



St. Petersburg State University
Department of System Programming

Implementation of log-structured block device in Linux kernel

Mikhail Gavrilenko , group 23.Б15-ММ

Chief scientist: engineer-researcher of the programming technologies laboratory of infrastructure solutions SPbSU Vasenina A. I.

St. Petersburg
2025

Log-Structured storage

Log-structured storage is a data storage technology based on the journal principle. Instead of writing to random locations, all data is being written sequentially.

- Increases speed of write operation
- Used in the implementation of Snapshot and Copy-On-Write

Existing solutions

- Sprite LFS
 - ▶ A file system, not a block device
 - ▶ Introduced the concept of log-structuring and implemented basic functionality
- Logical Disk
 - ▶ Optimization of file systems at the block level
 - ▶ Created as a tool for small systems
- ZFS:
 - ▶ An advanced file system
 - ▶ High resource usage
 - ▶ Difficult to integrate due to the project's size
 - ▶ No direct low-level access to data

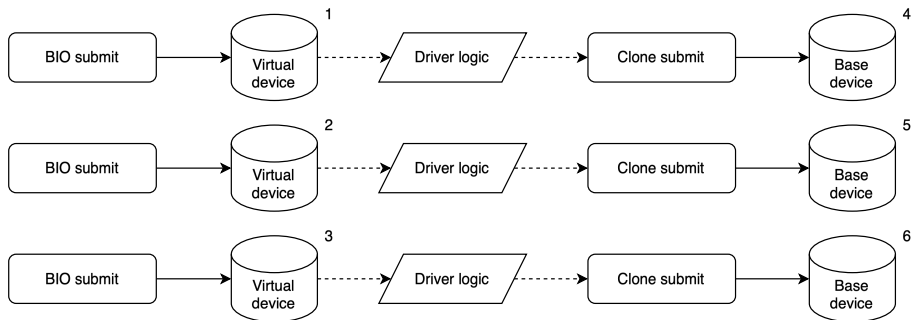
Task formulation

The goal of this work is to study the methods of implementing log-structured addressing in block devices.

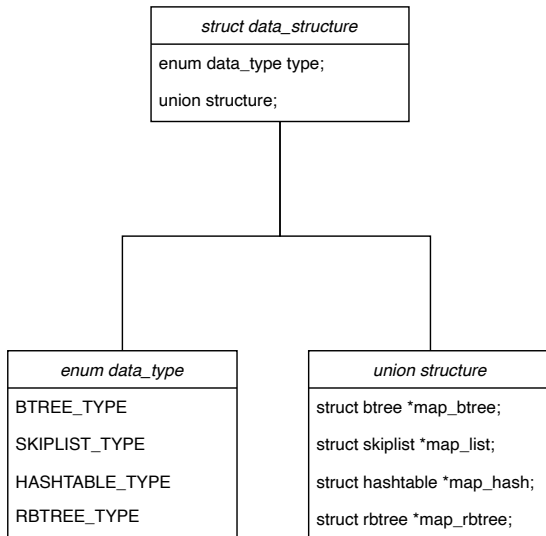
Tasks

- Develop a basic virtual block device with log-structured addressing
- Implement the log-structuring principle using various data structures
- Test the integrity of the virtual block device

Architecture of the Basic Virtual Block Device



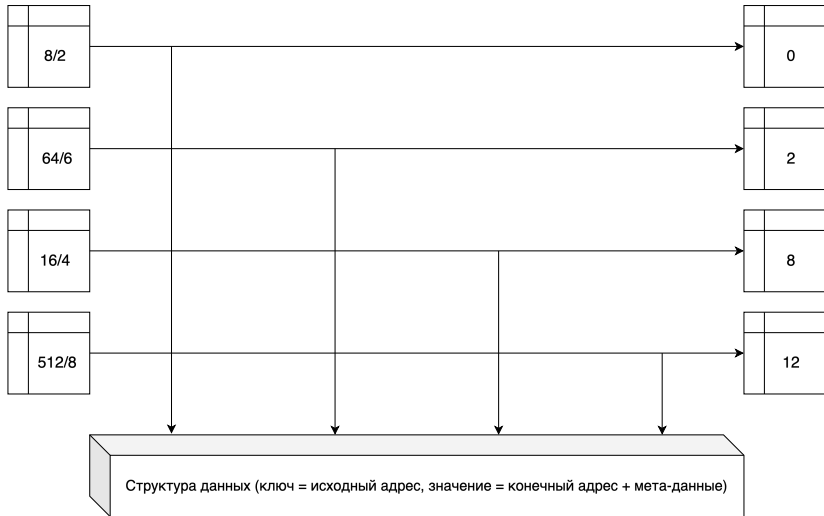
Integration of Various Data Structures



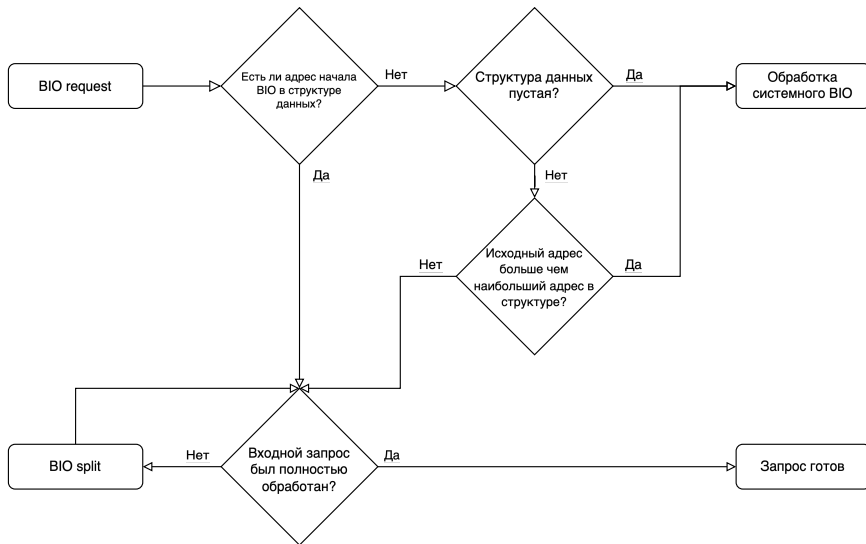
Implementation of Writing to a Block Device

Исходный адрес запроса
/длина запроса

Конечный адрес



Implementation of Reading from a Block Device



Integrity testing

Two tools were chosen for integrity testing:

- The `dd` utility
 - ▶ A basic GNU Linux utility
 - ▶ The `direct` option was used
 - ▶ A system of sequential requests was implemented
- The `fio` utility
 - ▶ A specialized tool for testing block devices
 - ▶ The `direct` option was used
 - ▶ Data verification based on a template was used

Results

- A basic virtual block device with log-structured addressing was developed
- The log-structuring principle was implemented based on various data structures, and a unified interface for interacting with data structures was added
- The integrity of the implemented virtual block device was tested

The implementation is available on GitHub ¹.

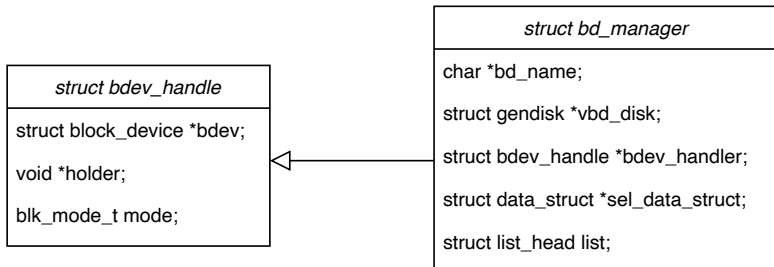
¹Implementation link: github.com/qrutyy/ls-bdd

Plans for the current semester

- Conduct performance testing and benchmarking on a production storage system
- Optimize the block device driver
- Analyze log-structured storage implementations based on different data structures
- Remake data structures from sync type to the concurrent in a lock-free way

Additional slide

Structure of block device descriptions



Additional slide

Test cases

