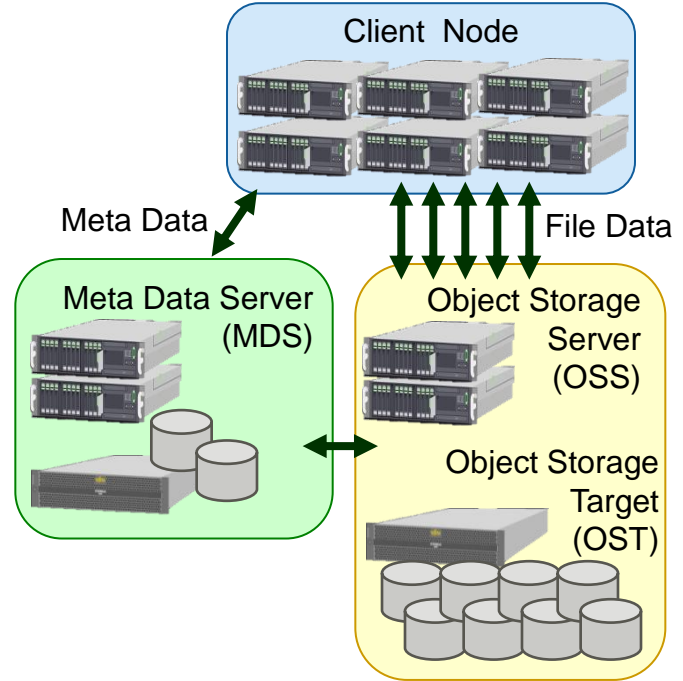


FEFS: Scalable Cluster File System

Features of FEFS

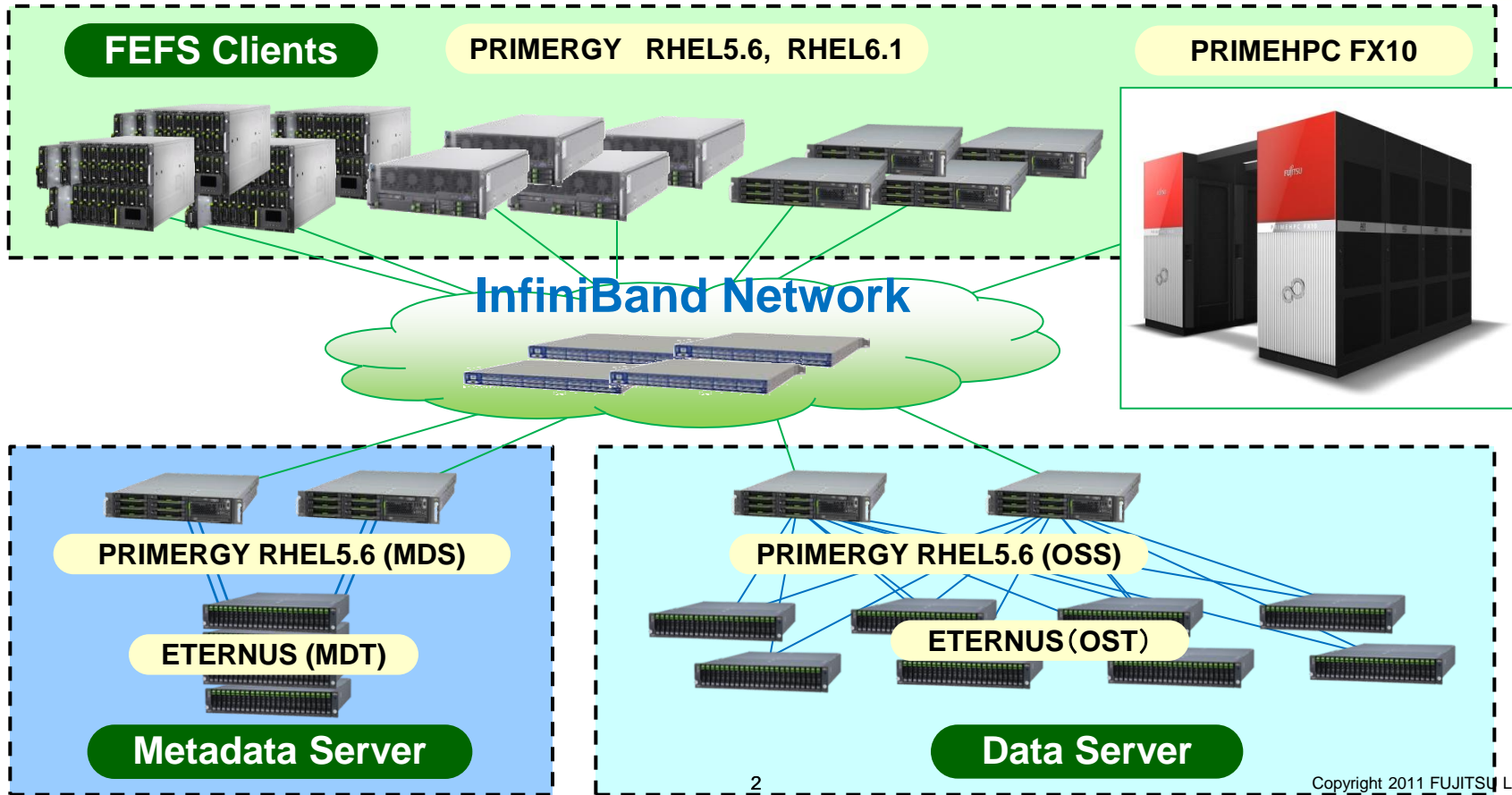
FEFS[†] is scalable parallel file system based on Lustre.

- High Performance & High Scalability
 - Scalable I/O performance (~1TB/s) & capacity (~8EB).
- I/O Bandwidth Guarantee
 - Fair share and Best effort QoS.
- High Reliability & High Availability
 - Failover with redundant hardware and continuing file system service.



† FEFS: Fujitsu Exabyte File System

System Configuration



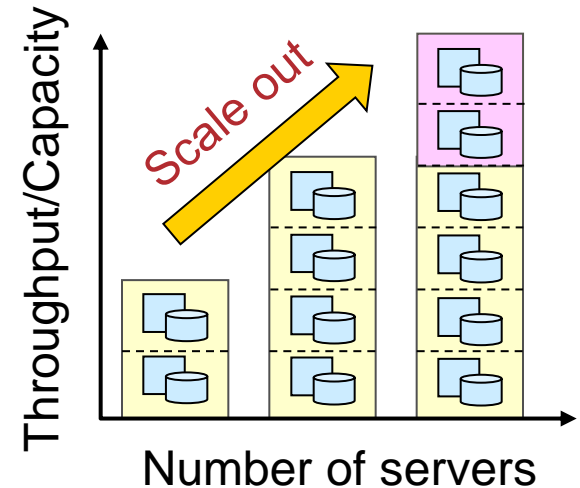
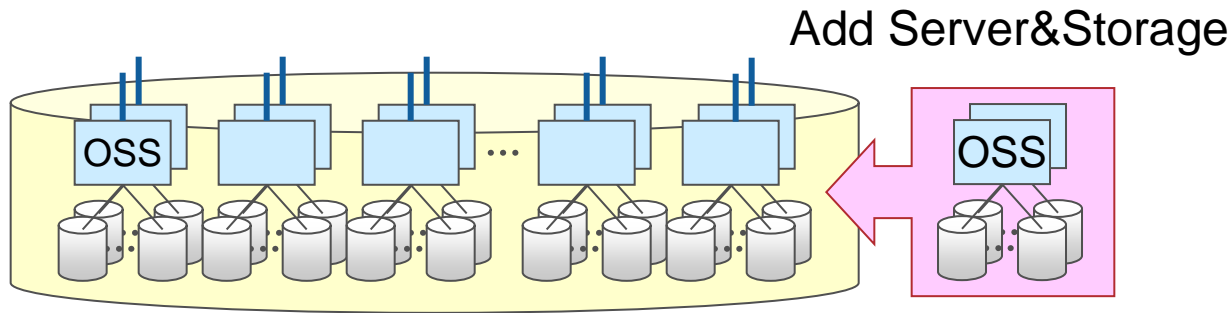
Specification of FEFS

- Fujitsu expand system limits and add new functions to Lustre.

Item	FEFS	Lustre
Max. file system size	8 EB	64 PB
Max. file size	8 EB	320 TB
Max. number of files	8×10^{18} files	4×10^9 files
Max. number of OST	20×10^3	8,150
Max. OST size	1 PB	16 TB
Max. number of clients	1×10^6 clients	128×10^3 clients
Max. block size	512 KB	4 KB
Max. number of stripes	20×10^3 stripes	160 stripes
QoS (Fair share/Best effort)	Yes	No
Directory Quota	Yes	No
InfiniBand Multi-rail	Yes	No

High Performance & High Scalability

- Achieved high-scalable I/O performance with multiple server.
 - Scale out throughput & capacity by adding server & storage.

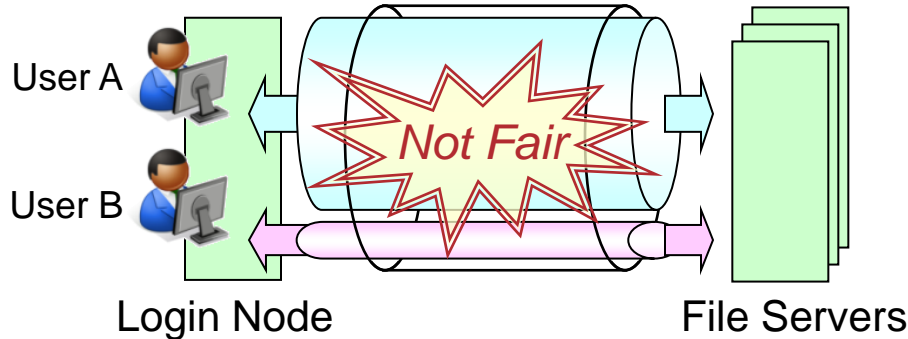


I/O Bandwidth Guarantee: Fair Share QoS

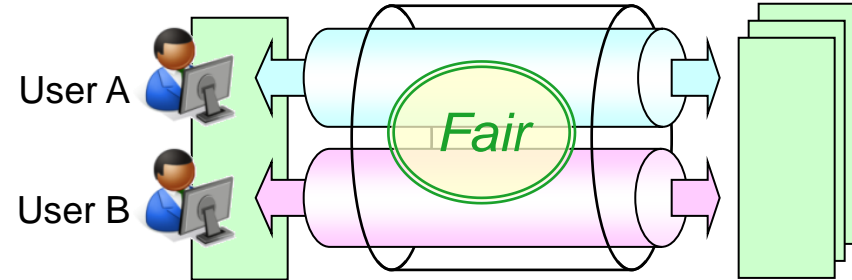
- Sharing IO bandwidth with all users.
 - Prevent slowdown from huge I/O from a user.
 - Prevent variability in job execution time.

Fujitsu extended function

Without Fair Share QoS



With Fair Share QoS

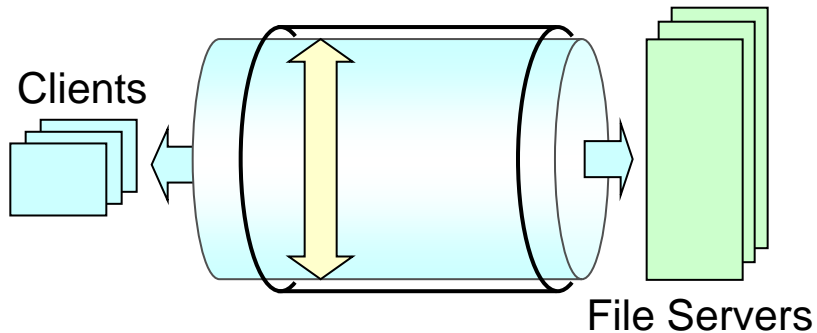


I/O Bandwidth Guarantee: Best Effort QoS

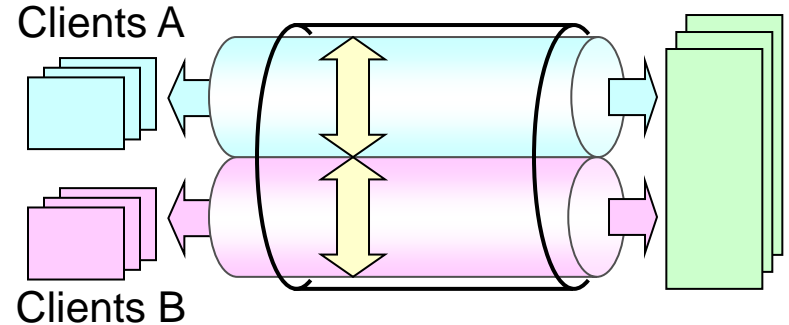
- Utilize all I/O bandwidth exhaustively.

Fujitsu extended function

Occupied by one client



Shared by all clients



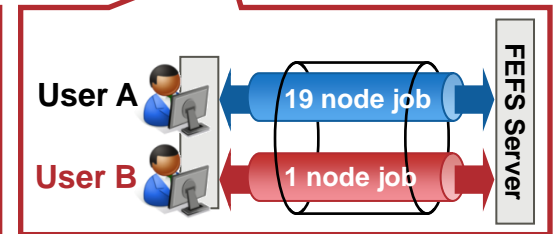
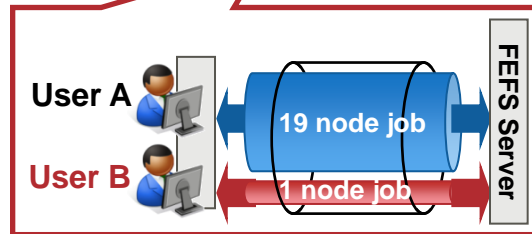
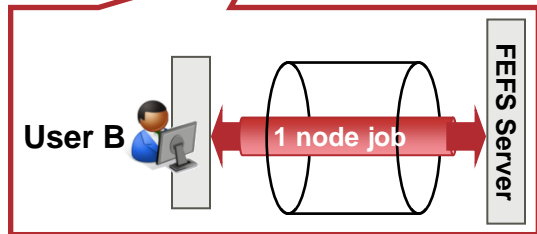
Effectiveness of Fair Share QoS

■ Sample Case:

- User A 19 node job
- User B 1 node job ⇒ Creation and removal time of 10,000 files.

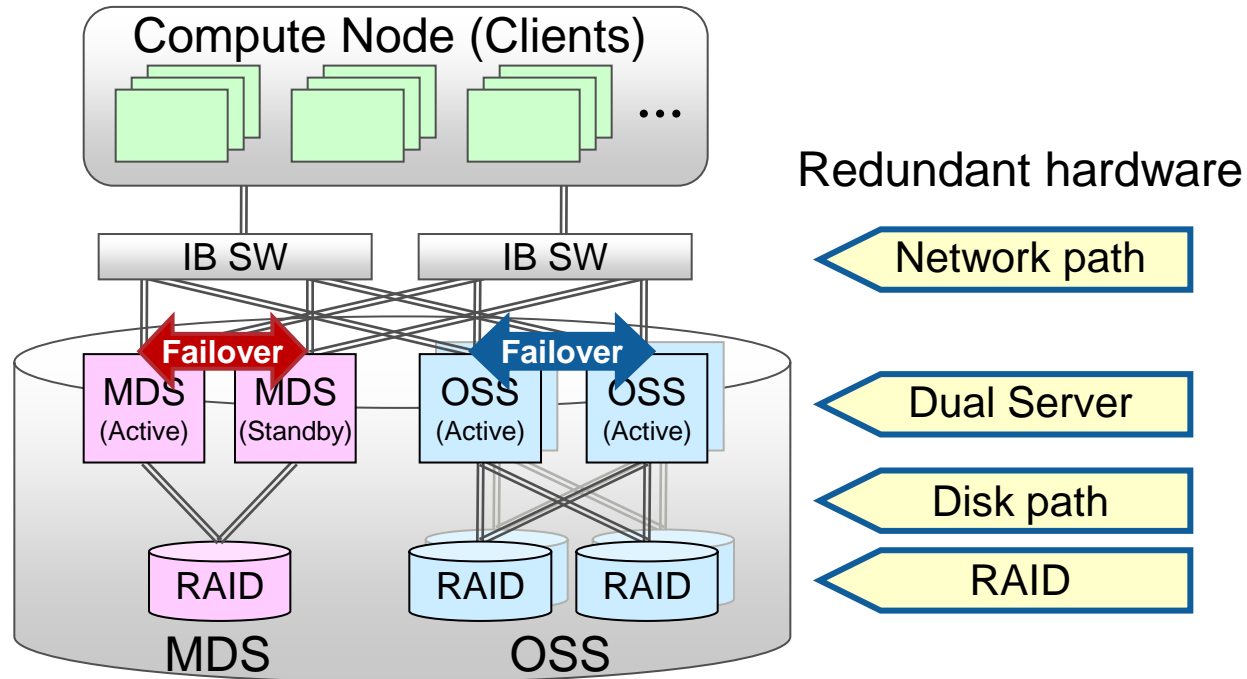
User B's processing time

User B 10,000 files	Without fair share Single user	Without fair share Multiple users	With fair share Multiple users
Create files	4.1 sec	10.1 sec	3.9 sec
Remove files	4.2 sec	14.0 sec	5.5 sec



High Reliability and High Availability

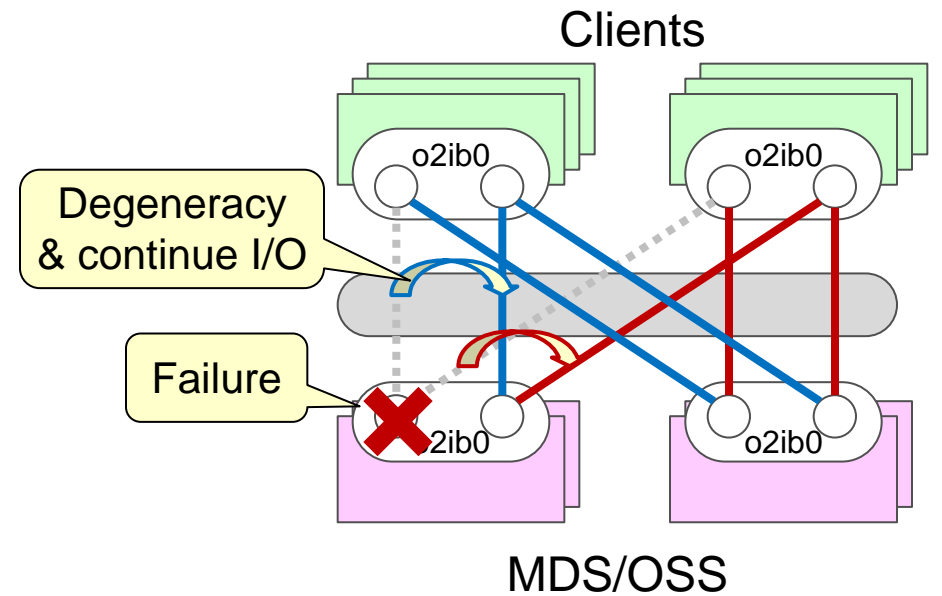
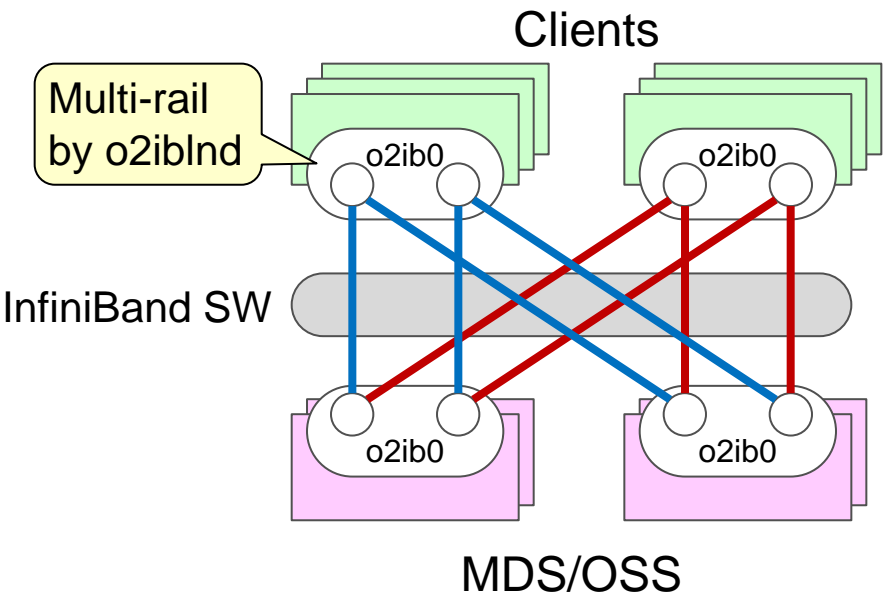
- Avoid out of service time caused by a single point of failure with redundant hardware and failover mechanism.



High Availability: InfiniBand Multi-rail

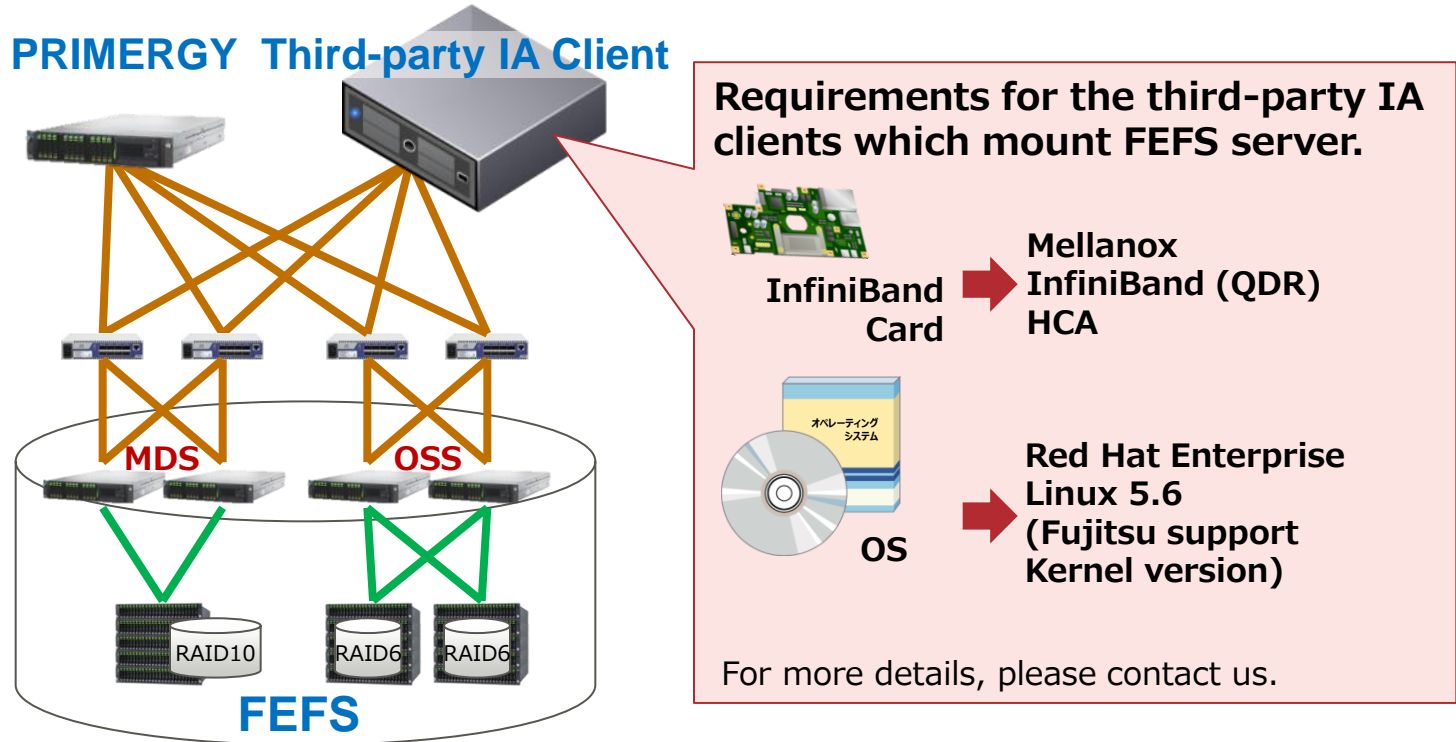
- Continue communication when single point of IB failure occurs.
 - All IB connections are used by round-robin order by each requests.

Fujitsu extended function




Third-party Clients Connectivity

- FEFS can be mounted on the third-party IA clients



■ Fujitsu will

- work with Lustre community,
- and merge our Lustre enhancements into the future version of Lustre 2.x community release.



FUJITSU

shaping tomorrow with you