

Backing up AFS Using TSM

March 24th, 2004

Xueshan Feng <sfeng@stanford.edu>

Overview

AFS at Stanford

Old Backup Systems

Top Requirements

Implementation using TSM

Summary

Questions and Comments

AFS at Stanford

RW Volumes	2,506,130	11,350,278	54,686	45.83
RO Volumes	150,964	1,744,689	6,108	24.72
BK Volumes	-	-	54,684	-
Total Volumes	2,657,094	13,094,967	115,478	23.01

Category	Used (MB)	Quota (MB)	Volumes	Avrg. (MB)
-----------------	------------------	-------------------	----------------	-------------------

	Total (MB)	Used (MB)	Free (MB)	% Used
Total Disk Space	3,910,413	2,644,824	1,265,590	67.64%

Accounts: 40,000 **Class: 4000** **Group & Research: 15,00** **Dept: 258**
Servers: 17 Sun Netra20 **Uptime: 99.973%**

Old Backup Systems – Before 1999

Old system was built around Legato software and traditional Exabyte tape library

Manual tape changes were required everyday for both backup and restore purposes; hardware failed often

Restore took days and many hours of staff time

Tape capacity was limited and non-automated procedures did not scale as AFS grows

Top Requirements

Automation

Capacity

AFS file level backup and restore; ACLs awareness

Use of centralized storage management infrastructure, if possible

Implementation Using TSM

Hardware

- Two AIX RS 6000 H50 servers
- IBM 3494 tape library – 60 TB capacity
- 200 GB EMC CLARiion disk

Software and internal developed tools

- Tivoli's TSM server software
- Tivoli's TSM Unix backup-archive client
- Set of tools to optimize the process and schedule concurrent backup jobs
- OpenAFS client on TSM server

Backup Procedures

1. **AFS file servers generate <volume>.backup**
2. **A script generates 2 lists – one for each TSM server - containing volume name, .backup creation time, and modification time.**
3. **Volumes with .nb suffix are not in the lists**
4. **TSM server reads the list, only processes a volume when its modification time is newer than the last TSM backup time**
5. **Concurrent jobs are scheduled to backup the volumes to TSM system**

Issues to be Aware of

Backup retention policy

Special backup principal to read AFS files

Excluding data you don't want to backup

No crossing of mount point

Reporting

Summary

Optimization and concurrency reduce unneeded processing and speed up the backup process

Only 10% of the 55,000 volumes are processed

Daily backup data average – 50 GG

Restore can be done from command line

The system has been in production for 5 years, very stable and scalable

Questions and Comments
