

# AFS (Andrew File System) on Linux for S/390 and zSeries

Kris Van Hees – Sine Nomine Associates

Session 5510

March 8<sup>th</sup>, 2002

# Agenda



- What is AFS?
- Virtue, Vice and The Devil: benefits and drawbacks.
- Where do you get it?
- Planning an AFS install.
- Why on S/390 and zSeries?
- Questions and (hopefully) answers.

# What is AFS?



- AFS is a distributed file system:
  - Component servers rather than a monolith
    - Fileserver, volume server, salvager
    - Authentication server
    - Volume location server
    - Cache manager (client)
  - Not just distribution, also redundancy.

# AFS Implementation Concepts



- **Cell**

A group of machines (servers and clients) under a single administrative control organisation.

- **Server**

A physical machine providing AFS functions as part of a cell.

- **Server Process**

An internal component of the AFS software.

# AFS Data Concepts



- **Partition**

Physical storage for one or more volumes.

- **Volume**

A group of files managed as a single unit.

- **Replication**

Keeping copies of popular or important volumes on multiple servers and/or multiple partitions.

- **Caching**

Storing parts of files in use on a local disk.

# Virtue, Vice and The Devil



- Benefits
  - Strong authentication model
  - Automatic survival of file system access for replicated files
  - Global file space naming
- Drawbacks
  - Invasive install
  - Complexity of back end server functions
  - Authentication issues with applications (e.g. ticket expiration)

# Versions of AFS



- CMU
- Transarc/IBM
- University of Michigan (OS/390)
- OpenAFS
- ARLA/milko

# OpenAFS



- Based on Transarc AFS:
  - Kernel interface module
  - Deep ties into the internal kernel structures
  - Large amount of old and ugly legacy code
- Actively maintained by a large set of volunteers in a true Open Source model.



# Where do you get it?



- Current OpenAFS 1.2.3 code is available at:
  - [www.openafs.org](http://www.openafs.org)
- Mailing lists:
  - [openafs-announce@openafs.org](mailto:openafs-announce@openafs.org)
  - [openafs-devel@openafs.org](mailto:openafs-devel@openafs.org)
  - [openafs-info@openafs.org](mailto:openafs-info@openafs.org)

# Where do you get it?



- Documentation:
  - Included in the OpenAFS 1.2.3 distribution
  - <http://www.ibm.com/developerworks/oss/afs>
- READ THE DOCUMENTATION!!!
- READ THE DOCUMENTATION!!!
- READ THE DOCUMENTATION!!!

# Planning an AFS install...



- To practically implement AFS, you need a minimum of two machines (server and client).
- For testing purposes, server and client can be the same machine.
- In production, plan to separate AFS functions on separate machines.

# Planning an AFS install: Roles and processes



- Seven major roles in an AFS cell:
  - System control server
    - Basic overseer server process (bosserver)
  - Database server
    - Protection server process (ptserver)
    - Volume location server process (vlserver)
  - Authentication server
    - AFS authentication server process (kaserver)
      - Could also be a Kerberos 4 or 5 authentication server.

# Planning an AFS install: Roles and processes



- File server
  - File server process (bosserver)
  - Volume server process (volserver)
  - Salvager process (salvager)
- Backup server
  - Backup server process (buserver)
- Binary distribution server
  - Update server process (upserver)
- Client system
  - Cache manager (afsd + kernel module)

# Planning an AFS install: Location and distribution of data



- Before installing any machines in the given roles, plan the location and distribution of your data carefully:
  - Allocate file server machines on all major segments of the network.
  - Use a minimum of 3 separate file server machines per cell for performance and redundancy.

# Planning an AFS install: File servers



- Optimize system for I/O performance.
- RAID is desirable but not necessary.
- Large RAM desirable (256 MB or greater).
- Network performance is critical.
- Separate partitions for AFS data.
- Plan to dedicate at least one file server to each major group of users.
- Don't forget to plan to replicate each volume to two other file servers.

# Planning an AFS install: Database servers



- Often combines authentication server role, database server role and backup server role.
- RAID with snapshot capability recommended.
- Multiple channel RAID recommended.
- Maximum RAM recommended.
- First database server installed should have the lowest IP address available! Otherwise performance will be a lot worse than it ought to be.



# Planning an AFS install: System control and binary distribution



- Often combined with the first database server.
- It is worth replicating the system control and binary distribution processes on all database servers. 18 MB of disk space isn't much of a sacrifice to gain enhanced reliability.

# Setting up AFS...



- Before you do *ANYTHING*:
  - Read the Quick Start Guide for UNIX from start to finish. Do not skip anything, do not collect \$200.
  - Some steps are a point-of-no-return.
- Don't try this unless you have a reasonable amount of time to do it. The compilation of the AFS source code is s...l...o...w...!
- The good news is that it compiles right out of the box.

# Why on S/390 and zSeries?



- Usual limiting factors for AFS:
  - I/O throughput (disks)
    - S/390 and zSeries are strong in the I/O arena with DASDs
  - Network throughput (especially for backups)
    - High-speed networking between virtual machines.
- Management:
  - A few consoles rather than 10-12 physical machines.

# Why on S/390 and zSeries?



- High reliability and stability of S/390 and zSeries:
  - Replicate volumes between data centers for increased availability, and performance.
  - Replicate volumes within a single box for increased performance.
  - Use Linux cloning mechanisms to bring up another file server:
    - ‘Almost’ instantaneous.
    - No downtime for clients.

# Questions...?



And hopefully answers, as well...

# Contact information



Kris Van Hees

Sine Nomine Associates

[kvanhees@sinenomine.net](mailto:kvanhees@sinenomine.net)

+1 703 723 6673