High Availability Planning for Tier-2 Services

Shawn McKee/AGLT2

USATLAS Tier-2/Tier-3 Workshop

Ann Arbor, May 27th, 2008

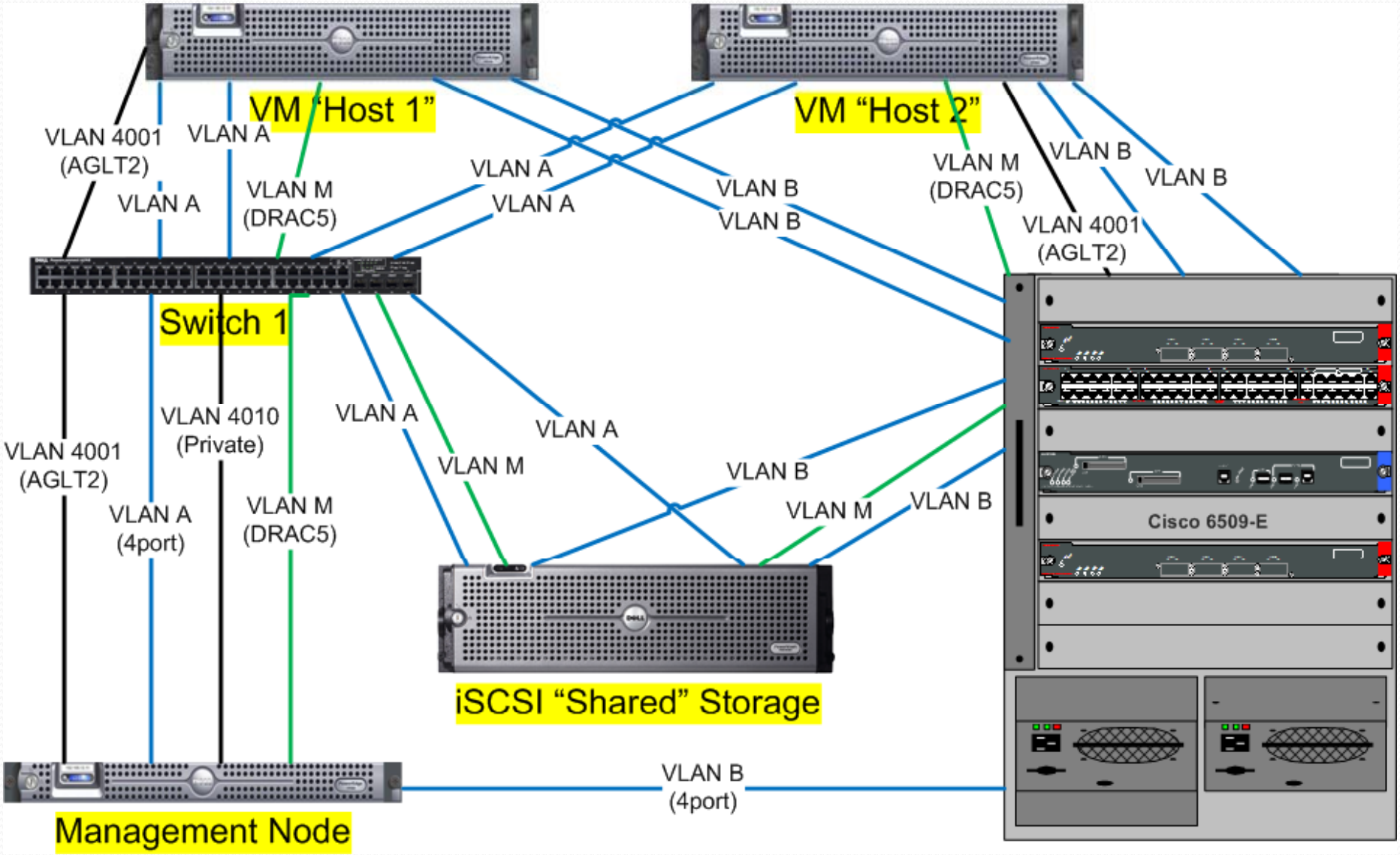
Motivation

- Within the USATLAS facilities we have ever increasing Tier-2 capability, soon to average >1000 processors / Tier-2
- Loss of critical services can total disrupt running jobs and the overall system
- Wouldn't it be nice to have a more "robust" system for enabling our critical services?
- At AGLT2 we are looking at both **VMware ESX** and **Xen Enterprise** as possible solutions for a base system
- **Robust service options could be layered on top…**

Hardware Level

- We need to build a resilient(redundant) hardware basis to support the higher-level services
- To do this we ordered a matched set of servers, an iSCSI backend storage system and a management node
- All systems have redundant power supplies, network connections, network switches, RAID5 or RAID 1 disk configurations and UPS based power.
- High-availability is easiest. Loss of one physical server causes the other to instantiate any missing servers.
- Fault-tolerance (“resilience”) is more difficult…

Diagram of AGLT2 Test Platform



Concepts for Fault Tolerance

- Make sure stateful services store information in "shared" storage areas
- Use techniques like DRBD (Distributed Replicated Block Device) to distribute critical information between independent locations
- Evaluate/modify server applications for their ability to fail-over to another running system instance without losing their context.
- **GOAL: Run a resilient Globus Gatekeeper instance which will not lose its running jobs if the underlying physical node crashes or goes offline.**

Summary

- Need to test/evaluate both VMware and Xen to determine which can provide us the most effective environment
- The FNAL folks have started evaluating resilient grid services (see: http://www2.twgrid.org/event/isgc2008/Presentation%20Meterial/Operation%20&%20Management/Operation%20&%20Management_Eileen%20Berman.pdf) We need to test these configurations on our systems.
- We want a very robust basis to build our critical grid services on: **GUMS, Globus Gatekeeper, Condor headnode, Kerberos, NIS, AFS, dCache, etc**.