

# SAMBA

## EXPERIENCE

## New Transports in Samba

quic + highspeed with io\_uring and smbdirect ...

Stefan Metzmacher <metze@samba.org>

Samba Team / SerNet

2026-04-20

<https://samba.org/~metze/presentations/2026/SambaXP/>

# Topics

- ▶ Existing Transports: NBT, TCP
- ▶ New Transports: SMBDIRECT, QUIC
- ▶ QUIC support using quic.ko on Linux
- ▶ How QUIC is tested without kernel support
- ▶ SMBDIRECT support for Linux and Samba
- ▶ How we test without kernel support
- ▶ Generic improvements using io\_uring
- ▶ Performance boost with io\_uring
- ▶ smb2.bench.read performance examples
- ▶ Samba is back in the game
- ▶ Questions? Feedback!

# Last Status Updates for io\_uring and smbdirect

- ▶ smbdirect support within Samba and Linux (SDC 2018):
  - ▶ See <https://samba.org/~metze/presentations/2018/SDC/>
  - ▶ It explains the initial research and prototype phase (in detail)
- ▶ io\_uring status update within Samba (SDC 2020):
  - ▶ See <https://samba.org/~metze/presentations/2020/SDC/>
  - ▶ It explains the milestones and design up to Samba 4.13 (in detail)
- ▶ io\_uring status update within Samba (SDC 2021):
  - ▶ See <https://samba.org/~metze/presentations/2021/SDC/>
  - ▶ It explains the milestones and updates up to Samba 4.15 (in detail)
- ▶ io\_uring status update within Samba (SambaXP 2023):
  - ▶ See <https://samba.org/~metze/presentations/2023/SambaXP/>
  - ▶ It explains the milestones and updates up to Samba 4.19 (in detail)
- ▶ new transports in Samba, quic and smbdirect support (SDC 2025):
  - ▶ See <https://samba.org/~metze/presentations/2025/SDC/>
  - ▶ It explains the milestones and updates up to Samba 4.23 (in detail)

## Existing Transports: NBT, TCP

- ▶ SMB Versions 1, 2 and 3 operate on message based transports
  - ▶ Each transport message can transport one or more SMB PDUs
- ▶ The oldest transport supported by Samba is NBT over TCP
  - ▶ This uses tcp port 139
  - ▶ Message header: 8 bits type, 7 bits reserved, 17 bits length
  - ▶ Simple exchange of netbios names at the start of a connection
  - ▶ Payload header (type=0): 4 bytes big endian, max ~128KiB (0x1FFFF)
  - ▶ Keepalive messages (type=85)
- ▶ Direct TCP is also (almost forever) available in Samba
  - ▶ This uses tcp port 445
  - ▶ Message header: 8 bits zero, 24 bits length
  - ▶ Just payload messages: 4 bytes big endian, max ~16MiB (0xFFFFFFFF)
- ▶ Both operate just on BSD stream sockets for TCP
  - ▶ They are identical to implement for the payload messages

# New Transports: SMBDIRECT

- ▶ SMBDIRECT (a tiny layer on top of RDMA/IB Verbs APIs)
  - ▶ Infiniband, RoCE (v1 and v2) and iWarp are supported protocols
  - ▶ A message based transport is implement using SEND[\_INV] verbs
  - ▶ SEND verbs are also message based but only allow small messages
  - ▶ Payload messages are reassembled, typically up to 1MiB (0x100000)
  - ▶ RDMA Read/Write verbs are used as out of band memory copy
- ▶ Windows 2012 introduced it to be used via SMB3 multi-channel
  - ▶ Clients start with TCP and ask for server interface IP addresses
  - ▶ The client connects via SMBDIRECT is the interface is RDMA capable
  - ▶ But the connection only uses IP addresses, so Infinibad requires IPoIB
- ▶ SMB2/3 Read/Write can offload data transfer to RDMA Read/Write
  - ▶ The client tells the server what client memory the server can read/write
  - ▶ The server is the initiator for RDMA Read/Write
  - ▶ SMB2 Read results in RDMA Write from server to client
  - ▶ SMB2 Write results in RDMA Read from client to server

# New Transports: QUIC

- ▶ QUIC a rich stream protocol based on TLS 1.3 and UDP
  - ▶ Designed for HTTP/3 with security, performance and robustness
  - ▶ Used UDP port 443 and should go through routers and firewalls
  - ▶ Supports Application-Layer Protocol Negotiation (ALPN)
  - ▶ ALPN: "h3" for HTTP/3, "smb" for SMB3
  - ▶ Multiple logical stream over a single QUIC connection
- ▶ Windows 2022 introduced SMB3 over QUIC
  - ▶ Clients try TCP first and fallback to QUIC
  - ▶ Only a single logical stream is used per QUIC connection
  - ▶ Message header: 8 bits zero, 24 bits length (Same as Direct TCP!)
  - ▶ Clients need to access the server via a name
  - ▶ Via ip address the certificate would not match
  - ▶ SMB3 encryption is done in addition to the QUIC encryption
- ▶ Windows 2025 introduced new features
  - ▶ The double encryption of QUIC and SMB3 layer is optional
  - ▶ SMB2\_ACCEPT\_TRANSPORT\_LEVEL\_SECURITY can be negotiated
  - ▶ Server can required client certificates

# QUIC support using quic.ko on Linux

- ▶ quic.ko for Linux provides a BSD stream socket
  - ▶ Standalone: <https://github.com/lxin/quic>
  - ▶ It uses `socket(AF_INET[6], SOCK_STREAM, IPPROTO_QUIC)`;
  - ▶ Userspace needs to call `quic_handshake()` until connection is ready
  - ▶ From there it is just a stream socket similar to TCP
- ▶ The changes to Samba are not very complex
  - ▶ We just need to replace `IPPROTO_TCP` with `IPPROTO_QUIC`
  - ▶ Then setup the gnutls bases tls structures
  - ▶ And call `quic_handshake()`
  - ▶ The message processing is not changed at all
  - ▶ Only `smbstatus -json` output was adjusted to show the transport
  - ▶ New options "client smb transports" and "server smb transport"
- ▶ Linux quic.ko will hopefully mainlined soon
  - ▶ Review for v11 of the patches:
  - ▶ <https://lore.kernel.org/quic/cover.1774410440.git.lucien.xin@gmail.com/>

# QUIC support with only UDP sockets from the kernel

- ▶ On systems without IPPROTO\_QUIC support we use libngtcp2
  - ▶ It provides userspace QUIC support on top of UDP sockets
  - ▶ It is relatively easy to use with gnutls
  - ▶ We can use the same handshake code as with quic.ko
  - ▶ It can operate non-blocking and has callback apis
- ▶ It is easy to use in Samba on the client side
  - ▶ It plugs nicely into our tevent and tstream abstractions
  - ▶ We just use it as fallback if quic.ko is not available
  - ▶ Should also work on non Linux systems
- ▶ Server support is not possible
  - ▶ We rely stream sockets and fd-passing for multichannel
  - ▶ Sharing a single UDP socket would be very complex

# SMBDIRECT support for Linux and Samba

- ▶ The Linux kernel already has some support for SMBDIRECT
  - ▶ The client (cifs.ko) has an "rdma" mount option
  - ▶ The server (ksmbd.ko) listens on any available rdma interface
  - ▶ Both had their own implementation (partly copied)
- ▶ The recent 7.1 merge windows included a new smbdirect.ko
  - ▶ After a long journey and ~420 small patches
  - ▶ cifs.ko and ksmbd.ko use a common smbdirect.ko
- ▶ There are IPPROTO\_SMBDIRECT patches for in kernel use
  - ▶ cifs.ko and ksmbd.ko use the generic socket layer where possible
  - ▶ <https://lore.kernel.org/linux-cifs/cover.1775571957.git.metze@samba.org/>
- ▶ IPPROTO\_SMBDIRECT has a working prototype for userspace
  - ▶ <https://git.samba.org/?p=metze/linux/wip.git;a=shortlog;h=refs/heads/master-ippROTO-smbdirect>
  - ▶ It basically works and but still needs a bit cleanup for upstream
- ▶ The Samba part is basically finished
  - ▶ [https://gitlab.com/samba-team/samba/-/merge\\_requests/4343](https://gitlab.com/samba-team/samba/-/merge_requests/4343)
  - ▶ It works and shows very good results (examples follow :-)

## How we test without kernel support

- ▶ We use `socket_wrapper` in Samba testing
  - ▶ It emulates TCP and UDP sockets using unix domain sockets
  - ▶ It uses `LD_PRELOAD` in order to inject itself into syscalls
  - ▶ `socket_wrapper_ipproto_quic_socket()` uses `SOCK_SEQPACKET`
  - ▶ `socket_wrapper_ipproto_smbdirect_socket()` uses `SOCK_STREAM`
- ▶ `quic_ko_wrapper` was created in order to emulate `quic.ko`
  - ▶ It works on top of `socket_wrapper_ipproto_quic_socket()`
  - ▶ `SOCK_SEQPACKET` on unix sockets provides the connection semantic
  - ▶ The QUIC crypto is implemented using `libngtcp2` too
- ▶ `smbdirect_ko_wrapper` was created in order to emulate `smbdirect.ko`
  - ▶ It works on top of `socket_wrapper_ipproto_smbdirect_socket()`
  - ▶ `process_vm_readv()/process_vm_writev()` for RDMA Read/Write
  - ▶ Thanks, Volker!
- ▶ They are good enough for regression testing in Samba
  - ▶ We run a few tests to explore the related code paths

## samba\_io\_uring abstraction (Part 1)

API glue to tevent:

```
void samba_io_uring_ev_register(void);

const struct samba_io_uring_features *samba_io_uring_system_features(void);

struct samba_io_uring *samba_io_uring_ev_context_get_ring(struct tevent_context *ev);

const struct samba_io_uring_features *samba_io_uring_get_features(
    const struct samba_io_uring *ring);

ev = tevent_context_init_byname(mem_ctx, "samba_io_uring_ev");
```

- ▶ samba\_io\_uring abstraction factored out of vfs\_io\_uring:
  - ▶ samba\_io\_uring\_ev\_hybrid tevent backend (glued on epoll backend)
  - ▶ It means every layer getting the tevent\_context can use io\_uring
  - ▶ No #ifdef's just checking if the required features are available

## samba\_io\_uring abstraction ( Part 2)

generic submission/completion api:

```
void samba_io_uring_completion_prepare(struct samba_io_uring_completion *completion,
    void (*completion_fn)(struct samba_io_uring_completion *completion,
        void *completion_private,
        const struct io_uring_cqe *cqe),
    void *completion_private);

void samba_io_uring_submission_prepare(struct samba_io_uring_submission *submission,
    void (*submission_fn)(struct samba_io_uring *ring,
        struct samba_io_uring_submission *submission,
        void *submission_private),
    void *submission_private,
    struct samba_io_uring_completion *completion);

struct io_uring_sqe *samba_io_uring_submission_sqe(struct samba_io_uring_submission *
    submission);

size_t samba_io_uring_queue_submissions(struct samba_io_uring *ring,
    struct samba_io_uring_submission *submission);

int samba_io_uring_submit_sqes(struct samba_io_uring *ring);
```

- ▶ Using it ...
  - ▶ convert vfs\_io\_uring
  - ▶ use it in smb2\_server.c
  - ▶ In future use it in other performance critical places too.

## performance boost with io\_uring (Part 1)

- ▶ `sendmsg()` and `recvmsg()` are performance killers
  - ▶ They block the main process burning cpu doing memcpy
  - ▶ Even on highspeed networks only a few GBytes/s are possible
- ▶ Refactoring of `smb2_server.c`
  - ▶ add optional `IORING_OP_SENDMSG`, `IORING_OP_RECVMSG` support
  - ▶ using `IOSQE_ASYNC` in order to memcpy in helper threads
  - ▶ improves multichannel between 30% and 100% depending on the setup
- ▶ With `IORING_OP_SENDMSG_ZC` only 1 one copy is used:
  - ▶ It is able to avoid copying to the socket
  - ▶ We get an extra completion once the buffers are not needed anymore
  - ▶ Only with real hardware, not on loopback in an upstream kernel
  - ▶ improves again between 30% and 60% depending on the setup
  - ▶ reduces the server cpu usage by 100% and more
  - ▶ currently tcp only, but smbdirect will follow

## performance boost with io\_uring (Part 2)

- ▶ Usage of IORING\_OP\_SPLICE socket ->pipe ->filesystem
  - ▶ IORING\_OP\_RECVMSG only for the header
  - ▶ IORING\_OP\_SPLICE from socket to a pipe (zero-copy)
  - ▶ IORING\_OP\_SPLICE from pipe to file (copy)
  - ▶ currently tcp only, but would work without problems
  - ▶ rdma uses RDMA READ offload instead
- ▶ Usage of IORING\_OP\_SPLICE filesystem ->pipe ->socket
  - ▶ IORING\_OP\_SPLICE from file to pipe (zero-copy from page cache)
  - ▶ IORING\_OP\_SENDMSG[\_ZC] only for the header
  - ▶ IORING\_OP\_SPLICE from pipe to socket (for tcp)
  - ▶ SMBDIRECT\_URING\_CMD\_BUFFER\_DO\_REMOTE\_SPLICE\_WRITE
  - ▶ (pipe to RDMA WRITE offload)
- ▶ There are structural problems with splice from a file
  - ▶ I had a discussion with the Linux developers about it:
  - ▶ The page content from the page cache may change unexpectedly
  - ▶ <https://lists.samba.org/archive/samba-technical/2023-February/thread.html#137945>
  - ▶ So this may not be the default, but maybe there will be tricks to handle it

# performance testsetup

- ▶ 2 Servers with a AMD Ryzen 9 9950X3D 16-Core Processor
  - ▶ 96 GB DRAM (2 x 48 GB DDR5 5600 MT/s)
  - ▶ PCIe 5.0
  - ▶ 2x SAMSUNG MZ7L3960 SATA-SSD (Windows2025/Ubuntu26.04)
  - ▶ Ubuntu 26.04 used a custom 7.0 kernel and custom samba
  - ▶ 2x Samsung NVMe SSD 9100 PRO (DATA: Windows/Linux)
  - ▶ Mellanox ConnectX-7 (RoCEv2) with 2x 200Gbit/s ports
  - ▶ The 200Gbit/s links are connected via direct cables
- ▶ Additional RDMA cards not always connected
  - ▶ For additional correctness testing we can add them
  - ▶ Mellanox ConnectX-7 also supports InfiniBand
  - ▶ Intel-E810-CQDA2 2x 100Gbit/s (iWarp and RoCEv2)
  - ▶ Chelsio T520-BT 2x 10Gbit/s (iWarp)

# smb2.bench.read io\_size=2MB tcp splice (Samba)

single client, 8 connections, qdepth=2, ~20 GBytes/s, ~700% client cpu, 2x200Gbit/s interfaces

```
root@rdmatest0410:~/samba.git# bin/smbtorture //10.123.201.5/torture -UUser1%a1B2c3D4 --option=clientsmbtransports=tcp --option=torture:timeLimit=600 --option=torture:rdma_threshold=1 --option=torture:io_size=0x200000 --option=torture:qdepth=2 --option=torture:nprocs=8 smb2.bench.read --option="libsmb:client_guid=6112f7d3-9528-4a2a-8861-0ca129aae6c6" --unclist=./unclist-20X.txt
smbtorture 4.25.0pre1-DEVELOPERBUILD
Using seed 1776364306
time: 2026-04-16 18:31:46.264739
test: read
time: 2026-04-16 18:31:46.265018
Opening 8 connections
Opened 8 connections with qdepth=2 => 16 loops
Running for 600 seconds
72.00 second: read[num/s=9,509;bytes/s=19,941,818,368;avslat=0.001681;minlat=0.001154;maxlat=0.010515]
```

```
top - 18:32:57 up 3:12, 2 users, load average: 7.54, 4.17, 2.38
Tasks: 554 total, 5 running, 549 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.2 us, 20.8 sy, 0.0 ni, 66.2 id, 0.0 wa, 0.0 hi, 12.8 si, 0.0 st
MiB Mem: 91818.2 total, 14673.6 free, 3027.8 used, 75060.3 buff/cache
MiB Swap: 0.0 total, 0.0 free, 0.0 used, 88790.4 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
14662	root	20	0	889336	810552	35280	R	710.7	0.9	8:20.40	smbtorture
85	root	20	0	0	0	0	S	27.3	0.0	0:20.96	ksoftirqd/11
73	root	20	0	0	0	0	R	27.0	0.0	0:19.33	ksoftirqd/9
194	root	20	0	0	0	0	S	24.3	0.0	0:59.17	ksoftirqd/29
109	root	20	0	0	0	0	R	24.0	0.0	0:31.37	ksoftirqd/15
49	root	20	0	0	0	0	R	21.0	0.0	0:28.65	ksoftirqd/5
134	root	20	0	0	0	0	S	18.7	0.0	0:13.63	ksoftirqd/19
152	root	20	0	0	0	0	S	18.7	0.0	0:12.84	ksoftirqd/22
31	root	20	0	0	0	0	S	18.3	0.0	0:13.39	ksoftirqd/2
72	root	rt	0	0	0	0	S	0.3	0.0	0:00.49	migration/9

```
[0] 0:dmesg 1:top- 2:top* 3:bash 4:bash 5:bash 6:bash
```

```
"rdmatest0410" 18:32 16-Apr-26
```

# smb2.bench.read io\_size=2MB tcp splice (Samba top)

single smbd + iou-wrk threads, 8 connections, qdepth=2, ~20 GBytes/s, ~70% server cpu, 2x200Gbit/s interfaces

```
top - 20:36:20 up 38 min, 2 users, load average: 0.61, 7.61, 12.44
Threads: 543 total, 3 running, 540 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.3 us, 1.3 sy, 0.0 ni, 96.9 id, 0.0 wa, 0.0 hi, 1.5 si, 0.0 st
MiB Mem : 91816.0 total, 87495.8 free, 1975.2 used, 3187.1 buff/cache
MiB Swap: 0.0 total, 0.0 free, 0.0 used, 89840.8 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
72482	root	20	0	107020	23600	19384	R	22.9	0.0	1:03.12	smbd[10.123.201
72542	user1	20	0	107020	23600	19384	S	9.0	0.0	0:13.11	iou-wrk-72482
72561	user1	20	0	107020	23600	19384	S	9.0	0.0	0:06.18	iou-wrk-72482
72577	user1	20	0	107020	23600	19384	S	8.6	0.0	0:03.42	iou-wrk-72482
72578	user1	20	0	107020	23600	19384	R	8.6	0.0	0:02.68	iou-wrk-72482
72583	user1	20	0	107020	23600	19384	S	4.7	0.0	0:01.68	iou-wrk-72482
71044	root	20	0	0	0	0	I	0.3	0.0	0:00.44	kworker/23:0-events
71895	root	20	0	0	0	0	I	0.3	0.0	0:00.03	kworker/u130:0-events_power_efficient
72441	root	20	0	0	0	0	I	0.3	0.0	0:00.15	kworker/1:0-events
72547	root	20	0	0	0	0	I	0.3	0.0	0:00.18	kworker/u128:1-mlx5_cmd_0000:01:00.1
72571	root	20	0	13656	5752	3560	R	0.3	0.0	0:00.14	top
1	root	20	0	25856	16396	11368	S	0.0	0.0	0:01.80	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.01	kthreadd
3	root	20	0	0	0	0	S	0.0	0.0	0:00.00	pool_workqueue_release
4	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-rcu_gp
5	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-sync_wq
6	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-kvfree_rcu_reclaim
7	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-slab_flushwq
8	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-netns
10	root	20	0	0	0	0	I	0.0	0.0	0:00.23	kworker/0:1-slab_flushwq
11	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/0:0H-kblockd
13	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-mm_percpu_wq
14	root	20	0	0	0	0	S	0.0	0.0	0:00.53	ksoftirqd/0
15	root	20	0	0	0	0	I	0.0	0.0	0:01.22	rcu_preempt
16	root	20	0	0	0	0	S	0.0	0.0	0:00.00	rcu_exp_par_gp_kthread_worker/1
17	root	20	0	0	0	0	S	0.0	0.0	0:00.00	rcu_exp_gp_kthread_worker
18	root	rt	0	0	0	0	S	0.0	0.0	0:00.16	migration/0
19	root	20	0	0	0	0	S	0.0	0.0	0:00.00	kprobe-optimizer
20	root	-51	0	0	0	0	S	0.0	0.0	0:00.00	idle_inject/0

```
[0] 0:bash- 1:bash 2:bash 3:bash 4:top* 5:bash
```

```
"rdmatest0510" 20:36 16-Apr-26
```

# smb2.bench.read io\_size=2MB tcp splice (Samba perf top)

single smbd + iou-wrk threads, 8 connections, qdepth=2, ~20 GBytes/s, no copy-<sub>{to,from}</sub>\_iter, 2x200Gbit/s interfaces

Samples: 15M of event 'cpu/cycles/P', 4000 Hz, Event count (approx.): 196943551866 lost: 0/0 drop: 0/0

Overhead	Shared Object	Symbol
9.79%	[kernel]	[k] amd_iommu_iotlb_sync_map
5.31%	[kernel]	[k] acpi_os_read_port
4.27%	[kernel]	[k] pt_iommu_amdvl_iova_to_phys
3.80%	[kernel]	[k] pt_iommu_amdvl_unmap_pages
3.64%	[kernel]	[k] pt_iommu_amdvl_map_pages
2.76%	[kernel]	[k] _raw_spin_lock_irqsave
2.12%	[kernel]	[k] _unmap_range
1.97%	[kernel]	[k] fq_codel_dequeue
1.90%	[kernel]	[k] iommu_pgsz
1.70%	[kernel]	[k] page_cache_pipe_buf_release
1.59%	[kernel]	[k] native_irq_return_iret
1.40%	[kernel]	[k] __tcp_transmit_skb
1.35%	[kernel]	[k] page_cache_pipe_buf_confirm
1.31%	[kernel]	[k] io_idle
1.21%	[kernel]	[k] skb_release_data
1.10%	[kernel]	[k] _iommu_dma_unmap
1.06%	[kernel]	[k] skb_splice_from_iter
1.01%	[kernel]	[k] mlx5e_select_queue
1.01%	[kernel]	[k] read_tsc
0.97%	[kernel]	[k] splice_folio_into_pipe
0.97%	[kernel]	[k] mlx5e_completion_event
0.92%	[kernel]	[k] filemap_get_read_batch
0.89%	[kernel]	[k] skb_shift
0.84%	[kernel]	[k] native_queued_spin_lock_slowpath
0.76%	[kernel]	[k] _raw_spin_unlock_irqrestore
0.75%	[kernel]	[k] iov_iter_extract_pages
0.73%	[kernel]	[k] iommu_map_nosync
0.70%	[kernel]	[k] net_rx_action
0.65%	[kernel]	[k] __rcu_read_lock
0.58%	[kernel]	[k] iommu_dma_free_iova
0.58%	[kernel]	[k] alloc_iova_fast
0.58%	[kernel]	[k] _iommu_unmap
0.50%	[kernel]	[k] make_range_ul

For a higher level overview, try: perf top --sort comm,dso

[0] 0:bash- 1:bash 2:bash 3:bash 4:perf\* 5:bash

"rdmatest0510" 20:34 16-Apr-20

# smb2.bench.read io\_size=2MB rdma offload (Windows)

single client, 2x8 connections, qdepth=2, ~46 GBytes/s, 2x200Gbit/s interfaces, RoCEv2

```
root@rdmatest04l0:~# cd samba.git/
root@rdmatest04l0:~/samba.git# bin/smbtorture //10.123.201.5/torture -User%1a182c304 --option=clientsmbtransports=smbdirect_ib --option=torture:timelimit=600 --op
tion=torture:rdma_threshold=1 --option=torture:io_size=0x200000 --option=torture:qdepth=2 --option=torture:nprocs=8 smb2.bench.read --option="libsmb:client_guid=61
12f7d3-9528-4a2a-8861-0ca129aae6c6" --unclist=./unclist-20X.txt
smbtorture 4.25.0pre1-DEVELOPERBUILD
Using seed 1776260218
time: 2026-04-15 15:50:18.312331
test: read
time: 2026-04-15 15:50:18.312480
Opening 8 connections
Opened 8 connections with qdepth=2 => 16 loops
Running for 600 seconds
600.03 second: read[num/s=11,039;bytes/s=23,150,460,928;avslat=0.001449;minlat=0.000768;maxlat=0.062323]
600.03 seconds
time: 2026-04-15 16:00:20.513577
success: read
root@rdmatest04l0:~/samba.git#
```

```
root@rdmatest04l0:~# cd samba.git/
root@rdmatest04l0:~/samba.git# bin/smbtorture //10.123.201.5/torture -User%1a182c304 --option=clientsmbtransports=smbdirect_ib --option=torture:timelimit=600 --op
tion=torture:rdma_threshold=1 --option=torture:io_size=0x200000 --option=torture:qdepth=2 --option=torture:nprocs=8 smb2.bench.read --option="libsmb:client_guid=61
12f7d3-9528-4a2a-8861-0ca129aae6c6" --unclist=./unclist-20X.txt
smbtorture 4.25.0pre1-DEVELOPERBUILD
Using seed 1776268234
time: 2026-04-15 15:50:34.653039
test: read
time: 2026-04-15 15:50:34.653180
Opening 8 connections
Opened 8 connections with qdepth=2 => 16 loops
Running for 600 seconds
600.02 second: read[num/s=11,018;bytes/s=23,106,420,736;avslat=0.001452;minlat=0.000857;maxlat=0.062320]
600.02 seconds
time: 2026-04-15 16:00:36.257598
success: read
root@rdmatest04l0:~/samba.git#
```

```
[0] 0:bash- 1:bash 2:bash* 3:bash 4:bash 5:bash
```

```
"rdmatest04l0" 19:21 15-Apr-26
```

# smb2.bench.read io\_size=2MB rdma splice (Samba)

single client, 2x8 connections, qdepth=2, ~46 GBytes/s, 2x200Gbit/s interfaces, RoCEv2

```
root@rdmatest04l0:~/samba.git# bin/smbtorture //10.123.201.5/torture -User1%a1B2c3D4 --option=clientsmbtransports=smbdirect_ib --option=torture:timelimit=600 --op
tion=torture:rdma_threshold=1 --option=torture:io_size=0x200000 --option=torture:qdepth=2 --option=torture:nprocs=8 smb2.bench.read --option="libsmb:client_guid=61
12f7d3-9528-4a2a-8861-0ca129aae6c6"
smbtorture 4.25.0pre1-DEVELOPERBUILD
Using seed 1776363051
time: 2026-04-16 18:10:51.506457
test: read
time: 2026-04-16 18:10:51.506591
Opening 8 connections
Opened 8 connections with qdepth=2 => 16 loops
Running for 600 seconds
600.02 second: read[num/s=10,986;bytes/s=23,039,311,872;avslat=0.001451;minlat=0.000474;maxlat=0.012254]
600.02 seconds
time: 2026-04-16 18:22:04.189317
success: read
root@rdmatest04l0:~/samba.git# [

^C.00 second: read[num/s=10,994;bytes/s=23,056,089,088;avslat=0.001455;minlat=0.000457;maxlat=0.002464]
root@rdmatest04l0:~/samba.git# bin/smbtorture //10.123.202.5/torture -User1%a1B2c3D4 --option=clientsmbtransports=smbdirect_ib --option=torture:timelimit=600 --op
tion=torture:rdma_threshold=1 --option=torture:io_size=0x200000 --option=torture:qdepth=2 --option=torture:nprocs=8 smb2.bench.read --option="libsmb:client_guid=61
12f7d3-9528-4a2a-8861-0ca129aae6c6"
smbtorture 4.25.0pre1-DEVELOPERBUILD
Using seed 1776363049
time: 2026-04-16 18:10:49.486423
test: read
time: 2026-04-16 18:10:49.486638
Opening 8 connections
Opened 8 connections with qdepth=2 => 16 loops
Running for 600 seconds
600.02 second: read[num/s=11,022;bytes/s=23,114,809,344;avslat=0.001451;minlat=0.000459;maxlat=0.019519]
600.02 seconds
time: 2026-04-16 18:22:04.136384
success: read
root@rdmatest04l0:~/samba.git#
[0] 0:dmesg 1:bash- 2:bash* 3:bash 4:bash 5:bash 6:bash "rdmatest04l0" 18:23 16-Apr-26
```

# smb2.bench.read io\_size=2MB rdma splice (client top)

single client, 2x8 connections, qdepth=2, ~46 GBytes/s, 2x ~63% client cpu, 2x200Gbit/s interfaces, RoCEv2

```
top - 18:29:27 up 3:09, 2 users, load average: 1.66, 1.66, 1.38
Tasks: 534 total, 3 running, 531 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.5 us, 3.2 sy, 0.0 ni, 96.2 id, 0.0 wa, 0.0 hi, 0.1 si, 0.0 st
MiB Mem : 91818.2 total, 14003.6 free, 3705.9 used, 75050.4 buff/cache
MiB Swap: 0.0 total, 0.0 free, 0.0 used, 88112.4 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
14550	root	20	0	781672	703168	35544	R	63.8	0.7	0:32.12	smbtorture
14534	root	20	0	897096	818192	35120	R	63.5	0.9	0:37.58	smbtorture
14217	root	0	-20	0	0	0	I	2.0	0.0	0:08.66	kworker/u133:3-smbdirect-refill
13986	root	0	-20	0	0	0	I	1.7	0.0	0:12.46	kworker/u133:2-smbdirect-immediate
14445	root	0	-20	0	0	0	I	1.7	0.0	0:01.88	kworker/u133:0-smbdirect-immediate
14088	root	0	-20	0	0	0	I	1.3	0.0	0:10.50	kworker/u133:5-smbdirect-immediate
14484	root	20	0	13692	5892	3688	R	0.3	0.0	0:00.24	top
1	root	20	0	25736	16632	11384	S	0.0	0.0	0:01.25	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.01	kthreadd
3	root	20	0	0	0	0	S	0.0	0.0	0:00.00	pool_workqueue_release
4	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-rcu_gp
5	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-sync_wq
6	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-kvfree_rcu_reclaim
7	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-slub_flushwq
8	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-nets
11	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/0:0H-kblockd
13	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-mm_percpu_wq
14	root	20	0	0	0	0	S	0.0	0.0	0:00.36	ksoftirqd/0
15	root	20	0	0	0	0	I	0.0	0.0	0:01.01	rcu_preempt
16	root	20	0	0	0	0	S	0.0	0.0	0:00.00	rcu_exp_par_gp_kthread_worker/1
17	root	20	0	0	0	0	S	0.0	0.0	0:00.00	rcu_exp_gp_kthread_worker
18	root	rt	0	0	0	0	S	0.0	0.0	0:00.33	migration/0
19	root	20	0	0	0	0	S	0.0	0.0	0:00.00	kprobe-optimizer
20	root	-51	0	0	0	0	S	0.0	0.0	0:00.00	idle_inject/0
21	root	20	0	0	0	0	S	0.0	0.0	0:00.00	cpuhp/0
22	root	20	0	0	0	0	S	0.0	0.0	0:00.00	cpuhp/1
23	root	-51	0	0	0	0	S	0.0	0.0	0:00.00	idle_inject/1
24	root	rt	0	0	0	0	S	0.0	0.0	0:00.24	migration/1
25	root	20	0	0	0	0	S	0.0	0.0	0:00.37	ksoftirqd/1

```
[0] 0:dmesg 1:top* 2:bin/smbtorture- 3:bash 4:bash 5:bash 6:bash
```

```
"rdmatest0410" 18:29 16-Apr-26
```

# smb2.bench.read io\_size=2MB rdma splice (server top)

single smb2 + iou-wrk threads, 2x8 connections, qdepth=2, ~46 GBytes/s, ~250% server cpu, 2x200Gbit/s interfaces, RoCEv2

```
top - 20:25:12 up 27 min, 2 users, load average: 17.54, 15.85, 14.20
```

```
Tasks: 520 total, 2 running, 518 sleeping, 0 stopped, 0 zombie
```

```
%Cpu(s): 0.9 us, 5.6 sy, 0.0 ni, 92.9 id, 0.0 wa, 0.0 hi, 0.6 si, 0.0 st
```

```
MiB Mem : 91816.0 total, 87372.4 free, 2070.5 used, 3214.8 buff/cache
```

```
MiB Swap: 0.0 total, 0.0 free, 0.0 used, 89745.5 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
72048	user1	20	0	107288	24024	19288	R	251.5	0.0	2:00.64	smbd[10.123.201
70852	root	0	-20	0	0	0	I	1.0	0.0	0:04.25	kworker/u132:1-smbdirect-refill
257	root	0	-20	0	0	0	I	0.7	0.0	0:02.98	kworker/u133:0-smbdirect-refill
70855	root	0	-20	0	0	0	I	0.7	0.0	0:02.43	kworker/u133:1-smbdirect-refill
70908	root	0	-20	0	0	0	I	0.7	0.0	0:02.45	kworker/u133:4-smbdirect-refill
72089	root	0	-20	0	0	0	I	0.7	0.0	0:00.24	kworker/u132:2-smbdirect-refill
67	root	20	0	0	0	0	S	0.3	0.0	0:00.52	ksoftirqd/8
73	root	20	0	0	0	0	S	0.3	0.0	0:01.66	ksoftirqd/9
79	root	20	0	0	0	0	S	0.3	0.0	0:02.23	ksoftirqd/10
121	root	rt	0	0	0	0	S	0.3	0.0	0:00.08	migration/17
146	root	20	0	0	0	0	S	0.3	0.0	0:00.07	ksoftirqd/21
188	root	20	0	0	0	0	S	0.3	0.0	0:00.55	ksoftirqd/28
205	root	rt	0	0	0	0	S	0.3	0.0	0:00.10	migration/31
256	root	0	-20	0	0	0	I	0.3	0.0	0:03.59	kworker/u132:0-smbdirect-refill
71246	root	20	0	0	0	0	I	0.3	0.0	0:00.04	kworker/26:1-mm_percpu_wq
72133	root	20	0	13592	5792	3608	R	0.3	0.0	0:00.04	top
1	root	20	0	25856	16396	11368	S	0.0	0.0	0:01.78	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.00	kthreadd
3	root	20	0	0	0	0	S	0.0	0.0	0:00.00	pool_workqueue_release
4	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-rcu_gp
5	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-sync_wq
6	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-kvfree_rcu_reclaim
7	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-slub_flushwq
8	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-nets
10	root	20	0	0	0	0	I	0.0	0.0	0:00.23	kworker/0:1-events
11	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/0:0H-kblockd
13	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-mm_percpu_wq
14	root	20	0	0	0	0	S	0.0	0.0	0:00.39	ksoftirqd/0
15	root	20	0	0	0	0	I	0.0	0.0	0:00.77	rcu_preempt

```
[0] 0:bash- 1:bash 2:bash 3:bash 4:top* 5:bash
```

```
"rdmatest0510" 20:25 16-Apr-26
```

# smb2.bench.read io\_size=2MB rdma splice (server topH)

single smb2 + iou-wrk threads, 2x8 connections, qdepth=2, ~46 GBytes/s, ~82% server cpu smb2, 2x200Gbit/s interfaces, RoCEv2

```
top - 20:26:55 up 28 min, 2 users, load average: 25.12, 19.20, 15.59
Threads: 580 total, 3 running, 577 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.8 us, 5.8 sy, 0.0 ni, 92.6 id, 0.0 wa, 0.0 hi, 0.7 si, 0.0 st
MiB Mem : 91816.0 total, 87370.6 free, 2072.3 used, 3214.9 buff/cache
MiB Swap: 0.0 total, 0.0 free, 0.0 used, 89743.7 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
72048	user1	20	0	107288	24028	19288	R	82.4	0.0	2:01.79	smbd[10.123.201
72083	user1	20	0	107288	24028	19288	D	5.3	0.0	0:06.48	iou-wrk-72048
72072	root	20	0	107288	24028	19288	D	5.0	0.0	0:06.86	iou-wrk-72048
72113	root	20	0	107288	24028	19288	D	5.0	0.0	0:06.71	iou-wrk-72048
72134	user1	20	0	107288	24028	19288	S	5.0	0.0	0:05.79	iou-wrk-72048
72137	user1	20	0	107288	24028	19288	D	5.0	0.0	0:05.16	iou-wrk-72048
72140	user1	20	0	107288	24028	19288	D	5.0	0.0	0:04.84	iou-wrk-72048
72151	user1	20	0	107288	24028	19288	D	5.0	0.0	0:02.14	iou-wrk-72048
72050	root	20	0	107288	24028	19288	D	4.7	0.0	0:06.46	iou-wrk-72048
72077	user1	20	0	107288	24028	19288	S	4.7	0.0	0:06.92	iou-wrk-72048
72079	user1	20	0	107288	24028	19288	D	4.7	0.0	0:06.76	iou-wrk-72048
72080	user1	20	0	107288	24028	19288	D	4.7	0.0	0:06.86	iou-wrk-72048
72088	user1	20	0	107288	24028	19288	D	4.7	0.0	0:06.70	iou-wrk-72048
72117	user1	20	0	107288	24028	19288	R	4.7	0.0	0:06.21	iou-wrk-72048
72119	user1	20	0	107288	24028	19288	S	4.7	0.0	0:06.73	iou-wrk-72048
72120	user1	20	0	107288	24028	19288	D	4.7	0.0	0:06.51	iou-wrk-72048
72123	user1	20	0	107288	24028	19288	D	4.7	0.0	0:06.43	iou-wrk-72048
72124	user1	20	0	107288	24028	19288	D	4.7	0.0	0:06.57	iou-wrk-72048
72135	user1	20	0	107288	24028	19288	D	4.7	0.0	0:05.78	iou-wrk-72048
72141	user1	20	0	107288	24028	19288	D	4.7	0.0	0:04.51	iou-wrk-72048
72148	user1	20	0	107288	24028	19288	D	4.7	0.0	0:02.82	iou-wrk-72048
72150	user1	20	0	107288	24028	19288	D	4.7	0.0	0:02.20	iou-wrk-72048
72154	root	20	0	107288	24028	19288	S	4.7	0.0	0:01.74	iou-wrk-72048
72161	user1	20	0	107288	24028	19288	D	4.7	0.0	0:00.31	iou-wrk-72048
72163	user1	20	0	107288	24028	19288	D	4.7	0.0	0:00.30	iou-wrk-72048
72049	root	20	0	107288	24028	19288	S	4.3	0.0	0:06.77	iou-wrk-72048
72070	user1	20	0	107288	24028	19288	D	4.3	0.0	0:05.86	iou-wrk-72048
72071	user1	20	0	107288	24028	19288	D	4.3	0.0	0:06.40	iou-wrk-72048
72145	user1	20	0	107288	24028	19288	D	4.3	0.0	0:03.07	iou-wrk-72048

```
[0] 0:bash-1:bash 2:bash 3:bash 4:top* 5:bash
```

```
"rdmatest0510" 20:26 16-Apr-20
```

# smb2.bench.read io\_size=2MB rdma splice (perf top)

single smb2 + iou-wrk threads, 2x8 connections, qdepth=2, ~46 GBytes/s, no copy\_{to,from}\_iter, 2x200Gbit/s interfaces, RoCEv2

```
Samples: 4M of event 'cpu/cycles/P', 4000 Hz, Event count (approx.): 285561376650 lost: 0/0 drop: 0/0
```

Overhead	Shared Object	Symbol
6.32%	[kernel]	[k] page_cache_pipe_buf_release
4.46%	[kernel]	[k] acpi_os_read_port
4.06%	[kernel]	[k] kernel_init_pages
3.78%	[kernel]	[k] pt_iommu_andv1_map_pages
3.76%	[kernel]	[k] smbdirect_splice_to_bvecs_actor
3.49%	[kernel]	[k] page_cache_pipe_buf_confirm
1.89%	[kernel]	[k] filemap_get_read_batch
1.60%	[kernel]	[k] splice_folio_into_pipe
1.51%	[kernel]	[k] generic_pipe_buf_get
1.45%	[kernel]	[k] ep_poll_callback
1.35%	[kernel]	[k] select_task_rq_fair
1.31%	[kernel]	[k] io_idle
1.18%	[kernel]	[k] smbdirect_sk_logging_needed
1.14%	[kernel]	[k] _switch_to
1.05%	[kernel]	[k] mlx5_ib_cq_comp
1.04%	[kernel]	[k] iommu_pgsize
1.02%	[kernel]	[k] iommu_map_nosync
0.96%	[kernel]	[k] read_tsc
0.93%	[kernel]	[k] free_pipe_info
0.88%	[kernel]	[k] _raw_spin_lock_irqsave
0.86%	[kernel]	[k] enqueue_task_fair
0.78%	[kernel]	[k] psi_group_change
0.73%	[kernel]	[k] dma_iova_link
0.71%	[kernel]	[k] mlx5_ib_poll_cq
0.66%	[kernel]	[k] menu_select
0.64%	[kernel]	[k] _splice_from_pipe
0.60%	[kernel]	[k] native_irq_return_iret
0.55%	[kernel]	[k] tctx_task_work
0.54%	[kernel]	[k] _raw_spin_lock
0.53%	[kernel]	[k] _update_load_avg_cfs_rq
0.52%	[kernel]	[k] smbdirect_sk_release_cb
0.50%	[kernel]	[k] _map_single_page2.isra.0
0.48%	[kernel]	[k] switch_mm_irqs_off

```
For a higher level overview, try: perf top --sort comm,dso
```

```
[0] 0:bash- 1:bash 2:bash 3:bash 4:perf* 5:bash
```

```
"rdmatest0510" 20:39 16-Apr-26
```

# smb2.bench.read io\_size=8MB rdma splice (Samba)

single client, 2x2 connections, qdepth=4, ~46 GBytes/s, 2x200Gbit/s interfaces, RoCEv2

```
time: 2026-04-17 13:18:05.374841
Opening 2 connections
Opened 2 connections with qdepth=4 => 8 loops
Running for 600 seconds
^C00 second: read[num/s=2,745;bytes/s=23,026,728,960;avslat=0.002915;minlat=0.002184;maxlat=0.003894]
root@rdmatest04l0:~/samba.git#
root@rdmatest04l0:~/samba.git# bin/smbtorture //10.123.201.5/torture -User%1a1B2c3D4 --option=clientsmbtransports=smbdirect_ib --option=torture:timelimit=600 --op
tion=torture:rdma_threshold=1 --option=torture:io_size=0x800000 --option=torture:qdepth=4 --option=torture:nprocs=2 smb2.bench.read --option="libsmb:client_guid=61
12f7d3-9528-4a2a-8861-0ca129aae6c6"
smbtorture 4.25.0pre1-DEVELOPERBUILD
Using seed 1776431897
time: 2026-04-17 13:18:17.707679
test: read
time: 2026-04-17 13:18:17.707813
Opening 2 connections
Opened 2 connections with qdepth=4 => 8 loops
Running for 600 seconds
541.10 second: read[num/s=2,759;bytes/s=23,144,169,472;avslat=0.002900;minlat=0.002074;maxlat=0.003845]

Opening 2 connections
Opened 2 connections with qdepth=4 => 8 loops
Running for 600 seconds
^C00 second: read[num/s=2,689;bytes/s=22,556,966,912;avslat=0.002950;minlat=0.002157;maxlat=0.018467]
root@rdmatest04l0:~/samba.git#
root@rdmatest04l0:~/samba.git# bin/smbtorture //10.123.202.5/torture -User%1a1B2c3D4 --option=clientsmbtransports=smbdirect_ib --option=torture:timelimit=600 --op
tion=torture:rdma_threshold=1 --option=torture:io_size=0x800000 --option=torture:qdepth=4 --option=torture:nprocs=2 smb2.bench.read --option="libsmb:client_guid=61
12f7d3-9528-4a2a-8861-0ca129aae6c6"
smbtorture 4.25.0pre1-DEVELOPERBUILD
Using seed 1776431895
time: 2026-04-17 13:18:15.601119
test: read
time: 2026-04-17 13:18:15.601398
Opening 2 connections
Opened 2 connections with qdepth=4 => 8 loops
Running for 600 seconds
543.11 second: read[num/s=2,758;bytes/s=23,135,780,864;avslat=0.002900;minlat=0.002043;maxlat=0.003898]

[0] 0:dmesg 1:bash- 2:bin/smbtorture* 3:bash 4:bash 5:bash 6:bash
```

"rdmatest04l0" 13:27 17-Apr-26

# smb2.bench.read io\_size=8MB rdma splice (client top)

single client, 2x2 connections, qdepth=4, ~46 GBytes/s, 2x ~38% client cpu, 2x200Gbit/s interfaces, RoCEv2

```
top - 13:21:36 up 22:01, 2 users, load average: 1.27, 1.17, 1.05
Tasks: 528 total, 2 running, 526 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.1 us, 2.0 sy, 0.0 ni, 97.8 id, 0.0 wa, 0.0 hi, 0.1 si, 0.0 st
MiB Mem : 91818.2 total, 13714.0 free, 3628.0 used, 75419.6 buff/cache
MiB Swap: 0.0 total, 0.0 free, 0.0 used, 88190.2 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
18264	root	20	0	791296	712600	35320	D	38.2	0.8	1:14.01	smbtorture
18251	root	20	0	797780	718932	35200	R	35.9	0.8	1:10.86	smbtorture
17348	root	0	-20	0	0	0	I	0.7	0.0	0:07.14	kworker/u133:0-smbdirect-refill
18270	root	0	-20	0	0	0	I	0.7	0.0	0:00.96	kworker/u133:1-smbdirect-refill
15141	root	0	-20	0	0	0	I	0.3	0.0	0:11.97	kworker/u133:6-smbdirect-refill
18287	root	20	0	13592	5768	3584	R	0.3	0.0	0:00.07	top
1	root	20	0	25736	16632	11384	S	0.0	0.0	0:02.34	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.04	kthreadd
3	root	20	0	0	0	0	S	0.0	0.0	0:00.00	pool_workqueue_release
4	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-rcu_gp
5	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-sync_wq
6	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-kvfree_rcu_reclaim
7	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-slub_flushwq
8	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-nets
11	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/0:0H-kblockd
13	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/R-mm_percpu_wq
14	root	20	0	0	0	0	S	0.0	0.0	0:14.33	ksoftirqd/0
15	root	20	0	0	0	0	I	0.0	0.0	0:02.27	rcu_preempt
16	root	20	0	0	0	0	S	0.0	0.0	0:00.00	rcu_exp_par_gp_kthread_worker/1
17	root	20	0	0	0	0	S	0.0	0.0	0:00.00	rcu_exp_gp_kthread_worker
18	root	rt	0	0	0	0	S	0.0	0.0	0:00.72	migration/0
19	root	20	0	0	0	0	S	0.0	0.0	0:00.00	kprobe-optimizer
20	root	-51	0	0	0	0	S	0.0	0.0	0:00.00	idle_inject/0
21	root	20	0	0	0	0	S	0.0	0.0	0:00.00	cpuhp/0
22	root	20	0	0	0	0	S	0.0	0.0	0:00.00	cpuhp/1
23	root	-51	0	0	0	0	S	0.0	0.0	0:00.00	idle_inject/1
24	root	rt	0	0	0	0	S	0.0	0.0	0:00.55	migration/1
25	root	20	0	0	0	0	S	0.0	0.0	0:11.46	ksoftirqd/1
27	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/1:0H-kblockd

```
[0] 0:dmesg 1:top* 2:bin/smbtorture- 3:bash 4:bash 5:bash 6:bash
```

```
"rdmat0410" 13:21 17-Apr-26
```

# smb2.bench.read io\_size=8MB rdma splice (server top)

single smb2 + iou-wrk threads, 2x2 connections, qdepth=4, ~46 GBytes/s, ~110% server cpu, 2x200Gbit/s interfaces, RoCEv2

```
top - 15:16:07 up 19:18, 2 users, load average: 13.32, 10.53, 8.69
Tasks: 521 total, 1 running, 520 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.2 us, 2.6 sy, 0.0 ni, 97.1 id, 0.0 wa, 0.0 hi, 0.1 si, 0.0 st
MiB Mem : 91816.0 total, 86812.0 free, 2047.5 used, 3810.2 buff/cache
MiB Swap: 0.0 total, 0.0 free, 0.0 used, 89768.5 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
79767	user1	20	0	107420	23536	19392	S	110.6	0.0	1:59.46	smbd: client [10.123.201.4]
85	root	20	0	0	0	0	S	0.3	0.0	0:01.38	[ksoftirqd/11]
70852	root	0	-20	0	0	0	I	0.3	0.0	0:11.48	[kworker/u132:1-smbdirect-immediate]
72649	root	0	-20	0	0	0	I	0.3	0.0	0:06.81	[kworker/u132:3-smbdirect-refill]
77630	root	0	-20	0	0	0	I	0.3	0.0	0:03.80	[kworker/u133:2-smbdirect-refill]
77687	root	0	-20	0	0	0	I	0.3	0.0	0:03.46	[kworker/u133:3-smbdirect-immediate]
79804	root	20	0	0	0	0	I	0.3	0.0	0:00.01	[kworker/9:2-events]
1	root	20	0	25856	16520	11432	S	0.0	0.0	0:03.06	/sbin/init
2	root	20	0	0	0	0	S	0.0	0.0	0:00.04	[kthreadd]
3	root	20	0	0	0	0	S	0.0	0.0	0:00.00	[pool_workqueue_release]
4	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	[kworker/R-rcu_gp]
5	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	[kworker/R-sync_wq]
6	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	[kworker/R-kvfree_rcu_reclaim]
7	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	[kworker/R-slub_flushwq]
8	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	[kworker/R-netns]
11	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	[kworker/0:0H-kblockd]
13	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	[kworker/R-mm_percpu_wq]
14	root	20	0	0	0	0	S	0.0	0.0	0:01.19	[ksoftirqd/0]
15	root	20	0	0	0	0	I	0.0	0.0	0:04.76	[rcu_preempt]
16	root	20	0	0	0	0	S	0.0	0.0	0:00.00	[rcu_exp_par_gp_kthread worker/1]
17	root	20	0	0	0	0	S	0.0	0.0	0:00.00	[rcu_exp_gp_kthread worker]
18	root	rt	0	0	0	0	S	0.0	0.0	0:01.06	[migration/0]
19	root	20	0	0	0	0	S	0.0	0.0	0:00.00	[kprobe-optimizer]
20	root	-51	0	0	0	0	S	0.0	0.0	0:00.00	[idle_inject/0]
21	root	20	0	0	0	0	S	0.0	0.0	0:00.00	[cpuhp/0]
22	root	20	0	0	0	0	S	0.0	0.0	0:00.00	[cpuhp/1]
23	root	-51	0	0	0	0	S	0.0	0.0	0:00.00	[idle_inject/1]
24	root	rt	0	0	0	0	S	0.0	0.0	0:01.07	[migration/1]
25	root	20	0	0	0	0	S	0.0	0.0	0:00.62	[ksoftirqd/1]

```
[0] 0:bash- 1:bash 2:bash 3:bash 4:top* 5:bash
```

```
"rdmatetest0510" 15:16 17-Apr-26
```

# smb2.bench.read io\_size=8MB rdma splice (server topH)

single smb2 + iou-wrk threads, 2x2 connections, qdepth=4, ~46 GBytes/s, ~21% server cpu smb2, 2x200Gbit/s interfaces, RoCEv2

```
top - 15:14:58 up 19:16, 2 users, load average: 9.95, 9.58, 8.28
Threads: 580 total, 4 running, 576 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.2 us, 2.7 sy, 0.0 ni, 97.0 id, 0.0 wa, 0.0 hi, 0.1 si, 0.0 st
MiB Mem : 91816.0 total, 86801.8 free, 2057.7 used, 3810.2 buff/cache
MiB Swap: 0.0 total, 0.0 free, 0.0 used, 89758.3 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
79767	user1	20	0	107420	23532	19392	S	21.9	0.0	0:08.31	smbd[10.123.201
79775	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.72	iou-wrk-79767
79776	user1	20	0	107420	23532	19392	S	4.3	0.0	0:01.72	iou-wrk-79767
79779	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.71	iou-wrk-79767
79780	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.73	iou-wrk-79767
79785	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.64	iou-wrk-79767
79793	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.62	iou-wrk-79767
79794	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.62	iou-wrk-79767
79795	user1	20	0	107420	23532	19392	R	4.3	0.0	0:01.63	iou-wrk-79767
79796	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.60	iou-wrk-79767
79797	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.61	iou-wrk-79767
79798	user1	20	0	107420	23532	19392	D	4.3	0.0	0:01.62	iou-wrk-79767
79777	user1	20	0	107420	23532	19392	D	4.0	0.0	0:01.72	iou-wrk-79767
79778	user1	20	0	107420	23532	19392	D	4.0	0.0	0:01.73	iou-wrk-79767
79781	user1	20	0	107420	23532	19392	D	4.0	0.0	0:01.33	iou-wrk-79767
79783	user1	20	0	107420	23532	19392	S	4.0	0.0	0:01.71	iou-wrk-79767
79784	user1	20	0	107420	23532	19392	D	4.0	0.0	0:01.70	iou-wrk-79767
79800	user1	20	0	107420	23532	19392	D	4.0	0.0	0:01.62	iou-wrk-79767
79768	root	20	0	107420	23532	19392	D	3.7	0.0	0:01.38	iou-wrk-79767
79782	user1	20	0	107420	23532	19392	R	3.7	0.0	0:01.35	iou-wrk-79767
79769	root	20	0	107420	23532	19392	S	3.3	0.0	0:01.31	iou-wrk-79767
79801	user1	20	0	107420	23532	19392	S	3.3	0.0	0:01.20	iou-wrk-79767
79799	user1	20	0	107420	23532	19392	D	3.0	0.0	0:01.23	iou-wrk-79767
72649	root	0	-20	0	0	0	I	0.3	0.0	0:06.68	kworker/u132:3-smbdirect-refill
77687	root	0	-20	0	0	0	I	0.3	0.0	0:03.33	kworker/u133:3-smbdirect-refill
79024	root	20	0	0	0	0	I	0.3	0.0	0:00.04	kworker/11:0-events
1	root	20	0	25856	16520	11432	S	0.0	0.0	0:03.06	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.04	kthreadd
3	root	20	0	0	0	0	S	0.0	0.0	0:00.00	pool_workqueue_release

```
[0] 0:bash- 1:bash 2:bash 3:bash 4:top* 5:bash
```

```
"rdmatetest0510" 15:14 17-Apr-26
```

# smb2.bench.read io\_size=8MB rdma splice (perf top)

single smb2 + iou-wrk threads, 2x2 connections, qdepth=4, ~46 GBytes/s, no copy\_{to,from}\_iter, 2x200Gbit/s interfaces, RoCEv2

Samples: 2M of event 'cpu/cycles/P', 4000 Hz, Event count (approx.): 137334428876 lost: 0/0 drop: 0/0

Overhead	Shared Object	Symbol
13.02%	[kernel]	[k] page_cache_pipe_buf_release
8.04%	[kernel]	[k] pt_iommu_andv1_map_pages
6.89%	[kernel]	[k] page_cache_pipe_buf_confirm
5.35%	[kernel]	[k] acpi_os_read_port
4.93%	[kernel]	[k] smbdirect_splice_to_bvecs_actor
3.20%	[kernel]	[k] splice_folio_into_pipe
3.01%	[kernel]	[k] generic_pipe_buf_get
2.17%	[kernel]	[k] iommu_map_nosync
1.80%	[kernel]	[k] iommu_pgsz
1.49%	[kernel]	[k] dma_iova_link
1.44%	[kernel]	[k] filemap_get_read_batch
1.41%	[kernel]	[k] kernel_init_pages
1.24%	[kernel]	[k] __splice_from_pipe
1.10%	[kernel]	[k] io_idle
0.91%	[kernel]	[k] rdma_rv_ctx_init_bvec
0.88%	[kernel]	[k] select_task_rq_fair
0.78%	[kernel]	[k] __map_single_page1.isra.0
0.76%	[kernel]	[k] __map_single_page3.isra.0
0.76%	[kernel]	[k] __map_single_page2.isra.0
0.75%	[kernel]	[k] __map_single_page4.isra.0
0.75%	[kernel]	[k] read_tsc
0.71%	[kernel]	[k] __switch_to
0.71%	[kernel]	[k] make_range_u1
0.66%	[kernel]	[k] __unmap_range
0.64%	[kernel]	[k] enqueue_task_fair
0.63%	[kernel]	[k] __map_single_page0.isra.0
0.61%	[kernel]	[k] ep_poll_callback
0.59%	[kernel]	[k] psi_group_change
0.57%	[kernel]	[k] free_pipe_info
0.54%	[kernel]	[k] smbdirect_sk_logging_needed
0.52%	[kernel]	[k] __update_load_avg_cfs_rq
0.52%	[kernel]	[k] switch_mm_irqs_off
0.49%	[kernel]	[k] menu_select

For a higher level overview, try: perf top --sort comm,dso

[0] 0:bash- 1:bash 2:bash 3:bash 4:perf\* 5:bash

"rdmat0510" 15:20 17-Apr-26

# Samba is back in the game

- ▶ This means Samba is able to handle highspeed networks
  - ▶ There's no longer the need to use proprietary smb stacks on Linux!
- ▶ Most if not all should be ready for the next Samba version
  - ▶ This will be released in September together with
  - ▶ Support for directory leases
  - ▶ Support for persistent handles

# Questions? Feedback!

- ▶ Stefan Metzmaker, [metze@samba.org](mailto:metze@samba.org)
- ▶ <https://www.sernet.com>
- ▶ <https://samba.plus>

Slides: <https://samba.org/~metze/presentations/2026/SambaXP/>