SambaXP 2023

Accessing files remotely from the smallest devices to the largest devices (and the cloud): SMB3.1.1 Improvements to the Linux client

Presented by Steve French

Principal Software Engineer

Microsoft Azure Storage

- This work represents the views of the author(s) and does not necessarily reflect the views of Microsoft Corporation
- Linux is a registered trademark of Linus Torvalds.
- Other company, product, and service names may be trademarks or service marks of others.

Who am ?

- Steve French smfrench@gmail.com
- Author and maintainer of Linux cifs vfs (for accessing Samba, Azure, Windows and various SMB3/CIFS based NAS appliances)
- Co-maintainer of the new kernel server (ksmbd)
- Also wrote initial SMB2 kernel client prototype
- Member of the Samba team
- coauthor of SNIA CIFS Technical Reference, former SNIA CIFS Working Group chair
- Principal Software Engineer, Azure Storage: Microsoft

Outline

- Overview of Linux FS activity
- Recent ksmbd (server) improvements
- Recent client improvements
- Coming soon … what to look forward to
- Testing improvements

A year ago and now ...

- Now: 6.3 “Hurr durr I’ma ninja sloth”

- Then: 5.18-rc4 “Superb Owl”

LSF/MM/eBPF summit is back in person too

- Picture from 2022 (2023 summit is going on same week as SambaXP)

Some Linux FS topics of interest from LSF and other recent discussions

- Folios, netfs, iov_iter, variable size pages, and the redesign of page cache and offline (fscache)
- Improvements to statx and fsinfo and to inotify/fanotify
- Idmapped mounts
- Updates to POSIX ACL internal API
- Extending in kernel encryption: TLS handshake (for NFS) and QUIC (SMB3.1.1 and other)
- io_uring (async i/o improvements)
- Shift to cloud
- Better support for faster storage (NVME) and net (RDMA/sm

Linux Filesystems Activity over past year (since 5.18-rc4)

- 5400 filesystems changesets (6.2% of total kernel changesets, one of the most watched parts of the kernel, and FS activity is up slightly)
- Linux kernel fs are 1.07 million lines of code (measured this week)

Most Active Linux Filesystems over the past year

- VFS (mapping layer) 420 changesets
- The top filesystems and VFS dominate the activity
- Most active are BTRFS 1216 (huge increase), XFS 553, ext4 386
- SMB3.1.1 (cifs.ko) 339 (activity up)
- Then NFSD (server) 279, and NFS (client) 208 (activity down)
  - cifs.ko had more than 3x more lines changed. It has been a VERY active year for cifs.ko
- Other:
  - ksmbd (new, added in the 5.15 kernel) (121), ntfs3 (178, added in 5.15), gfs2 (1

SMB3.1.1 Activity was strong this year

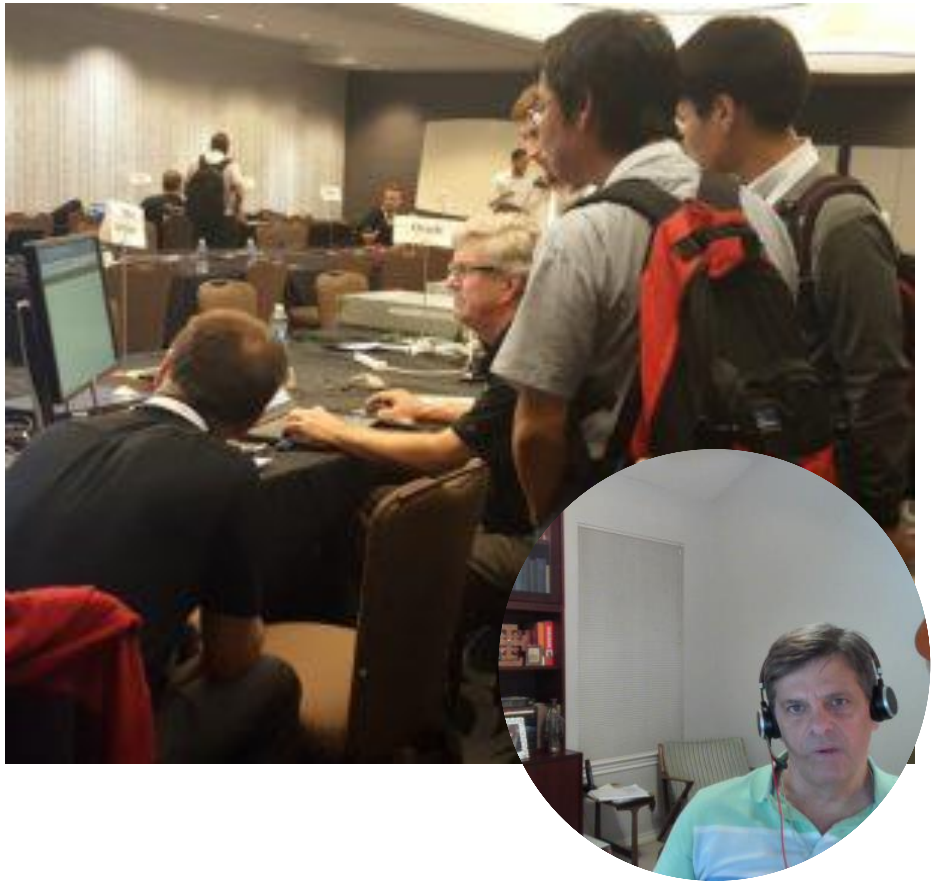
- cifs.ko activity was strong, 339 changesets
  - cifs is 60KLOC kernel code (not counting user space utilities)
- ksmbd activity was down
  - Introduced in the 5.15 kernel, 25KLOC kernel code, 310 changesets since its introduction
- Samba server (userspace) is over 3.5 million lines of code (orders of magnitude bigger than the kernel smbd server or any of the NFS servers) and is even more active

Goals and Actions for SMB3.1.1 on Linux

- Be the fastest, most secure general-purpose way to access file data, whether in the cloud or on premises or virtualized
  - Improve directory lease support
  - Keep improving compounding, multichannel
- Support more Linux/POSIX features – so apps don’t know they run on SMB3 mounts (vs. local)
  - SMB3.1.1 POSIX extensions, new FSCTLs
  - Use xfstests to locate new features to emulate
- As Linux evolves, quickly add features to Linux kernel client and Samba and ksmbd
  - More test automation and keep adding more tests

One of the strengths of SMB3.1.1 is broad interop testing

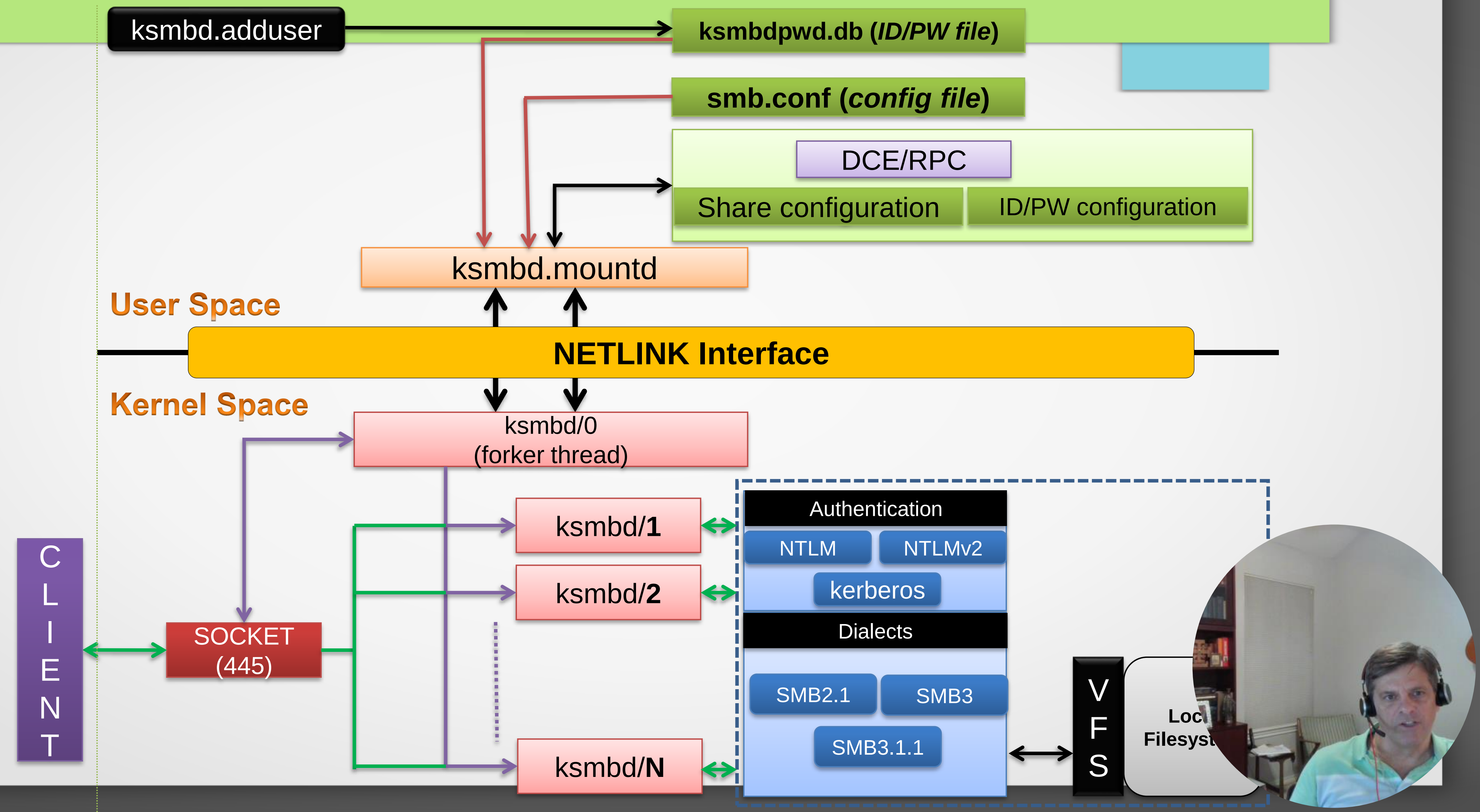
- In-person plugfests are back!
- SMB3.1.1 plugfest collocated with SDC last fall
- Hoping for much informal testing here
  - Contact Paulo and Enzo e.g.
- Many exciting things being tested



Progress and Status update for Linux Kernel Server (ksmbd)

Additional information provided by Namjae Jeon (linkinjeon@kernel.org)

Architecture



ksmbd.adduser

ksmbdpwd.db (*ID/PW file*)

smb.conf (*config file*)

DCE/RPC

Share configuration

ID/PW configuration

ksmbd.mountd

User Space

NETLINK Interface

Kernel Space

ksmbd/0
(forker thread)

ksmbd/**1**

ksmbd/**2**

ksmbd/**N**

CLIENT

SOCKET
(445)

Authentication

NTLM

NTLMv2

kerberos

Dialects

SMB2.1

SMB3

SMB3.1.1

VFS

Loc
Filesyst

Some examples of exciting recent progress

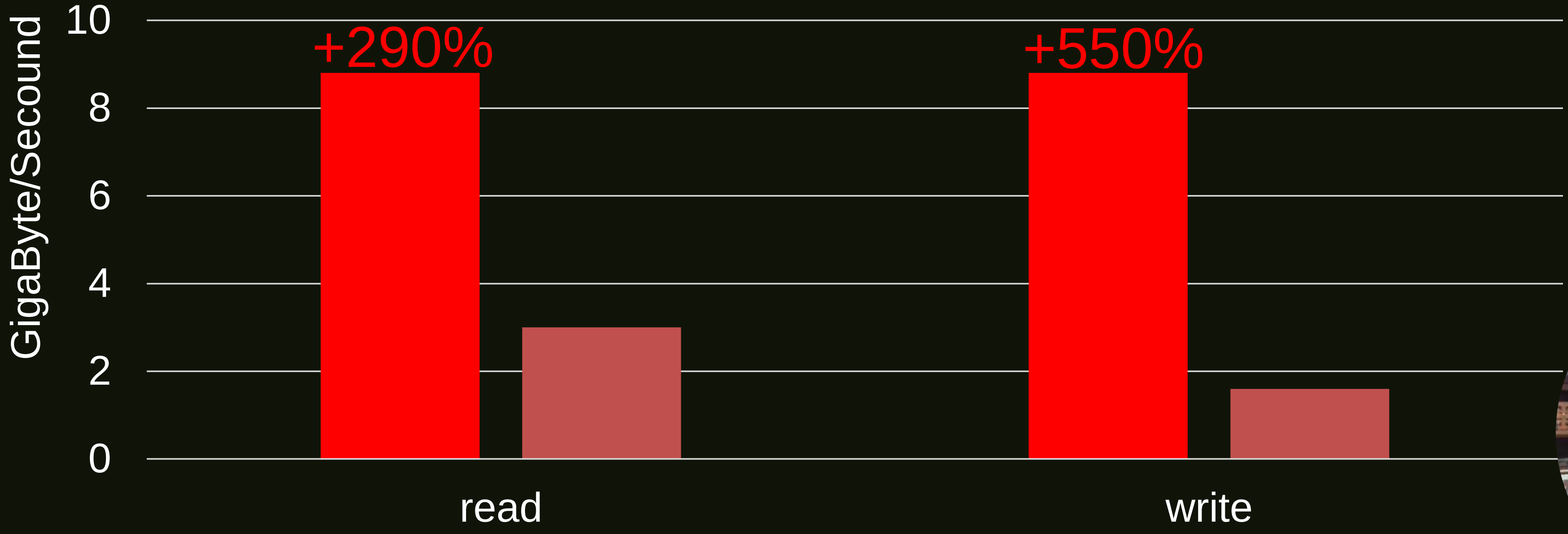
- POSIX extensions
  - Server supports SMB3.1.1 POSIX Extensions
  - Change the SID to the one Samba server using for POSIX Extensions
    - Samba set SIDOWNER and and SIDUNIX_GROUP in create posix context
    - And sets SIDUNIX
  - Set file permission to match Samba server POSIX extension behavior
  - Fill in SIDs in SMB_FIND_FILE_POSIX_INFO responses
- Fixes for various security issues
  - ZDI and others had reported several security issues
- Fix unlink and rename races (new Linux VFS helpers added to aid this)
- Multichannel and SMB Direct improvements
- Improve management of SMB3 credits (flow control improvement

RSS(Receive Side Scaling) mode support

- ksmbd now supports RSS mode
- Ziwei Xie(high-flyer) compared the performance samba and ksmbd on their test environment. Thanks Ziwei!
- In RSS mode, there is a performance difference of 3 times for read and 4 times for write on his setup.

Performance Comparison on Multichannel + RSS mode

*Test Environment
CPU : AMD EPYC 7H12 64-Core Processor
NIC : MCX653105A-HDAT
Client : Windows
Benchmark tool : FIO



Future Plan

- More Directory lease testing
  - Plan to turn on leases by default (instead of older oplocks, current default)
- SMB2 notify (WIP)
- Improve MacOS compatibility
- Add new FSCTLs to help Linux kernel client
  - e.g. rename exchange and any remaining fallocate corner cases
- Durable handle v1/v2 feature (WIP)
- Add ksmbd status option to show statistics using ksmbd.control
  - Processed requests, session info(user info, number of credits and more), session list, openfiles, NIC info.
- Config backend (Recently get a request, make the configuration in available remotely over the WINREG RPC interface)

Linux Kernel Server, KSMBD (continued)

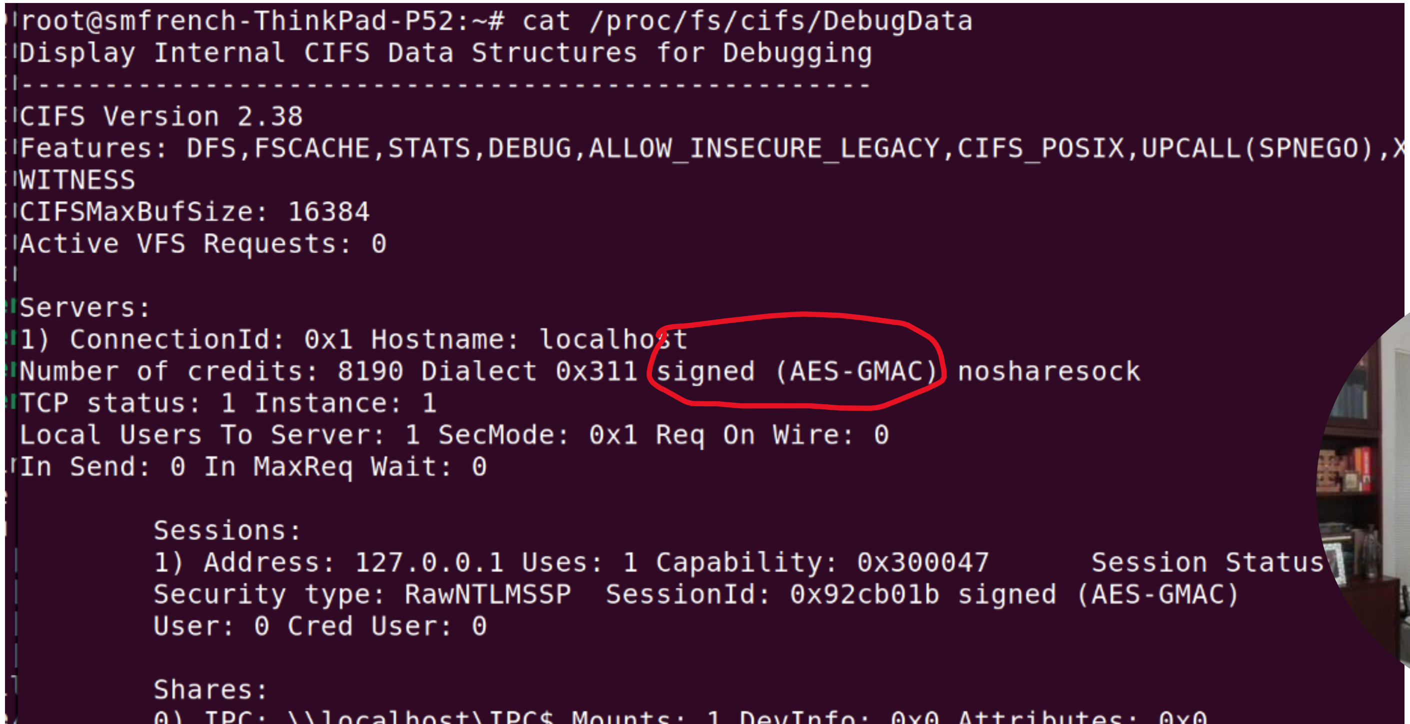
- If interested in contributing there are lots of cool features to work on, as well as improved integration with Samba (e.g. user space upcalls for additional features). The SMB3.1.1 family of protocols is huge!
- Roles: Namjae (the maintainer) has done a lot, but additional features or subcomponents could be delegated. I am managing the git merges, ensuring additional functional testing is done regularly, and reviewing patches as requested by Namjae (my focus is largely on the client)
- Namjae would welcome additional help with code reviews security auditing, testing and new features
- Very exciting time!

Recent improvements in the kernel client

(cifs.ko)

Signing algorithm negotiation – faster signing

# modinfo cifs | grep signing
parm: enable_negotiate_signing:Enable negotiating packet signing algorithm with server. Default: n/N/0 (bool)
“insmod cifs.ko enable_negotiate_signing”



```
root@smfrench-ThinkPad-P52:~# cat /proc/fs/cifs/DebugData
Display Internal CIFS Data Structures for Debugging
---------------------------------------------------
CIFS Version 2.38
Features: DFS,FSCACHE,STATS,DEBUG,ALLOW_INSECURE_LEGACY,CIFS_POSIX,UPCALL(SPNEGO),X
WITNESS
CIFSMaxBufSize: 16384
Active VFS Requests: 0

Servers:
1) ConnectionId: 0x1 Hostname: localhost
Number of credits: 8190 Dialect 0x311 signed (AES-GMAC) nosharesock
TCP status: 1 Instance: 1
Local Users To Server: 1 SecMode: 0x1 Req On Wire: 0
In Send: 0 In MaxReq Wait: 0

        Sessions:
        1) Address: 127.0.0.1 Uses: 1 Capability: 0x300047      Session Status
        Security type: RawNTLMSSP  SessionId: 0x92cb01b signed (AES-GMAC)
        User: 0 Cred User: 0

        Shares:
        0) IPC: \\localhost\IPC$ Mounts: 1 DevInfo: 0x0 Attributes: 0x0
```

Signing changes (WIP)

6.0 and earlier kernels /proc/fs/cifs/DebugData showed:

```
    Security type: RawNTLMSSP  SessionId: 0x5f08b08 signed
```

About 20% faster performance was demonstrated if workload not network constrained (thank you Enzo!). Still testing AES-GMAC

```
  Security type: RawNTLMSSP  SessionId: 0x5f08b08 signed (AES-GMAC)
```

Or if server doesn’t support GMAC will fall back to:

```
  Security type: RawNTLMSSP  SessionId: 0x5f08b08 signed (AES-CMAC)
```

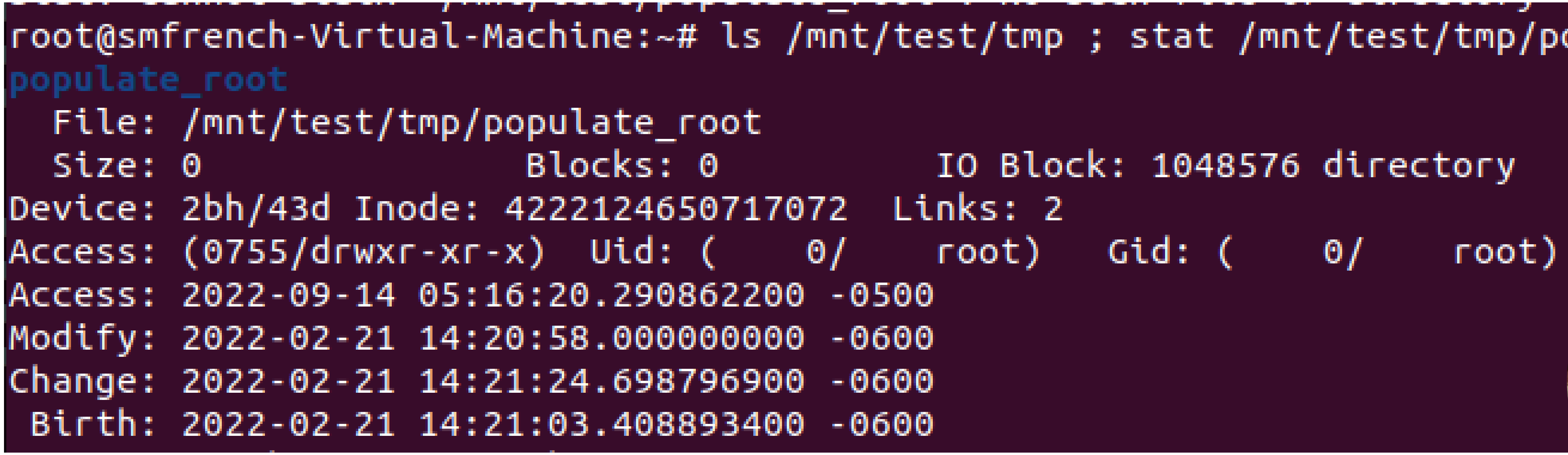
Example perf #s

- “dd if=/dev/zero of=/mnt/target bs=4M count=256”
- Signing (default prior to patches): 280MB/sec
- Signing (GMAC, with the experimental patches): 310MB/sec
- Encryption (vers=3.0, CCM): 170MB/sec
- Encryption (vers=3.1.1 GCM): 1.1GB/sec
- (testing on my laptop at SDC)

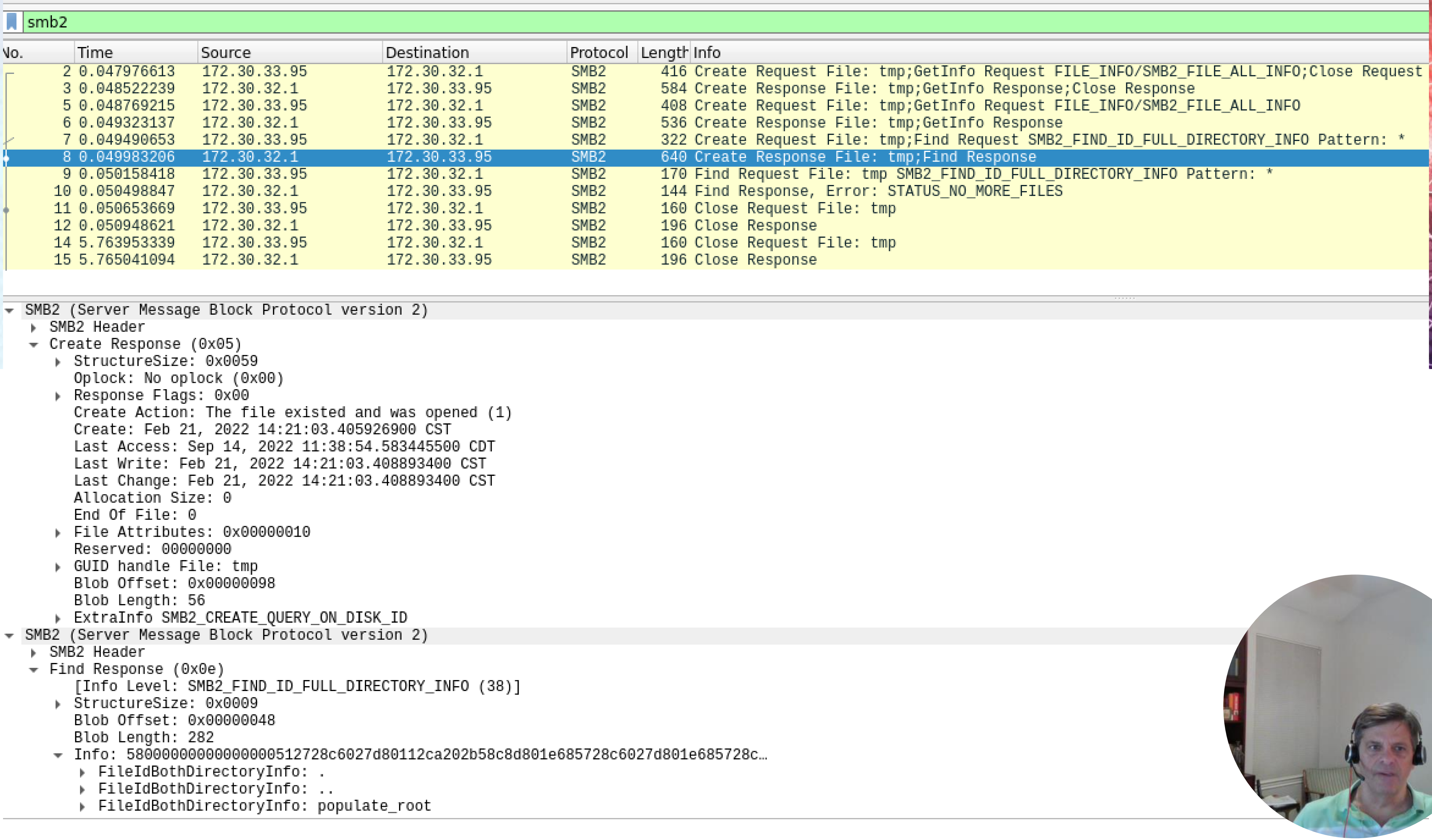
Directory Caching Improvements

- Thanks to Ronnie Sahlberg, directory caching and use of directory leases (to improve metadata caching even more, and safely) is MUCH improved
- Huge perf win!
- Continuing to optimize

Notice cached directory information with lease reduces requests needed (stat does not need to be sent)



```
root@smfrench-Virtual-Machine:~# ls /mnt/test/tmp ; stat /mnt/test/tmp/populate_root
populate_root
  File: /mnt/test/tmp/populate_root
  Size: 0               Blocks: 0          IO Block: 1048576 directory
Device: 2bh/43d Inode: 4222124650717072  Links: 2
Access: (0755/drwxr-xr-x)  Uid: (    0/    root)   Gid: (    0/    root)
Access: 2022-09-14 05:16:20.290862200 -0500
Modify: 2022-02-21 14:20:58.000000000 -0600
Change: 2022-02-21 14:21:24.698796900 -0600
 Birth: 2022-02-21 14:21:03.408893400 -0600
```



smb2

```
No.    Time         Source          Destination     Protocol  Length Info
     2 0.047976613  172.30.33.95    172.30.32.1     SMB2      416 Create Request File: tmp;GetInfo Request FILE_INFO/SMB2_FILE_ALL_INFO;Close Request
     3 0.048522239  172.30.32.1     172.30.33.95    SMB2      584 Create Response File: tmp;GetInfo Response;Close Response
     5 0.048769215  172.30.33.95    172.30.32.1     SMB2      408 Create Request File: tmp;GetInfo Request FILE_INFO/SMB2_FILE_ALL_INFO
     6 0.049323137  172.30.32.1     172.30.33.95    SMB2      536 Create Response File: tmp;GetInfo Response
     7 0.049490653  172.30.33.95    172.30.32.1     SMB2      322 Create Request File: tmp;Find Request SMB2_FIND_ID_FULL_DIRECTORY_INFO Pattern: *
     8 0.049983206  172.30.32.1     172.30.33.95    SMB2      640 Create Response File: tmp;Find Response
     9 0.050158418  172.30.33.95    172.30.32.1     SMB2      170 Find Request File: tmp SMB2_FIND_ID_FULL_DIRECTORY_INFO Pattern: *
    10 0.050498847  172.30.32.1     172.30.33.95    SMB2      144 Find Response, Error: STATUS_NO_MORE_FILES
    11 0.050653669  172.30.33.95    172.30.32.1     SMB2      160 Close Request File: tmp
    12 0.050948621  172.30.32.1     172.30.33.95    SMB2      196 Close Response
    14 5.763953339  172.30.33.95    172.30.32.1     SMB2      160 Close Request File: tmp
    15 5.765041094  172.30.32.1     172.30.33.95    SMB2      196 Close Response
```

```
SMB2 (Server Message Block Protocol version 2)
  SMB2 Header
  Create Response (0x05)
    StructureSize: 0x0059
    Oplock: No oplock (0x00)
    Response Flags: 0x00
    Create Action: The file existed and was opened (1)
    Create: Feb 21, 2022 14:21:03.405926900 CST
    Last Access: Sep 14, 2022 11:38:54.583445500 CDT
    Last Write: Feb 21, 2022 14:21:03.408893400 CST
    Last Change: Feb 21, 2022 14:21:03.408893400 CST
    Allocation Size: 0
    End Of File: 0
    File Attributes: 0x00000010
    Reserved: 00000000
    GUID handle File: tmp
    Blob Offset: 0x00000098
    Blob Length: 56
    ExtraInfo SMB2_CREATE_QUERY_ON_DISK_ID
SMB2 (Server Message Block Protocol version 2)
  SMB2 Header
  Find Response (0x0e)
    [Info Level: SMB2_FIND_ID_FULL_DIRECTORY_INFO (38)]
    StructureSize: 0x0009
    Blob Offset: 0x00000048
    Blob Length: 282
    Info: 580000000000000000512728c6027d80112ca202b58c8d801e685728c6027d801e685728c…
      FileIdBothDirectoryInfo: .
      FileIdBothDirectoryInfo: ..
      FileIdBothDirectoryInfo: populate_root
```

Deferred close improvements

- Currently used for i/o patterns like open/read/close/open/read/close
- Extending to cover many more scenarios, greatly improving performance
- Handle cache (deferred close time) now configurable with new mount parm “closetimeo” (thank you Bharath!)
- Improvements to lease break corner cases recently added

SMB1/CIFS deprecation

- SMB3.1.1 rocks …
- Gradually move the old, insecure dialects out of the default module used for SMB2.1/SMB3/SMB3.1.1, so easier to deprecate SMB1/CIFS

Multichannel improvements

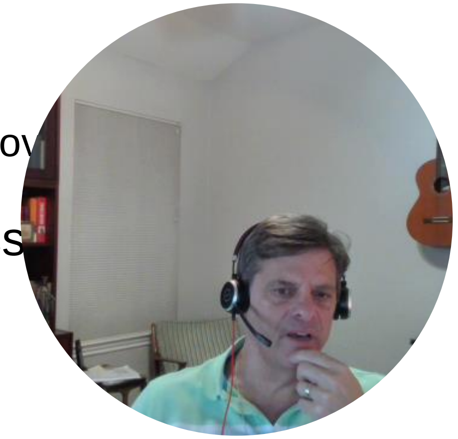
- Requerying network interfaces, dynamically adjusting
- Reconnect improvements
- Performance improvements (thank you Shyam Prasad!)
- Soon will be enabled by default (when server supports multiple interfaces or RSS)

SMB Direct (thanks to Tom Talpey and Metze)

- Reduce SGE usage, and decrease maximum fragment size
  - Fails to operate on SoftiWARP provider
  - Needless memory usage
  - High SGE usage impacts performance
- Fix RDMA “responder resources”, which do NOT apply to RDMA Writes
  - Significant performance limiter for bulk reads
- Fix sends to not wait for completion before returning
  - Stalls the pipeline, and costs significant context switching
- Use RDMA post-multiple to improve compound send efficiency
- Ensure packet kmem cache optimal packing (3x1364 == 4092)
- Review protocol parsing and state validation
  - E.g. ksmbd allows renegotiate (?), reassembles oversize segments (?)
- Hangs when shutting down with connection held
- Merge the two implementations: fs/cifs/smbdirect.[ch] and fs/ksmbd/transport_rdma.[ch]
  - Either refactor and merge, or consider metze’s alternative “smbdirect socket” driv

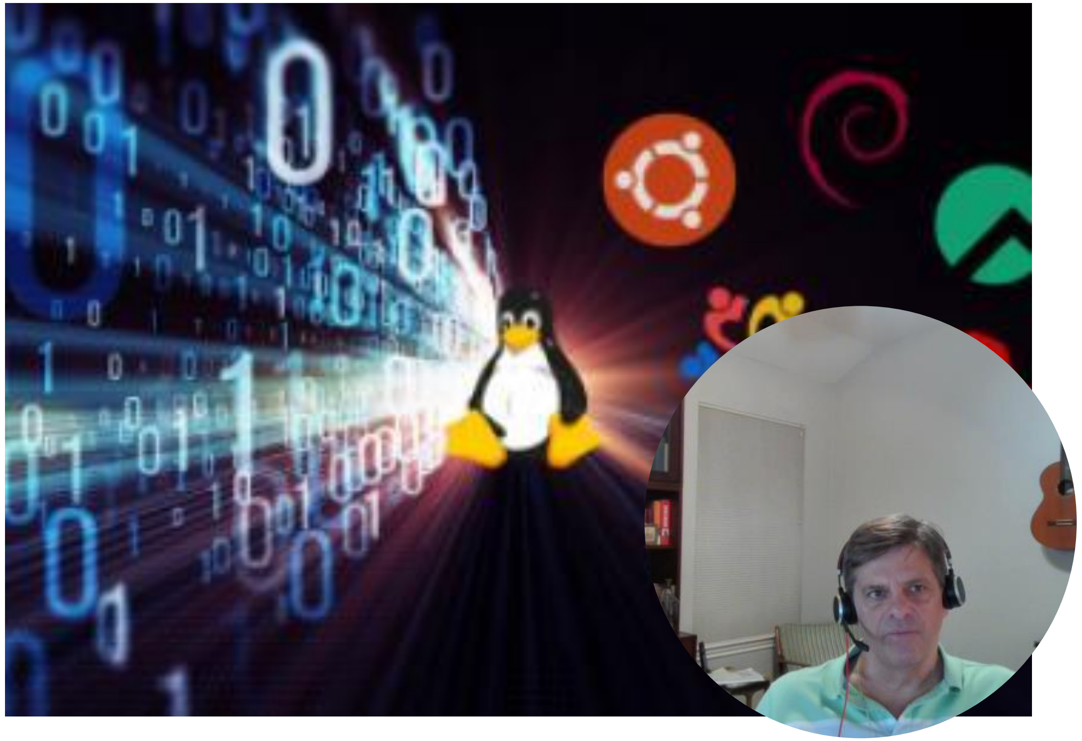
SMBDIRECT Transport Improvements – RDMA for the world

- SMBDIRECT is an abstraction layer for making RDMA useable more broadly. It has no SMB dependencies (SMB3 was just the first consumer of this generic transport layer, but it applies more broadly)
- Longer term plan is to:
  - Bring common from cifs.ko and ksmbd for RDMA into smbdirect.ko
  - Enable user space access to RDMA through smbdirect.ko so user space applications can benefit from the performance gains of RDMA
  - Improvements to this common module will benefit both client and server (and userspace)
- smbdirect.ko will provide
  - PF_SMBDIRECT sockets
  - Send message and receive message will get MSG_OOB messages for read and write offload, greatly improv and reducing CPU overhead
- (SMB independent) “echo server client” smbdirect tests under development to improve regress requiring SMB
- Thanks to Metze for this work. Feedback and review and testing welcome.



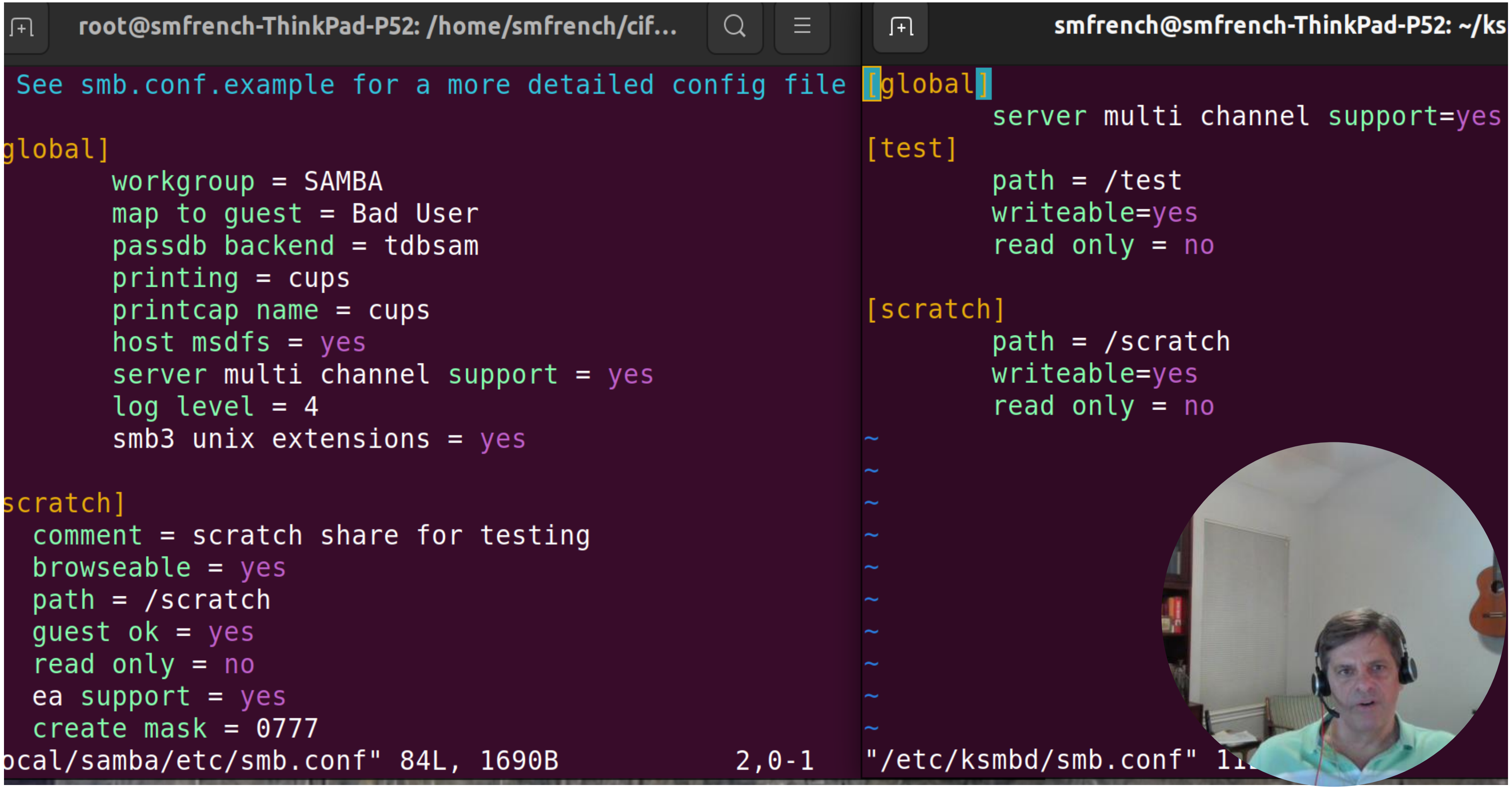
SMB3.1.1 POSIX Extensions

See my other talk at SambaXP and Volker’s talks
Has been in Linux client for years & is simpler than SMB1 Unix extensions
Great progress on Samba (client and server) and ksmbd server
Testing now possible with three servers and at least two clients



Setting up Samba and ksmbd shares are easy

Samba requires “smb3 unix extensions = yes” in smb.conf

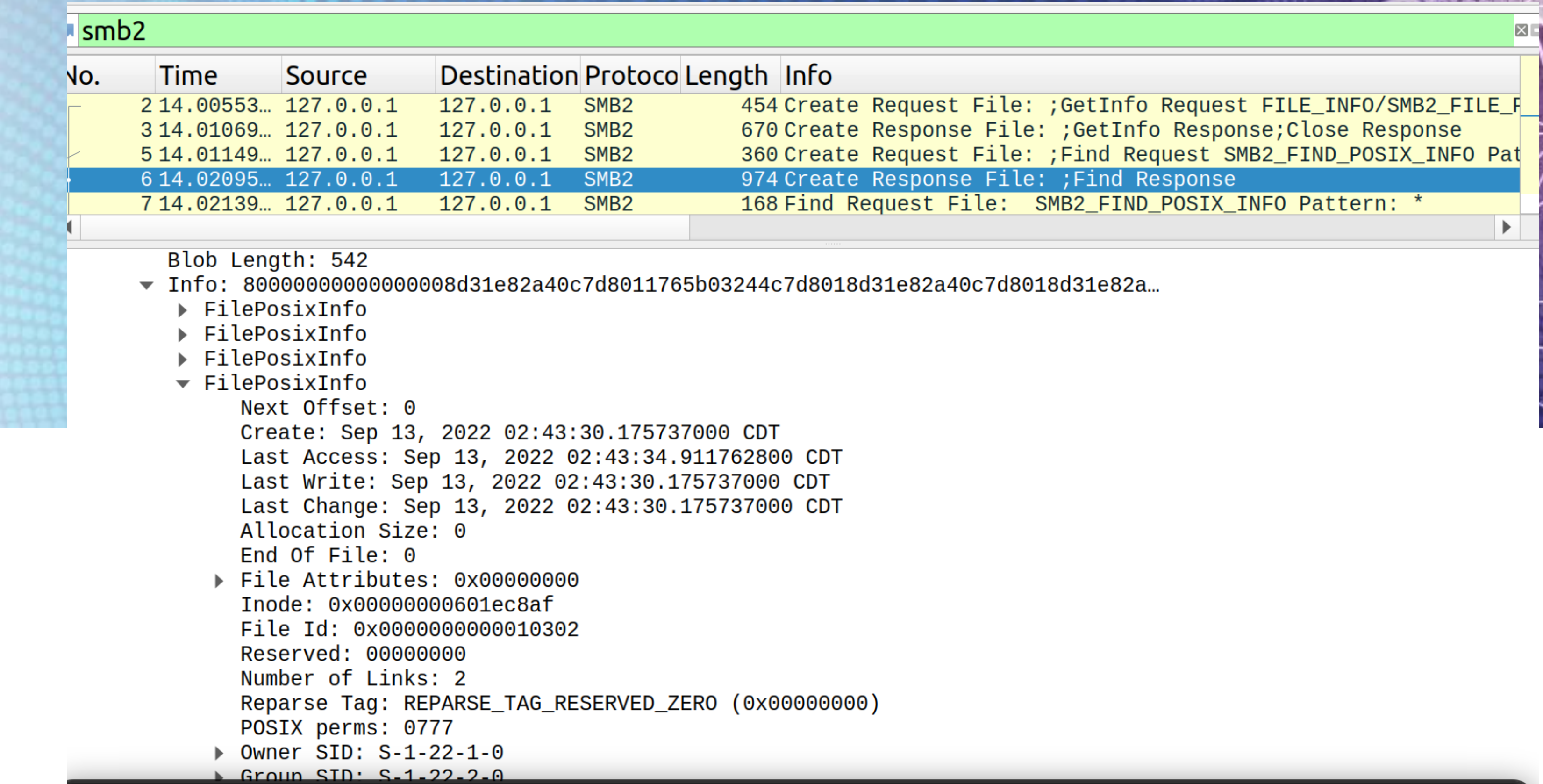


root@smfrench-ThinkPad-P52: /home/smfrench/cif...
smfrench@smfrench-ThinkPad-P52: ~/ks

```
See smb.conf.example for a more detailed config file

global]
        workgroup = SAMBA
        map to guest = Bad User
        passdb backend = tdbsam
        printing = cups
        printcap name = cups
        host msdfs = yes
        server multi channel support = yes
        log level = 4
        smb3 unix extensions = yes

scratch]
  comment = scratch share for testing
  browseable = yes
  path = /scratch
  guest ok = yes
  read only = no
  ea support = yes
  create mask = 0777
ocal/samba/etc/smb.conf" 84L, 1690B                2,0-1
```

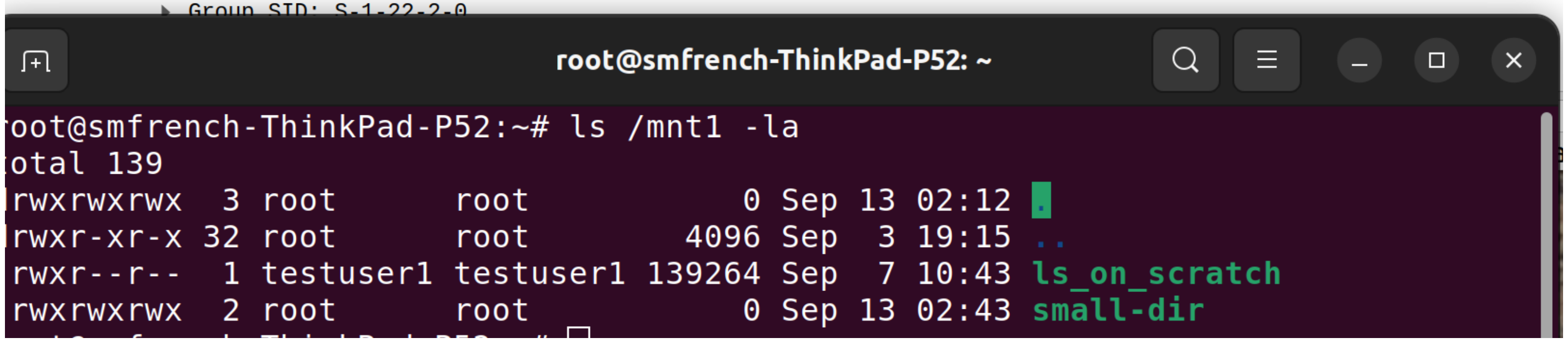
```
[global]
        server multi channel support=yes
[test]
        path = /test
        writeable=yes
        read only = no

[scratch]
        path = /scratch
        writeable=yes
        read only = no
~
~
~
~
~
~
~
~
~
~
"/etc/ksmbd/smb.conf" 11
```

Note exact mode bits and owner reported w/POSIX Extensions

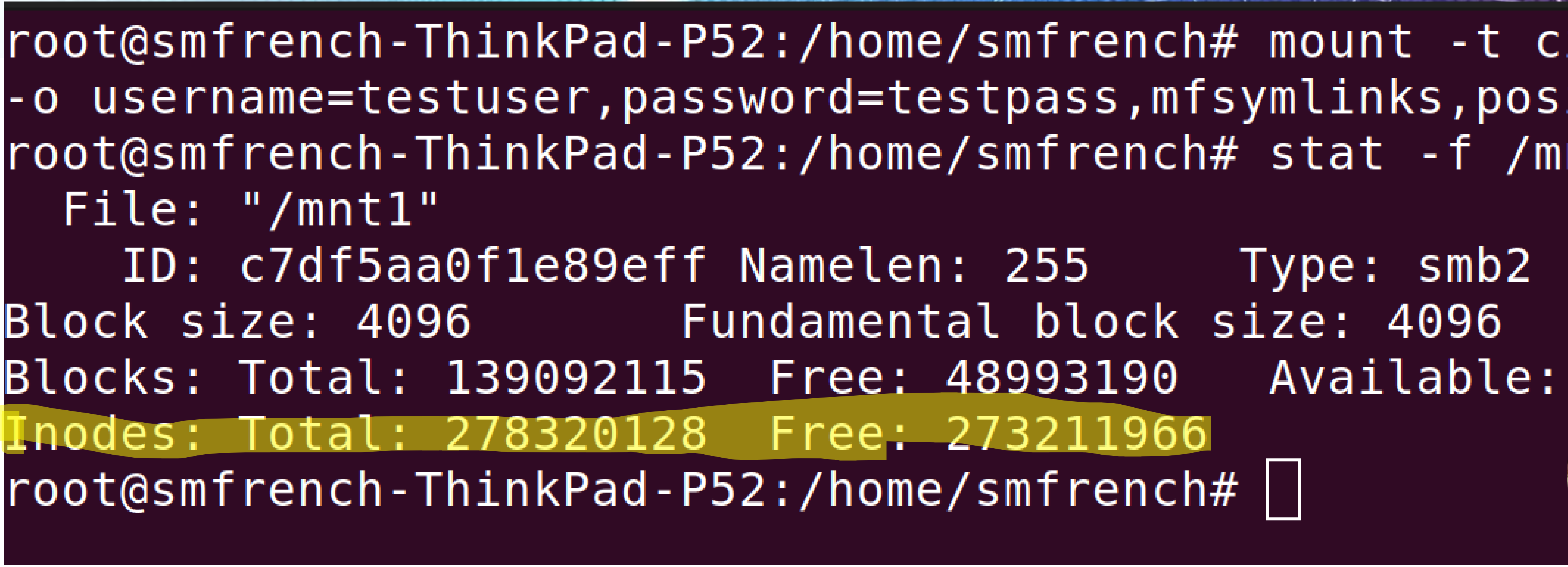


No.	Time	Source	Destination	Protoco	Length	Info
2	14.00553…	127.0.0.1	127.0.0.1	SMB2	454	Create Request File: ;GetInfo Request FILE_INFO/SMB2_FILE_F
3	14.01069…	127.0.0.1	127.0.0.1	SMB2	670	Create Response File: ;GetInfo Response;Close Response
5	14.01149…	127.0.0.1	127.0.0.1	SMB2	360	Create Request File: ;Find Request SMB2_FIND_POSIX_INFO Pat
6	14.02095…	127.0.0.1	127.0.0.1	SMB2	974	Create Response File: ;Find Response
7	14.02139…	127.0.0.1	127.0.0.1	SMB2	168	Find Request File: SMB2_FIND_POSIX_INFO Pattern: *

```
   Blob Length: 542
 ▾ Info: 800000000000000008d31e82a40c7d8011765b03244c7d8018d31e82a40c7d8018d31e82a…
   ▸ FilePosixInfo
   ▸ FilePosixInfo
   ▸ FilePosixInfo
   ▾ FilePosixInfo
       Next Offset: 0
       Create: Sep 13, 2022 02:43:30.175737000 CDT
       Last Access: Sep 13, 2022 02:43:34.911762800 CDT
       Last Write: Sep 13, 2022 02:43:30.175737000 CDT
       Last Change: Sep 13, 2022 02:43:30.175737000 CDT
       Allocation Size: 0
       End Of File: 0
     ▸ File Attributes: 0x00000000
       Inode: 0x00000000601ec8af
       File Id: 0x0000000000010302
       Reserved: 00000000
       Number of Links: 2
       Reparse Tag: REPARSE_TAG_RESERVED_ZERO (0x00000000)
       POSIX perms: 0777
     ▸ Owner SID: S-1-22-1-0
     ▸ Group SID: S-1-22-2-0
```

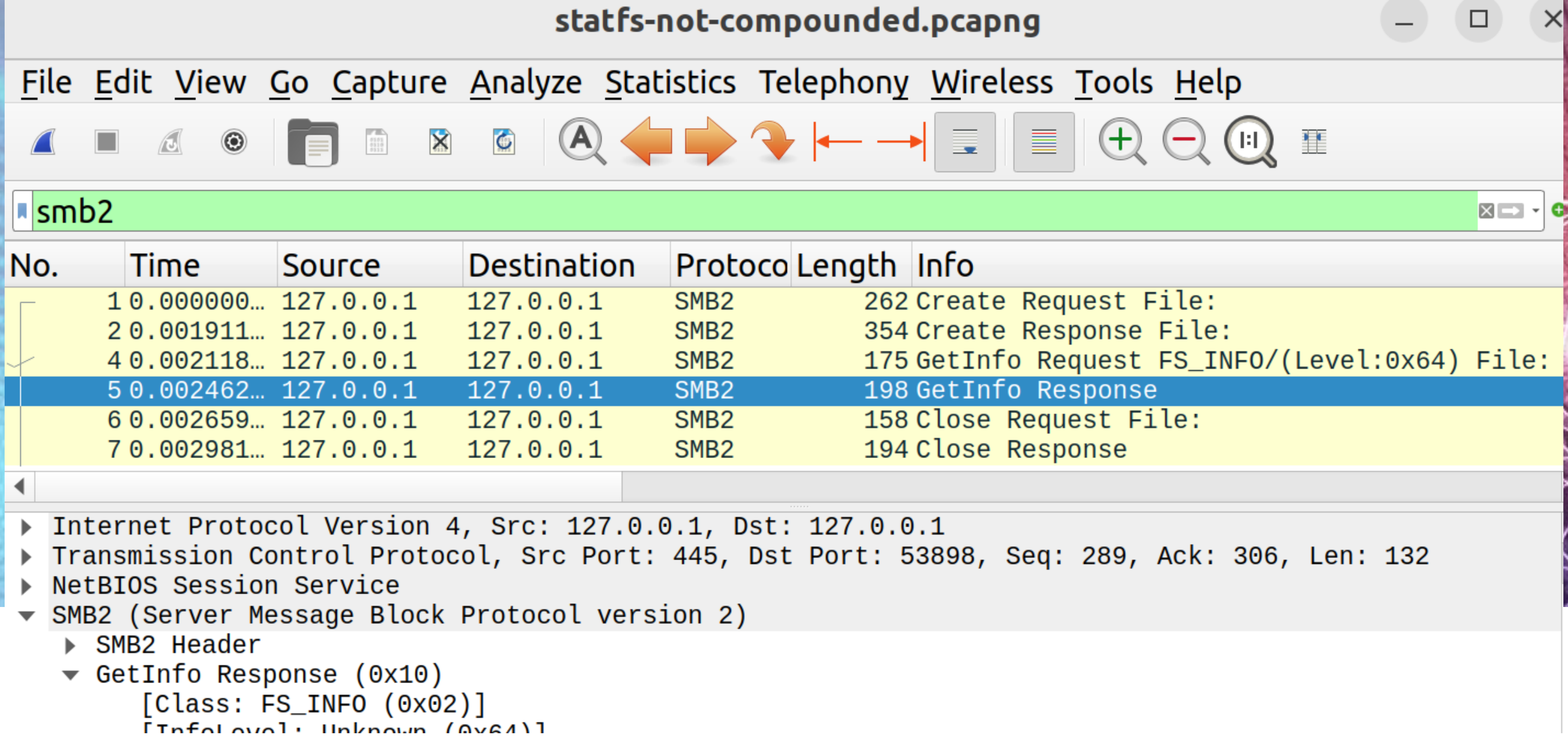


```
oot@smfrench-ThinkPad-P52:~# ls /mnt1 -la
otal 139
rwxrwxrwx  3 root       root           0 Sep 13 02:12 .
rwxr-xr-x 32 root       root        4096 Sep  3 19:15 ..
rwxr--r--  1 testuser1 testuser1 139264 Sep  7 10:43 ls_on_scratch
rwxrwxrwx  2 root       root           0 Sep 13 02:43 small-dir
```



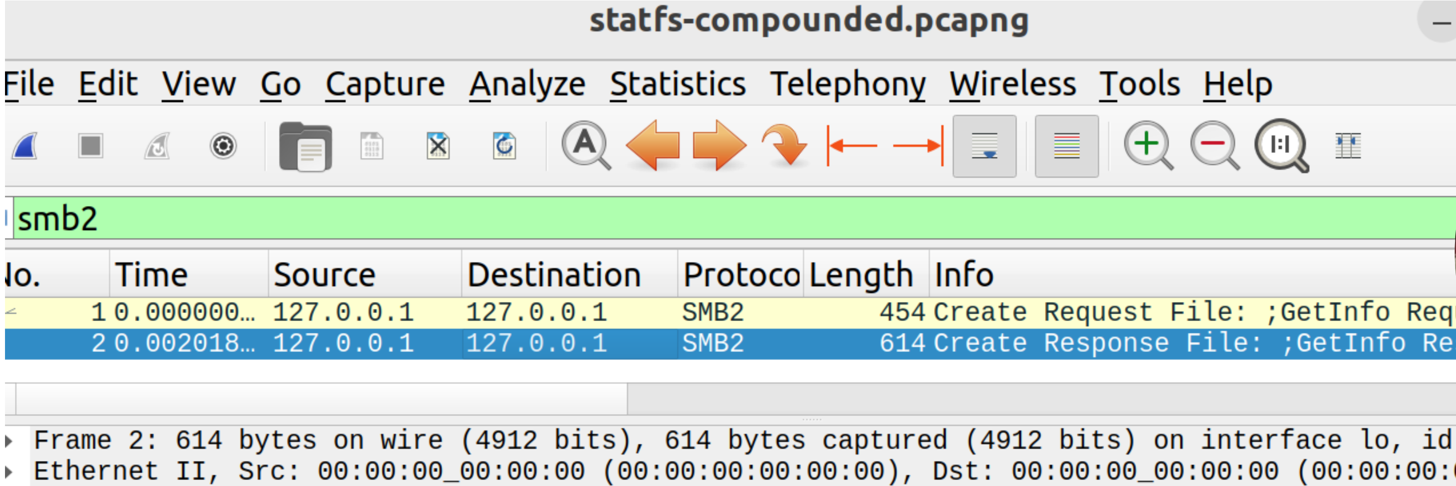
```
root@smfrench-ThinkPad-P52:/home/smfrench# mount -t cifs //local
-o username=testuser,password=testpass,mfsymlinks,posix
root@smfrench-ThinkPad-P52:/home/smfrench# stat -f /mnt1
  File: "/mnt1"
    ID: c7df5aa0f1e89eff Namelen: 255     Type: smb2
Block size: 4096       Fundamental block size: 4096
Blocks: Total: 139092115  Free: 48993190   Available:
Inodes: Total: 278320128  Free: 273211966
root@smfrench-ThinkPad-P52:/home/smfrench# 
```

Before:

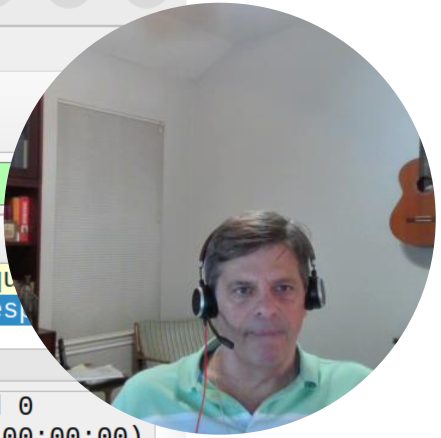

statfs-not-compounded.pcapng
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

```
smb2

No.  Time         Source     Destination  Protoco Length Info
   1 0.000000…   127.0.0.1  127.0.0.1    SMB2     262 Create Request File:
   2 0.001911…   127.0.0.1  127.0.0.1    SMB2     354 Create Response File:
   4 0.002118…   127.0.0.1  127.0.0.1    SMB2     175 GetInfo Request FS_INFO/(Level:0x64) File:
   5 0.002462…   127.0.0.1  127.0.0.1    SMB2     198 GetInfo Response
   6 0.002659…   127.0.0.1  127.0.0.1    SMB2     158 Close Request File:
   7 0.002981…   127.0.0.1  127.0.0.1    SMB2     194 Close Response

▸ Internet Protocol Version 4, Src: 127.0.0.1, Dst: 127.0.0.1
▸ Transmission Control Protocol, Src Port: 445, Dst Port: 53898, Seq: 289, Ack: 306, Len: 132
▸ NetBIOS Session Service
▾ SMB2 (Server Message Block Protocol version 2)
  ▸ SMB2 Header
  ▾ GetInfo Response (0x10)
      [Class: FS_INFO (0x02)]
      [InfoLevel: Unknown (0x64)]
```

Now better performance (POSIX QFS Info now compounded)

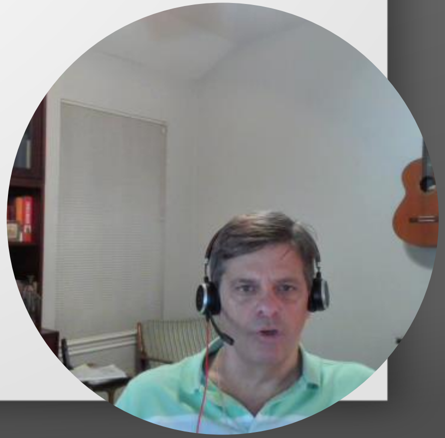

statfs-compounded.pcapng
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

```
smb2

No.  Time         Source     Destination  Protoco Length Info
   1 0.000000…   127.0.0.1  127.0.0.1    SMB2     454 Create Request File: ;GetInfo Requ
   2 0.002018…   127.0.0.1  127.0.0.1    SMB2     614 Create Response File: ;GetInfo Resp

▸ Frame 2: 614 bytes on wire (4912 bits), 614 bytes captured (4912 bits) on interface lo, id 0
▸ Ethernet II, Src: 00:00:00_00:00:00 (00:00:00:00:00:00), Dst: 00:00:00_00:00:00 (00:00:00:00:00:00)
```



5.17 kernel (March 20th), 51 changesets, cifs.ko ver 2.35

- Add support for new fscache (offline files caching mechanism)
- Send additional NTLMSSP info (including module and OS version) for improved debugging
- DFS and ACL fixes
- Key modefromsid fix (where client enforced mode bits retrieved from special ACE)
- Restructuring of multichannel code



5.18 kernel (May 22nd), 40 changesets, cifs.ko ver 2.36

- Important performance improvement (reuse cached file handle for various common operations like stat and statfs if available), greatly reducing metadata operations (like open/close)
- Important fscache (offline file caching) and DFS improvements
- cross mount reflink now supported, which can dramatically improve copy performance from one share to another (on the same server) if they support duplicate extents.

5.19 kernel (July 31st, 2022)

- Important performance optimization for directory searches, now we cache the root directory content (to the many servers which support directory leases) reducing amount of network traffic for queries in the root directory
- Multichannel reconnect improvements (e.g. when address or interfaces change)
- RDMA (smbdirect) improvements
- New mount parm “nosparse” to optionally disable use of spars

6.0 kernel (October 2nd, 2022) (cifs module version 2.39)

- Fallocate improvements (insert and collapse range)
- Module size shrunk significantly when SMB1/CIFS (insecure legacy) disabled
- New mount parm “closetimeo” allows extending deferred closes (handle leases) longer or even disabling the feature (and default increased to 5 seconds from 1 sec)
- Important deferred close fix
- Multichannel perf (locking) improvements

6.1 kernel (Dec 11th, 2022) (cifs module ver: 2.40)

- Performance improvement for path revalidation (metadata ops perf better) by using cached dentry for subdirectory if lease held on it
  - Expanding cached directories to include subdirectories (thanks Ronnie!)
- New ioctl for change notify added that returns the name(s) of any changed files in the directory (not just that the directory has changed)
  - e.g. so app can do their own offline caching of files and sync with server
- Improve symlink handling (avoid an extra roundtrip when symlink detected via STOPPED_ON_SYMLINK message)
- RDMA (smbdirect) improvements (thanks Tom Talpey and M

6.2 kernel (February 19th, 2023) (cifs module 2.41)

- Important SMB3.1.1 POSIX extensions improvement (parse owner and group SIDs to improve stat output)
- DFS performance improvements (reducing roundtrips) and DFS fixes
- Multichannel improvements
- Integration with the new kernel page caching infrastructure, folios, iov_iter and memory management layering cleanup

6.3 kernel (April 23rd, 2023) (cifs module 2.42)

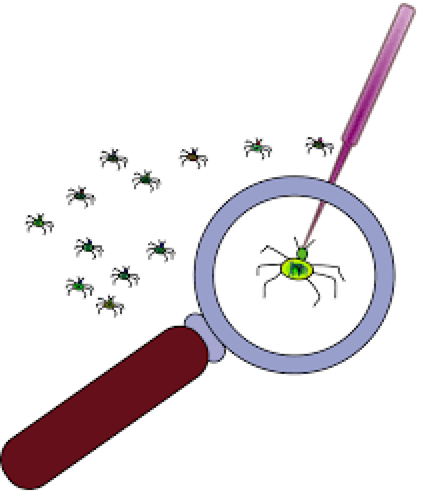
- Kernel idmapping improvements
- Improvements to use folios (better mm integration and cached writes)
- RDMA (smbdirect) improvements (thanks Metze and David)
- Many multichannel improvements (including using least loaded channel for sending I/O, and improvements for reconnect). Thanks Shyam!
- Various DFS fixes
- Lower default deferred close timeout

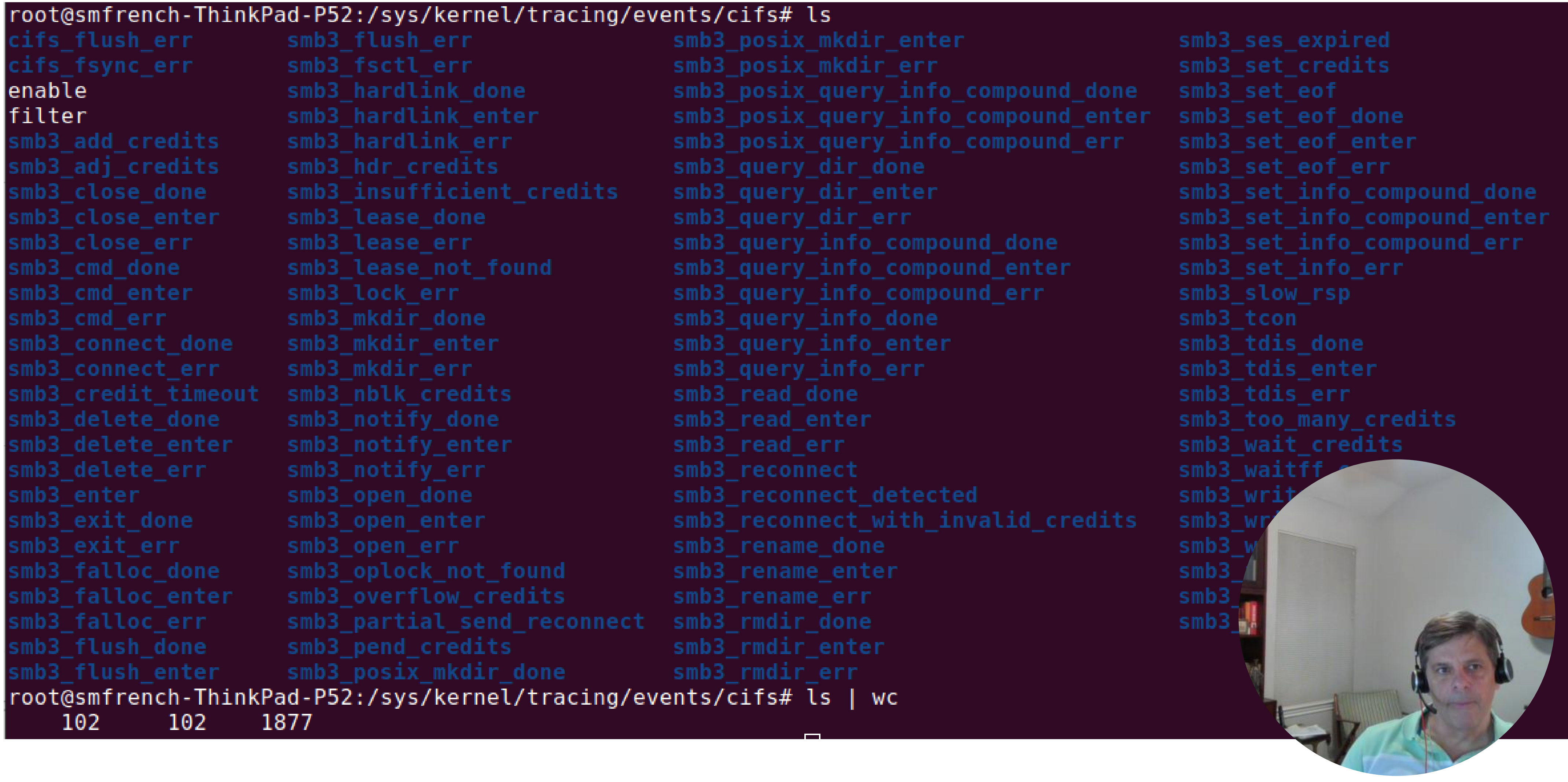
6.4-rc kernel (expected early July) (cifs.ko version: 2.43)

- Important deferred close (lease break corner case) fixes
- Reconnect and DFS fixes
- Important crediting improvements expected
- Compounding improvements expected

Tracing continues to improve ...

- Added 4 additional dynamic tracepoints





```
root@smfrench-ThinkPad-P52:/sys/kernel/tracing/events/cifs# ls
cifs_flush_err       smb3_flush_err               smb3_posix_mkdir_enter                smb3_ses_expired
cifs_fsync_err       smb3_fsctl_err               smb3_posix_mkdir_err                  smb3_set_credits
enable               smb3_hardlink_done           smb3_posix_query_info_compound_done   smb3_set_eof
filter               smb3_hardlink_enter          smb3_posix_query_info_compound_enter  smb3_set_eof_done
smb3_add_credits     smb3_hardlink_err            smb3_posix_query_info_compound_err    smb3_set_eof_enter
smb3_adj_credits     smb3_hdr_credits             smb3_query_dir_done                   smb3_set_eof_err
smb3_close_done      smb3_insufficient_credits    smb3_query_dir_enter                  smb3_set_info_compound_done
smb3_close_enter     smb3_lease_done              smb3_query_dir_err                    smb3_set_info_compound_enter
smb3_close_err       smb3_lease_err               smb3_query_info_compound_done         smb3_set_info_compound_err
smb3_cmd_done        smb3_lease_not_found         smb3_query_info_compound_enter        smb3_set_info_err
smb3_cmd_enter       smb3_lock_err                smb3_query_info_compound_err          smb3_slow_rsp
smb3_cmd_err         smb3_mkdir_done              smb3_query_info_done                  smb3_tcon
smb3_connect_done    smb3_mkdir_enter             smb3_query_info_enter                 smb3_tdis_done
smb3_connect_err     smb3_mkdir_err               smb3_query_info_err                   smb3_tdis_enter
smb3_credit_timeout  smb3_nblk_credits            smb3_read_done                        smb3_tdis_err
smb3_delete_done     smb3_notify_done             smb3_read_enter                       smb3_too_many_credits
smb3_delete_enter    smb3_notify_enter            smb3_read_err                         smb3_wait_credits
smb3_delete_err      smb3_notify_err              smb3_reconnect                        smb3_waitff_
smb3_enter           smb3_open_done               smb3_reconnect_detected               smb3_writ
smb3_exit_done       smb3_open_enter              smb3_reconnect_with_invalid_credits   smb3_wr
smb3_exit_err        smb3_open_err                smb3_rename_done                      smb3_w
smb3_falloc_done     smb3_oplock_not_found        smb3_rename_enter                     smb3_
smb3_falloc_enter    smb3_overflow_credits        smb3_rename_err                       smb3_
smb3_falloc_err      smb3_partial_send_reconnect  smb3_rmdir_done                       smb3_
smb3_flush_done      smb3_pend_credits            smb3_rmdir_enter
smb3_flush_enter     smb3_posix_mkdir_done        smb3_rmdir_err
root@smfrench-ThinkPad-P52:/sys/kernel/tracing/events/cifs# ls | wc
    102     102    1877
```

eBPF is amazing ...

- See Brendan Gregg’s website
- Also see e.g. https://wiki.samba.org/index.php/LinuxCIFS_troubleshooting
- Can be as simple to do as “trace-cmd record -e cifs”
  - And then “trace-cmd show” in another window
- Let us know if suggestions on other debugging tracepoints that would be helpful
- And don’t forget about *proc/fs/cifs/Stats, proc/fs/cifs/open_files and* *proc*/fs/cifs/DebugData ...

Recent improvements – cifs-utils

Userspace tools

Improved user space tools (cifs-utils)

Speaker Photo Will Be Placed Here

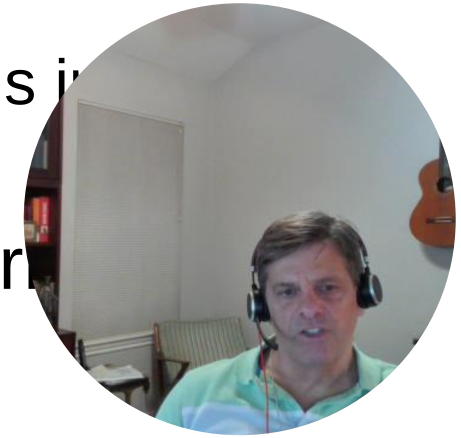
- cifs-utils 7 released in August
  - Add support for gss-proxy (improving krb5 credential retrieval)
  - Misc. bug fixes
- Contributions welcome – lots of cool opportunities for tooling
  - e.g. to leverage the new notify ioctl or to make snapshot mounts easier (one step instead of two) or to improve backup tooling

Coming soon ...

New features under development for SMB3.1.1 on Linux

What features can you expect in next few releases?

- Analyze cases where use of directory leases, deferred close operations could better optimize network traffic while caching safely
- Add use of compounding in more cases or extend it (e.g. open/querydir/querydir instead of open/querydir), and better use existing file leases for compound reqs which include SMB3 open
- Improvements to performance when low on SMB3 credits
- Continued focus on multichannel performance improvements
  - Dynamically adding channels better, and picking optimal channels i special cases
- SMB3.1.1 compression support (allow compressing networ based on the SMB3.1.1 compress mount parm)



What features can you expect in next few releases?

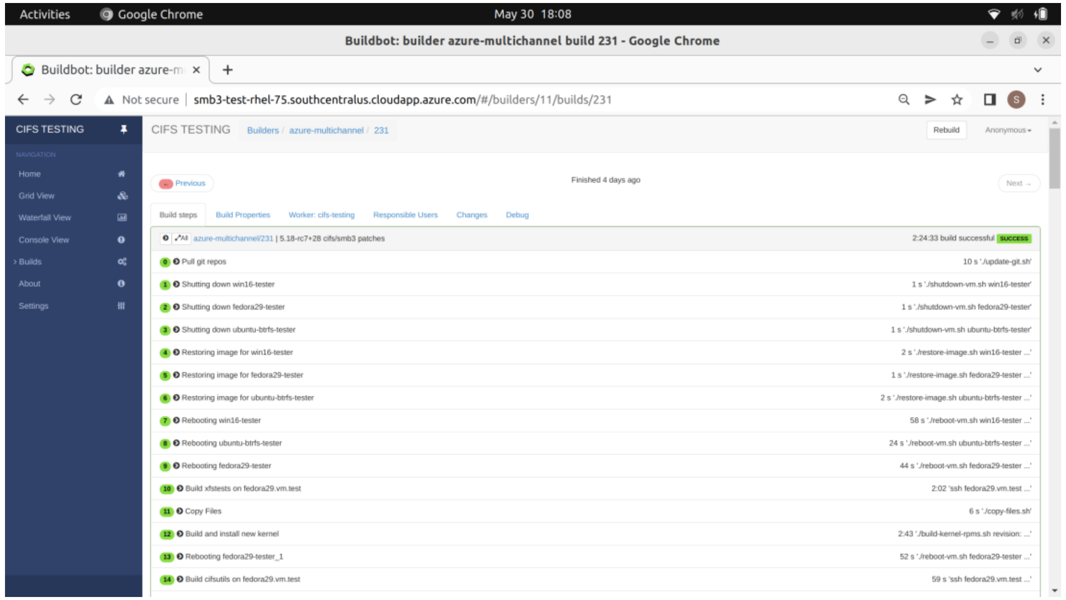
- Packet signing performance improvements
- Reenabling support for swapfile over SMB3.1.1 mounts
- Support for creating with O_TMPFILE
- statx to return additional SMB3.1.1 attributes like “offline”
- Improvements to enable fanotify/inotify over SMB3.1.1 mounts (currently requires a private SMB3.1.1 specific ioctl)
- Prototype of SMB3.1.1 over QUIC (new encrypted network tra)
- More perf improvements for folios, cache, parallel i/o, multic (thank you Dave Howells, Matthew Wilcox et al)
- More testing of the SMB3.1.1 POSIX with new Samba serve

Testing Improvements

Section Subtitle

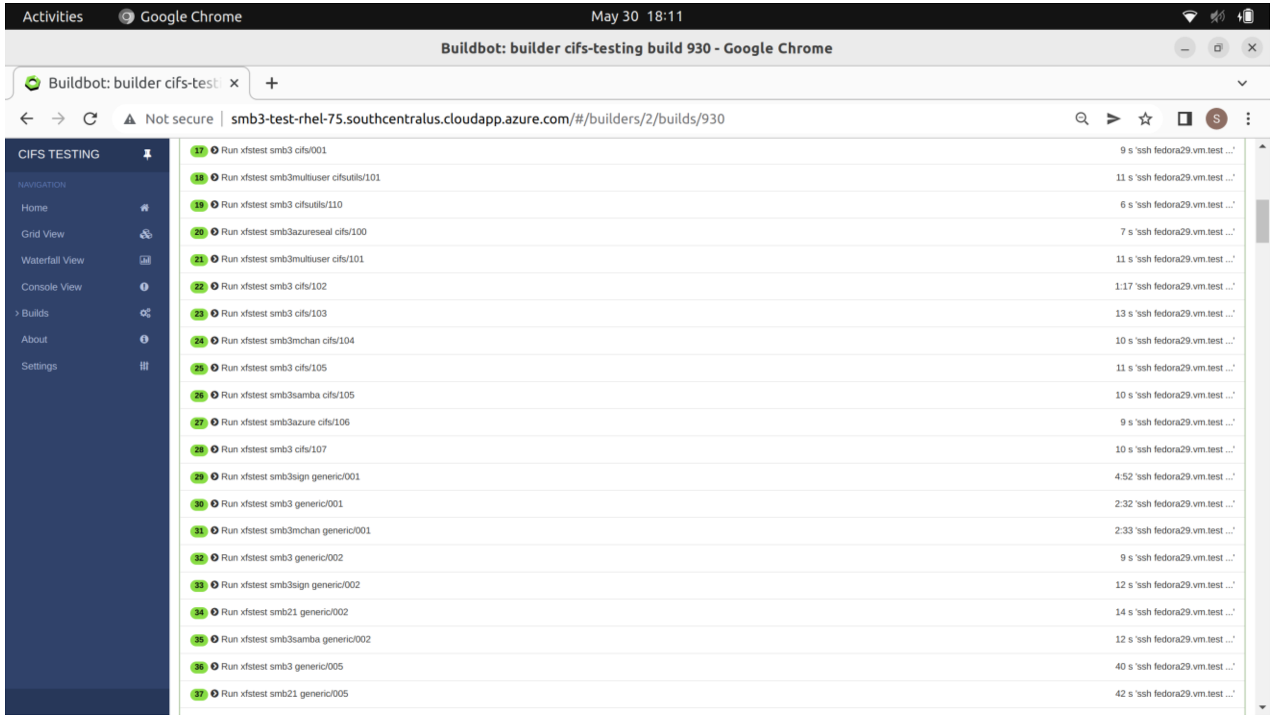
Automated testing has greatly improved

- Historically SMB3.1.1 plugfests multiple times a year have hel
- The ‘buildbot’ continues to improve, more tests added, reducing regressions and improving quality, migrating to new better hosts
- Test groups for different server types and a general “cifs-testing” one



Finished 4 days ago

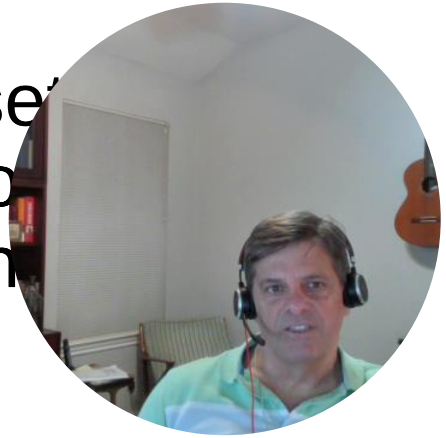
```
azure-multichannel/231 | 5.18-rc7+28 cifs/smb3 patches          2:24:33 build successful SUCCESS
0  Pull git repos                                   10 s './update-git.sh'
1  Shutting down win16-tester                       1 s './shutdown-vm.sh win16-tester'
2  Shutting down fedora29-tester                    1 s './shutdown-vm.sh fedora29-tester'
3  Shutting down ubuntu-btrfs-tester                1 s './shutdown-vm.sh ubuntu-btrfs-tester'
4  Restoring image for win16-tester                 2 s './restore-image.sh win16-tester ...'
5  Restoring image for fedora29-tester              1 s './restore-image.sh fedora29-tester ...'
6  Restoring image for ubuntu-btrfs-tester          2 s './restore-image.sh ubuntu-btrfs-tester ...'
7  Rebooting win16-tester                           58 s './reboot-vm.sh win16-tester ...'
8  Rebooting ubuntu-btrfs-tester                    24 s './reboot-vm.sh ubuntu-btrfs-tester ...'
9  Rebooting fedora29-tester                        44 s './reboot-vm.sh fedora29-tester ...'
10 Build xfstests on fedora29.vm.test               2:02 'ssh fedora29.vm.test ...'
11 Copy Files                                       6 s './copy-files.sh'
12 Build and install new kernel                     2:43 './build-kernel-rpms.sh revision: ...'
13 Rebooting fedora29-tester_1                      52 s './reboot-vm.sh fedora29-tester ...'
14 Build cifsutils on fedora29.vm.test              59 s 'ssh fedora29.vm.test ...'
```


Activities
Google Chrome
May 30 18:11
Buildbot: builder cifs-testing build 930 - Google Chrome
Buildbot: builder cifs-test
Not secure
smb3-test-rhel-75.southcentralus.cloudapp.azure.com/#/builders/2/builds/930
CIFS TESTING
NAVIGATION
Home
Grid View
Waterfall View
Console View
Builds
About
Settings

```
17 Run xfstest smb3 cifs/001                        9 s 'ssh fedora29.vm.test ...'
18 Run xfstest smb3multiuser cifsutils/101          11 s 'ssh fedora29.vm.test ...'
19 Run xfstest smb3 cifsutils/110                   6 s 'ssh fedora29.vm.test ...'
20 Run xfstest smb3azureseal cifs/100               7 s 'ssh fedora29.vm.test ...'
21 Run xfstest smb3multiuser cifs/101               11 s 'ssh fedora29.vm.test ...'
22 Run xfstest smb3 cifs/102                        1:17 'ssh fedora29.vm.test ...'
23 Run xfstest smb3 cifs/103                        13 s 'ssh fedora29.vm.test ...'
24 Run xfstest smb3mchan cifs/104                   10 s 'ssh fedora29.vm.test ...'
25 Run xfstest smb3 cifs/105                        11 s 'ssh fedora29.vm.test ...'
26 Run xfstest smb3samba cifs/105                   10 s 'ssh fedora29.vm.test ...'
27 Run xfstest smb3azure cifs/106                   9 s 'ssh fedora29.vm.test ...'
28 Run xfstest smb3 cifs/107                        10 s 'ssh fedora29.vm.test ...'
29 Run xfstest smb3sign generic/001                 4:52 'ssh fedora29.vm.test ...'
30 Run xfstest smb3 generic/001                     2:32 'ssh fedora29.vm.test ...'
31 Run xfstest smb3mchan generic/001                2:33 'ssh fedora29.vm.test ...'
32 Run xfstest smb3 generic/002                     9 s 'ssh fedora29.vm.test ...'
33 Run xfstest smb3sign generic/002                 12 s 'ssh fedora29.vm.test ...'
34 Run xfstest smb21 generic/002                    14 s 'ssh fedora29.vm.test ...'
35 Run xfstest smb3samba generic/002                12 s 'ssh fedora29.vm.test ...'
36 Run xfstest smb3 generic/005                     40 s 'ssh fedora29.vm.test ...'
37 Run xfstest smb21 generic/005                    42 s 'ssh fedora29.vm.test ...'
```

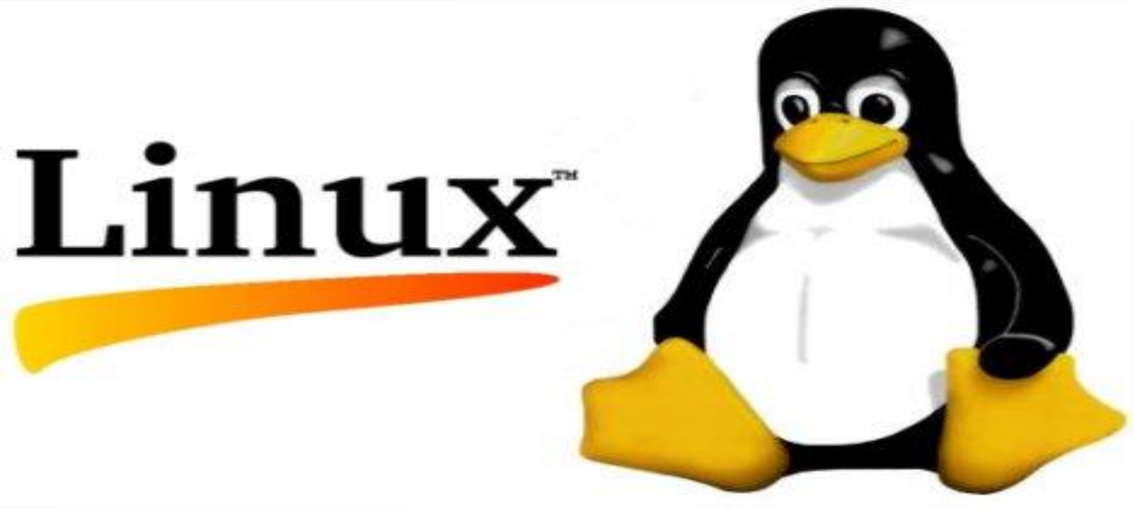
Additional tests are encouraged

- Xfstests are the standard Linux filesystem functional tests
- “Buildbot” is in the process of being upgraded/migrated to new hosts
- Last year added 21 to the main “cifs-testing” regression testing group (up to 245 tests run on every checkin from this group)
- Various server specific groups have added even more
  - Azure SMB3.1.1 multichannel: up 25% more tests, now includes 133 tests
  - Ksmbd (Linux kernel server target) up 15%, now includes 144 tests
- Detailed wiki pages on wiki.samba.org go through how to se xfstests with cifs.ko, and what features need to be added to more tests (tests that currently skip or fail so aren’t run in th



Thank you for your time

- Future is very bright!



+ S M B 3

Additional Resources to Explore for SMB3 and Linux

- https://msdn.microsoft.com/en-us/library/gg685446.aspx
  - In particular MS-SMB2.pdf at https://msdn.microsoft.com/en-us/library/cc246482.aspx
- https://wiki.samba.org/index.php/Xfstesting-cifs
- Linux CIFS client https://wiki.samba.org/index.php/LinuxCIFS
- Samba-technical mailing list and IRC channel
- And various presentations at http://www.sambaxp.org and Microsoft channel 9 and of course SNIA … http://www.snia.org/events/storage-developer
- And the code:
  - https://git.kernel.org/cgit/linux/kernel/git/torvalds/linux.git/tree/fs/cifs
  - For pending changes, soon to go into upstream kernel see:
    - https://git.samba.org/?p=sfrench/cifs-2.6.git;a=shortlog;h=refs/heads/for-next
  - Kernel server code: https://git.samba.org/ksmbd.git/?p=ksmbd.git (ksmbd-for-next branch)