

European AFS meeting 2009

<mailto:afscon09@dia.uniroma3.it>

<http://www.dia.uniroma3.it/~afscon09/>

Wolfgang A. Gehrke

DIA

Welcome and Overview

Contents

- 1 Introduction
 - General Notes
 - Historical Notes
- 2 Overview of AFS Technology
 - Technology Stack
 - Local and Distributed File Systems
- 3 Organizational Hints
 - Technical Hints
 - Logistical Hints

Welcome to the workshop

WiFi connectivity, documents, requests

The meeting:

- 13 years AFS workshop using German language
- 2008 first time using English language in Graz, Austria
- 2009 second workshop in Rome, Italy

Our department:

- long term AFS user but not a power user (mail, web, home)
- outsourced cell vn.uniroma3.it (initially Unix unification)
- in-house cell dia.uniroma3.it (laboratory setup, exam support)

Participants: Austria, Czech Republic, Germany, Great Britain, Italy, USA

Industry: E4 company (close to Modena, Italy)
SUN Microsystems

Who made it possible

CASPUR coffee breaks

Andrei Maslennikov, Federico Gorelli

DIA central organization, coffee break contribution, visit S. Paolo
Wolfgang Gehrke, Franco Milicchio

ENEA video streaming
Giovanni Bracco, Francesco Beone

E4 company coffee break contribution
Matteo Spatola, Ludovica Delpiano

INFN conference material, funding of guests, organization support
Sandro Angius, Claudio Bisegni, (and Fabrizio Manfredi)

ROMA3 locality and network
Paolo Cursi, Paolo Capozucca

SUN bus tour
Walter Moriconi

Invited guests

Jeffrey Altman

- elder, gatekeeper, developer
- *Windows* port

Derrick Brashear

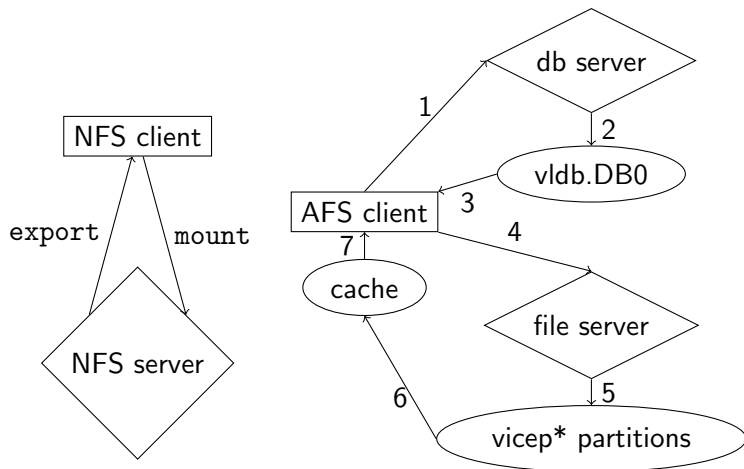
- elder, *gatekeeper*, developer
- AIX/Mac OS X/Irix/Solaris port

Since 2004 “AFS & Kerberos Best Practices Workshop” in the USA:
bring together users and administrators.

What is the Andrew File System?

- distributed file system
- client-server architecture
- servers offer homogeneous and location independent file space
- exist for wide range of UNIX platforms
- clients have local cache
- exist for Linux, Mac OS X, other Unix variants, Windows
- scalability, security, transparent data migration
- optional RO content replication
- developed at CMU as part of the project Andrew \approx 1984
- then product by Transarc Corporation, purchased by IBM 1994
- branch of the AFS source by IBM \rightsquigarrow OpenAFS \approx 2000
- AFS and DCE/DFS products withdrawn 2005

File space organization with volumes



Some volume operations

`vos create` ● reserve space on a fileserver
 ● remember the location in DB

`vos move` ● transport between fileserver
 ● change the location in DB

`vos remove` ● free space on the fileserver
 ● destroy the location in DB

`vos dump` to a file

`vos restore` from a file

and more ● status reports
 ● DB operations

Note: ① administration from arbitrary client!
 ② no down-time on server exchange!

More volume operations

vos backup root.afs prepare backup operations
vos copy root.afs fs1 vicepa fs2 vicepa other method
vos create fs1 a root.cell 5000 fresh volume for cell root
fs mkmount /afs/dia.uniroma3.it root.cell normal mount
fs mkmount /afs/.dia.uniroma3.it root.cell -rw RW mount
vos addsite fs1 a root.afs ; vos addsite fs2 a root.afs
vos addsite fs1 a root.cell ; vos addsite fs2 a root.cell
vos release root.afs ; vos release root.cell
touch /afs/dia.uniroma3.it/test should fail
touch /afs/.dia.uniroma3.it/test should work
vos changeaddr change IP of a file server
vos convertROtoRW repair method by lost RW volume

Some context

Managing AFS, The Andrew File System

Richard Campbell
Prentice Hall, Inc. 1998

Distributed Services with OpenAFS for Enterprise and Education

Franco Milicchio and Wolfgang A. Gehrke
Springer-Verlag 2007

Distributed Systems, Principles and Paradigms (second edition)

Andrew S. Tanenbaum and Maarten Van Steen
Pearson Education, Inc. 2007

in the following:

↑ distributed file systems

↔ related technology and AFS components

↓ local file systems

AFS clients

/afs	/dia.uniroma3.it	/user
historical AFS root dynamic root or <i>root.afs</i> other mount point possible		
	cell name (e.g. <i>ThisCell</i>) CellServDB file DNS <i>AFSDB</i> record	
		↗ db servers ↗ file servers ↙ cacheinfo

In case of replication: remain as long as possible on RO path,
 clients do load-balancing between replicas,
 explicit RO or RW mounts can interfere.

AFS servers

data base server	bossserver	e.g. 3 or 5, Ubik ptserver, vlserver, buserver (kaserver, upserver, runntp)
file server	bossserver	/vicep* partitions fileserver, volserver, salvager
backup host	bossserver	to tape or disk butc flexible <i>cron</i> <i>backup</i> volumes

bossserver \approx Solaris service manager
 \approx mysqld_safe

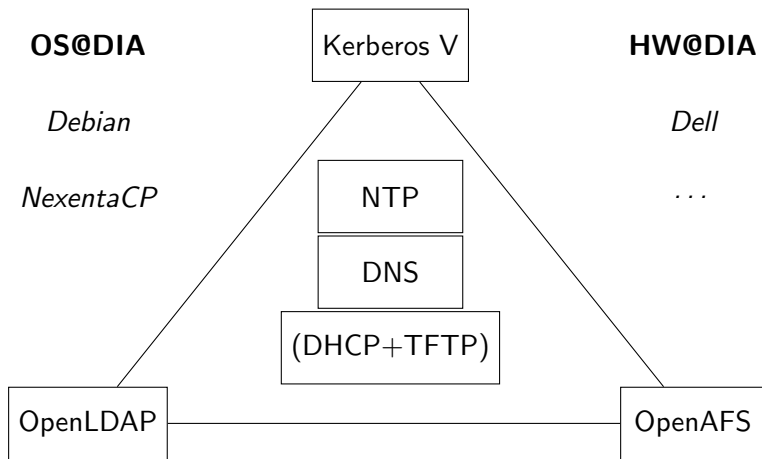
Ubik leader-election
sync site

Server and client pairing

afsd	fs: e.g. ACL, mount, quota management
bossserver	bos: e.g. create process instances
buserver / butc:	backup: e.g. volume dump or restore also DB
ptserver	pts: e.g. user and group management
vlserver / volservers	vos: e.g. volume and partition management

- daemons have a “-help” option
- daemons may have a “-noauth” option
- commands have a “help” subcommand
- commands have a “-noauth” option
- commands may have the “-localauth” option (KeyFile)
- commands have a “-cell” option

Related services with redundancy



Virtualization e.g. *XEN* or *VirtualBox*

Multiple possibilities

Kerberos authentication trusted central third party

- Active Directory
- Heimdal
- MIT

Naming service provide user info for AFS IDs

- LDAP
- NIS
- ptserver gateway

AFS implementations

- Arla
- OpenAFS
- Transarc legacy

Authentication process

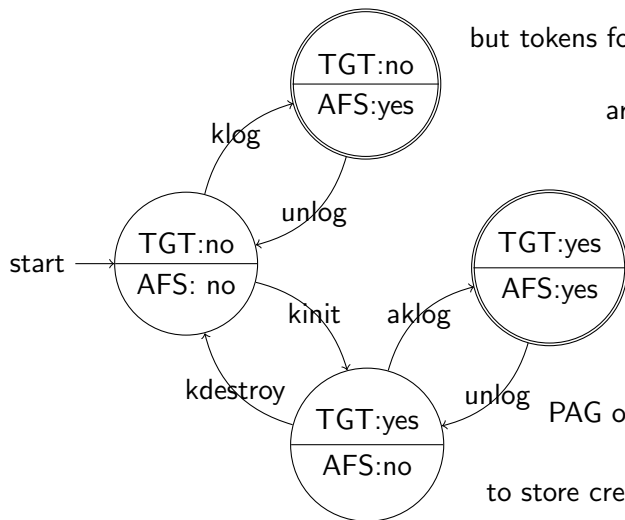
usually just one Kerberos ticket;

but tokens for different cells

are quite common

klist

tokens



PAG or keyring

to store credentials

Developments

Kerberos kserver \rightsquigarrow Kerberos 4 \rightsquigarrow Kerberos 5

- LDAP backend
- pkinit (e.g. grid context)

LDAP ptserver \rightsquigarrow LDAP integration?

- configuration inside the directory
- multi-master

OpenAFS UDP \rightsquigarrow TCP, IPv6 ?

- AFS3 protocol design
- Rx extensions

How to leverage AFS

Mail mbox vs. Maildir

Samba Kerberos and LDAP integration

Web WebDAV with Apache2 and subversion

- 1 service restart or automatic token renewal
- 2 single identity or user identity
- 3 redundancy possible

Further uses

Kerberos

- cross-realm authentication
- modauthkerb \rightsquigarrow webauth \rightsquigarrow Shibboleth
- RADIUS

OpenLDAP

- addressbook
- aliases for mail
- DNS backend

OpenAFS

- central configuration files
- files for daemons e.g. TFTP
- centralized backup

Things to be aware for the moment

- 1 directory limitations
- 2 only per-directory ACLs
- 3 only RO volume replication
- 4 still DES encryption
- 5 one single KeyFile
- 6 no POSIX semantics
- 7 no byte-range locks
- 8 no Unicode support
- 9 no fine-grained administrative rights
- 10 device, pipe, socket not available

Good practices

- 1 DNS domain name = afs cell name = Kerberos REALM name
- 2 Unix uid = Pts id through naming service
- 3 project group owned by project admin group
- 4 IP account has to be in a group for ACLs
- 5 firewall has to accept late arriving call backs
- 6 attention to MTU value in the network
- 7 use of @sys in a path which expands to client architecture
- 8 RO replication of seldom changing data
- 9 long batch jobs have to take care of credentials
- 10 PAM supported login acquires automatically credentials

Difference to UNIX permissions

UNIX permissions for a file

r read
w write
x execute

AFS permissions for a directory

a administer
d delete
i insert
k lock
l lookup
r read
w write

- root not special !
- negative rights ?
- owner from token

special rules

user r when missing nobody can read
user w when missing nobody can write
user x when missing nobody can execute

Debugging

cmdebug information about cache manager

- cache size
- cache type

rxdebug information about Rx RPC

- different ports for the client and servers
- get version of OpenAFS

udebug information about Ubik data base

- different ports for the client and servers
- indication of sync site

Attention everyone can use these commands

- over the net
- without authentication

Local choices on servers and clients

- OS
- AIX
 - HP/UX
 - Irix
 - Linux
 - Solaris

- file server
- inode
 - namei

- Solaris
- UFS
 - ZFS

- Linux
- ext3
 - xfs

- OS
- AIX
 - BSDs
 - HP/UX
 - Irix
 - Linux
 - Mac OS X
 - Solaris
 - Windows
- cache type
- UFS for Solaris
 - ext3 for Linux
 - in memory

Other distributed file systems

NFSv4 client can mount a server export everywhere

- got Kerberized
- now stateful
- parallel NFS

Lustre Linux only

- large scale cluster computing
- high scalability
- top HPC applications

Hadoop no direct host integration

- inspired by Google's MapReduce and file system
- Java implementation
- used by Yahoo!

Windows Dfs file replication service in a domain

Presentations

- Windows notebook
- streaming and projector integration
- WiFi and network connectivity
- Acrobat Reader, MS Office, OpenOffice
- OpenAFS client
- putty for ssh session

- send your material to the conference e-mail
- have site reports ready
- Wednesday 15 min slots

Lunch

Pensionato San Paolo, 00146 Roma - Google Maps

http://maps.google.com/maps?f=q&source=s_q&hl...

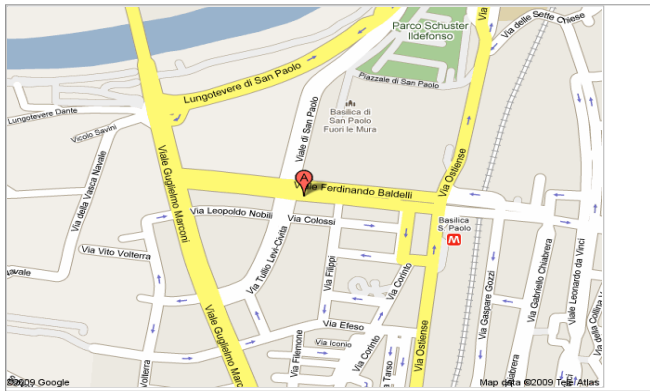


**Pensionato San Paolo, near 00146
Rome, Italy**

Notes : Location for lunch.

A. Pensionato San Paolo

Viale Ferdinando Baldelli, 41, 00146
Roma, Roma (Lazio), Italy - 06 54224867
23 reviews



Social events

bus tour September, 28

- 18:00
- Via Volterra 3
- sponsored

visit S. Paolo September, 29

- 17:30
- Basilica San Paolo
- sponsored

Please, have some *cash* with you.

My cell phone: +39 329 0552317

Dinner September, 28, 20:30

Pizzeria Biondo Tevere, 00145 Roma, Italy - Google...

http://maps.google.com/maps?f=q&source=s_q&hl...



Pizzeria Biondo Tevere, near **00145**
Rome, Italy

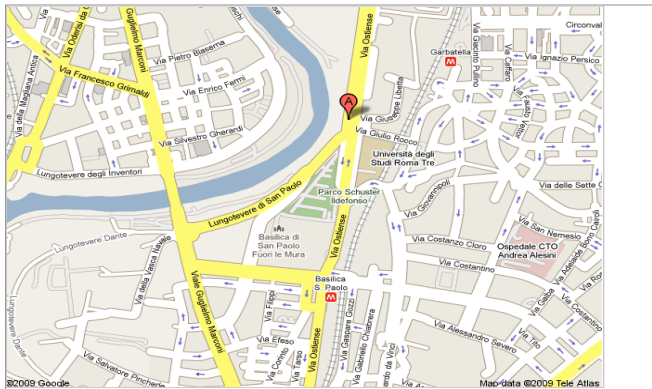
Notes : free choice including Pizza.

A. Ristorante Al Biondo Tevere

Ristorante - **Pizzeria**

178, Via Ostiense, Roma, RM 00154,
00146, Italy - 06 5741172

Business listings provided by PagineGialle.it



Dinner September, 29, 20:00

Trattoria dell'Omo Snc, 00146 Roma, Italy - Google...

http://maps.google.com/maps?f=q&source=s_q&hl...



Trattoria dell'Omo Snc, near **00146**
Rome, Italy

Notes :menu ca. 25 Euro + drinks.

A. **Trattoria Dell'Omo Di A.Dell'Omo Snc**

Viale Guglielmo Marconi, 475, 00146
 Roma, Rom (Lazio), Italy - 06 5593715
 1 review

Business listings provided by PagineGialle.it

