STARBLAZE

FMS
the Future of Memory and Storage

Distributed Software Architecture in SSD

Zicheng Pan

Contents

01/ Introduction to SSD

02/ SSD platform

03/ File system in SSD

04/ Database in SSD

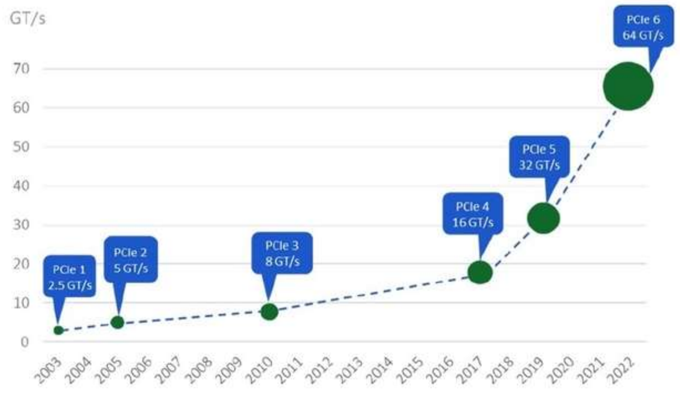
05/ Further more

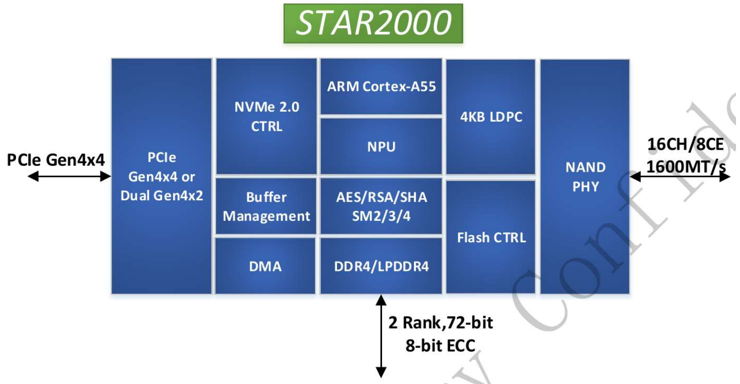
Introduction to SSD

Introduction to SSD

With the increase in PCIe speed, the capabilities of SSD controllers are becoming increasingly powerful:

- Powerful CPUs
- High speed DMA
- DDR with large capacity
- Hardware accelerator such as encryption, buffer management…





STAR2000

PCIe Gen4x4

PCIe Gen4x4 or Dual Gen4x2

NVMe 2.0 CTRL

ARM Cortex-A55

NPU

4KB LDPC

Buffer Management

AES/RSA/SHA SM2/3/4

DMA

DDR4/LPDDR4

Flash CTRL

NAND PHY

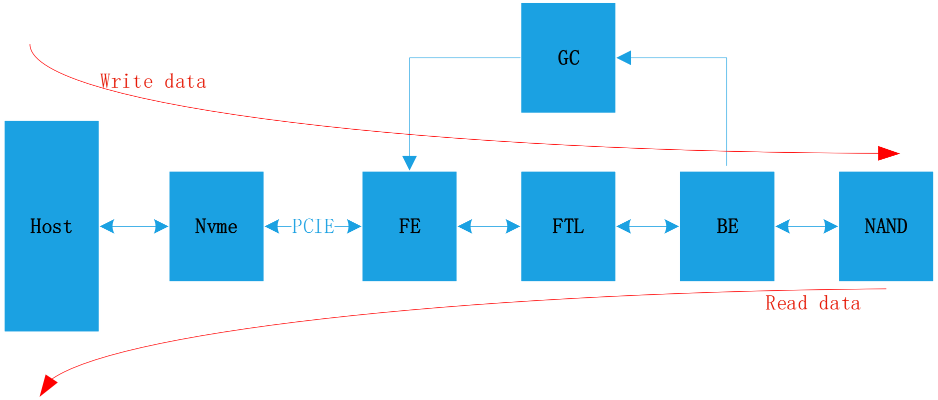
16CH/8CE 1600MT/s

2 Rank,72-bit 8-bit ECC

SSD platform

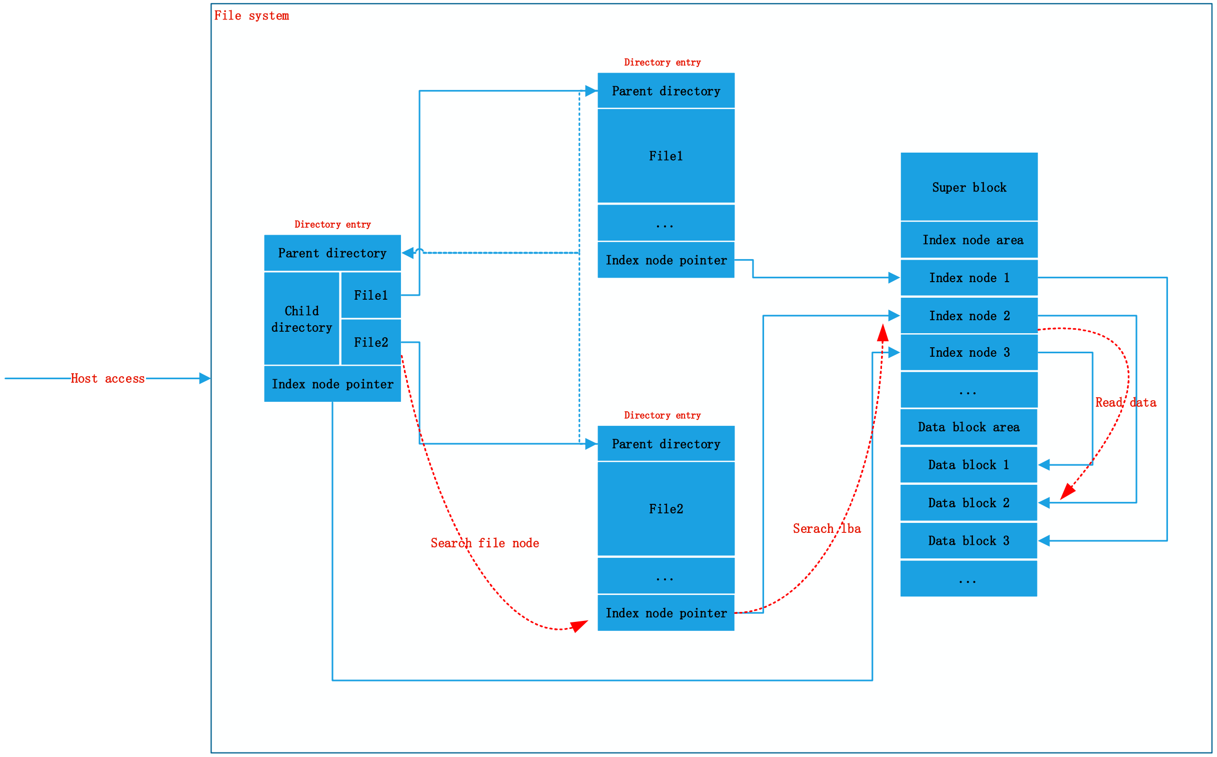
SSDs (Solid State Drives) enhance data storage performance through a combination of hardware and software components：

- Front-End: the interface between the SSD and the host system.
- Flash Translation Layer: handles tasks like wear leveling, bad block management, and logical to physical address translation.
- Back-End: management of data storage and retrieval on the NAND flash memory.
- Garbage Collection: identifies and merges fragmented data, freeing up space in the SSD by erasing blocks that contain both valid and deleted data



Distributed file system

File system overview



File system

Host access

Directory entry

Parent directory

Child directory

File1

File2

Index node pointer

Directory entry

Parent directory

File1

...

Index node pointer

Directory entry

Parent directory

File2

...

Index node pointer

Search file node

Serach lba

Super block

Index node area

Index node 1

Index node 2

Index node 3

...

Data block area

Data block 1

Data block 2

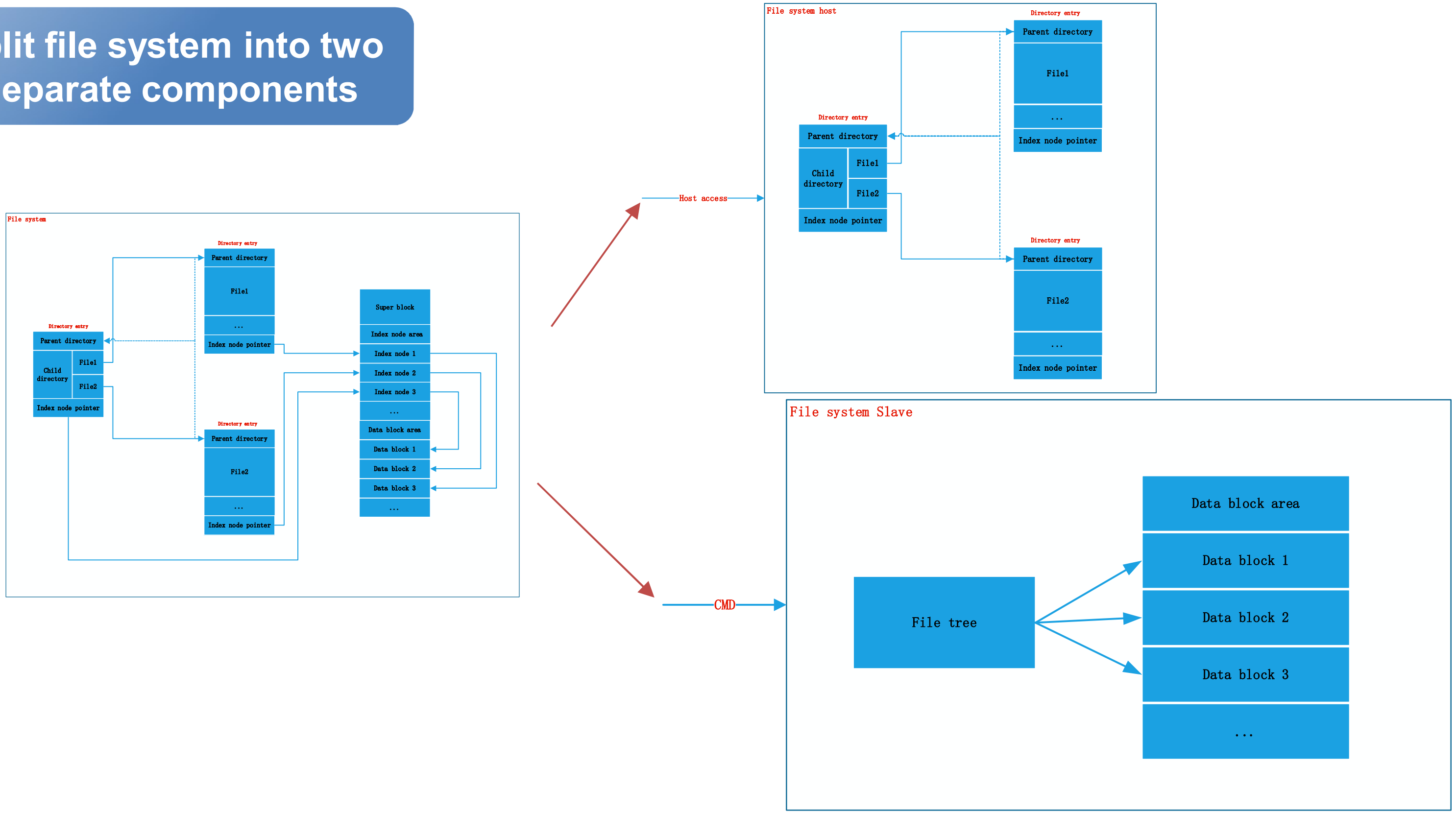
Data block 3

...

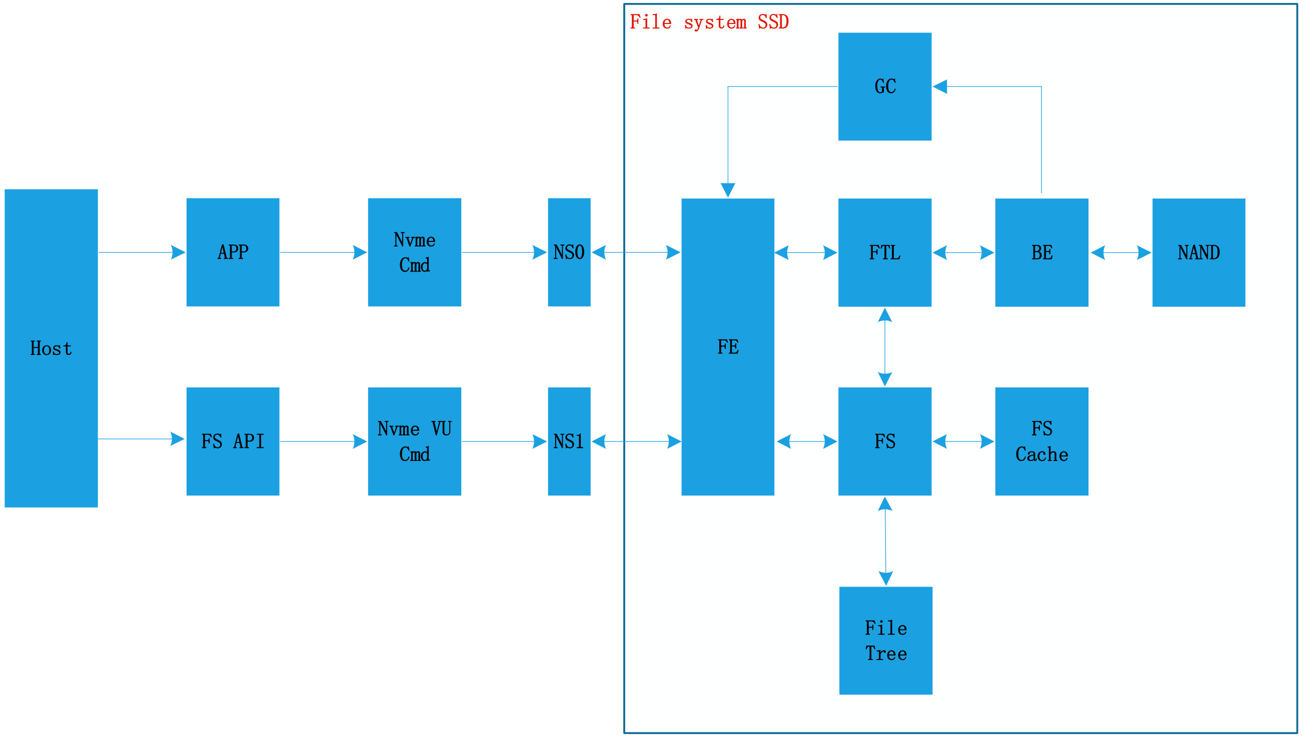
Read data

Distributed file system

Split file system into two separate components



File system in SSD – Dual namespace

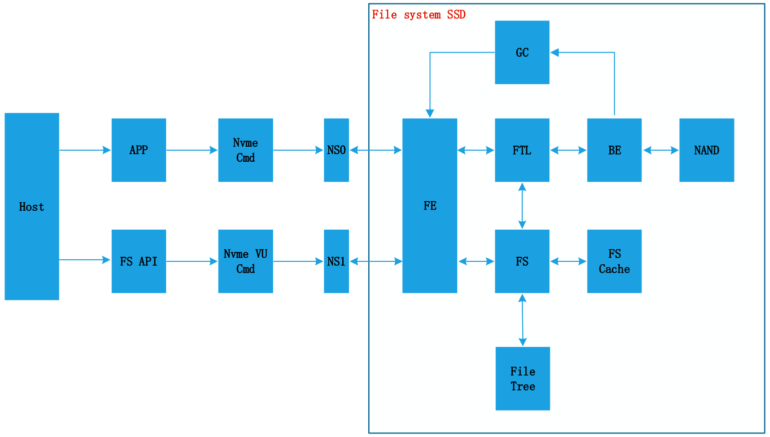


File system in SSD - Advantages

the Future of Memory and Storage

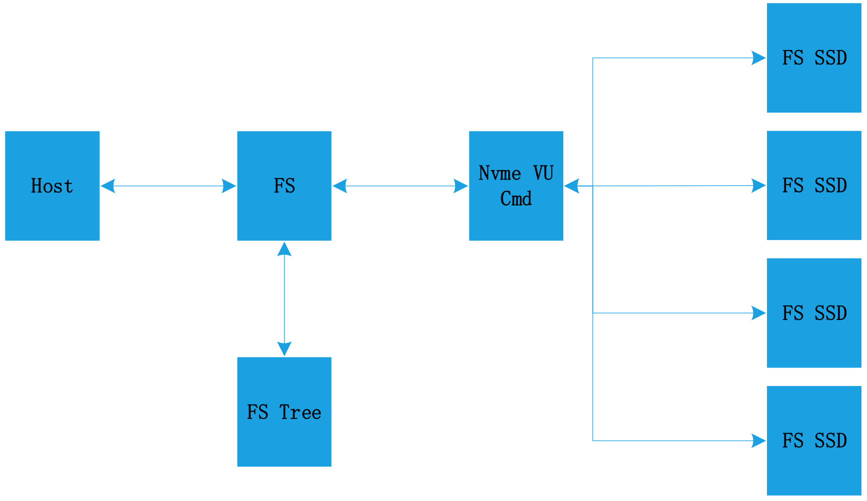
Improvement:

- The file system operates within an SSD, reducing the CPU time cost on the host.
- Data moved by high speed DMA instead of CPU
- Synchronize the management of file system and SSD cache space, reducing redundant computations and memory overhead.



File system in SSD – Distributed Storage

In a host system, running a simple management system, we can use a single file system to manage multiple Flash Storage Drives (SSDs). This not only improves performance but also makes it more convenient to implement protective measures such as RAID.



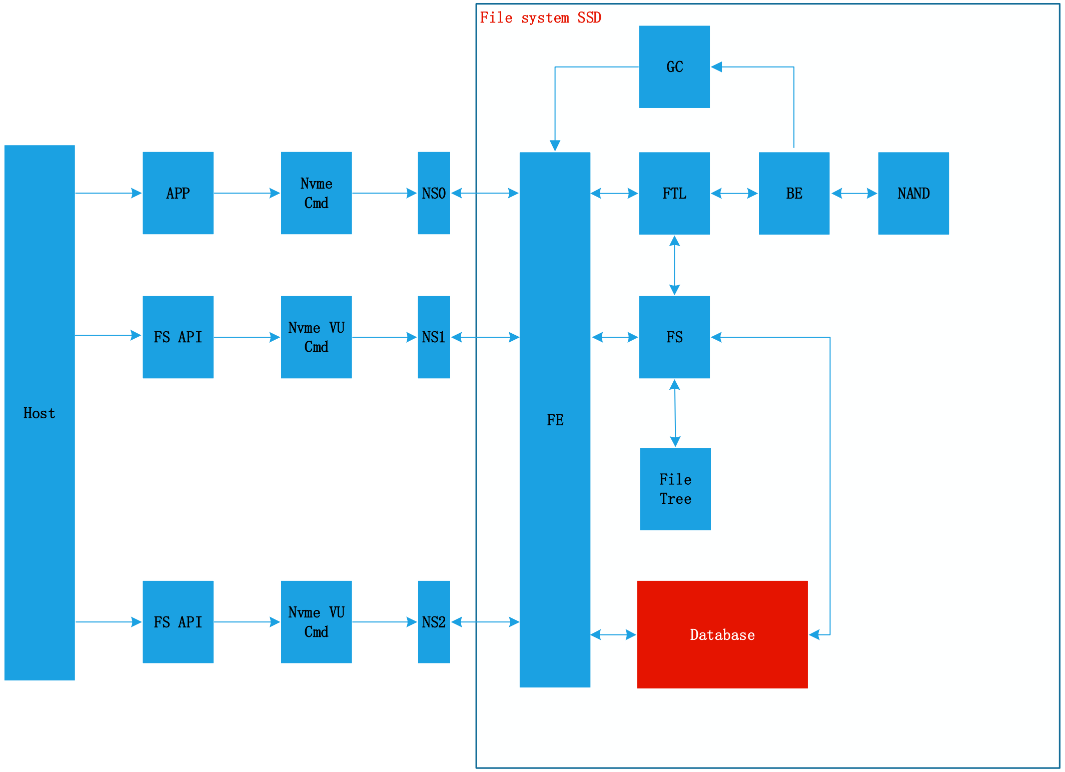
File system in SSD - Performance

The total time and memory overheads are difficult to estimate. We only provide a comparison of the space overhead and the tasks for the two schemes here.

	Ordinary file system	**File system SSD**
Buffer alloc per LBA	Three times	Two times （-33%）
Data copy per LBA	Two times	Once （-50%）
Host cpu task	File search, Entry load, Nvme cmd…	Nvme vu cmd

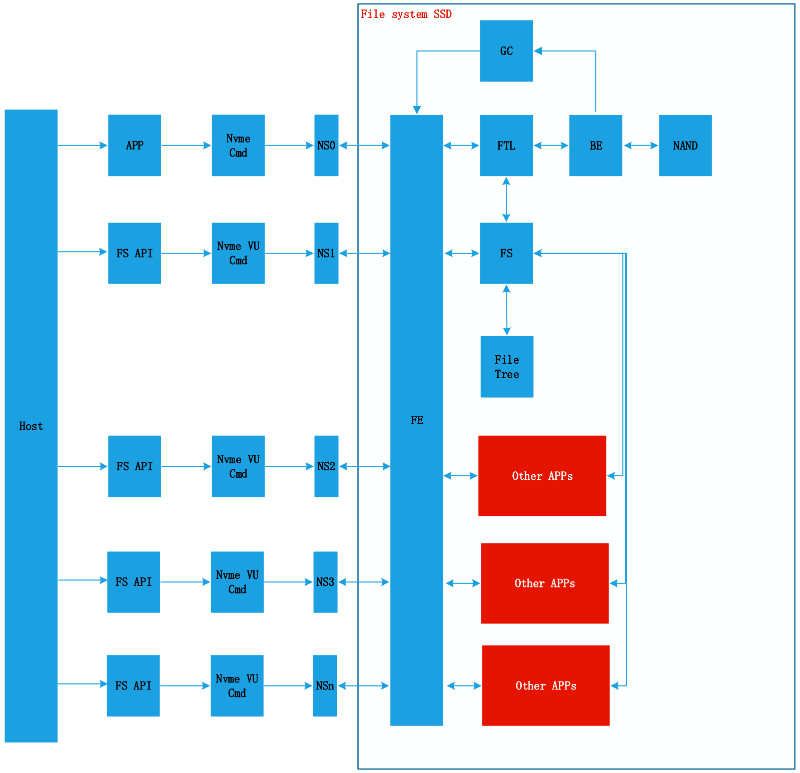
What's more?
Data copy is the most heavy work for cpu while it can be done by DMA in File system SSD.

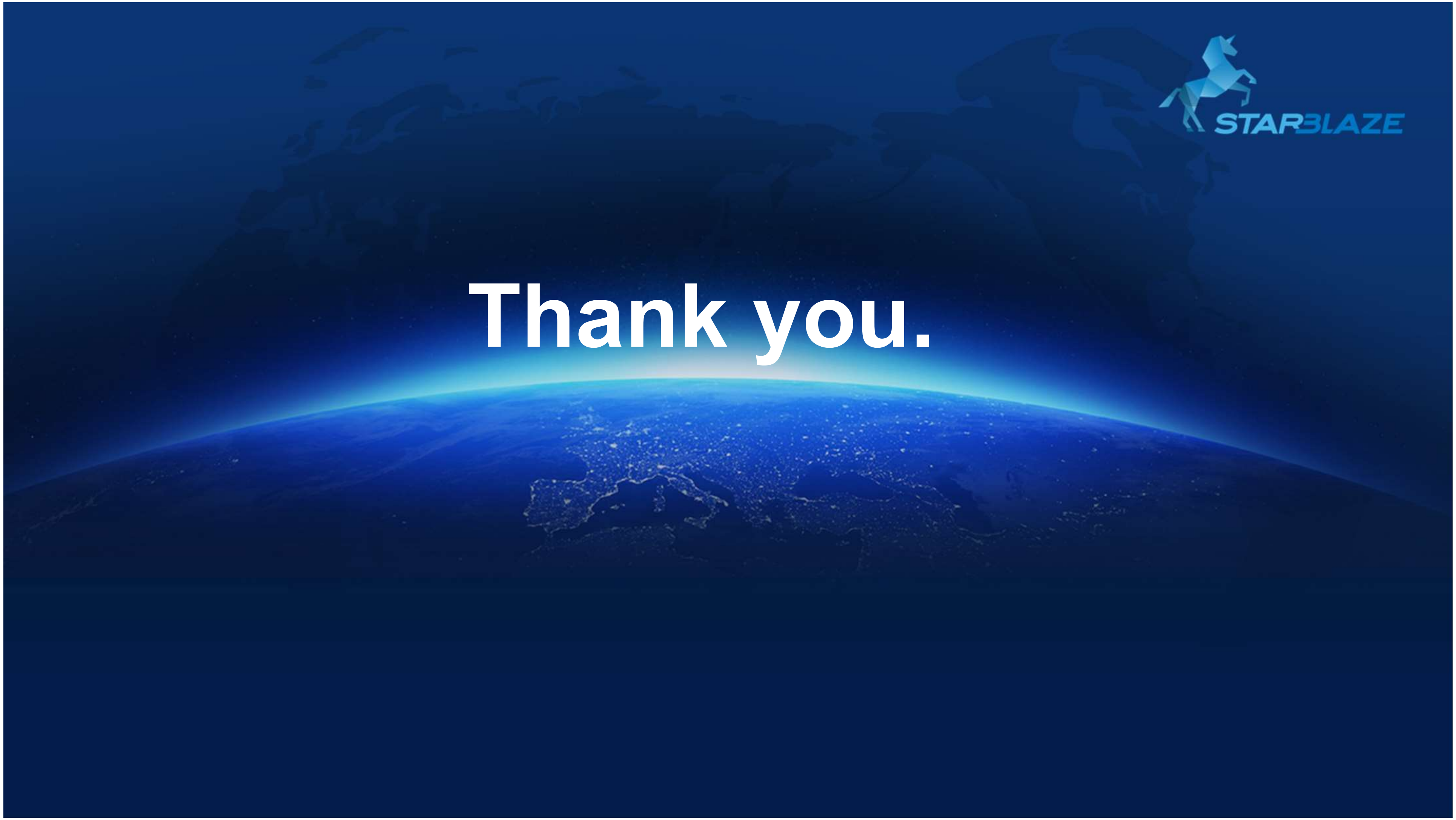
Further more – DB SSD



The host accesses the database via NS2, allowing the database to interact directly with the file system on the disk. This direct interaction reduces read and write latency.

Further more – Computational SSD





Thank you.