VMs as Site Services for T3s

Alden Stradling

May 19, 2009

Why you need DQ2 Site Services in a Tier3

(even though you don't plan to accept Grid jobs)

- To use a Tier3, you must have ntuples locally
- Data are in Grid storage. You have to get those data to the T3 to work with them!
- Options -- hire and train a grad student (this is beyond a lot of undergrads) or have a bunch of automated subscriptions via DQ2 Site Services

US ATLAS

UT ARLINGTON

ATLAS Configuration

- Not too onerous:
  - TiersOfAtlas entry (ToA)
  - FTS Channel at BNL (Contacting Hiro)
  - Host certificates for SRM and LFC (can be generated by Pat)
  - Fixed IPs at your institution (3 or so)
  - Entry in the Panda DB, if you want jobs to run

Basic Needs -- Boxen

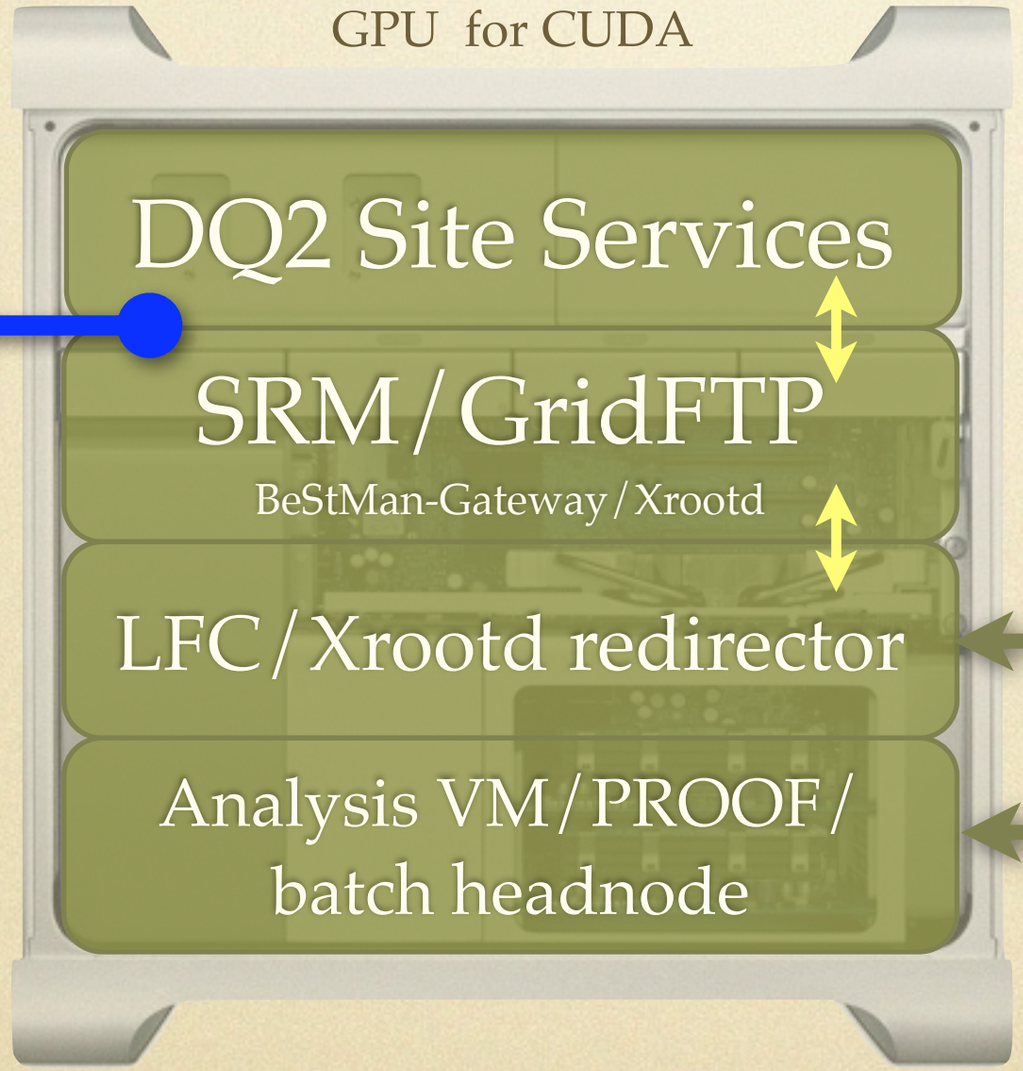
(small site assumption)

- DQ2 Site Services Installation: Heavy DB usage
- SRM (Storage Resource Manager): Heavy *other* DB usage (a completely different MySQL version)
- GridFTP (can be combined with the SRM, for now)
- LFC (Logical File Collection): Lighter and easier
- Storage Backend

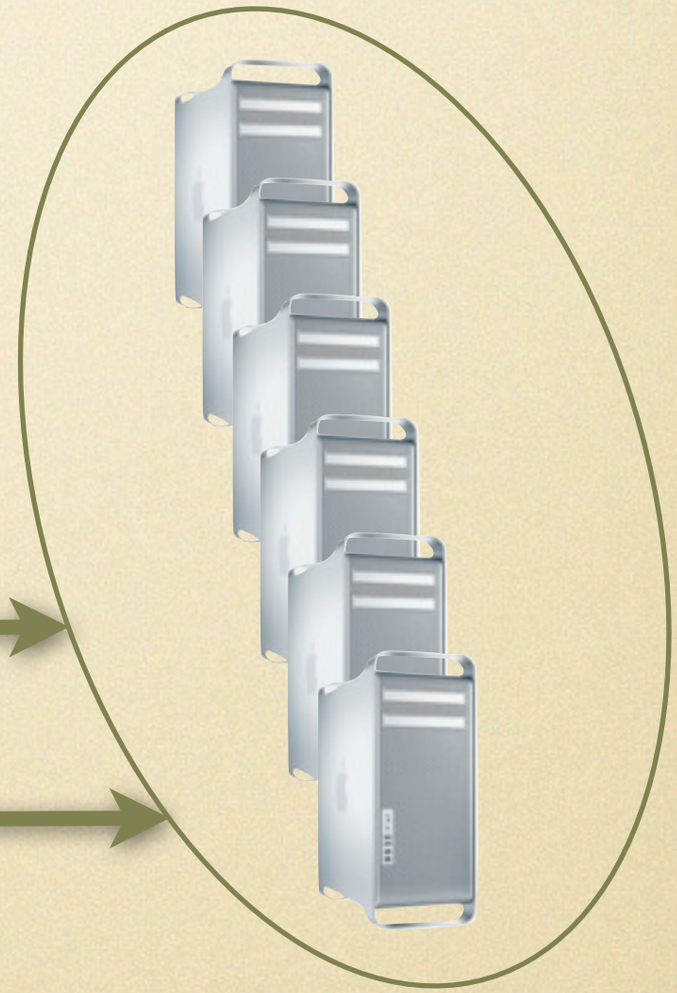
Pictures > Words

8 Core (i7) DDM Server

SSDs for Xrootd caching
GPU for CUDA



Cluster of standard worker machines
Whatever flavor (or mix) you want
Desktop or racked -- multicore helps
Stuffed with disks for Xrootd storage



6 TB local disk

Attractions of Virtualization

- Nice to do it in VMs:
  - **Prepackaged solutions**
  - Easy to recombine/break out services as needed
  - Easy to do handoffs on the same hardware
  - Hardware compatibility forever

SRM Storage Backend

- Starting with just normal file transfers (BeStMan) -- successfully done last night. Subscriptions are working!
  - Even that was a major effort, several hours with an expert doing the heavy lifting
- Will now upgrade to single-machine Xrootd filesystem (BeStMan Gateway), then n-machine Xrootd with multiple disks per machine
  - Substitute for inflexible RAIDs
  - Requires another box (redirector)
    - *But that's just another lightweight VM!*

LFC -- File Catalog

- Relatively easy to install and bring up (meaning, a few hours of work, trial and error with an expert)
- Mostly a database and a front end -- can probably be combined with an Xrootd redirector when the time comes

DQ2 Site Services

- It's now simpler to install: all connections to SRM and LFC are defined at the ToA level
- But what a monster!
  - Very finicky about various releases and RPM versions
  - Almost no good documentation
- It's a heck of a thing to install and maintain yourself
- Possible problem with a production role requirement -- under investigation

First Look -- Today!

- I have pulled together site services running on three VMs (VMware, custom SLC4 kernel) to cover the basic site services

- Very rough, but we subscribed to a dataset and it transferred!

- More refined (beta-like) version will be released for download in the next couple of weeks.

The Ideal

(what I intend to have running here)

- Site installable in 3 hours by a grad student following instructions
- Attach to an array of machines stuffed with disks, fill them with subscriptions and let ‘em be used by students with ideas
- Minimal maintenance -- wouldn’t that be nice
- VMs prevent a site from needing to buy and place a bunch of special boxen for site services
- Maintenance/upgrades a few hours instead of a few days.

Next Steps

- Apply an Xrootd backend, spreading the files across multiple nodes for easy student and faculty access (PROOF farm), and tune performance/topology
- Run Panda/pathena locally on my own T3
- Jump that T3 to a cloud provider (Nimbus or Amazon) and run there, with full DDM site services -- test practicality
- Accelerate file access with SSD cacheing in Xrootd
- Further accelerate access with GPUs (still a ways off)

Long-term

- DQ2 should allow subscriptions via client tools (as it has in the distant past) to get around needing all this monkeying around
  - But: not happening any time soon
- EVERY T3 will need this kind of access -- there is no other way to proceed

Acknowledgements

- Many thanks to Patrick McGuigan, Hironori Ito, Mark Sosebee and Amir Farbin