WLCG operations

ALICE T1-T2 Workshop
Strasbourg
May 3-5, 2017

Maarten Litmaath
CERN-IT

v1.1

Central services – 2016

Machine	Machine status			Machine type								Mem	
Machine	**Online**	**Uptime**	**Load**	**Kernel**	**OS**	**Machine model**	**CPU**	**CPUs**	**MHz**			**Total ▲**	**Used**
8. db6c		9d 14:17	5.09	4.4.0-72-...	16.04	ProLiant DL380 Gen9	Xeon E5-2667 v4 3.20GHz	32	3502	...		1.476 TB	179.6 GB
9. db6d		2d 14:24	1.23	4.4.0-75-...	16.04	ProLiant DL380 Gen9	Xeon E5-2687W v3 3.10GHz	40	1364			755.8 GB	154.1 GB
7. db6b		58d 6:34	1.35	3.19.0-82...	15.04	ProLiant DL380 Gen9	Xeon E5-2697A v4 2.60GHz	64	2089	...		755.8 GB	141.5 GB
12. db6g		50d 7:51	0	4.4.0-66-...	16.04	ProLiant DL380 Gen9	Xeon E5-2687W v3 3.10GHz	40	1200			755.8 GB	15.63 GB
11. db6f		53d 8:14	2.35	4.4.0-65-...	16.04	ProLiant DL380 Gen9	Xeon E5-2687W v4 3.00GHz	48	3131	...		755.8 GB	170.7 GB
15. db9		113d 4:50	0.3	3.13.0-10...	14.04	ProLiant DL380p Gen8	Xeon E5-2690 2.90GHz	32	2900			377.9 GB	25.16 GB
13. db7		114d 6:23	1.03	3.13.0-10...	14.04	ProLiant DL380p Gen8	Xeon E5-2690 2.90GHz	32	2900	...		377.9 GB	7.667 GB
37. pcaliweb02		114d 9:35	0.01	3.13.0-10...	12.04	ProLiant DL380p Gen8	Xeon E5-2690 2.90GHz	32	2901			377.9 GB	5.965 GB
41. pcalimonitor4		114d 9:35	0.85	4.4.0-57-...	16.04	ProLiant DL380p Gen8	Xeon E5-2690 2.90GHz	32	3363	...		377.9 GB	7.132 GB
39. pcalimonitor2		114d 9:34	0.76	4.4.0-57-...	16.04	ProLiant DL380p Gen8	Xeon E5-2690 2.90GHz	32	2007			377.9 GB	26.51 GB
38. pcalimonitor		114d 9:32	17.82	4.4.0-57-...	16.04	ProLiant DL380p Gen8	Xeon E5-2690 v2 3.00GHz	40	3000	...		377.9 GB	96.04 GB
29. alienvm1		114d 2:11	6.62	4.4.0-57-...	16.04	ProLiant DL380p Gen8	Xeon E5-2690 v2 3.00GHz	40	3000			377.9 GB	59.59 GB
10. db6e		114d 9:34	0.39	4.4.0-57-...	16.04	ProLiant DL380 G7	Xeon X5680 3.33GHz	24	3333			283.4 GB	25.11 GB

- Mostly stable, a few incidents
  - DB locks, connection leaks, MySQL hangs
  - API server slowdown due to catalog bug
    - Mostly fixed mid 2016, the remainder hit us early 2017 and got fixed then

- Unavailability Sep 14-15 due to big network intervention
  - All user and grid activity had to be stopped
  - In parallel the File Catalog was moved to a new, more powerful machine!
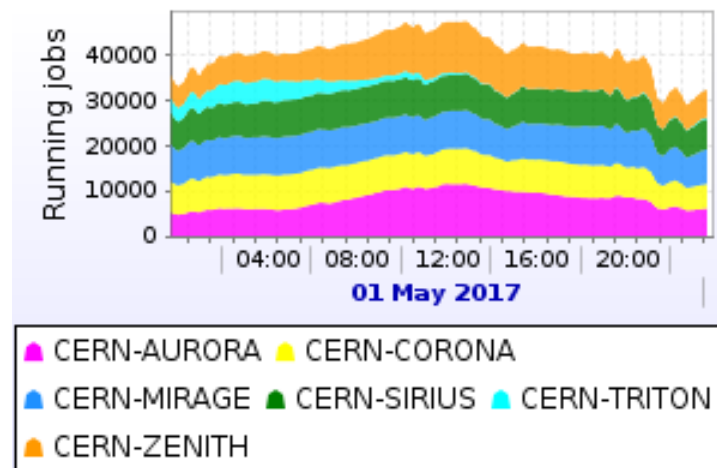
Central services – 2017

- Jan 4 – yearly scheduled power cut lasted much longer than usual, way beyond the UPS battery lifetime
  - It took ~2 days to deal with all the fall-out

- Jan 14 – AliEn CA certificates of various services expired
  - Now the CERN Grid CA is used, with automatic reminders

- Week of Feb 27
  - Authen instabilities due to remnant of catalog bug
  - Broken RAID controller slowing down the Task Queue
  - MySQL crashes due to a corruption – the master DB had to be restored from the slave

- Week of March 13 – catalog unexpectedly became slow in the compression of various tables
  - Not clear what caused this or how it got cured exactly, but the configuration ended up better tuned as a result

- April 14-15 – multiple MySQL crashes due to corruptions
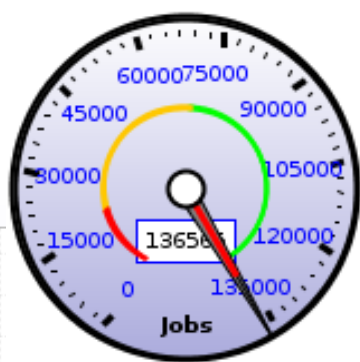 - Finally attributed to a broken expansion slot cache

Site evolution

- New site: UPB (Bucharest)

- Univ. of Derby: in preparation

- New CERN "sites" for redundancy and load sharing, submitting to the steadily increasing HTCondor resources (currently 60% vs. 40% LSF)
  - CERN-CORONA
  - CERN-MIRAGE

- FZK / KIT
  - Second "site" for ARC resources
  - Original site also switched to ARC after CREAM CE decommissioning

- Prague
  - Second "site" for ARC resources

- Torino
  - Opportunistic use of HPC cluster

- LBNL
  - New proof-of-concept cluster for future migration

- Altaria: a virtual site for various cloud resources
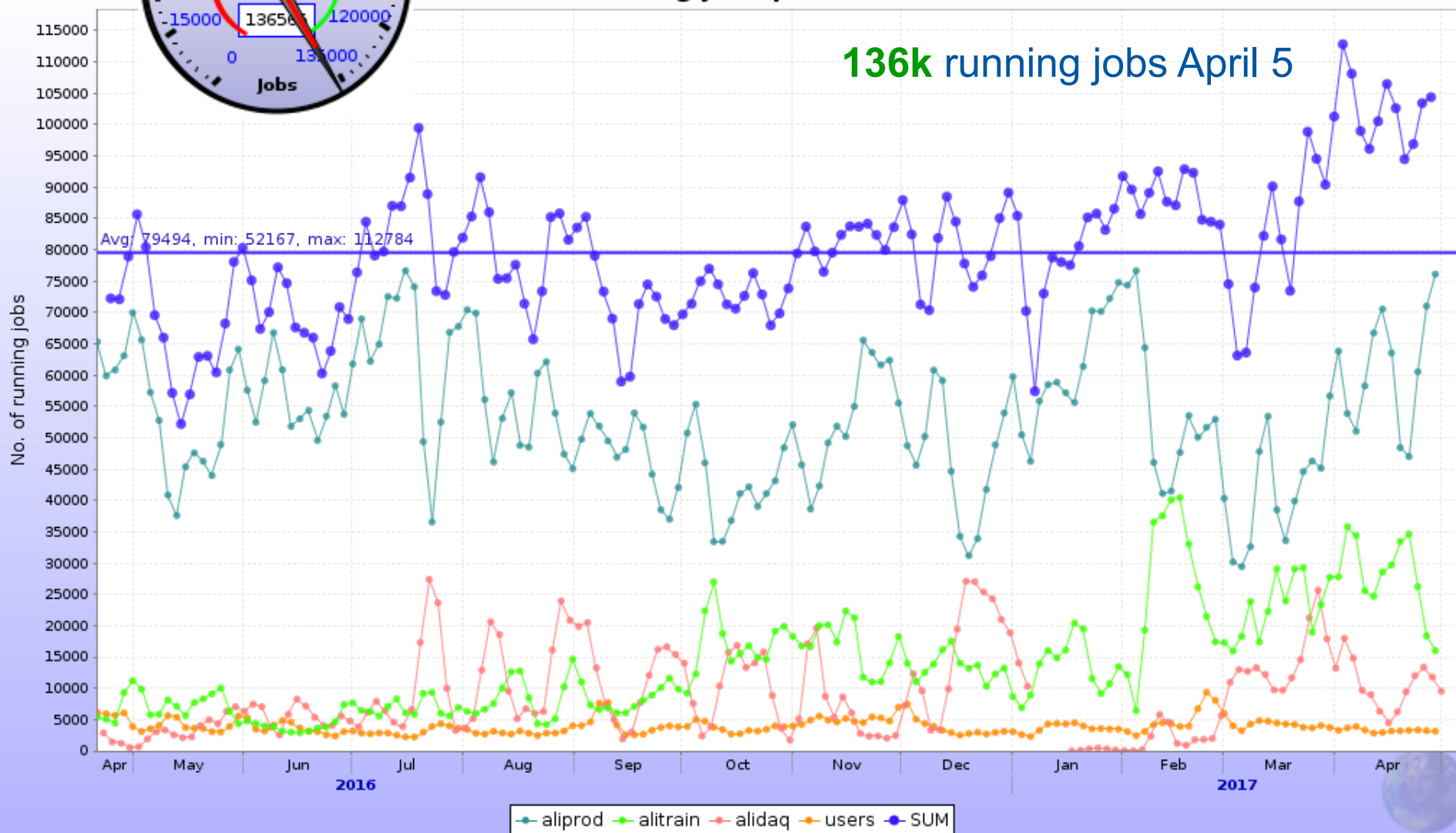 - Details further in this presentation

High activity

Running jobs per user

136k running jobs April 5

Avg: 79494, min: 52167, max: 112784

No. of running jobs

Legend: aliprod — alitrain — alidaq — users — SUM

Storage

- In general: growth and reliability – thanks!

- CERN: CASTOR
  - Mostly stable, a few incidents
  - Recently overloaded by reco jobs at CERN *plus* T1 sites
    - Cured by limiting the concurrent transfers + request queuing

- CERN: EOS
  - Mostly stable, a few incidents
  - Recent disk server additions to keep up with very high activity!

- New EOS instances at Hiroshima, NIPNE, ZA_CHPC

- Please consider EOS for your next SE solution
  - A first EOS workshop was held Feb 2-3
  - Check out the presentations!

Issues at sites or with jobs   (1)

- **CERN: multiple incidents affecting CREAM CEs**
  - Currently the gateways to 40% of the resources at CERN
  - One time it was our fault: the sandbox partitions filled up with debug logs (fixed)

- **CERN: multiple incidents affecting the new HTCondor CEs**
  - Currently the gateways to 60% of the resources at CERN
  - CE configurations were improved for robustness and scalability
  - Some bugs were encountered, worked around and investigated with help from the HTCondor developers

- **Asian sites: CVMFS issues involving the Stratum-1 at ASGC**
  - Worked around by overriding the default configuration
    - The ASGC Stratum-1 will be removed from it
  - And by the new AliEn version v2-19.395
    - The Job Agent checks the CVMFS status of the WN to try and ensure that the user payload dependencies will be available

Issues at sites or with jobs   (2)

- **CERN-AURORA upgrade to broken AliEn version**
    - Non-trivial upgrades will be announced better

- **KISTI: CVMFS Squid services incident**
    - "Random" job failures for 5 days

- **High rates of expired jobs for a particular MC campaign**
    - They were using much more memory than most jobs
    - Several sites complained
    - At CNAF we had to switch to the high-memory queue
        - With half the usual number of job slots
    - The campaign finished nonetheless and experts managed to reduce the memory consumption to normal levels for future cases
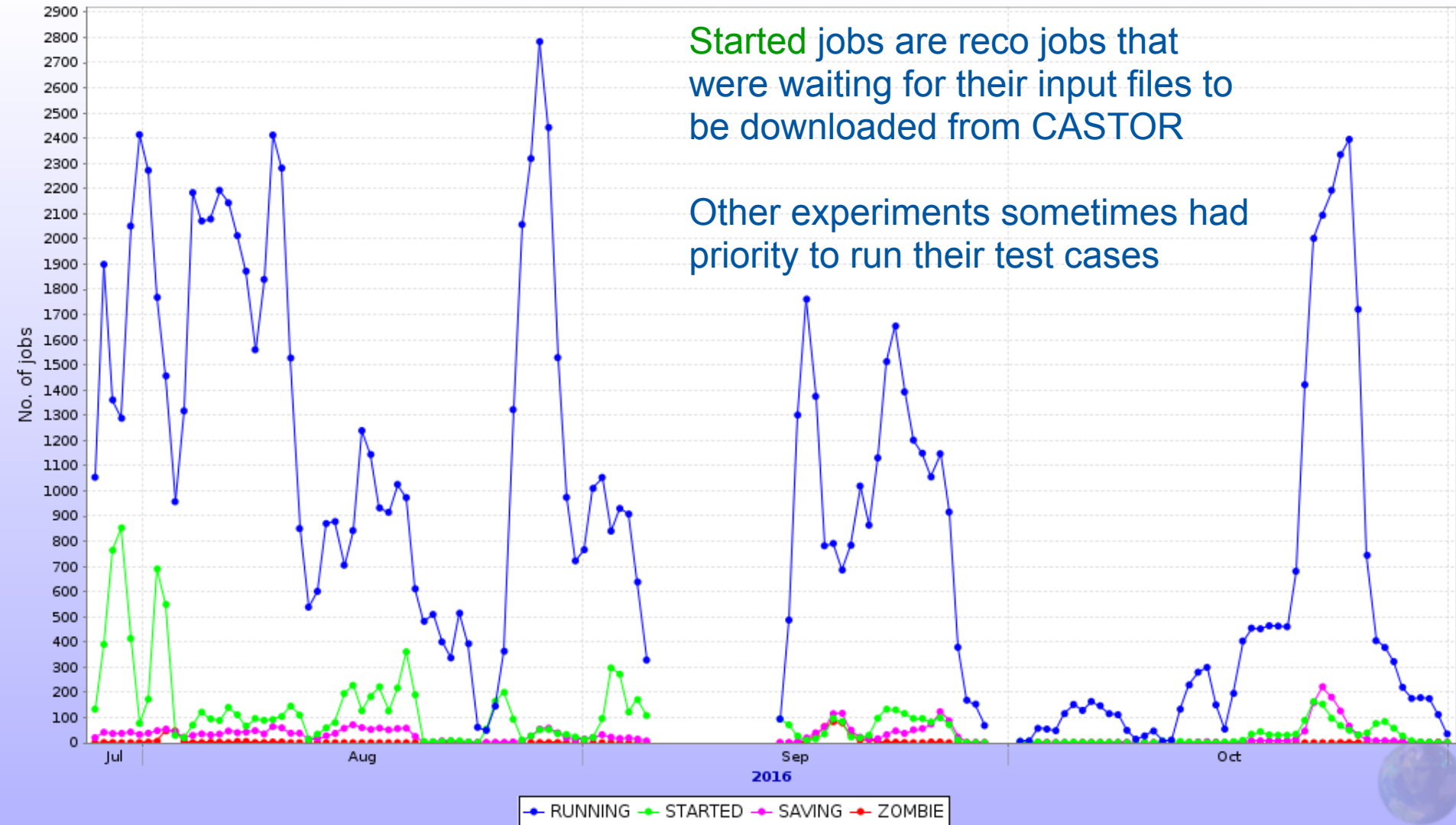
Altaria   (1)

- A virtual site to drive cloud resources

- First big use case: a temporary extension of CERN resources located in a cloud at T-Systems, Germany
  - A common project to test how cloud resources can be procured and used efficiently

- Up to 3500 cores were available, shared by the 4 experiments
  - ALICE were the first to start using them as of July 28, up to Oct 31 when the project ended

- The resources were used for reco, MC and user analysis
  - Trains were excluded to reduce the load on the 10-Gbit link to CERN

- The exercise has helped the IT cloud team gain useful insights
  - And allowed 630k ALICE jobs to complete successfully!
  - The error types and rates were compatible with those of the CERN "sites"

- The next, bigger phase will start in the coming months
  - Coordinated by **Helix Nebula – The Science Cloud**

Active jobs in Altaria

Started jobs are reco jobs that were waiting for their input files to be downloaded from CASTOR

Other experiments sometimes had priority to run their test cases

Altaria (2)

- Currently being used in a proof of concept for sites that want to provide their resources via cloud instead of traditional grid mechanisms

- In particular the UK T2 sites are moving to that model
  - Manchester (first ALICE jobs since 2008)
  - Liverpool (first ALICE jobs ever?)
  - Birmingham
  - Oxford

- Cloud VMs are configured such that they connect to an HTCondor pool at CERN to which Altaria submits its jobs
  - For monitoring and accounting it may be desirable to have an HTCondor pool per site, hosted on its own VOBOX

- The VMs are managed by the sites, AliEn just sees resources appear as if they were WN job slots
  - Managed e.g. through Vac or Vcycle

Middleware

- CentOS/EL7 on the rise
    - A number of service types already available in UMD-4
    - CREAM, EMI-UI and EMI-WN expected by May
        - WLCG VOBOX will follow

- SL6 still the default, but the experiments have been preparing for physical worker nodes running CentOS/EL7
    - Containers (or VMs) could still provide SL6
    - ALICE jobs have been tested OK on CentOS/EL7

- SL6 in UMD-4 only has officially supported products

- SL6 in UMD-3 may still be the easiest to use for certain services
    - Mind that new features only go into UMD-4

Containers

- An isolation paradigm much lighter than VMs
  - See the session on containers in the April GDB

- A new tool to launch containers is gaining momentum in our community and beyond: Singularity
  - US-CMS already use it at their T2 sites
  - Provide SL6 environment on CentOS/EL7 WN
  - And isolate each user payload from other processes → the plan is to let our Job Agents use it where available
    - No timeline yet
  - It currently needs to be *setuid root* because the kernel requires root for *mount namespace* operations
    - A future EL update (tentatively 7.4) is expected to have that requirement still by default, but configurable through sysctl

SAM

- New Availability / Reliability profile based on selected MonALISA metrics in use since 1 year

- So far no big issues were reported

- Reminder: SE test failures will reduce the A / R!

- Corrections have been applied as needed

- Test job submission to the HTCondor CE has been added to the preprod instance
  - Production to be updated soon

Xrootd reminder

- Sites should continue upgrades to Xrootd >= 4.1
  - Most sites have done that already, thanks!
  - Required for IPv6 support

- Mind that v4.6.0 has some issues (463, 465)
  - Should be fixed in v4.6.1

- Communication via LCG Task Force list as usual for expert advice

- ALICE add-ons are available through rpms
  - http://linuxsoft.cern.ch/wlcg/
  - Thanks to Adrian Sevcenco!

Networks

- ## WLCG IPv6 deployment plans

  - CERN + T1 sites should at least have some production storage dual-stack since April 2017

  - Most sites *ought* to have their production storage dual-stack by April 2018

  - ALICE needs ~all storage to be dual-stack before IPv6-only computing can be used

- ## 2nd Asia Tier Center Forum held Nov 30 – Dec 2

  - Hosted by SUT, Thailand – thanks!

  - Fostering network evolution in Asia – well done!

Tips for sites – thanks!

- Possible issues on VOBOX, CE, WN
  - CVMFS problem, CE not ready for jobs, myproxy running low, myproxy type wrong, …
  - Absence of "system" library
    - HEP_OSlibs rpm helps avoid that

- Jobs may fail due to SE problems

- Admins please check site issues page
  - http://alimonitor.cern.ch/siteinfo/issues.jsp

- Subscribe to relevant notifications
  - http://alimonitor.cern.ch/xml.jsp