ATOL: Global File System

Marek Grác
xgrac@fi.muni.cz

Red Hat Czech s.r.o. / Faculty of Informatics, Masaryk University

Advanced Topics of Linux Administration

What is GFS?

- ▶ Symmetric, shared-disk filesystem for cluster
- ▶ Relies on cluster infrastructure
  - ▶ Inter-machine locking (DLM) - cman, ccs, openAIS
  - ▶ I/O fencing and recovery coordination
- ▶ CLVMD to distribute LVM medata updates to cluster nodes
- ▶ Metadata journaling, data journaling on per-file or per-directory basis

What is GFS? II

- ▶ Shared file system
- ▶ 64-bit 'clean', POSIX compliant, ACL
- ▶ Direct I/O capable, Online filesystem management
- ▶ SELinux policy
- ▶ Avoids central data structures
- ▶ Run on mixed environment x86, ia64, amd64
- ▶ Limits:
  - ▶ 100+ GFS client nodes
  - ▶ 16TB on 32bit, 8EB on 64bit (support on multiple 8TB)

Shared (GFS) vs Distributed Filesystem (AFS)

- Whole file locking
- One common journal
- Save-on-close – write() only update the local cache copy on client
- UNIX mode bits are ignored for group and other (provided by ACL)

Create a GFS File system

- Lock manager (lock_nolock, lock_dlm)
- Lock file name (cluster_name:fs_name)
- Number of journals (one per cluster node is needed)
- *gfs_tool*, gfs_mount

Context Dependent Path Names

- Use of special directory link names (@hostname, @mach, @os, @uid)
- Symlink dependent on caller's context
- Example: ln -s /nfs/@hostname/sysinfo /nfs/sysinfo

Goal

- Create a GFS filesystem
- Extend it and add one more journal
- Use a CDPN

Themes

- Themes:
  - GFS vs OCF
- Format:
  - Short presentation (15–20 minutes; 5-7 slides)
  - Paper containing comparision (1.000 words)