Samba in Enterprise Environments

eXPeriences and problems encountered

Joachim Jordan
Michael Weisbach
IBM Linux Competence Center Boeblingen
Agenda

- Overview over some customers
- A customer in detail
- Features requested
- Discussion: Samba in large enterprise environments
Some customers

- **German power supplier:**
  - 110 servers, ~40,000 user IDs via Winbind, 3.5 TB data

- **German parliament:**
  - ~1,500 Clients, about 100 decentralized server

- **German insurance company:**
  - ~1,000 servers

- **Danish bank:**
  - ~600 servers

- **US insurance company:**
  - 12 servers, 3.6 TB data
A customer in detail - ongoing Proof of concept

- **German power supplier:**
  - currently using Windows NT servers and clients
  - 110 file servers in main location
  - in summary ~3.5 TB data on these servers; SCSI disks
  - ~30,000 active clients

- **now evaluating**
  - Fileserving with Samba on Linux for S/390 vs. W2k/WinXP

- **future**
  - Active Directory infrastructure
  - Samba/390 *and/or* W2k/Windows XP based file servers
  - Windows XP clients
Summary experiences

- easy administration of repetetive tasks
- large user databases drive winbind requirements
- large storage requirements and clustering issues
- high failure costs drive need for redundancy and replication in all major components
- large enterprises need to lower admin costs, especially in heterogenous enviroment - (Active) Directory coexistence/integration
- Enterprise wide security via addition of Kerberos
- large enterprises are multi-national - internationalization requirement
Features needed - Access Control

- **migrations from non-UNIX to UNIX**
  - from OS/2, Novell Netware and Windows based servers
  - to Samba on Linux (UNIX)

- **ACL support to match "old" access controls**
  - journaling filesystem required
    - possible with XFS or ext3 with ACL patches
    - JFS ACL support planned
    - not official supported yet (kernel 2.4.x nor any current distributions)
  - POSIX ACLs are slightly different to Windows implementation
    - missing "append only" and "write/edit but not delete"
    - Solution: EA based access control vs. Samba-internal ACL database?
  - Problem
    - different access paths: Samba or via UNIX shell
    - ACL performance for backup/recovery?
Features needed

- **Winbind mapping is not unique**
  - today dynamic allocated
  - Backup is done for Unix ID, how to map correctly to NT RID in case of restore?
  - missing import/export; ASCII, xml?
  - unique mapping also needed for
    - replication/mirroring for high availability
    - scalability via clustering and loadbalancing

- **Backup**
  - backup tools need to backup all necessary data
    - ACL information
    - winbind mappings

- **Virus scanner integration**
expectations

- customers w/ WindowsNT servers using (and expecting) auditing features
- n:1 migration (integration) of file servers requires detailed accounting

Logging, Auditing and Accounting

- functions to support logging and auditing of user activity
- accounting
  - Samba process runs under security of user, therefore process accounting may be used
- no text logfiles please :) - API layer preferred
Features needed - UserAdmin for LDAP

- challenge: German Parliament decided use OpenSource on all servers; not to use Active Directory but instead native LDAP!
  - large, decentralized Samba environment on Intel based servers
  - native LDAP directory
  - Windows XP clients

- problem: How to administer, create and delete users?
  - use of NT tool "Domain Administration for Domains"
    - how to assign UNIX UID to new LDAP (sambaAccount) user?
  - have a samba account admin frontend w/ LDAP support
  - ...

Features needed - High Availability

- **High Availability and Scaleability needs**
  - support for shared (or replicated) file systems
    - not only VFS support
    - maybe cluster filesystem (GFS, GPFS, ...)
    - use of OpenAFS filesystem across many file servers (i.e. for load balancing)

- **unique Winbindd - NT username to UNIX uid mapping required**
  - two or more nodes of failover cluster / high availability
  - two or more nodes of load balancing cluster / more than one path to file
(Known) Limitations

- **User ID's**
  - currently only ~65000 userids /groupids possible

- **Sockets: TCP/IP stack limitations per Linux instance**
  - solution: "Samba farm" w/ shared filesystem?!
  - inside Samba server box, additional sockets needed too (LDAP connects)

- **Processes / Linux box limited**
  - Linux process scheduler scalability
  - multiprocess / multithreaded approach like Apache 2.x?

- **Memory limitations**
  - solution: 64bit support
Contact

- Michael Weisbach, Linux IT architect, IBM Germany Global Services
  - Linux IBM IGS Germany: crlinux@ch.ibm.com

- Joachim Jordan, Linux IT architect, IBM Europe technical sales support
  - Europe Linux ATS: tmcc@de.ibm.com

Discussion

- Feel free to comment...