High Availability Technologies for Tier2 Services

June 16th 2006

Tim Bell
CERN IT/FIO/TSI

Agenda

- MoU Levels
- Procedures
- High Availability approaches

Tier 2 Service Levels in MoU

Function	Maximum delay – prime time	Maximum delay – other periods	Avail
End user analysis facility	2 hours	72 hours	95%
All other services	12 hours	72 hours	95%

Conclusions

- No weekend coverage
- No overnight piquet
- Administrators can go for lunch

Downtime from a failure

Failure Occurs	Something breaks
Failure Detected	Latencies due to polling of status
Failure Noticed	Console, E-Mail, Siren,…
Investigation Started	Login, have a look
Problem Identified	Root cause found
Procedure Found	How to solve it
Problem Solved	Execute the procedure
Restore Production	Cleanup

MoU is not very ambitious

- 95% uptime
  - 8 hours / week down
  - 17 days / year down
- Does not cover impact of failure
  - Black holes / Recovery / Retries
  - Problem Analysis
  - Glitch effects

Coverage

- Working Time = 60 hours / week
  - 08:00-18:00
- Dead time
  - Meetings / Workshops
  - No checks before morning status reviews and coffee
  - Illness / Holidays

Changes

- New release needed rapidly
  - Security patches
  - Interface changes
- Slow time to drain
  - 1 week for jobs to complete
  - 1 week proxy lifetime
- Many applications do not provide drain or migrate functionality
  - Continue to serve existing requests
  - Do not accept new requests

So why do anything special ?

- Deliver resources for lower costs
  - Lower availability means lower contribution to LCG
- Direct Site Comparisons
  - Publicity
- Administrator Stress
  - Avoid urgent interventions
- Change Handling without downtime
  - Draining
- Happy Users

People and Procedures

- Problem Handling
  - Use GGUS for tickets
  - Careful co-ordination across distributed Tier2s
  - Know where the problem guides are
- Full day coverage
  - Slot between 08:00-09:00 is more difficult than 17:00-18:00

Technical Building Blocks

- Minimal Hardware for Servers
- Load Balancing
- RAC Databases (see later talk)
- High Availability Toolkits
- Cluster File Systems

Server Hardware Setup

- Minimal Standards
  - Rack mounted
  - Redundant power supplies
  - RAID on system and data disks
  - Console access
  - UPS
  - Physical access control
- Batch worker nodes should not be used for servers

Other Approaches

- Spare machine
  - Make it REALLY spare
  - Share between several production servers
- Swap disks
  - Can cause asset problems with maintenance contracts

Good Hardware is not sufficient

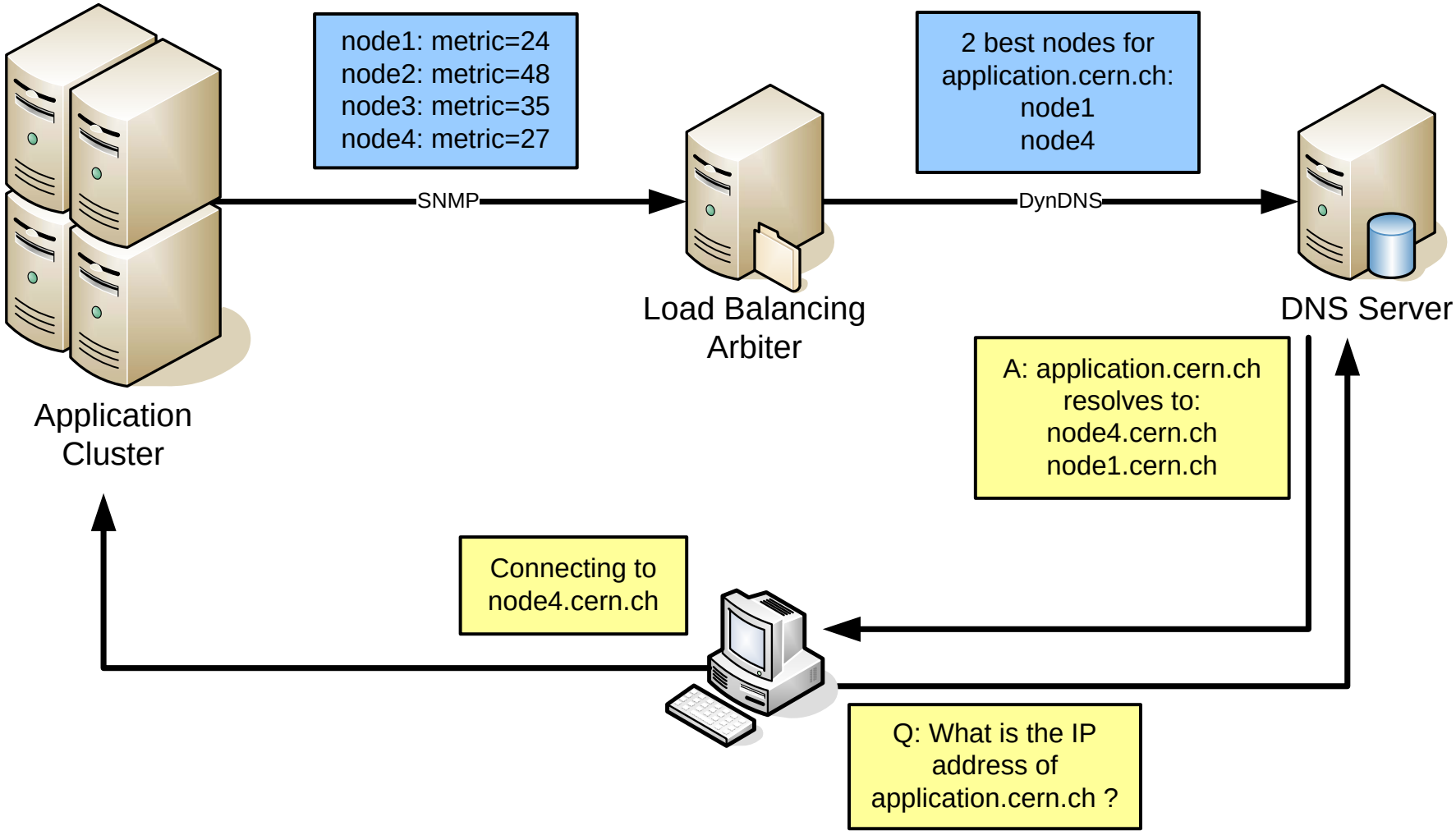
- Monitor the server status for problems
  - RAID
  - Power supplies
- Changes cannot be implemented while service is running
- Software failure is very common, even on reliable hardware

DNS Load Balancing

- Ask the DNS for the IP address of a machine
- DNS returns several
- Select one and connect

- Load Balancing can filter out down or overloaded servers
- Availability and Performance for little effort

DNS Load Balancing



DNS Load Balancing

- No application changes required
  - May benefit from using the list of IP addresses returned by DNS
- DNS caches can delay propagation so leave some buffer (e.g. 1 hour) when doing changes

Stateless vs State Aware

- System is not aware of the state of connections
- State Less Application
  - For any connection, any server will do
  - Only keeps the list of available hosts up-to-date
  - Example: WWW server serving static content
- State Aware Application
  - Initial connection to a server; subsequent connection to the same server
  - Our load balancing system can not help here
  - Solution: after the initial connection *the application* must indicate to the client where to connect
    - Effective bypass of the load balancing
    - Example: `ServerName` directive in Apache daemon

Example of kernel upgrade

A 2 server load balanced solution

1. Set server A to drain. DNS only returns server B.
2. Wait for DNS change worldwide
3. Reboot server A
4. Add server A back. Drain server B. DNS returns only server A.
5. Reboot server B
6. Add server B back. DNS now returns server A and B.

Upgrade performed with 0% user visible downtime

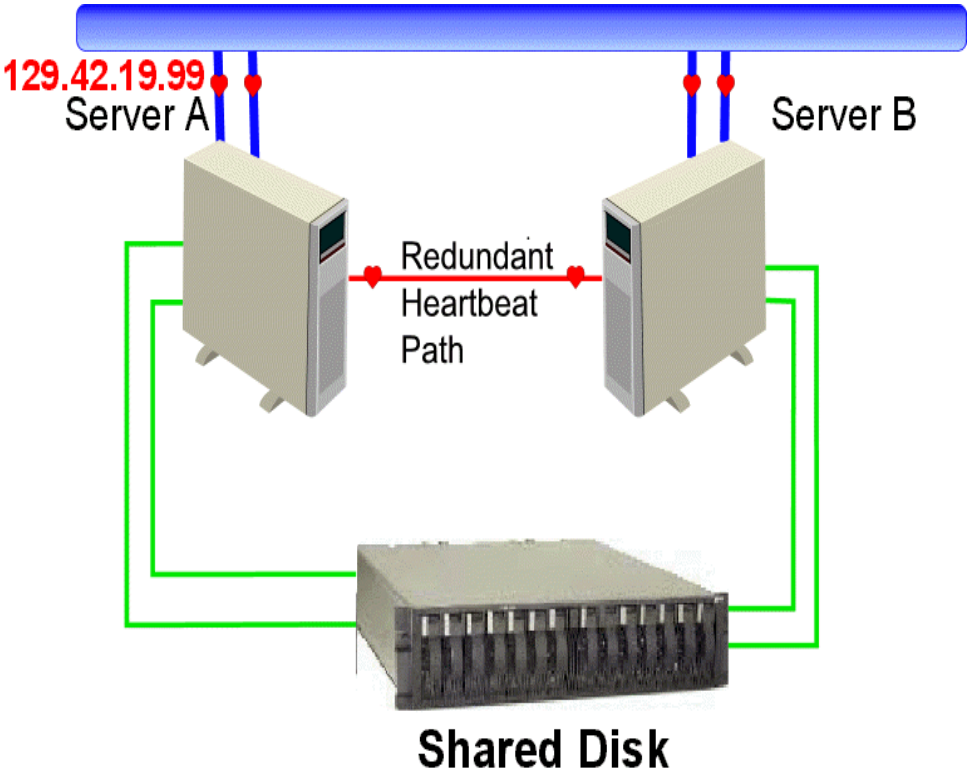
High Availability Toolkits

- Allows a few second switch from one server to another
- There is a real IP address for the service (and a real certificate)
- Switch when
  - Service goes down
  - Administrator request
- Switch with
  - IP Address of master machine
  - Shared disk (requires Fibre Channel)
  - Application specific procedures

- More effort to implement
- More administration skills required

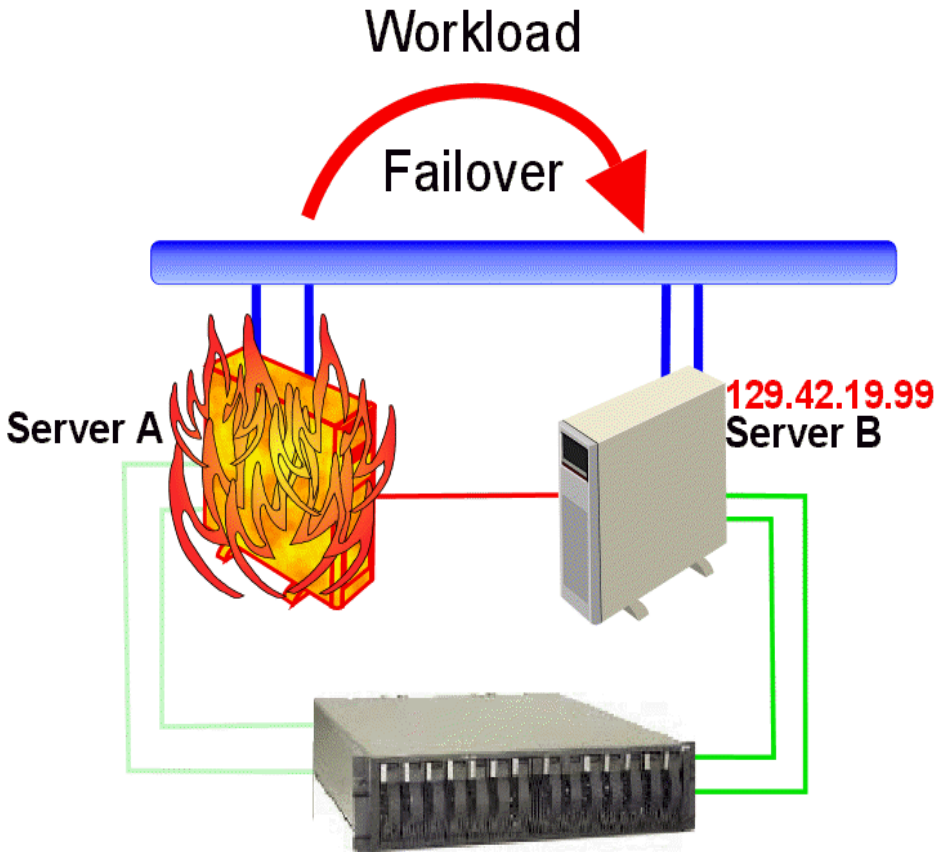
- CERN FIO is using Linux-HA
  - http://www.linux-ha.org/ running at 100s of sites on Linux, Solaris and BSD.

Typical Configuration with HA



- Redundancy eliminates Single Points Of Failure (SPOF)
- Monitoring determines when things need to change
- Can be administrator initiated for planned changes

Failure Scenario with HA



- Monitoring detects failures (hardware, network, applications)
- Automatic Recovery from failures (no human intervention)
- Managed restart or failover to standby systems, components

Cluster File Systems

- NFS does not work in production conditions under load
- CERN FIO has tested 7 different cluster file systems to try to identify a good shared highly available file system
- Basic tests (disconnect servers, kill disks) show instability or corruption
- No silver bullet as all solutions are immature in the high availability area
- Therefore, we try to avoid any shared file systems in the grid environment

Conclusions

- Focusing on Availability can bring improvements in the quality of service
- Biggest gains come from basic hardware hardening and procedures to solve problems quickly
- Planned changes are more frequent than hardware failures. Change automation reduces impact
- Load Balancing and High Availability can be implemented (relatively) easily for some applications

References

- CHEP '06 Load Balancing
  - http://indico.cern.ch/materialDisplay.py?contribId=4&sessionId=6&materialId=slides&confId=048
- Tier 0 Availability Implementation
  - http://indico.cern.ch/materialDisplay.py?contribId=s0t10&sessionId=s0&materialId=1&confId=a056628
- WLCG Workshop Mumbai '06 Tier 1 Availability
  - http://indico.cern.ch/getFile.py/access?contribId=s0t10&sessionId=s0&resId=0&materialId=1&confId=a056628