A Comparison Between the Samba 3 and Likewise Lwiod SMB File Servers

September 20, 2010

Outline

- Overview
- Threading Architecture
- Internal Model
- Feature Set
- Configuration

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Overview

- 18 years development
- GPLv3
- Personal copyright
- www.samba.org

- 2 years development
- GPLv2+
- Corporate copyright
- www.likewiseopen.org

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Comparison

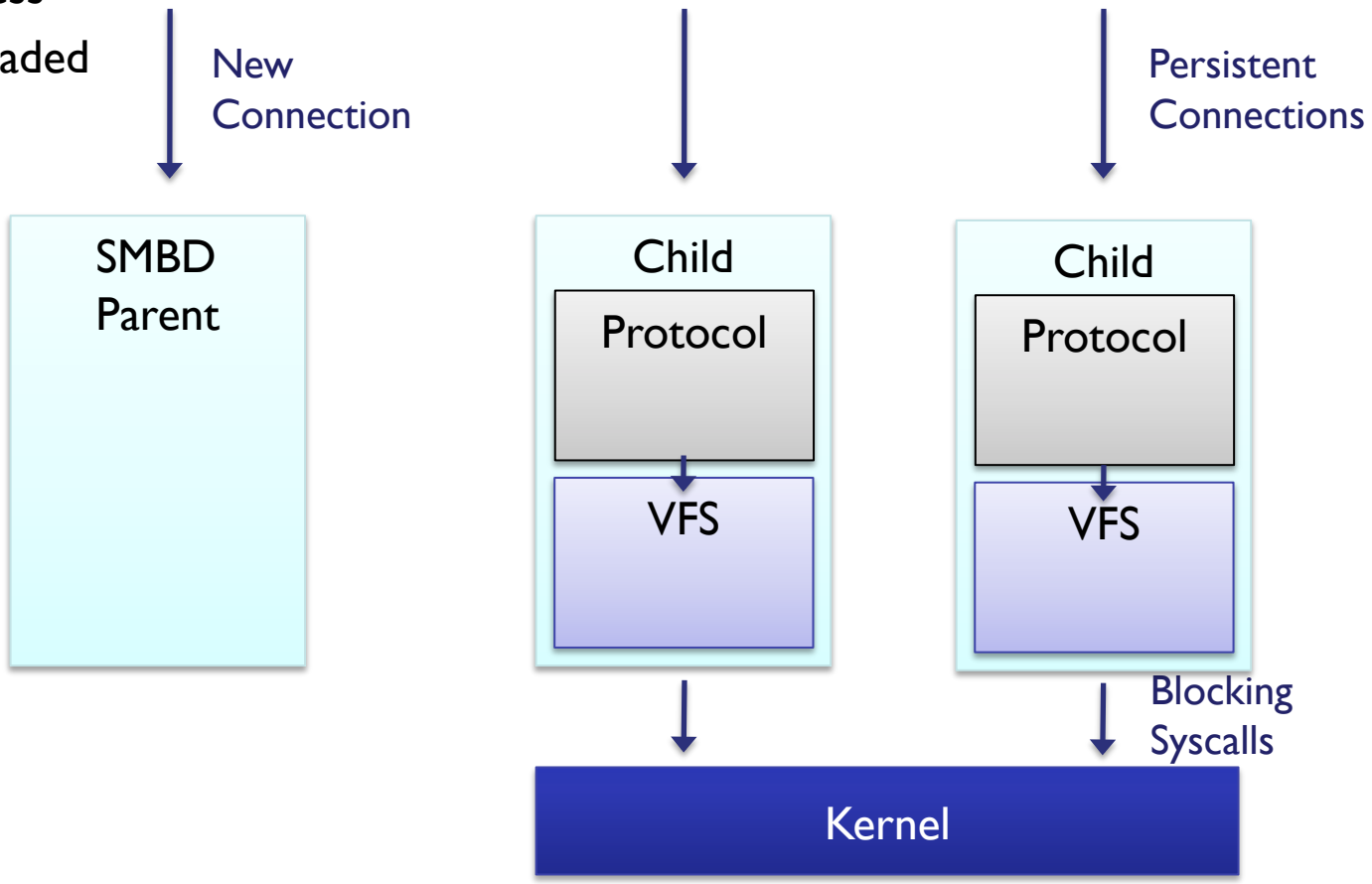
- Git "master" branches as of August 21, 2010
- Focus on file server functionality
- Focus on Linux platform
- Only small discussion of:
  - Samba 4
  - winbindd
  - lsassd
- No discussion of:
  - Clustering

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Threading Architecture

SMBD Architecture

- Multi-Process
- Single Threaded



2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Multi-process Single-threaded Pros

+ Simple, direct I/O
+ Compatible with POSIX per-process semantics
  + fcntl() – locking
  + setcred() – security credentials
+ No threading synchronization or context switches
+ Process crash only affects single connection

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

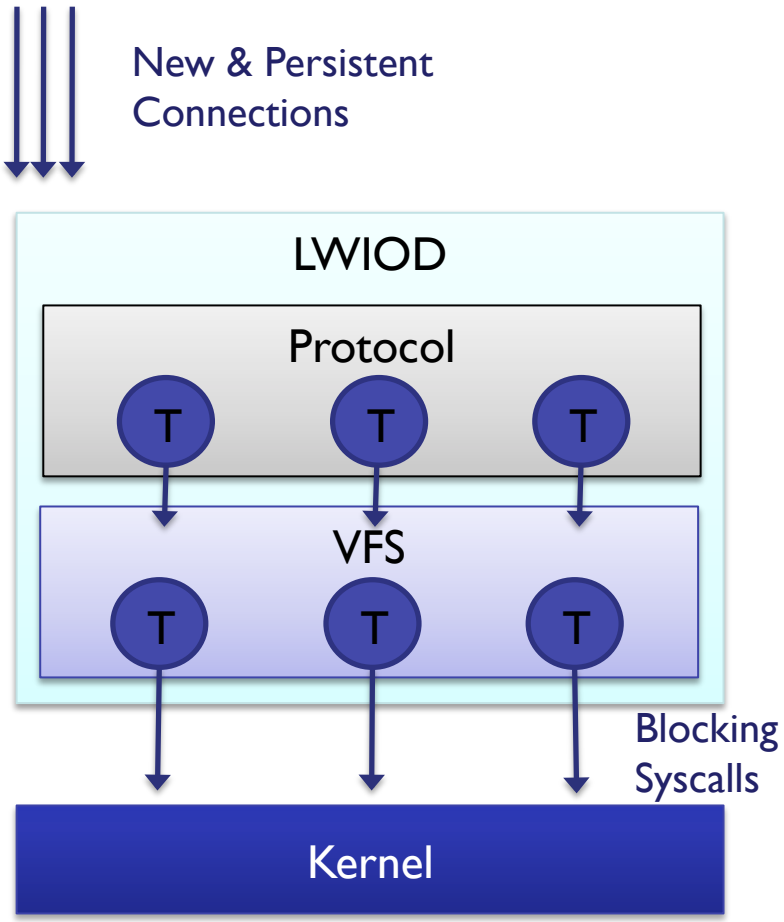
Multi-process Single-threaded Cons

- Slow kernel operation blocks new requests
  - Extreme case: ECHO packets go unreplied
  - Especially bad for multiplexed sessions
- New connection throttling all or nothing
  - "max smbd processes"
- Idle connections consume whole process
- IPC necessary for sharing connection state

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

LWIOD Architecture

- Single Process
- Multi-Threaded



2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Single-process Multi-threaded Pros

+ Pipelined network I/O written in parallel
+ Parallel syscalls: network I/O not blocked by file system
+ New connections limited by same thread pool as all other operations
+ Idle connections consume very little resources

Likewise®

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Single-process Multi-threaded Cons

- Incompatible with POSIX per-process semantics
  - Locking – must be implemented in user space
  - Credentials – access checks must be implemented in user space
- Thread context switch overhead
- Process crash affects ALL connections
- Hit OS limits faster: file descriptors

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Architecture Conclusion

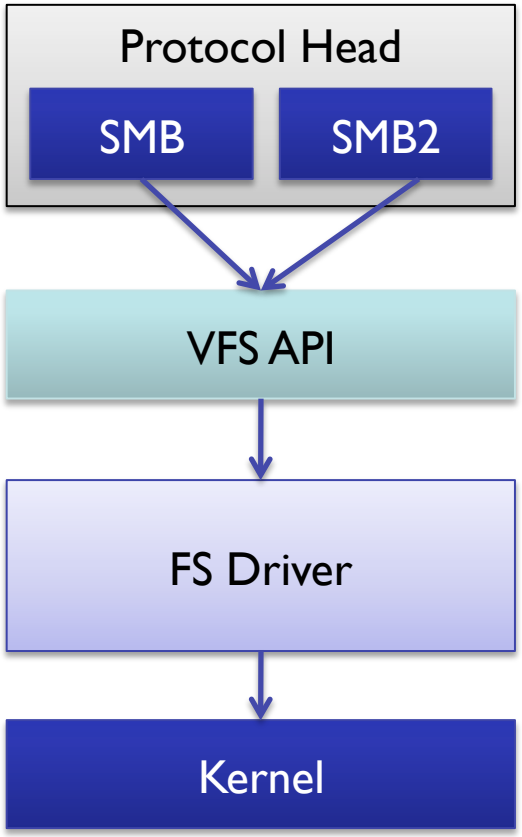
- For highly concurrent workflows / 100s clients:
  - LWIOD *should* scale better
    - More graceful degradation under load
  - LWIOD *should* use less total resources
    - Threads lighter weight / bounded thread pool
- Multi-threaded model forfeits **important** POSIX support
  - Kernel access checks
  - fcntl() locking

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Internal Model

User Space Process Model

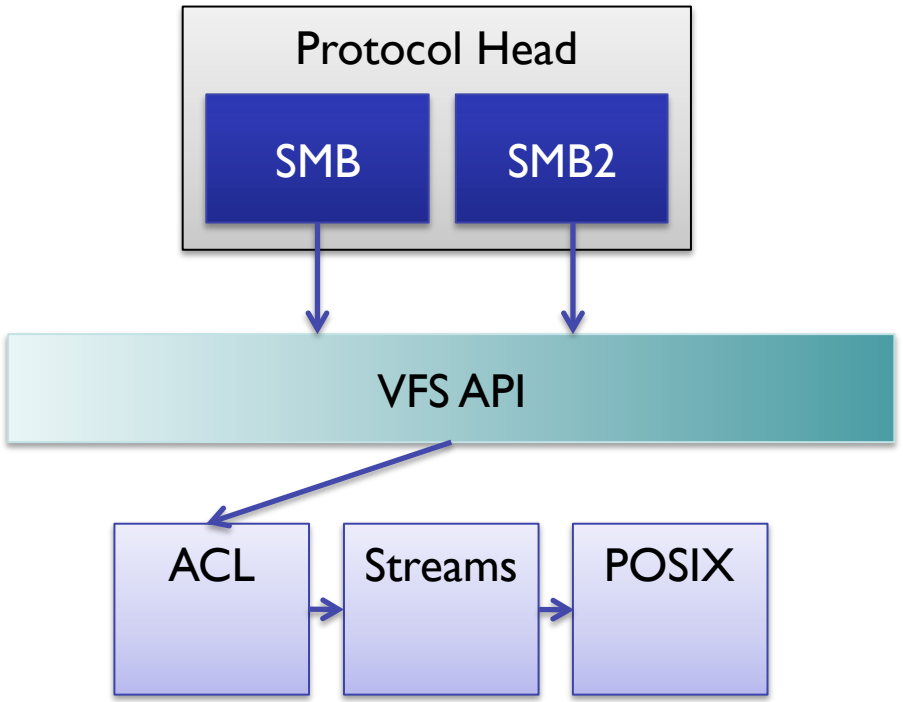
- Protocol Head
  - Marshall packets on/off socket
- VFS API
  - Abstract all file system calls
- FS Driver
  - Map NTFS FS semantics to native FS
  - Calls kernel syscalls()



2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

SMBD VFS

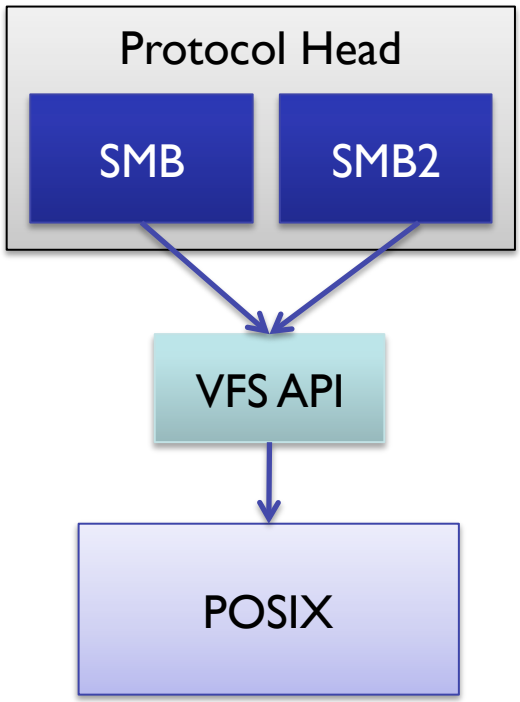
- Hybrid Interface
  - POSIX / NTFS / Other
- 113 functions
- Chained modules
- OS specific modules
  - irixacl, hpuxacl, zfsacl
- ~30 modules
  - Partial API



2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

LWIOD VFS

- IOMGR Interface
  - Abstract Device Driver API
- 16 functions
  - Includes IOCTL
  - Every call can go async
- No chaining or filtering
- Generic POSIX module
  - No OS specific functionality yet
  - Reference implementation
- 1 Module



2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

SMBD VFS Example - ACLs


```
/* NT ACL operations. */

        NTSTATUS fget_nt_acl(struct vfs_handle_struct *handle,
                             struct files_struct *fsp,
                             uint32 security_info,
                             struct security_descriptor **ppdesc);
        NTSTATUS get_nt_acl(struct vfs_handle_struct *handle,
                            const char *name,
                            uint32 security_info,
                            struct security_descriptor **ppdesc);
        NTSTATUS fset_nt_acl(struct vfs_handle_struct *handle,
                             struct files_struct *fsp,
                             uint32 security_info_sent,
                             const struct security_descriptor *psd);
```

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

LWIOD VFS Example - ACLs


```
/* IRP_TYPE_QUERY_SECURITY */

        NTSTATUS
        GetSecurityDescriptorFile(
            IN PPVFS_CCB pCcb,
            IN SECURITY_INFORMATION SecInfo,
            IN OUT PSECURITY_DESCRIPTOR_RELATIVE pSecDesc,
            IN OUT PULONG pSecDescLength
            )

/* IRP_TYPE_SET_SECURITY */

        NTSTATUS
        SetSecurityDescriptorFile(
            IN PPVFS_CCB pCcb,
            IN SECURITY_INFORMATION SecInfo,
            IN PSECURITY_DESCRIPTOR_RELATIVE pSecDesc,
            IN ULONG SecDescLength
            )
```

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

SMBD VFS Example – Dir Enum


```
/* Directory operations */

        SMB_STRUCT_DIR *opendir(struct vfs_handle_struct *handle,
                                const char *fname,
                                const char *mask,
                                uint32 attributes);
        SMB_STRUCT_DIRENT *readdir(struct vfs_handle_struct *handle,
                                SMB_STRUCT_DIR *dirp,
                                SMB_STRUCT_STAT *sbuf);
        void seekdir(struct vfs_handle_struct *handle, SMB_STRUCT_DIR *dirp, long offset);
        long telldir(struct vfs_handle_struct *handle, SMB_STRUCT_DIR *dirp);
        void rewind_dir(struct vfs_handle_struct *handle, SMB_STRUCT_DIR *dirp);
        int closedir(struct vfs_handle_struct *handle, SMB_STRUCT_DIR *dir);
```

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

LWIOD VFS Example – Dir Enum


```
/* IRP_TYPE_QUERY_DIRECTORY */

    NTSTATUS
    QueryDirInformation(
        OUT PVOID FileInformation;
        IN ULONG Length;
        IN FILE_INFORMATION_CLASS FileInformationClass;
        IN BOOLEAN ReturnSingleEntry;
        IN OPTIONAL PIO_MATCH_FILE_SPEC FileSpec;
        IN BOOLEAN RestartScan;
        )
```

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Developer Perspective

- LWIOD provides a cleaner abstraction for NTFS file system semantics
- SMBD provides more module implementations across a variety of existing file systems
- LWIOD inherently handles asynchronous operations within the VFS interface
- SMBD module chaining prevents code duplication

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Feature Set

Feature Set

	SMBD	LWIOD
SMB Dialect	<= NT LM 0.12	== NT LM 0.12
SMB2 Dialect	SMB 2.002	SMB 2.002
NetBios Transport	✓	✗
Share Level Security	✓	✗
Share Mode Locks	✓	✓
Oplocks	✓	✓
Change Notify	✓	✓
Byte Range Locks	○	○
Access Control Lists	○	○
Alternate Data Streams	✓	✗
Shadow Copy	✓	○

✓ Works ○ Mostly Works ✗ Doesn't Work

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

SMB Protocol Dialect / NetBios

+ Unix Extensions
+ Mac Extensions

+ SNIA Reference support
- Unsupported:
  - Port 139
  - NT4 clients
  - OS X <= 10.4 clients
- No password-less login

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

SMB2

- SMB v1 equivalent support in both.
  - SMB2 2.002 == Windows Vista / 2008 Server
- No SMB2 only features:
  - Durable / Resilient Handles
  - Leases
- LWIOD provides statically configurable credits accounting

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Share Mode / Oplock / Change Notify

+ Linux kernel oplock partial support
+ Linux kernel notify support

- Both: Functionality implemented in user space.
- No cross-protocol support without kernel implementation.

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Byte Range Locks

+ Can back with POSIX locks

- Both: Implement Windows locking semantics in user space.
- Both: Use user space queue that doesn't guarantee fairness.
- Both: Provide strict locking semantics for SMB only traffic.

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Access Control Lists

	SMBD	LWIOD
Access Checks	Kernel	User Space
Storage in xattr	✓	✓
Map to Mode Bits	✓	✓
Map to POSIX ACL	✓	✗
Map to NFSv4 ACL	○	✗
NT4 Style ACL	✓	✗

- SMBD: xattr storage *AND* POSIX ACL
- LWIOD: xattr storage *OR* mode bits

✓ Works ○ Mostly Works ✗ Doesn't Work

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Access Control Lists

- Best choice:
  - + Perfect POSIX support
  - + Perfect Windows
    - + get/set
  - – Imperfect Windows
    - – access check

- Best choice:
  - + Perfect Windows support
  - – No POSIX support

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Access Control Lists

- Both: Lossy conversion between NTFS ACL and mode bits / POSIX ACL
  - RWX != Full Control
  - Groups owning files
  - Deny aces and canonical order
- Both: Rely on SID to Unix ID mapping
  - Unmappable SIDs can't be stored
- NTFS ACLs to POSIX sucks no matter what you do!

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Alternate Data Streams

+ ADS storage in hidden subdirectory tree
+ ADS storage in xattrs

– Support for ::$DATA stream only

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Shadow Copy

+ Allows enumeration / restore
+ Conversion of wire name to directory name

– POSIX VFS module lacks enumeration / restore

- Both: Support parsing of @GMT path format in protocol head.
- Both: Rely on underlying file system to implement snapshots.

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Feature Set (continued)

	SMBD	LWIOD
Srvsvc	○	✓
DOS File Attributes	✓	✓
Distributed File System	✓	✗
Signing / Sealing	✓	✓
Privileges	✓	✗
Auditing	○	✗
IPv6	✓	✓
Copy Chunk	✗	✗

✓ Works ○ Mostly Works ✗ Doesn't Work

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Non-File Server Functionality

	SMBD	LWIOD
Printing / spoolss	✓	✗
NetBios Name Server	✓	✗
NT4 Domain Server	✓	✗
Active Directory Server	○	✗

✓ Works ○ Mostly Works ✗ Doesn't Work

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Feature Conclusion

- SMBD far more mature in number of features.
- No technical reason for lack of features in LWIOD simply a matter of development time and effort.
- LWIOD supports newer clients only.
- Some features cannot be perfectly duplicated on POSIX regardless of server implementation.
  - ACLs
  - BRL

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Configuration

Administrating / Configuration

- Stored in .ini text file
- 374 Total Parameters
- 146 Share Parameters
- Alternate registry configuration storage

- Stored in registry
- 22 Parameters (lwiod only)
+ MMC support by default
– Examples of missing config:
  – Access based enum
  – Share path substitution
  – Name mangling

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.

Questions?

Contact

Steven Danneman
sdanneman@isilon.com
steven@samba.org

2010 Storage Developer Conference. Isilon Systems. All Rights Reserved.