IllinoisHEP Tier3gs

David Lesny

Senior Research Physicist

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

IllinoisHEP T3gs Site

- IllinoisHEP T3gs Site
  - Associated with Atlas
  - Mark Neubauer, Assistant Professor, member of Tier 3 Task Force
  - Started building about 1.5 years ago
  - Most of site is re-tasked hardware from CDF
  - Funding provided by DOE, Campus Research Board, NCSA Fellowship
  - Campus Research Network provided by CITES
  - Began Atlas Production on March 18, 2009

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

IllinoisHEP T3gs Site Configuration

- Full configuration of Site Services
    - CE, SE, DDM, Worker, GUMS, Service nodes
    - CE is OSG 1.0.0, VDT 1.10.1y
    - SE is dCache 1.8 installed with VDT
    - DDM is LFC, DQ2 site services with a MySQL server
    - Eight Worker nodes with OSG 1.0.0 client
    - GUMS also installed using VDT
    - OS on all nodes is Scientific Linux 4.7
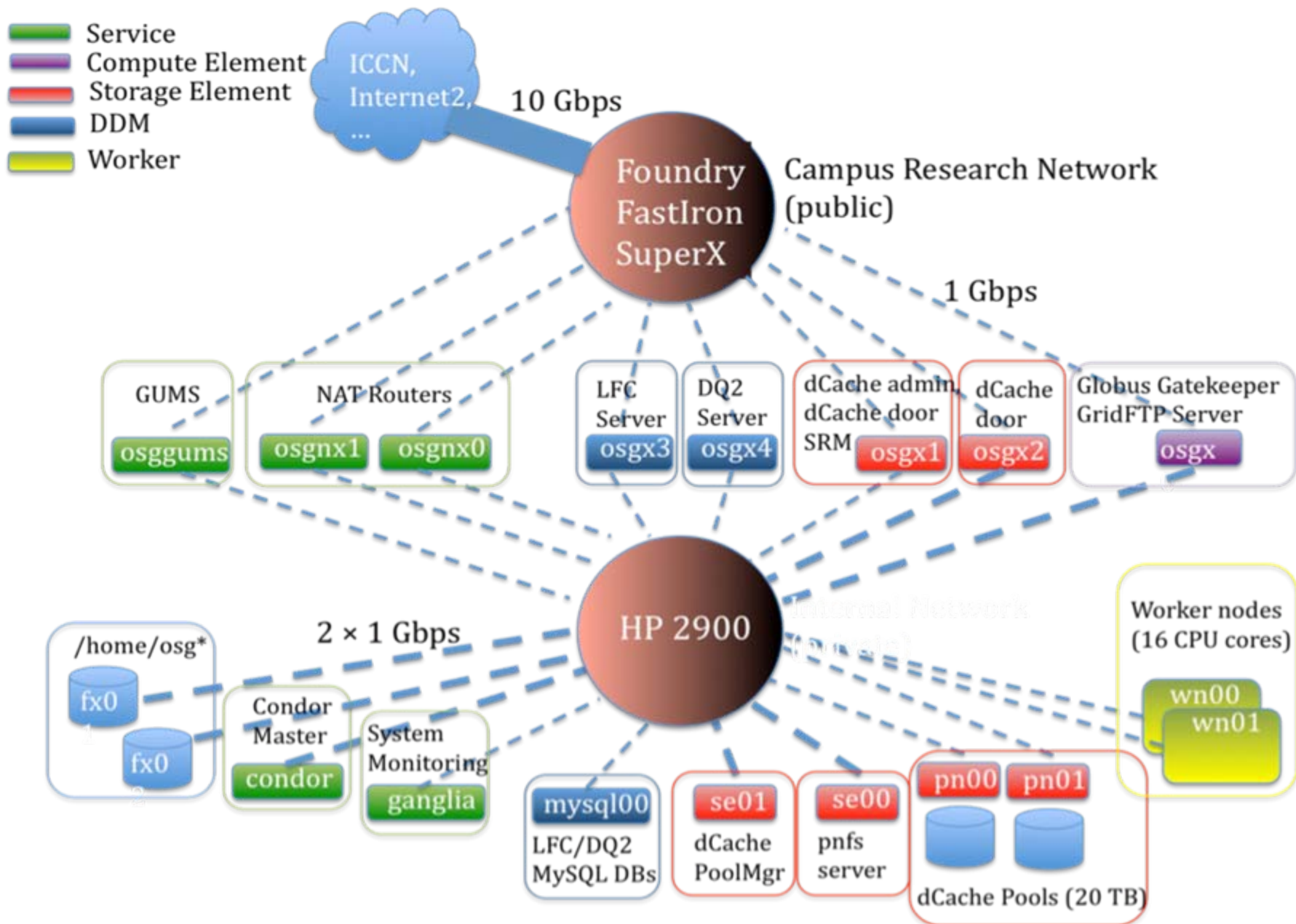
IllinoisHEP T3gs dCache Installation

- dCache 1.8 installed via VDT 1 year ago
  - Hope to update to 1.9 soon

- Picked dCache because
  - Mark has dCache experience with CDF
  - 1 year ago, BeStMan did not appear ready
  - Large vs small storage needs
  - Seems easy to add dCache pools

- Currently 6 nodes in use
  - Two doors on public via 1Gb, dual 1Gb bonded to private
  - Poolmanager and pNFS servers dual 1Gb bonded to private
  - Doors, managers and servers are dual Xeon 2.0Ghz, 6GB
  - Pools nodes are old PIII 1Ghz with 1Gb memory

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

IllinoisHEP T3gs Pool Nodes

- Two Pool Nodes…..
    - Pentium III 1Ghz, 1GB memory, 1Gb NIC
    - Dual host DAS Promise VTrak M610p via Adaptec 29160
    - Had most of this on hand
    - Got 20TB online pretty cheap
    - M610p not fast, but neither is PIII or 1Gb NIC (seems balanced)
    - Tops out at about 50MB/sec (ganglia)
    - Not the best for management, especially if one has many

I L L I N O I S
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

IllinoisHEP T3gs Pool Usage

- Current usage is 100% PanDA
  - 10TB for PRODDISK, 4TB for USERDISK
  - 1TB for MCDISK and DATADISK (not used)
  - How are these areas cleaned (PROD manual, USER automatic?)
  - Production is using about 98% of our site

- Performance
  - With 64 worker node cores, almost no load on SE
  - Could replicate pool nodes very easily to get more storage but….
  - Expect bottle necks when we add more cores and more user analysis
  - Bottle neck most likely will be doors and then pool nodes
  - With network and server upgrades that would move to Raid system
  - Looking at other storage solution (iSCSI, Fiberchannel, other DAS)

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

IllinoisHEP T3gs Networking

- Our access to Foundry FastIron SuperX 1Gb is very limited
  - CITES will not allow bonding
  - Only given 8 IPs on the Campus Research Network
  - Restricts us on doors, speed to CRN, etc
  - All Worker nodes access dCache via NATs and Doors
  - Available 10Ge port on SuperX
  - Checking into adding our own 10Ge switch

- HP 5400 or 6400 to get more 10Ge CX4 ports
  - Upgrade CE, Doors and NATS to newer servers
  - 10Ge ports to each server
  - Vlans on switch for public and private networks
  - Connect to our HP2900 with 10Ge port as well

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

IllinoisHEP T3gs Monitoring

- How do I monitor our site
  - Use as many tools as I can find
  - MyOSG
  - Gratia
  - RSV on CE with SE probes
  - dCache web
  - Panda Monitor
  - Logs, Logs and more Logs
  - Email from others when there is a problem

- Troubleshooting
  - Twiki pages (documentation changing for the better)
  - OSG and Panda distribution lists and ticket system
  - Lean on others (I have been a pest)

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Summary

- Huge learning curve for the whole T3
    - When started, documentation poor, but much better now
    - Had to lean on a lot of people to get operational
    - Lots of pieces are still black boxes

- Know that some pieces of hardware inadequate
    - Is dCache the right choice or should we switch to BeStMan
    - Amount of storage and worker node cores limited by budget
    - Expect to have around 256 cores later this year
    - Unknown how much disk, but answer is as much as we can

Thank you to all who have helped me get our T3gs operational

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN