Samba 4.2

FFG 2015
Stuttgart

Volker Lendecke

Samba Team / SerNet

2015-03-27

SerNet

- SLA based support for more than 650 customers
  - firewalls, VPN, certificates, audits
  - based on open standards wherever possible
- Support for many OS: Linux, Cisco IOS, Windows etc.
- Compliant with BSI Grundschutz and ISO 27001 and other international regulations
- SerNet and Samba
  - technological leadership of SerNet worldwide
    - SerNet distributes up-to-date Samba packages
  - samba eXPerience
    - May 19-21 in Göttingen, www.sambaxp.org

Samba 4.2

- Many new features
  - 4.1 was mainly a bugfix release
- SMB2.1 leases
  - Huge performance benefit for office workloads
- Improvements for OS/X clients
  - vfs_fruit and AAPL extensions
- ctdb included
  - No separate cluster daemon build required
- Performance improvements
  - tdb optimized
  - Messaging tuned
- The end of Samba 3

End of Samba 3.x

- Regular release cycle is nine months
- Current release fully supported (4.2)
  - Bug fixes, some new features
- Previous release (4.1)
  - Only bug fixes
- Next to previous (4.0)
  - Security fixes only
- Samba 3.6 went out of security fix support with 4.2
- All code from 3.6 continues to live
  - File server, print server, NT-style DC

SMB2.1 leases

- Oplocks done right
- Huge improvement for higher latency networks
- Clients can request caching permission
  - Server will make a client correctly flush caches
  - Conflicting file access will drop caches
  - Excel is heavily stepping on its own toes
- Once an oplock is gone, no way to get it back
- Leases can be re-acquired
  - Roundtrips and server load reduced

winbind changes

- source3 winbind used in AD Domain Controller
  - More flexible idmapping
  - consolidated code base for Domain Controller and Member
- SMB signing now required
  - Prevents MITM between winbind and DC
  - Authentication was protected before with SCHANNEL
- winbind does not list group memberships anymore

SMB3 support

- SMB3 support since Samba 4.0
- All mandatory SMB3 features
- Improved security
- Secure negotiate
  - No downgrade attack (SMB1) possible
- AES based signing
  - Hardware acceleration (Windows client only atm)
- Transport encryption
  - smb encrypt = yes

OS/X clients

- Modern OS/X prefers SMB2 over AFP
- Ralph Böhme's netatalk is the smbd of AFP
- Netatalk maps OS/X specifics to Unix
  - OS/X special characters like "*" and "/"
  - Mac Metadata (Finder info / resource fork)
- Ralph implemented vfs_fruit
  - Samba with vfs_fruit compatible with netatalk
- AAPL extension speed up directory listing

Clustering

- Samba provides active-active SMB
  - Scale-Out SMB Export of clustered file systems
  - Gluster, Ceph, OCFS, GFS, GPFS, StoreNext, ...
- CTDB (Clustered Trivial DataBase)
  - Clustering component of Samba
  - Lock coherency, Service Monitoring, Cluster Membership
- Up to Samba 4.1 ctdb was maintained independently
- Versioning of ctdb not always coherent
- CTDB now integrated into Samba
  - Better code sharing
  - No version confusion

Performance improvements

- Samba relies on tdb key/value stores (nosql ... :-))
- tdb uses Unix file locks via fcntl for coordination
  - fcntl scales very badly on all Unixes
  - Single global locks to go through
- Multi-threading provides scalable locking: Mutexes
- Linux provides "process shared robust mutexes"
  - No global contention, better scalability, less CPU
- tdb internally fragments
  - New mechanisms for reducing fragmentation
  - Smaller databases, better cache locality

Faster Inter Process Communication

- Samba processes talk to each other
  - Hey, smbd, please give up the oplock
- Messaging based on tdb files and signals
  - tdb files have to store a lot of messages
  - Unix Signals suck
- New messaging based Unix Domain Datagram sockets
  - Less CPU overhead
  - Better scalability
- Basic infrastructure to move to different programming paradigms
  - Actor Model parallelism?
- Merge ctdb messaging with Datagrams
  - Remove load from ctdbd

FileChangeNotify

- A client can request information about all changes in a share
- Current implementation tdb-based
  - tdb holds all change notify requests
  - High contention on some tdb records
  - Every change needs to read hot records "/"
- New implementation based on scalable messaging
- "inotify proxy" in a cluster

What's next

- SMB3 support
  - Multi-channel (prototype available)
  - Persistent file handles
  - SMBDirect (RDMA)
  - `https://wiki.samba.org/index.php/SMB3`
- Active Directory
  - Domain Trusts (available soon)
- Improved clustering
  - Make ctdb play well with other clustering components
- Performance tuning

Questions?

vl@samba.org / vl@sernet.de

http://www.sambaxp.org/