16.76	13	342.07	
4.	434196	KISTI	9.03	7	120.83	
5.	1385322	LBL	8.39	17	152.89	
6.	1384997	Trujillo	8.39			
7.	1385028	Bari	8.38	18	326.59	
8.	888336	CSIR	8.38	16	537.85	
9.	1384978	KISTI-CREAM	8.38	15	92.17	
10.	1457792	LLNL	8.38	14	145.70	
11.	1385036	Oxford	8.38	17	290.49	
12.	1384989	Subatech	8.38	18	316.62	
13.	1385023	Kolkata-CREAM	8.37	20	367.91	
14.	1385038	SUT	8.37	13	171.32	
15.	1407821	Talca	8.37			
16.	452428	CESGA	3.06	16	356.79	
17.	347924	Muenster	1.70	22	323.84	
18.	436378	Kolkata	0.69	14	501.93	
19.	1385014	Athens		16	328.33	
20.	1360590	Birmingham		16	307.37	
21.	1385002	BITP		23	339.92	1
22.	1384995	Bologna				
23.	1384954	Bratislava				
24.	1384928	Cagliari				
25.	1388109	Catania				
26.	1384959	CCIN2P3				
27.	1253750	CCIN2P31				
28.	1385008	CERN-CREAM				
29.	1384944	CERN-L				
30.	1384980	Clermont				
31.	1384992	CNAF				
32.	1384917	COMSATS				
33.	1385012	CSC				
34.	1385037	CyberSar_Ca				
35.	1384927	Cyfronet				
36.	1385006	DCSC_KU				
37.	632384	Dortmund				
38.	1385016	FZK				
39.	1384922	Grenoble				
40.	1384983	GRIF_IPNO				
41.	1384931	GRIF_IRFU				
42.	1384949	GSI				
43.	1188731	GSI-CREAM				
44.	774092	GSI-SCLAB				1

Similar tests in the past

Tests from **CERN-L** to **Hiroshima**

No.	ID	Speed (Mbps)	Hops	RTT (ms)	Streams
1.	1385027	58.66			1
2.	1360378	75.43			1
3.	1353280	16.76			1
4.	1347201	83.81			1
5.	1336042	8.39			1
6.	1330539	8.38			1
7.	1325945	33.56			1
8.	1320611	41.90			1
9.	1315129	75.43			1
10.	1310273	67.05			1
11.	1305259	75.50			1
12.	1300335	50.28			1

Tracepath for test 1384959

Tracepath from **Hiroshima** to **CCIN2P3**

Hop	IP	RTT (ms)	Domain	AS
0	202.13.220.47	0	ac.jp	as2505
1	202.13.220.62	0.18		
2	202.13.220.250	0.65		
3	202.13.223.49	0.69		
4	202.13.223.114	22.36		
5	202.13.223.113	29.06		
6	150.99.190.181	34.31	ad.jp	as2907
7	150.99.2.50	34.13	ad.jp	as2907
8	150.99.2.74	204.25	ad.jp	as2907
9	150.99.188.202	279.43		
10	62.40.112.9	298.54	geant2.net	as20965
11	62.40.112.162	330.19		
12	no_reply			
13	193.51.189.150	313	renater.fr	as2200
14	no_reply			
15	193.51.186.181	324.09	renater.fr	as2200
16	134.158.224.1	312.24	in2p3.fr	as789
17	134.158.108.110	324.06	in2p3.fr	as789

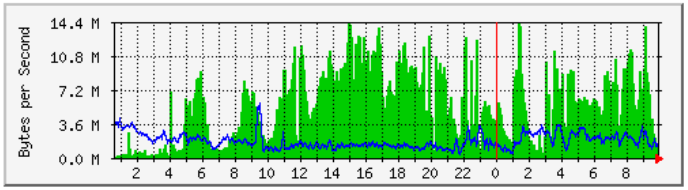
Tracepath OK, target reached

port13 -- hepfw

System: hepfw in
Maintainer:
Description: port13
ifType: ethernetCsmacd (6)
ifName: port13
Max Speed: 125.0 MBytes/s
Ip: 202.13.220.249 (hepfw.hepl.hiroshima-u.ac.jp)

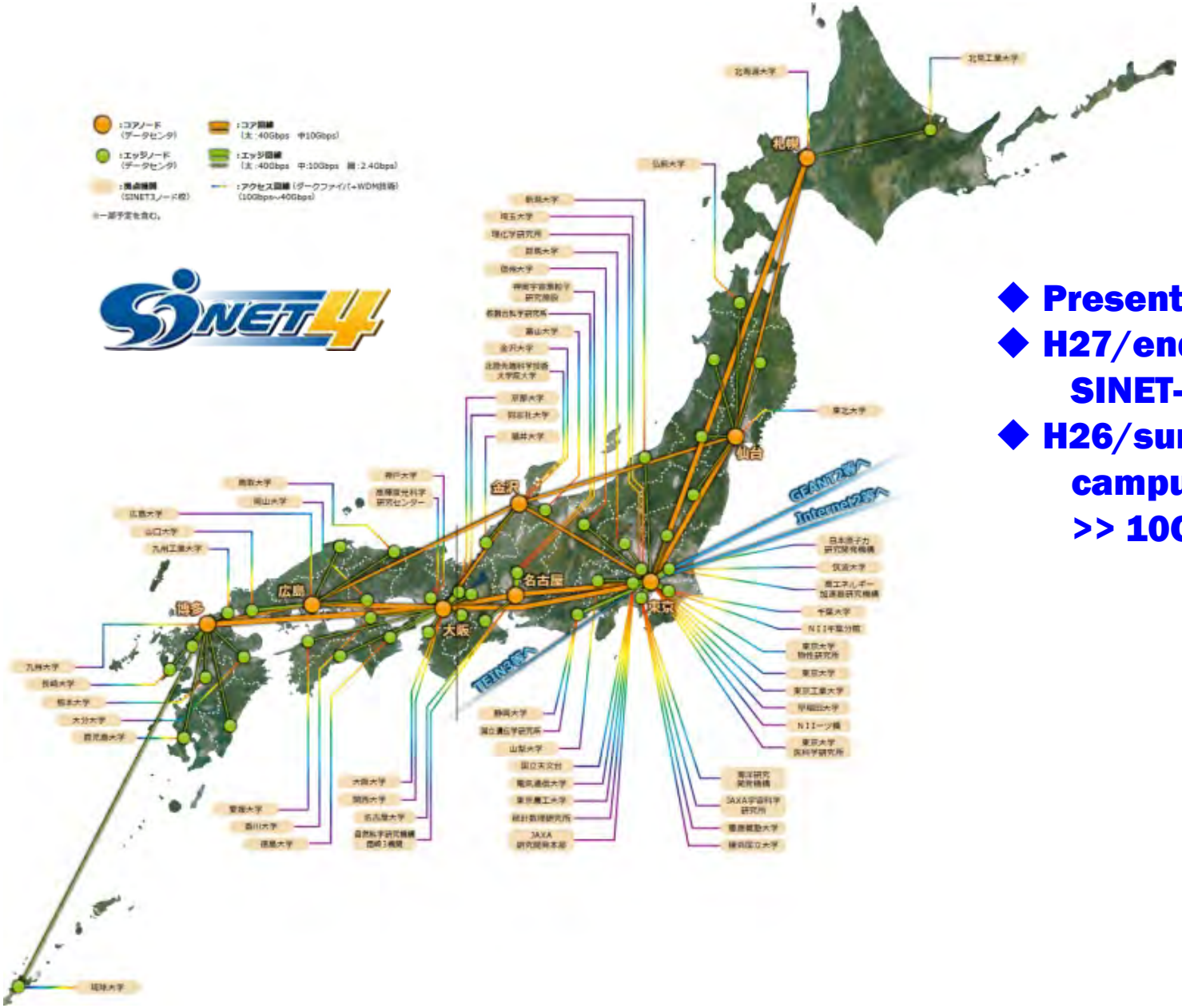
The statistics were last updated **Saturday, 16 February 2013 at 9:55**, at which time **'hepfw'** had been up for **24 days, 4:21:43**.

`Daily' Graph (5 Minute Average)



- Network speed varies from time to time?
- Slow connection from/to the world, especially the out-going traffic?
- Appreciate any help to tune/train the network routing/performance!!

SINET-4 to SINET-X



- **Present: 1Gbps-MPLS**
- **H27/end:**
  SINET-4 >> SINET-X
- **H26/summer:**
  campus LAN replace
  >> 10Gbps to the DC

Computing Resources Scrutiny Group 2012

ALICE has historically faced a lack of computing resources. Partly because of insistence by the CRSG, ALICE lowered its requests in order to better reflect the anticipated resources. However, in 2012 .....

	Fair share (based on M&O A sharing)			Pledged resources (2012 data from REBUS)		
	CPU (K HEPSPEC06)	Disk (TB)	Tape (TB)	CPU (K HEPSPEC06)	Disk (TB)	Tape (TB)
Member States	162	16223	11300	164	13059	11523
Δ (required to pledged)				1%	-24%	2%
Non Member States	102	10231	0	121	6845	0
Δ (required to pledged)				16%	-49%	-
Sub Total	264	26454	11300	285	19904	11523
Δ (required to pledged)				7%	-33%	2%
CERN	0	0	0	90	8100	20000
Sum T1	120	10800	21000	76	6197	11523
Δ (required to pledged)				-58%	-74%	-82%
Sum T2	145	15800		214	14203	-
Δ (required to pledged)				32%	-11%	-
Sum T1&T2	265	26600	11300	290	20400	11523
Δ (required to pledged)				9%	-30%	2%

- **T2s rescue T1s for CPU !**
- **Large disk and tape deficit in T1s**
- **After LS2 :**
  - **Tape +163 PB (× 3)**
  - **Disk +125 PB (× 2)**

Fair Share of Computing Resources

MB has approved "The Funding Agency will provide computing resources (CPU and disk) in a quantity greater than or equal to the fraction of the total resources required, minus the pledged CERN contribution, in proportion to its M&O-A contribution relative to the total ALICE M&O-A minus the CERN M&O-A.… The computing shares approved by the Computing RRB will then become the minimal resource requirements for each institution."

Country	Fair share (based on M&O A sharing)			Pledged resources (2012 data from REBUS)			Status
	CPU (K HEPSPEC06)	Disk (TB)	Tape (TB)	CPU (K HEPSPEC06)	Disk (TB)	Tape (TB)	
Armenia	1,0	97	0	0	0	0	No pledges
Brazil	2,4	244	0	1	140	0	Under pledged
CERN	0,0	0	0	90	8100	20000	#DIV/0!
China	1,9	195	0	2	6	0	Under pledged
Croatia	2,4	244	0	0	0	0	No pledges
Czech Republic	6,3	633	0	5	450	0	OK
France CEA + IN2P3	26,7	2679	4529	28	2445	800	OK
Germany	28,6	2874	2965	47	3200	5250	OK
Greece	1,5	146	0	0	0	0	Under pledged
Hungary	1,9	195	0	1	72	0	Under pledged
India	19,9	1997	0	6	240	0	Under pledged
Italy INFN + Centro Fermi	52,4	5262	8894	49	4100	3000	OK
Japan	4,9	487	0	4	118	0	Under pledged
Korea NRF	5,8	585	0	26	1050	0	OK
Mexico	5,8	585	0	8	500	0	OK
Netherlands	5,3	536	906	6	279	292	Under pledged
Nordic	18,0	1803	3047	15	1480	1761	OK
Pakistan	1,5	146	0	0	50	0	Under pledged
Peru	0,5	49	0	0	0	0	No pledges
Poland	9,7	974	0	5	325	0	Under pledged
Romania	5,3	536	0	16	1240	0	OK
Russia	24,3	2436	0	18	1301	0	Under pledged
Slovak Republic	4,9	487	0	4	320	0	Under pledged
South Africa	3,4	341	0	12	100	0	Under pledged
Spain/Cuba	2,4	244	0	3	208	0	OK
Thailand	0,0	0	0	2	100	0	#DIV/0!
Ukraine KIPT + Kiev	1,5	146	0	1	150	0	OK
United Kingdom	3,9	390	659	2	180	420	Under pledged
USA DOE+NSF	22,8	2290	0	24	1850	0	OK

ALICE Upgrade Plan

ALICE Upgrade: target LS2 (2018)

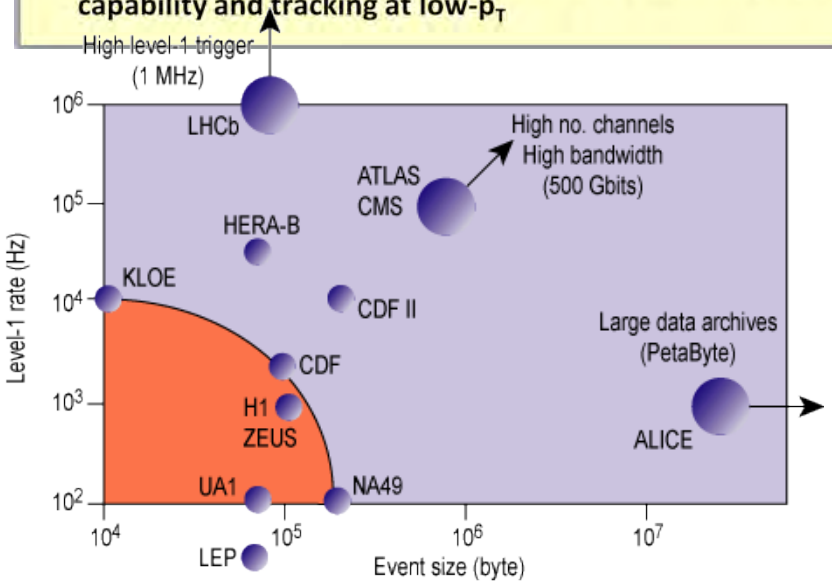
Ugrade ALICE for the last 3 years of the approved program and extend it for about three more, after LS3

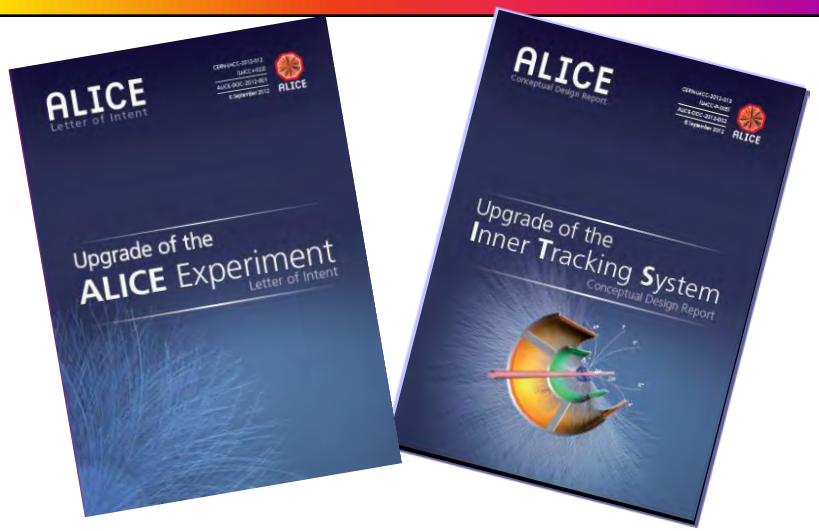
- Primary scope:
  - precision studies of charm and beauty mesons and baryons and charmonia
  - low mass lepton pairs and thermal photons
  - gamma-jet and jet-jet with particle identification from low momentum up to 30 GeV.
  - heavy nuclear states

➢ low-transverse momentum observables (complementary/orthogonal to the general-purpose detectors)
  - not triggerable => need to examine full statistics.

- **Operate ALICE at high rate while preserving its uniqueness, superb tracking and PID, and enhance its secondary vertex capability and tracking at low-p_T**

19





Experimental Strategy

- run ALICE at 50kHz Pb-Pb (i.e. L = 6×10^{27} $cm^{-1}s^{-1}$), with minimum bias (pipeline) readout (max readout with present ALICE set-up ~500Hz)
  - Gain a factor of 100 in statistics over current program: x 10 integrated luminosity, $1nb^{-1}$ => 10 nb^{-1}, x 10 via pipelined readout allowing inspection of all collisions, namely inspect O(10^{10}) central collisions instead of O(10^8)
- improve vertexing, and tracking at low p_t
- This entails :
  - New, smaller radius beam pipe
  - New inner tracker (ITS) (scope and rate upgrade)
  - High-rate upgrade for the readout of the TPC, TRD, TOF, CALs, DAQ/HLT, Muons and Trigger detectors
- Furthermore, three major proposals are under consideration by the collaboration to extend the scope of ALICE (decision in September): VHMPID, MFT, and FoCal
  - new high momentum PID capabilities
  - b-tagging for J/ψ, low-mass di-muons
  - low-x physics with identified γ/π^0

By PG in CB in Oct. 2012

22

Summary

- ALICE GRID
  - Process 50k jobs in 7 T1's and about 80 T2's
  - Need more CPU/Disks (x2) /Tapes (x3)
    - Eg. to Japan; 180TB >> 500TB. << pledge later
  - Pressure to Japan to sign up WLCG << pending
  - Another pressure to migrate a Torrent type software << Refused!!
- Hiroshima Tier-2
  - Accepts over 800 jobs stably, and process around 5000 jobs a day,
  - corresponding to about 2% contribution to the entire ALICE.
  - Hiroshima is now under migration of EMI-2 on SL64, except VOBOX.
  - Trace network and tune up speed to increase productivity, but…
    - Merits of 1st Asian Tier-1 unseen yet
  - Need "expertized" manpower / understanding inside collabration / cooperation with outside groups
  - Declare a 10Gbps connection to the DC in SINET-X
- LCG operation
  - Thanks to EGI for their frequent requests to up-to-date.
    - SLC4 >> SLC5 >> SL6
    - gLite 3.1 >> gLite 3.2 >> EMI-1 >> EMI-2
    - Many security patch requests
  - Thanks to 創夢（株） for their remote SE.

Sites under SAM/GGUS by EGI

MyEGI
Service Availability Monitoring Portal

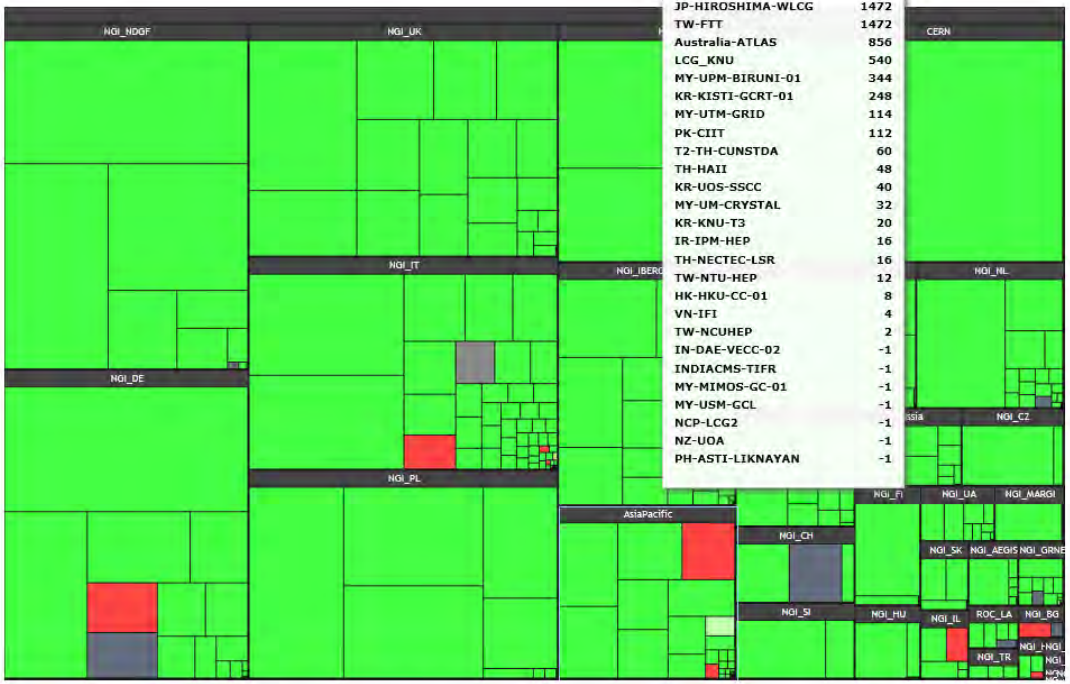
Home | Status | Availability & Reliability | Treemap | Monthly Reports | Topology

VO: ops
Profile: ch.cern.sam-WLCG_CREAM_LCGCE_CRITICA
Data View: Status
Topology View: Region
Size By: Logical CPUs
Date/Time: 05-04-2013 07:58:09
Resolution: Ho … Monthly

hide filters
Update Page

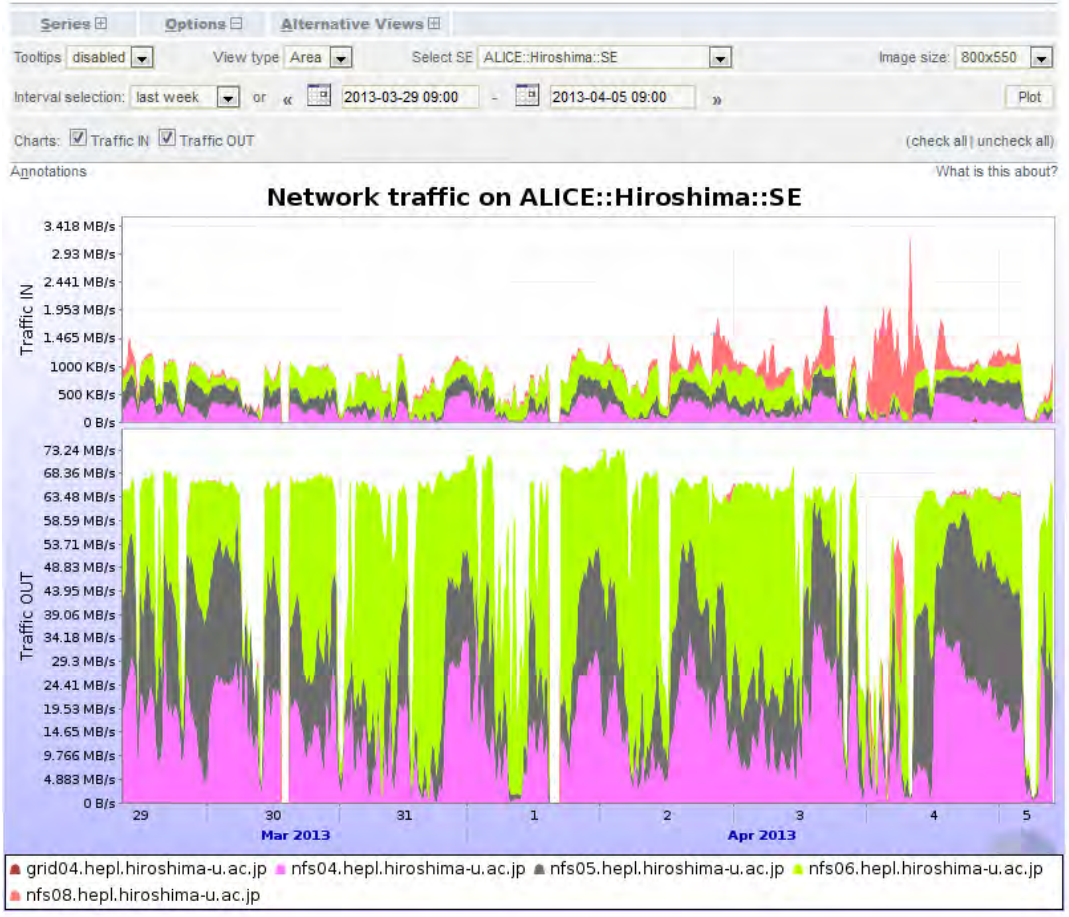
AsiaPacific
Logical CPUs: 16504

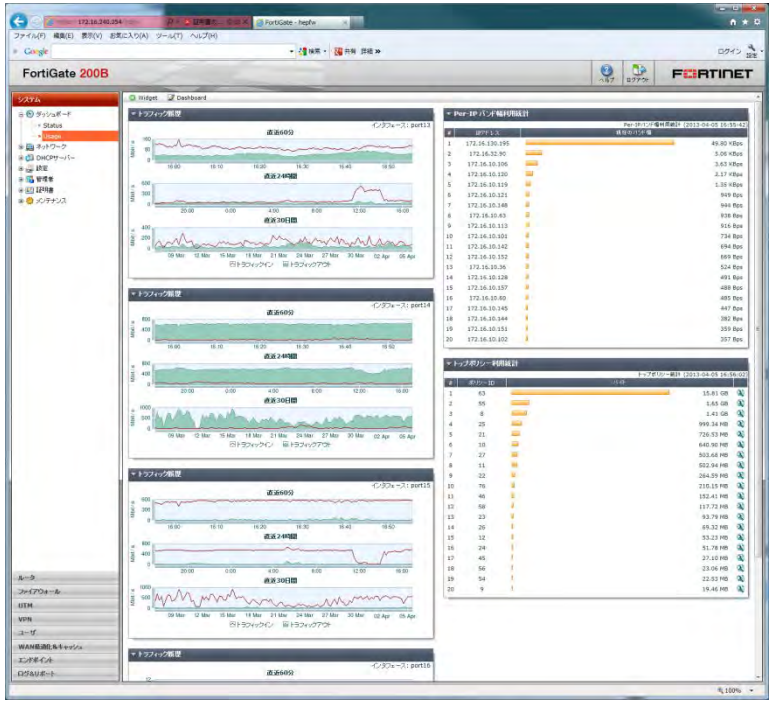
JP-KEK-CRC-02	2988
TOKYO-LCG2	2560
Taiwan-LCG2	2208
TW-eScience	1808
KR-KISTI-GSDC-01	1512
JP-HIROSHIMA-WLCG	1472
TW-FTT	1472
Australia-ATLAS	856
LCG_KNU	540
MY-UPM-BIRUNI-01	344
KR-KISTI-GCRT-01	248
MY-UTM-GRID	114
PK-CIIT	112
T2-TH-CUNSTDA	60
TH-HAII	48
KR-UOS-SSCC	40
MY-UM-CRYSTAL	32
KR-KNU-T3	20
IR-IPM-HEP	16
TH-NECTEC-LSR	16
TW-NTU-HEP	12
HK-HKU-CC-01	8
VN-IFI	4
TW-NCUHEP	2
IN-DAE-VECC-02	-1
INDIACMS-TIFR	-1
MY-MIMOS-GC-01	-1
MY-USM-GCL	-1
NCP-LCG2	-1
NZ-UOA	-1
PH-ASTI-LIKNAYAN	-1



Legend: OK | Warning | Critical | Downtime | Unknown | Missing | N/A | Removed

Documentation | Web Services | Release Notes | Tests | Test Web Services | Bug Report | About
Operations Dashboard | GSTAT | Accounting Portal | GGUS


Series
Options
Alternative Views
Tooltips disabled
View type Area
Select SE ALICE::Hiroshima::SE
Image size: 800x550
Interval selection: last week or
2013-03-29 09:00
2013-04-05 09:00
Plot
Charts: Traffic IN Traffic OUT
(check all | uncheck all)
Annotations
What is this about?
Network traffic on ALICE::Hiroshima::SE
Traffic IN
3.418 MB/s
2.93 MB/s
2.441 MB/s
1.953 MB/s
1.465 MB/s
1000 KB/s
500 KB/s
0 B/s
Traffic OUT
73.24 MB/s
68.36 MB/s
63.48 MB/s
58.59 MB/s
53.71 MB/s
48.83 MB/s
43.95 MB/s
39.06 MB/s
34.18 MB/s
29.3 MB/s
24.41 MB/s
19.53 MB/s
14.65 MB/s
9.766 MB/s
4.883 MB/s
0 B/s
29
30
31
1
2
3
4
5
Mar 2013
Apr 2013
grid04.hepl.hiroshima-u.ac.jp
nfs04.hepl.hiroshima-u.ac.jp
nfs05.hepl.hiroshima-u.ac.jp
nfs06.hepl.hiroshima-u.ac.jp
nfs08.hepl.hiroshima-u.ac.jp


FortiGate 200B
FORTINET