

KISTI Site Report

2nd Asia Tier Center Forum @ SUT in Thailand

30 Nov 2016 ~ 2 Dec 2016

Sang-Un Ahn

Global Science experimental Data hub Center

Korea Institute of Science and Technology Information

Contents

- KISTI Tier-1 Center
- Operations
- Plan

CPU

- Batch system
 - Torque w/ Maui scheduler
 - 3 different sub-clusters (3 CREAM-CE)
- Job capacity
 - 3,488 concurrent job slots
 - 84 machines w/ 32 cores (HT) + 20 machines w/ 40 cores (HT)
 - 3 or 3.2 GBytes per job

Disk

- Management system
 - XRootD : 1 redirector + 10 servers
 - SAN (Storage Area Network) storage is attached as block devices through multipath
- Storage capacity
 - 1.5 PBytes
 - EMC VNX (migrated from EMC Clariion)

Tape

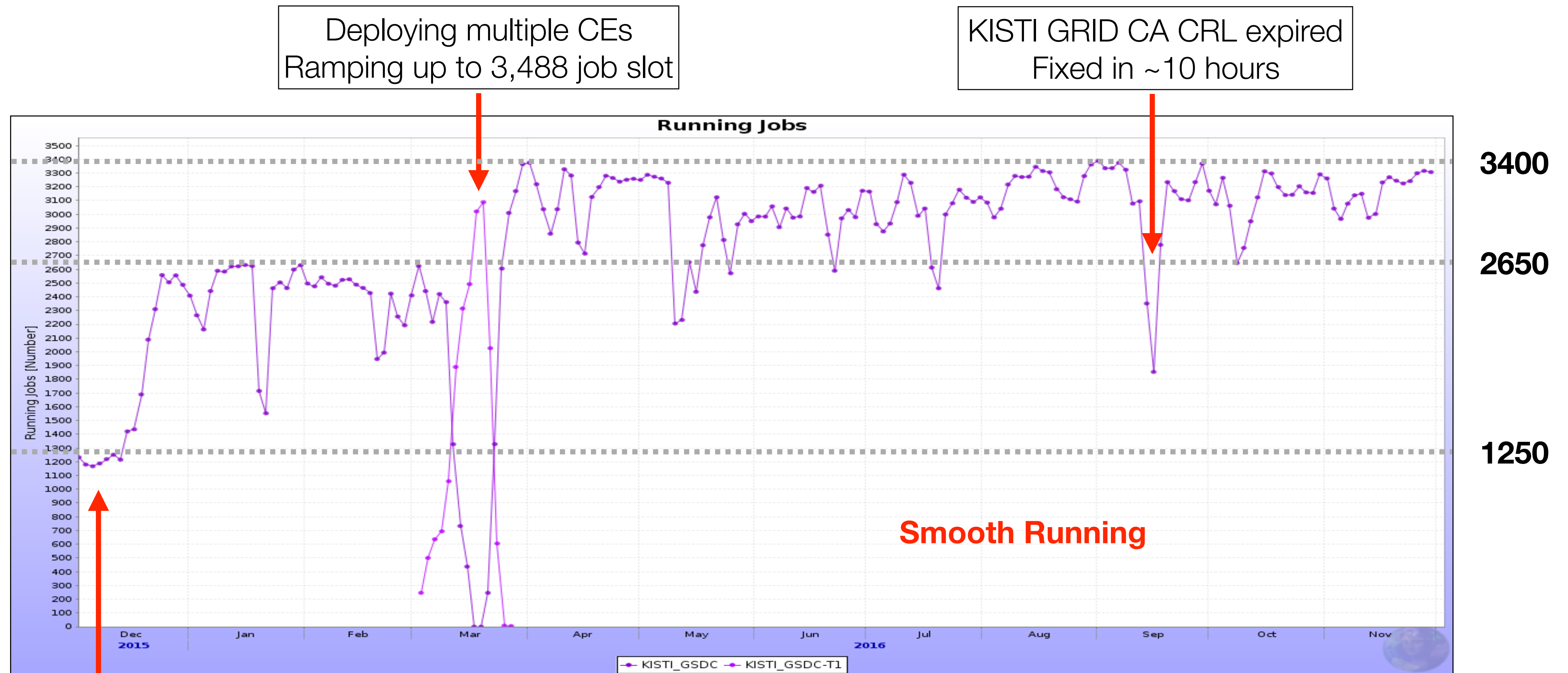
- Capacity : 3.0 PBytes
- IBM TS3500
- Tape frames have been extended this year to host more cartridges up to 3 PBytes
 - Total 5 tape frames + 2 robot arms

Network

- LHCOPN
 - Dedicated 10Gbits/s link to CERN via GLORIAD-KR and SURFnet
 - 10Gbits/s bandwidth is provided by GLORIAD-KR up to Chicago
 - See Dr. B. Cho's talks for remarkable upgrade of GLORIAD-KR
- IPv6
 - Fully enabled in backbone networks
 - All services for Tier-1 will be operated in dual stack mode by this year

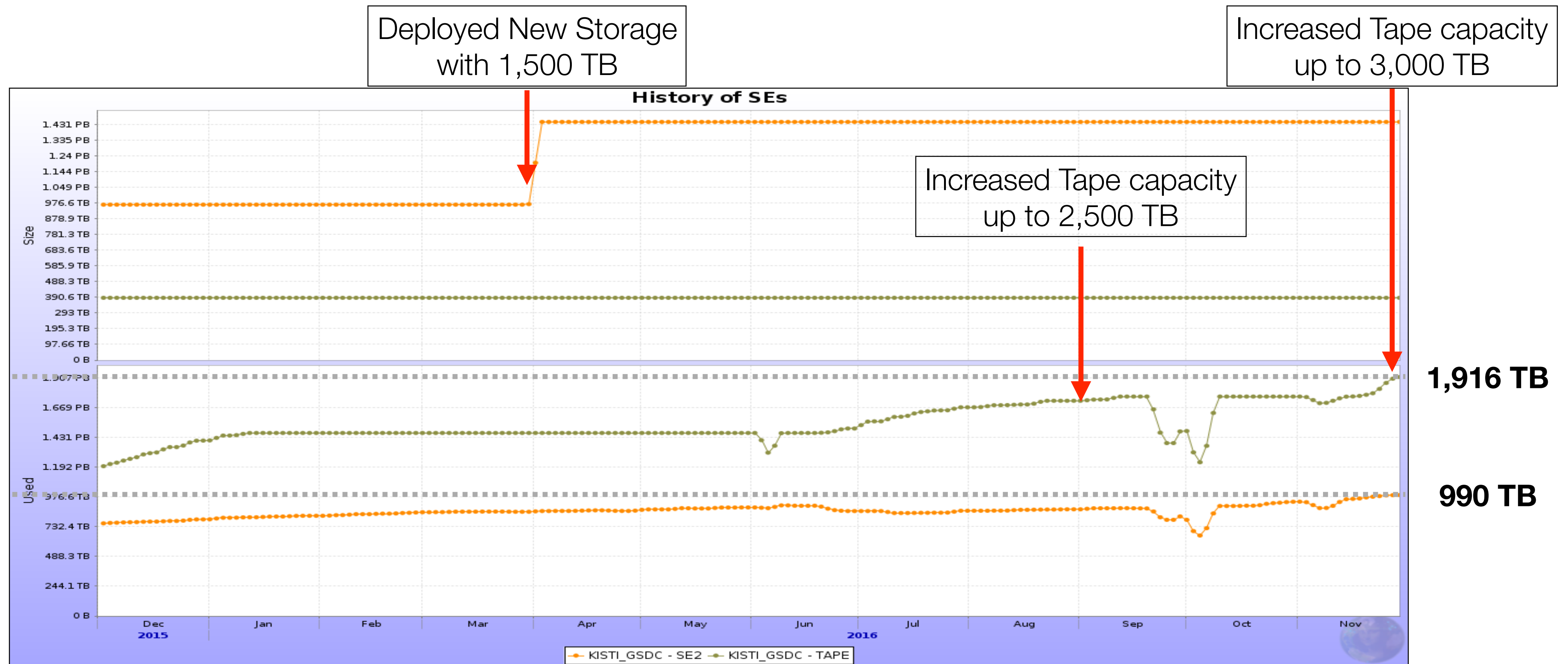
Operations

Resource Usage (1/2)



Providing 6GBytes per job
for HI-run

Resource Usage (2/2)



Maintenance

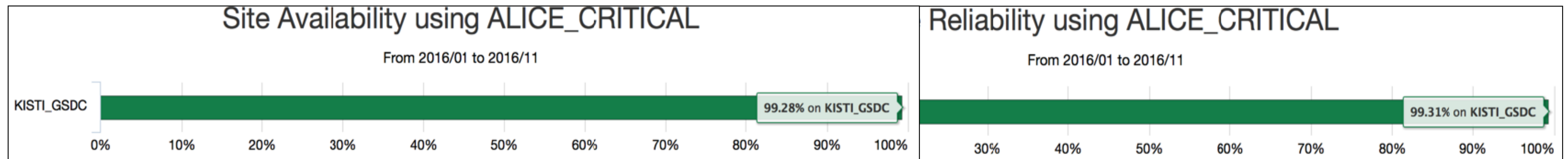
- Disk storage migration
 - Almost transparent with few hours of downtime
 - Snapshot of storage was taken through rsync from old to new machines (~1 week)
 - Last minute update was done with downtime (~1 hour)
- Tape frame extension
 - 36 hours downtime was declared for tape redirector
 - Additional tape frame was installed and ~500 TBytes cartridges were activated

Incident

- LHCOPN link was unavailable in February
 - Switching to backup link was failed for few hours
 - Normally the process is done within timeout; operators were absent due to national holidays (summoned and fixed quickly)
- KISTI GRID CA CRL (Certificate Revocation List) was expired in September
 - Manual CA CRL (lifetime ~ 30days) update was failed due to overlapped national holidays (human error)
 - KISTI GRID CA was rejected for ~10 hours
- Scheduled downtime for tape redirector has not been taken into account
 - Reported to SAM experts and requested A/R corrections (thanks to Maarten), Manual fix placed

Availability/Reliability

- Overall availability in 2016
 - Over 99% despite of a few incidents
 - Thanks to
 - Prompt actions of operators and administrators
 - Wonderful migration planning for storage



Plan

Batch System

- HTCondor
 - In the framework of Batch System Integration project:
 - GSDC RIM (Resource Integration Management)
 - Moving from Torque to HTCondor (One batch system for all)
 - MESOS based computing resource management
 - HTCondor-CE will be investigated
 - Support from ALICE will definitely be needed

Stratum-1

- For CernVM-FS and BDII
 - Currently relying on non-Asian repositories: CERN, FNAL, RAL, etc.
 - Adding more (and reliable) replica for Asian sites

Pledges

- For RUN2
 - Recent C-RRB approved 20-30% increased resources requested by experiments due to exceptional operations of LHC at RUN2
 - WLCG sites are asked to act on that by the end of November
- For RUN3 and HL-LHC
 - We will try to do our best for WLCG

	2015	2016	2017
CPU(HS06)	28,000	31,000	36,000
DISK	1,500	1,500	1,500
TAPE	1,500	3,000	3,000


 The logo for the ALICE experiment, consisting of the word "ALICE" in white capital letters inside a brown rectangular box.

CMS

- Tier-2
 - We will start setting up Tier-2 for CMS in 2017
 - Size of site will be similar to the current CMS Tier-2 in Korea (KNU)
 - CPU = 10,000 HS06, Disk = 500 TB
 - MSIP and NRF decided to replace the current one by KISTI

Thank you