Samba 3
Cleaning up

SambaXP

April 25, 2007

Volker Lendecke

Samba Team SerNet

Network Service in a Service Network

Volker Lendecke

- Co-founder SerNet - Service Network GmbH
  - Free Software as a successful business model
  - Network Security for the industry and the public sector
  - Samba-Support/Development in Germany
- For 15 years concerned with Free Software
- First patches to Samba in 1994
- Consultant for industry in IT questions
- Co-founder emlix GmbH (Embedded Systems)

Cleaning up - what's that?

- Samba 3 is essentially still the first implementation
- You can track some parts of the code to the mid-90s
- Modern demands can't be matched with that
- Samba 4 has created a very good blueprint of how Samba should look like
- We want to step by step migrate Samba 3

Overview

- Things we've done
  - Async winbind
  - Major rewrites of the open code (opening a file can be surprisingly subtle)
  - RPC rewrites
  - Merges with Samba4
- Things to do
  - Get rid of global variables (current_user!!)
  - Make spoolss use the registry, not vice versa

Async winbind

- Scalability in one direction: winbind had to trust almost 100 domains
- 3.0.14 would wait for the DC in Kasachstan to enumerate users before it would authenticate local users
- 3.0.20: Per domain a synchronous child, all can operate simultaneously
  - 4.0 libs can do that in one process, maybe later..

Convert share modes from DOS to NTCreate

- Historically we had DENY_READ, DENY_WRITE and friends coming from the Open&X calls
- Then came NTCreate&X: FILE_SHARE_READ/WRITE/etc
- Jeremy tried to map NTCreate&X onto DENY_READ etc, but this could not work
- DENY_READ etc is mappable onto NTcreate as Tridge found out
- Jeremy „stole" this from Samba4 into Samba3, we now pass all the share mode tests

Oplock rewrite

- 3.0.20 had an awful piece of code: After sending an oplock break request deep inside the open code it recursed into the main loop
  - Only one oplock break request outstanding possible
  - Only very specific SMBs allowed while waiting for the break response
  - Hand-crafted singalling mechanism: UDP sockets
- 3.0.21 switched to the usual messaging and removed that inner loop (Some become_root is necessary here ;-))

Trans2 rewrite

- Trans2/NTTrans calls ship arbitrary blobs of data client->server and back
- Trans calls can be split up
- After the first of those we **only** allowed the secondaries, nothing intermediate
  - For single-threaded clients it's ok (OS/2 is not...)
- Removed that restriction, receive_next_smb is gone now, we only have one central processing loop

TDB clear-if-first

- Some TDB databases need to be cleaned on startup
- How do we reliably find out we're the first user of a tdb?
  - Everybody gets and holds a fcntl lock
- Well, fcntl locks are typically held in a linked list
- Thousands of fcntl locks kill performance
  - Solaris was particularly bad, Linux also is not optimal
- Newer tdb only holds one lock

New talloc implementation

- Ages ago Samba introduced a memory pool concept called talloc
  - Many memory blocks freed at once
  - One pool (MEM_CTX), many blocks
- The 4_0 talloc implementation makes this hierarchical: Each block is a pool
  - Contains destructors: Functions called free()
- The hierarchies can lead to **very** compact and also extremely subtle code if not used with care

Valid users to use SIDs

- For 3.0.23 there was a customer demanding to get nested local groups to really work
- Access controls were based on names, this made nested groups extremely confusing
- 3.0.23 changed all the access controls to match what Windows does
- Samba access checks are no longer Name- but SID based.

Things remaining

- loadparm.c: Exposing the internal array data structure needs to go, this has been started
  - Necessary for more flexible config backends
- The printing subsystem needs to use the registry, not the other way round
- Global variables need to be removed, the hardest ones are the global inbuf/outbuf structures and the current_user structure

Things remaining II

- Winbind is a constant construction site, maybe we need Samba4 libs here
- Convert RPC pipes to PIDL
- We need to get better at internal abstractions
- ... And then there's certainly all new features, we need at least 2-3 attempts to get stuff right :-)

Questions/comments?

Volker Lendecke, VL@SerNet.DE

SerNet – Service Network GmbH
Bahnhofsallee 1b
37081 Göttingen

Tel:	+49 551 370000 0
Fax:	+49 551 370000 9

http://www.SerNet.DE

http://Samba.SerNet.DE