INTRODUCE IN-KERNEL SMB3 SERVER CALLED CIFSD
About me

- Linux kernel contributor since 2011
- Co-Creator of Samsung internal NTFS Filesystem
- Introduce collapse and insert range syscall
- Creator and maintainer of linux-cifsd project
Topic

- Introduction
- Architecture
- Components
- Performance/Stability/Compatibility
- Plan
- Proposal
What is cifsd?

- SMB Server for Linux kernel
- Kernel and Userspace daemons
- All SMB Ver. (SMB1 ~ SMB3.1.1)
- Authentication
  - NTLM
  - NTLMv2
- Performance feature
  - Oplock/lease
  - compound request
  - Copy offload
- Security feature
  - Signing
  - encryption
linux-cifsd project

- Github Repo
  - https://github.com/cifsd-team/cifsd
  - https://github.com/cifsd-team/cifsd-tools

- Mailing-list
  - linux-cifsd-devel@lists.sourceforge.net

- 5 active developers
  - SMB2 notify (In progress, Yunjae Lim)
  - SMBDirect (In progress, Hyunchul Lee)
**Key Concepts**

- **Can gain the performance in kernel?**
  - No system call (less TLB miss, less context switching)
  - Shorter path to use VFS and network functions in kernel
  - No duplicate memory allocation for inode and superblock
Key Concepts

- Optimized SMB over RDMA support
### Key Concepts

- **Simple/light file share for embedded device**

  Reported by Andy Walsh (OpenWRT)

<table>
<thead>
<tr>
<th>Binary Size</th>
<th>Main</th>
<th>Extra</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>cifsd</td>
<td>128KB (cifsd kmod, tools)</td>
<td>61KB (crypto kmods) + 872KB (glib2)</td>
<td>1061KB</td>
</tr>
<tr>
<td>samba4</td>
<td>6MB (samba libs, server package)</td>
<td>64KB (libtirpc, etc)</td>
<td>6064KB</td>
</tr>
</tbody>
</table>
Key Concepts

- Oplock/lease better handling (page 20)
Architecture

- **Separate kernel space and user space daemon**
  - Works related to performance in kernel space
  - Works related to non-performance in user space

- **Co-work cifsd and kcifsd in each space**
  - When cifsd is launched, kcifsd is activated
  - They exchange information necessary for each other
Architecture

- How to communicate between kernel and userspace
  - Use Netlink interface
  - Specify a few commands

<table>
<thead>
<tr>
<th>Name</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIFSD_EVENT_HEARTBEAT_REQUEST</td>
<td>Monitor cifsd is alive</td>
</tr>
<tr>
<td>CIFSD_EVENT_STARTING_UP</td>
<td>Transfer the initial information necessary for the start and shutdown</td>
</tr>
<tr>
<td>CIFSD_EVENT_SHUTTING_DOWN</td>
<td></td>
</tr>
<tr>
<td>CIFSD_EVENT_LOGIN_REQUEST</td>
<td>Transfer the user account / password information necessary for login</td>
</tr>
<tr>
<td>CIFSD_EVENT_LOGIN_RESPONSE</td>
<td></td>
</tr>
<tr>
<td>CIFSD_EVENT_SHARE_CONFIG_REQUEST</td>
<td>Transfer the share configuration</td>
</tr>
<tr>
<td>CIFSD_EVENT_SHARE_CONFIG_RESPONSE</td>
<td></td>
</tr>
<tr>
<td>CIFSD_EVENT_TREE_CONNECT_RESPONSE</td>
<td>Transfer the tree connect info</td>
</tr>
<tr>
<td>CIFSD_EVENT_TREE_DISCONNECT_REQUEST</td>
<td></td>
</tr>
<tr>
<td>CIFSD_EVENT_RPC_REQUEST</td>
<td>Transfer DCERPC requests</td>
</tr>
<tr>
<td>CIFSD_EVENT_RPC_RESPONSE</td>
<td></td>
</tr>
</tbody>
</table>
Architecture

User Space

- cifsadmin
- cifsd

Kernel Space

- SOCKET (445)
- kcifsd/0 (forker thread)
- kcifsd/1
- kcifsd/2
- kcifsd/N

NETLINK/SYSFS Interface

- cifspwd.db (ID/PW file)
- smb.conf (config file)
- DCE/RPC
- Share configuration
- ID/PW configuration

Authentication
- NTLM
- NTLMv2
- Kerberous

Dialects
- SMB1
- SMB2
- SMB2.1
- SMB3
- SMB3.1.1
Architecture

- KCIFSD Components

- NETLINK INTERFACE

- KCIFSD
  - TRANSPORT
  - IPC
  - SERVER
  - TRANSPORT
  - TCP
  - SMB Engine
  - Oplock/lease
  - AUTH
  - VFS CACHE (File, INODE)
  - VFS ABSTRACTION

- Virtual Filesystem
  - EXT4
  - XFS
  - NTFS
Architecture

- Minimum DCERPC implementation

- The parameter format of smb.conf
  - compatible with samba's one
  - Minimum implementation
  - List up of supported parameters in smb.conf.example

- SMB1 is disable at default
  - Smart phone apps (ES File Explorer) support only SMB1
  - Can easily remove it when merging into mainline
Performance comparison

- Tool: Iozone, fileop, bench-oplock(smbtorture)
- Mount share tmpfs
- Direct connection on two PC
- Oplock / lease is disable
- SMB client is a kernel cifs
Performance comparison

**Single Writer Iozone Throughput**

- Kilobyte/second vs. Record length (KB)

- Graph showing throughput for Samba write (blue line) and CifsD write (red line)

- Data points for record lengths 4, 8, 16, 32, 64, 128 KB
Performance comparison

Single Reader Iozone Throught

Kilobyte/second vs. Record length (KB)

- Samba read
- Cifsd read
Fileops Result

![Bar chart showing file operations per second for different operations and file systems. The x-axis represents various file system operations, including mkdir, chdir, rmdir, create, open, read, write, close, stat, access, chmod, readdir, link, unlink, and delete. The y-axis represents the number of operations per second. There are two sets of bars for each operation, one in blue for Samba and one in red for Cifs. The chart shows a comparison of the performance of these operations between the two file systems.](chart.png)
File lookup Performance (ls -l)

- **samba**
- **cifsd**

<table>
<thead>
<tr>
<th>Number of files</th>
<th>Time (millisecond)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0</td>
</tr>
<tr>
<td>5000</td>
<td>100</td>
</tr>
<tr>
<td>10000</td>
<td>300</td>
</tr>
<tr>
<td>50000</td>
<td>1400</td>
</tr>
</tbody>
</table>
## Compatibility

<table>
<thead>
<tr>
<th>SMB CLIENT VERSIONS</th>
<th>CIFSD SUPPORTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows XP (SMB 1.0)</td>
<td>✔</td>
</tr>
<tr>
<td>Windows Vista (SMB 2.0)</td>
<td>✔</td>
</tr>
<tr>
<td>Windows 7 (SMB 2.1)</td>
<td>✔</td>
</tr>
<tr>
<td>Windows 8 (SMB 3.0)</td>
<td>✔</td>
</tr>
<tr>
<td>Windows 10 (SMB 3.1.1)</td>
<td>✔</td>
</tr>
<tr>
<td>MacOS (~ High Sierra)</td>
<td>✔</td>
</tr>
<tr>
<td>Ubuntu File Explorer</td>
<td>✔</td>
</tr>
<tr>
<td>Linux CIFS Client (linux 4.16)</td>
<td>✔</td>
</tr>
</tbody>
</table>
Stability

**SMB TORTURE(SMB2/3)**
- **PASS**: 95%
- **FAIL**: 5%

**XFSTESTS(SMB2/3)**
- **PASS**: 97%
- **FAIL**: 3%
Plan

- SMB Direct Support (~ Sep)
  - Writing by Hyunchul Lee (LGE)
  - Share the status at SDC 2019

- Oplock/Lease is enable at default

- Send the patch-set to LKML (~ Sep)
Proposal

- How about make kcifsd & samba running together?
Proposal

- Define new parameter in smb.conf
- Kcifsd can be a kernel helper of samba
- Use SMBDirect in kcifsd through ioctl or netlink
Thank you!