Tivoli Storage Manager backup client for OpenAFS

Anders Magnusson
KRB&AFS Conference
Plzen, 2010-09-13

The northernmost University of Technology in Scandinavia
World-class research and education

The northernmost University of Technology in Scandinavia
World-class research and education

Background

TSM is one of the most common backup solutions in use today.

When LTU started to use AFS as filesystem for general use 1.5 years ago it was necessary to continue to use TSM.

TSM principles

All backups are at file-level, nothing like full/incremental dumps.

A number of versions of each file are kept in storage. There is only one copy of each file.

The newest file is labeled "active", all older files are "inactive"

Requirements

Must understand AFS ACL/mount points/...

Be able to recover individual files `the TSM way'

No host-specific backup

Should work at a reasonable speed

Existing ways to do AFS backups

- IBM AFS backup client
  - Binaries only available for AIX
  - Not supported since TSM 5.1

- buta/tsmpipe/...
  - Uses volume dumps, don't care about objects
  - Treats TSM just as a tape station

TSM Interfacing

Reasonably well specified C API for adding new backup clients to TSM

Objects backed up via API can be restored using the normal TSM dsmc tools.

There are API bugs :-/

The northernmost University of Technology in Scandinavia
World-class research and education

AFS Interfacing

- Use OpenAFS libraries to directly access volumes, do not go via cache manager

- Transarc documentation from -91 do still work

- Only backup BK volumes to avoid callbacks

- Cache manager handles tokens, should decouple tabackup entirely from CM.

TSM internal logic mapping

	Traditional	*OpenAFS*
Nodename	Server name	Well...
Filespace	Partition	Cell or so
High-level	Path	Volname + path
Low-level	filename	filename

Backup client (tabackup)

-f filespace (unspecified takes cell name)

-v volume name (can use wildcards)

-s server

-p partition

User friendly - if no arguments are given tabackup will backup your entire cell :-)

Restore client (tarestore)

-f filespace (must be given on restore)

-v volume name (must be given on restore)

-i (interactive, as the usual unix restore client)

Additional args may be paths to files/directories

Tarestore needs more love...

Installation

- Setup TSM parts (dsm.{opt,sys}, Nodename, password=generate...)
- Fetch tokens for something with enough power
- Vos backupsys
- Tarestore

Performance notes

- Profiling shows that 75% of the time is spent in the TSM routines. Best performance if tabackup runs on the TSM backup server.

- TSM 5.3 & 5.5 are reasonable fast. 6.1 are *much* slower and needs much more RAM.

- More than 200 filespaces degrades performance in TSM. With more than 8000 TSM is useless.

Use case at LTU

- 80k RW volumes, 4 file servers
- Vos backupsys starts at 21.30
- 4 parallel tabackup
- Finished in 4-8 hours an average day

The northernmost University of Technology in Scandinavia
World-class research and education