

ksmbd status update

Namjae Jeon
LGE / SAMBA team

Who am I

- **Maintainer of exfat filesystem**
- **Creator and Maintainer of ksmbd kernel server maintainer**
- **Member of Samba team.**
- **Reviewer of NTFS filesystem**

Topic

- **ksmbd overview**
- **Status update between 5.15 ~ 6.8 kernel**
- **User reports**
- **Next works**

ksmbd overview

- **SMB3 kernel server**
- **Project target**
 - **SMB 2.1 ~ SMB 3.1.1**
 - **Optimize performance for Linux workload**
 - **GPLv2 SMB Server**
 - **Better oplock/lease handling**
 - **Easy to develop new features(e.g. RDMA aka smb-direct)**
- **Very optimal for embedded device**
 - **Small memory footprint and binary size**
 - **Low CPU usage and better performance**
 - **Easy cross compile without special handling**

ksmbd overview – In use

- **Embedded opensource projects like openWRT, dd-wrt**
- **Commercial product like Network camera, Smart white board, etc**
- **Linux distribution like Ubuntu, Debian, openSUSE.**

Status update – security issues

- **ZDI(Zero Day Initiative), Huawei, the others report many vulnerabilities of ksmbd kernel server.**
- **Attempts various efforts to find the vulnerabilities**
 - **Add ksmbd to syzbot**
 - **Test using own fuzzing tool.**
- **Most issues have been clear very quickly within one or two weeks.**
- **Current reports have decreased significantly.**

Status update – move cifs and ksmbd to /fs/smb/

- Linus suggest cifs and ksmbd directory to fs/smb/
- And each directory name is also changed to fs/smb/client and server.
- Common helper function and macros headers to fs/smb/common/

fs/cifs -> fs/smb/client

fs/ksmbd -> fs/smb/server

fs/smbfs_common -> fs/smb/common

Status update – No longer Experimental

- No longer experimental module
- Many fixes and much testing for 2 years

```
-1,5 +1,5 @@
onfig SMB_SERVER
    tristate "SMB3 server support (EXPERIMENTAL)"
    tristate "SMB3 server support"
    depends on INET
    depends on MULTIUSER
    depends on FILE_LOCKING

ff --git a/fs/smb/server/server.c b/fs/smb/server/server.c
dex 801cd0929209c..5ab2f52f9b355 100644
- a/fs/smb/server/server.c
+ b/fs/smb/server/server.c
-590,8 +590,6 @@ static int __init ksmbd_server_init(void)
    if (ret)
        goto err_crypto_destroy;

    pr_warn_once("The ksmbd server is experimental\n");
```


Status update – LTS kernel support

- **ksmbd bug fix patches were not applied well to the LTS kernel for various reasons.**
 - Cleanup patches
 - New features
 - Directory name changes
- **Due to a lack of background knowledge, Users complain of difficulty in backporting them directly.**
- **Backport all ksmbd patches to LTS 5.15, 6.1, 6.6 kernels.**
- **stable kernel maintainers to apply patches without hunk failure in the future, all patches, not just bug fix patches, were backported.**

Status update – MacOS compatibility

- **ksmbd has verified compatibility with the Mojave version (SMB 3.0).**
- **As OS upgrade and supports SMB3.1.1, Various compatibility issues have been reported from Mac's users.**
 - **Connection failure(AES256 encryption by default)**
 - **Volume format show “SMB (Unknown)”**
 - **Read failure from compound (open-read-close)**

Status update – directory lease(v2 lease)

- **v2 lease allows the client to cache metadata operations in directory**
- **ksmbd supported directory leases now**
- **Demonstrate operations with smbtoriture tests.**

Status update – smb2 lease by default

- **Currently smb2 leases default disable now**
- **Enable “smb2 leases = yes” in ksmbd.conf**
- **Ready to make it by default enable after fixing a few issues**
- **Plan to enable it after more checking it with cifs.ko**

Status update – Durable handles v1/v2

- **Add support durable handle version 1 and 2 as defined in the specification.**
- **ksmbd can handle temporary loss of connection when a file is still opening now**
- **Demonstrated the operation of this feature with smb torture durable open tests.**
- **Continuous Availability is not implemented yet. (TODO)**

Status update – compound read support

- **ksmbd supported compound request**
- **Missing read compound case in implementation**
- **The patch consider the following cases**
 - **read()-read()-read() in compound**
 - **open()-read()-close() in compound**

Status update – big file copy issue on ksmbd + exfat

- **PROBLEM:** big file copy speed is very slow on ksmbd + exfat(same with samba + exfat)
- **Windows client send SMB_FILE_END_OF_FILE_INFORMATION smb2 setinfo request to pre-allocate space as much as the file size before writing the file.**
- **Ftruncate of Non-sparse filesystem like fat or exfat zero out blocks and cause low performance.(Looks like stuck on windows client)**
- **Exfat has valid-size field and add a half of sparse support to exfat on Linux 6.8 kernel**

Status update – credits management

- **Grant at least one credit to client although credit request in request is zero**
- **Refuse request process if :**
 - **all max credits in ksmbd are consumed**
 - **Client's granted credits are insufficient**
- **Add smb2 max credits parameter to ksmbd.conf**

Status update – SMB 3.1.1 posix extensions

- **Mount from cifs.ko(Linux kernel client) with “posix” option**
- **Fill in SIDs in SMB_FIND_FILE_POSIX_INFO, SMB_2_CREATE_POSIX_CONTEXT and SMB_FIND_FILE_POSIX_INFO responses**
- **All posix extensions implementation supported:**
 - **SMB2_CREATE_POSIX_CONTEXT in smb2 create**
 - **SMB_FIND_FILE_POSIX_INFO in smb2 query directory**
 - **FS_POSIX_INFORMATION in smb2 get info filesystem**
 - **SMB_FIND_FILE_POSIX_INFO in smb2 get info file**

Status update – Fully support SMB-Direct

- **Handle large RDMA read/write size(bulk data) supported by SMB Direct multi-descriptors(It supported single descriptor with 512KB size before)**
 - **8MB RDMA read/write size by default**
 - **Control read/write size through ksmbd configuration.(e.g. smb io size = 16MB)**
- **Improve the compatibility with various RDMA types of NICs**
 - **Tested smb-direct working with iWARP(Chelsio, soft-iWARP), Infiniband(Mellanox NICs, Connect x3 ~ x5), ROCE(soft-ROCE).**
- **Auto-detection of RDMA NIC without configuration.**
 - **Server should send RDMA NIC info to client.**
 - **No need to specify RDMA NIC information to smb.conf.**

Status update – rename and unlink

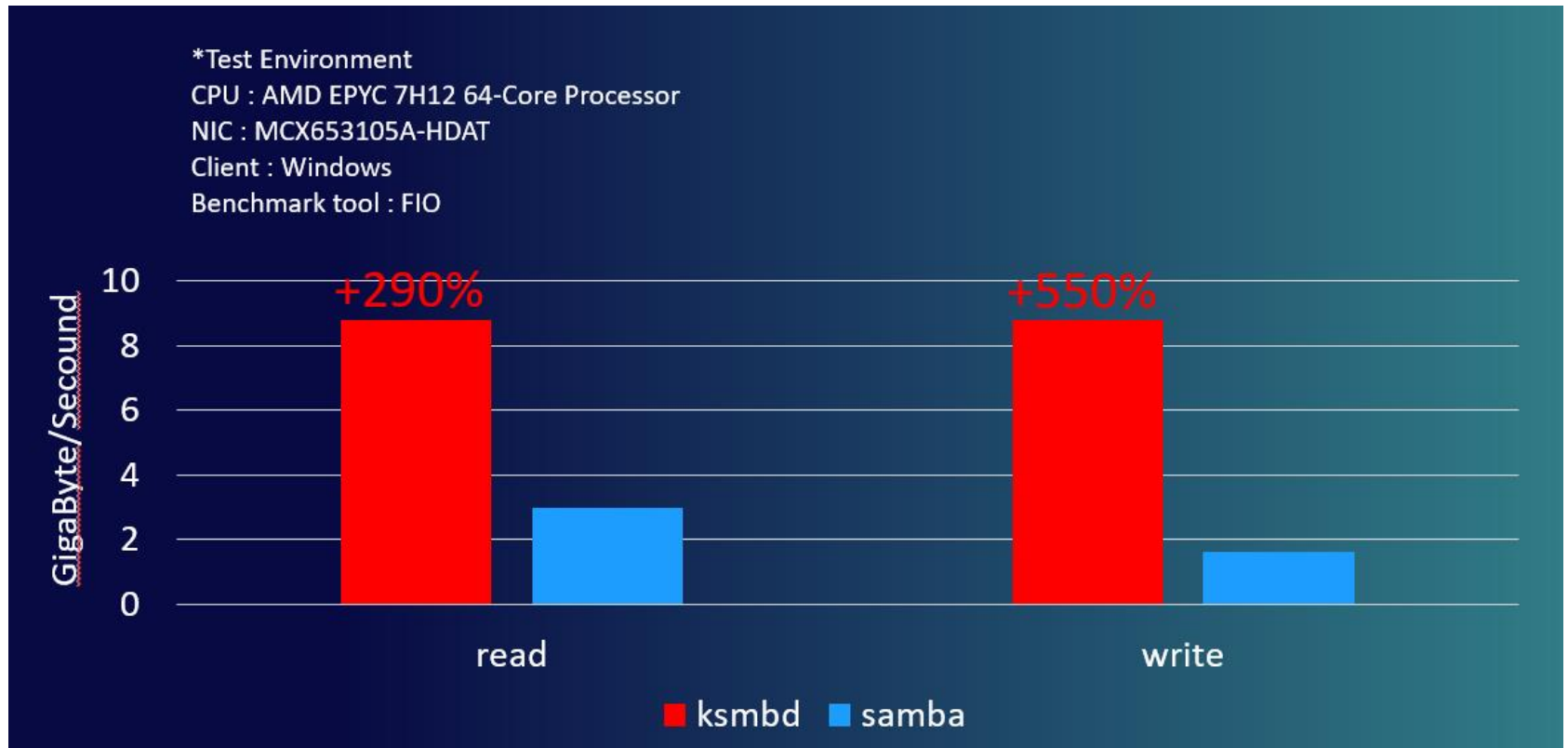
- **Rename and unlink racy issue**
- **Add two vfs helpers**
 - Add `vfs_path_parent_lookup()` helper
 - Add `lock_rename_child()` helper

Status update – user report and feedback

- **‘ksmbd can replace samba?’ DJ Ware ([link](#))**
- **ksmbd performance test ([link](#))**
- **Symlink support**
- **Config backend (Recently get a request, make the configuration interface available remotely over the WINREG RPC interface)**

Status update – Multichannel user report

- Ziwei Xie(high-flyer) give this test reports
- performance difference of 3 times for read and 4 times for write on his setup



Status update – SMB-Direct user report

- Yufan Chen give test results for SMB-Direct
- Benchmark tool
 - Framtest (<https://support.dvsus.com/hc/en-us/articles/212925466-How-to-use-frametest>)

- **Server**

CPU: intel Silver 4114 x 2

DRAM: 512GB

NVMe SSD: Kioxia CM6 1.9T x 9(mdadm raid0 with XFS)

NIC: MCX516A-CCAT 100GbE

- **Client**

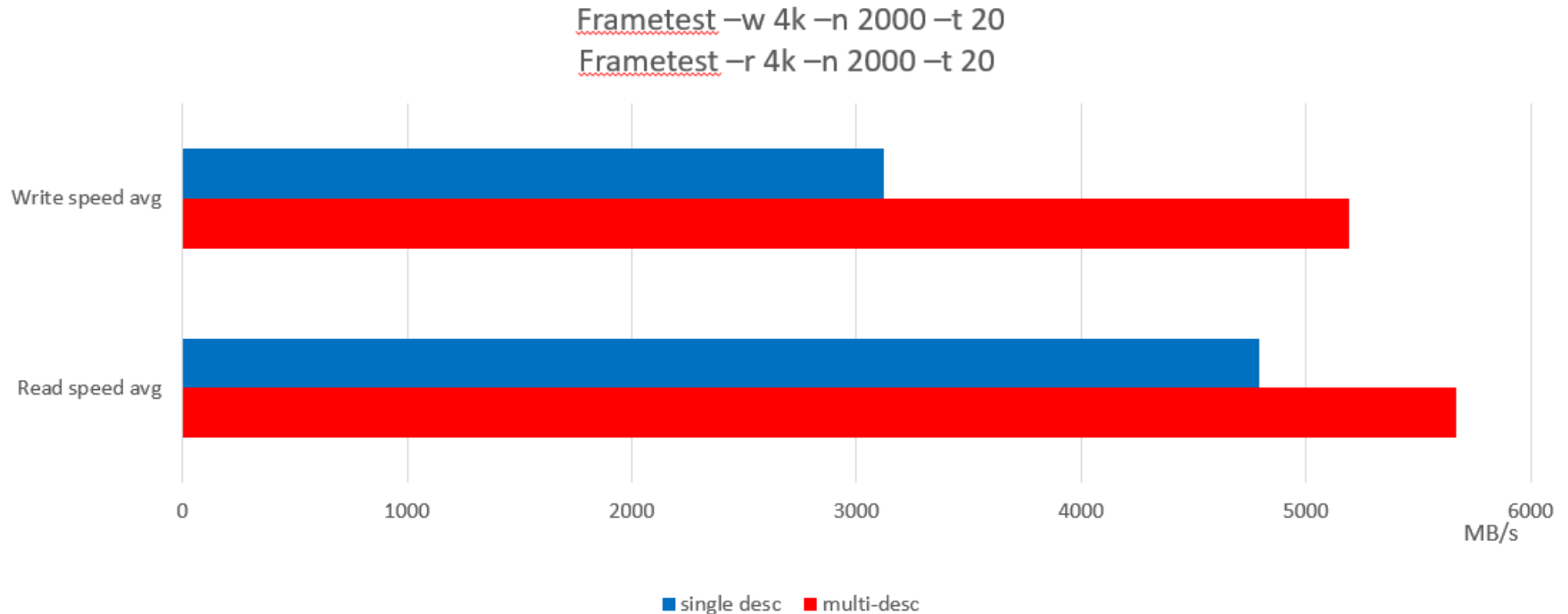
CPU: intel Silver 4215 x 2

DRAM: 64GB

NIC: MCX516A-CCAT

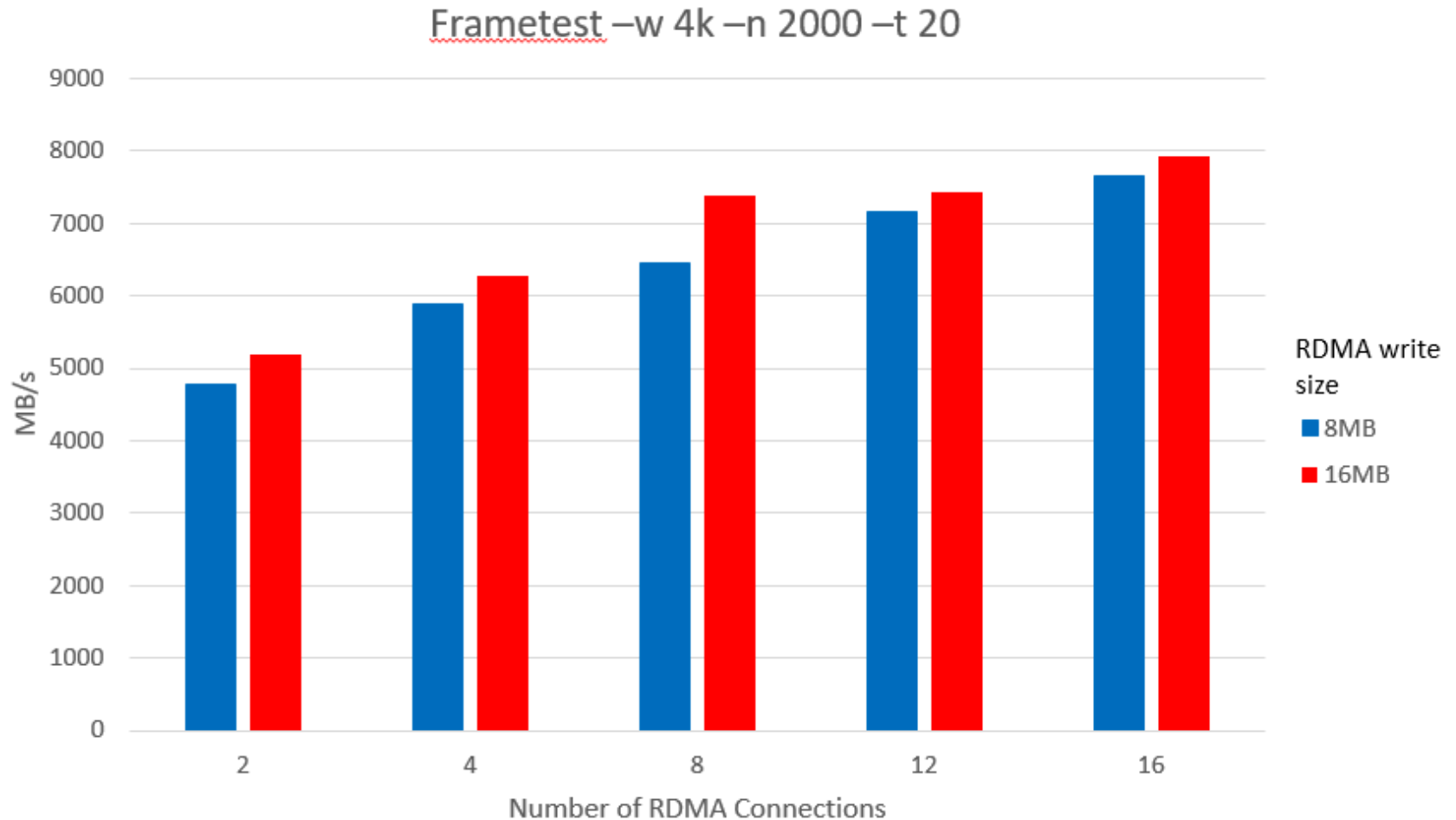
Status update – SMB-Direct user report

- Performance comparison between single and multi-descriptor



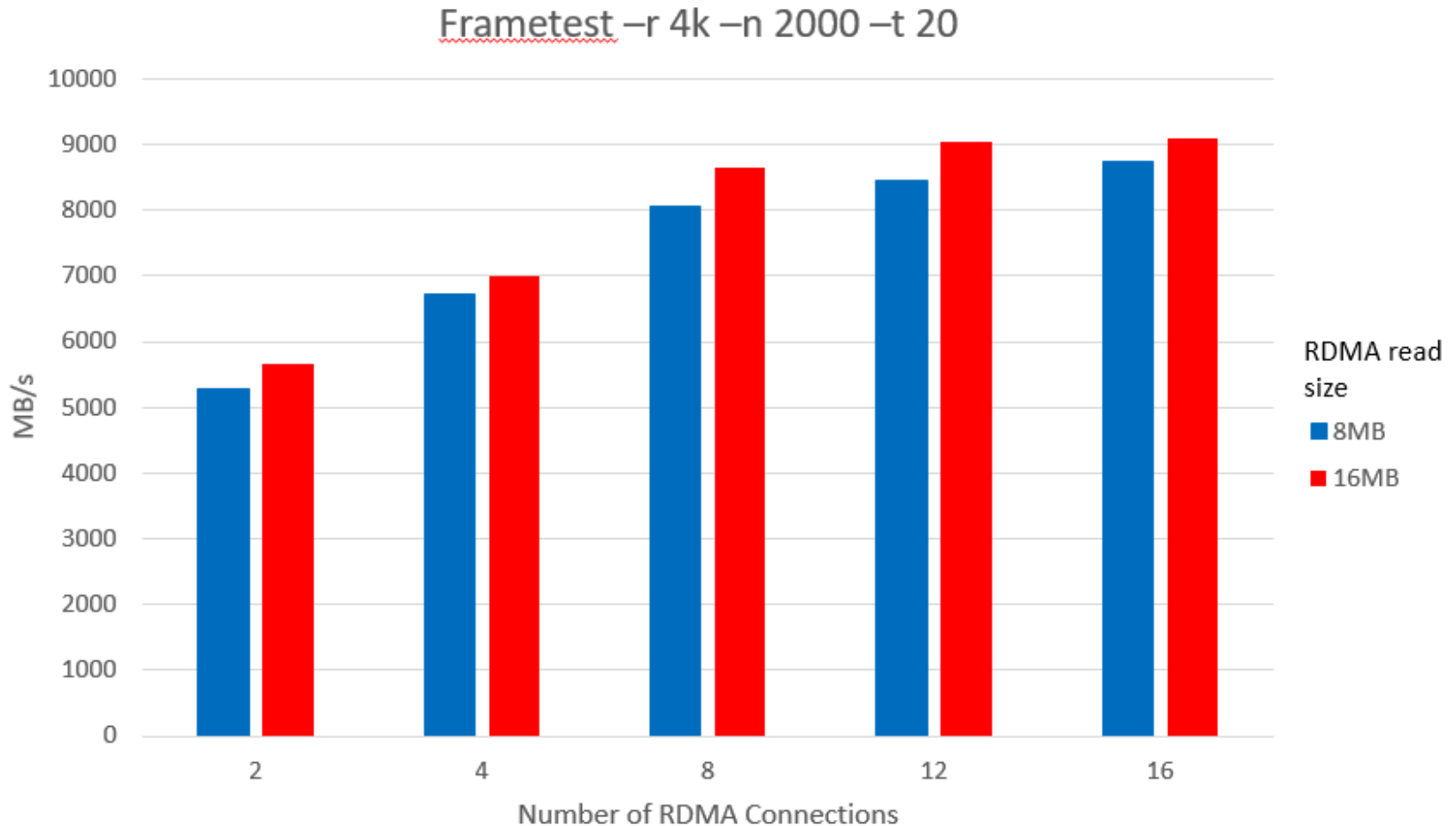
Status update – Fully support SMB-Direct

- RDMA write performance per number of connections



Status update – Fully support SMB-Direct

- RDMA read performance per number of connections



Status update – ksmbd-tools

- **Rename smb.conf to ksmbd.conf**
- **Atte Heikkilä refactored ksmbd-tools codes**
- **Add manpages for all utils and ksmbd.conf**
- **Added new parameters:**
 - **durable handles**
 - **crossmnt**
 - **smb3 encryption**
 - **smbd io size**
 - **smb2 max credits**

Next works

- **ksmbd over QUIC introduced by Xin Long**
- **Reply operations support as next step of durable handle**

Thank you!