Direct-FUSE: Removing the Middleman for High-Performance FUSE File System Support

Yue Zhu*, Teng Wang*, Kathryn Mohror[+],
Adam Moody[+], Kento Sato[+], Muhib Khan*,
Weikuan Yu*

Florida State University*
Lawrence Livermore National Laboratory[+]

Introduction

- An efficient file system is important for high-performance computing (HPC) systems in supporting large-scale scientific applications.

    - User-level file systems are more designed for particular I/O workloads with special-purpose, due to development complexity, reliability, and portability.
    - Different file systems are used for different kinds of data in a single job.

- Filesystem in Userspace (FUSE)

    - A software interface for Unix-like computer operating systems.
    - It allows non-privileged users to create their own file systems without editing kernel code.
    - User defined file system run as a separate process in user-space.

Breakdown of Metadata & Data Latency

- The create() and close(), and write() are taken as examples to show the percentage of real operation time in a complete FUSE metadata and data operation, respectively.

  ➢ Tests are on tmpfs and FUSE-tmpfs.

  ➢ **Real Operation** in metadata operation: the time of conducting operation.

  ➢ **Data Movement**: the actual time of write in a complete write function call.

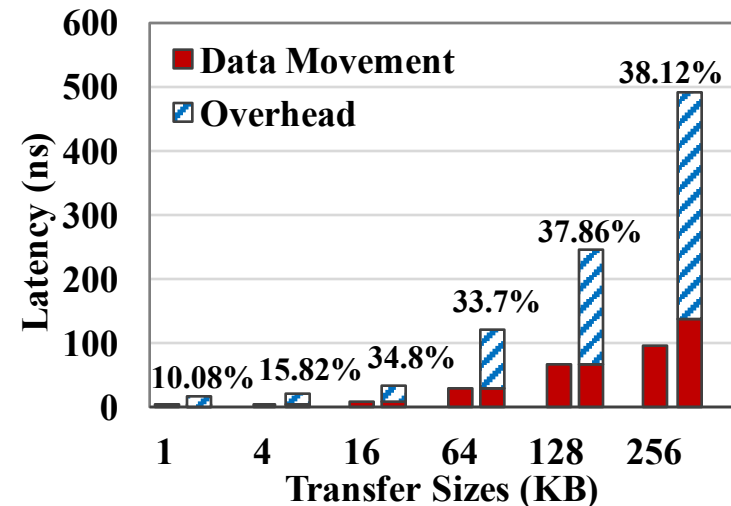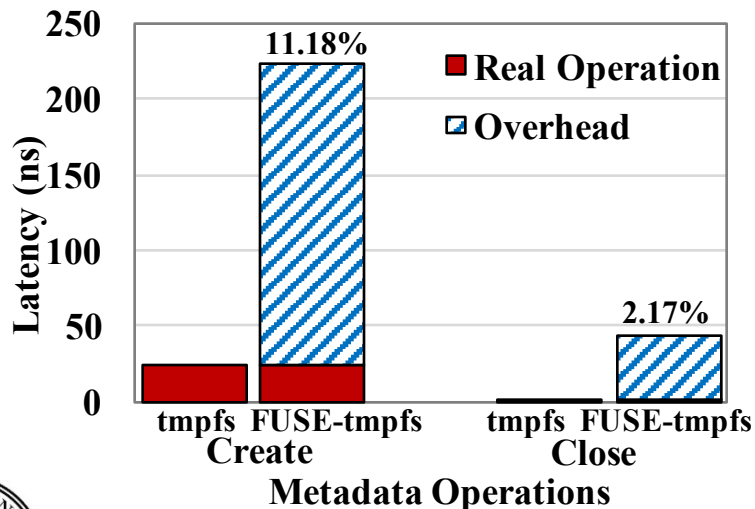 ➢ **Overhead**: the cost besides the above two, e.g. the time of context switches.

Fig. 1. The time Expense in Metadata Operations   Fig. 2. The time Expense in Data Operations

The Overview of Direct-FUSE

- *Direct-FUSE* contains the adapted *libsysio, lightweight-libfuse*, and *backend services*.
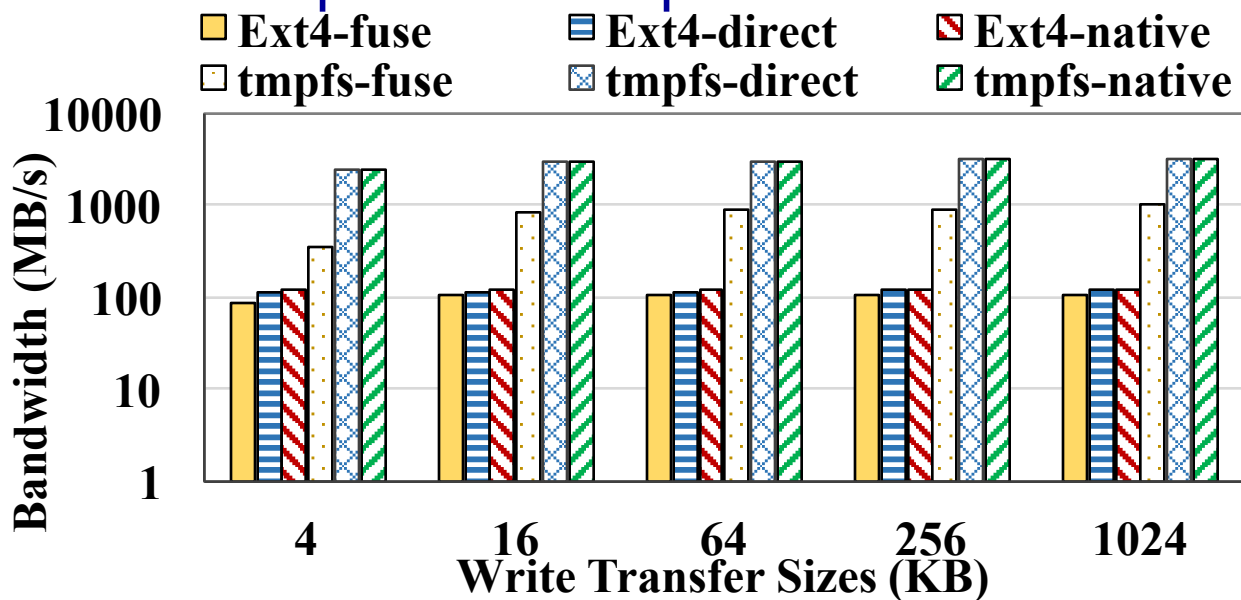
  - ➢ Adapted-libsysio
    - ○ Support multiple backends
  - ➢ lightweight-libfuse
    - ○ Not real *libfuse*
    - ○ Exposes file system operation to under layer backend services with supporting FUSE library.
  - ➢ Backend services
    - ○ Provide defined FUSE operations.

Application Program
Direct-FUSE
adapted-libsysio
lightweight-libfuse
Backend Services

FUSE-ext4	SSHFS	FTPFS
FusionFS	GlusterFS	…
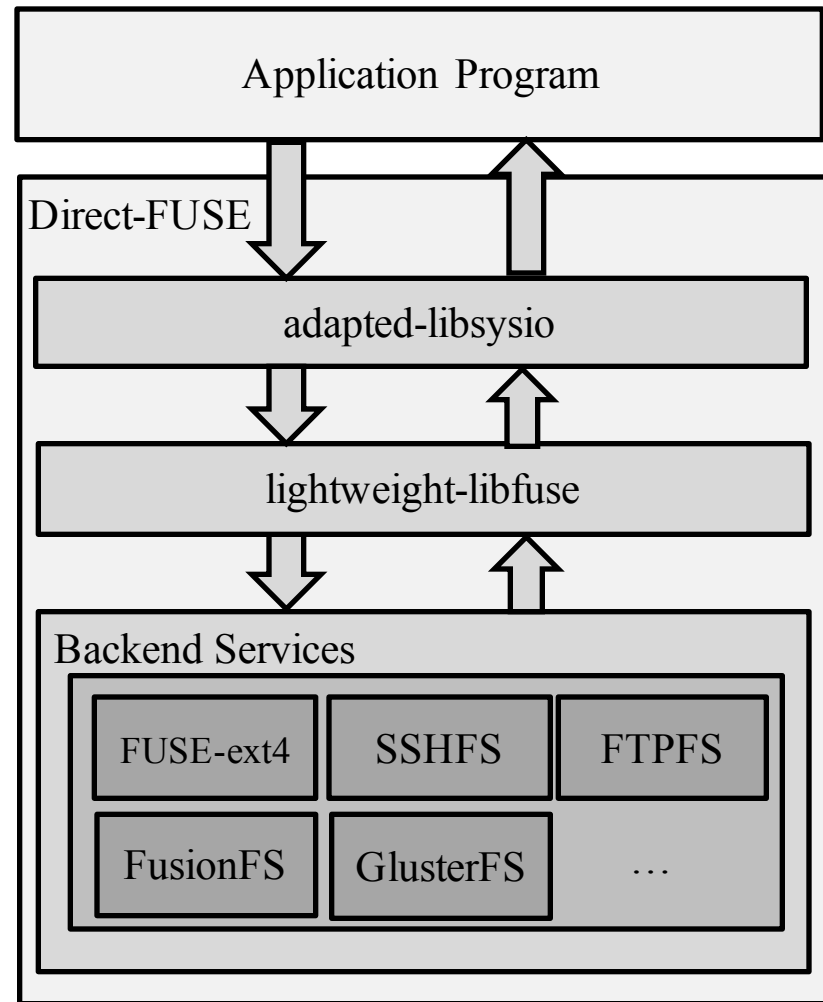
Sequential Write Bandwidth

- The bandwidth of Direct-FUSE is very close to the native file system.

  - **Ext4(tmpfs)-fuse**: FUSE file system overlying Ext4 (tmpfs); **Ext4(tmpfs)-direct**: Direct-FUSE on Ext4 (tmpfs); **Ext4(tmpfs)-native**: original Ext4 (tmpfs).

  - Ext4-direct outperforms Ext4-fuse by 11.9% on average

  - tmpfs-direct outperforms tmpfs-fuse at least 2.26x.

Sponsors of This Research