Simulation of the cache hit rate for data readout at the Tokyo Tier-2 center

Tomoe Kishimoto, Junichi Tanaka, Tetsuro Mashimo, Michiru Kaneda, Nagataka Matsui

ICEPP, The University of Tokyo

Mar. 15 2019

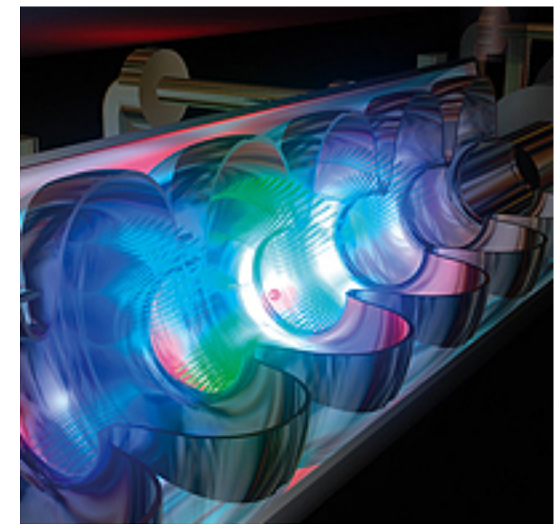
International Center for Elementary Particle Physics

- ✓ Main research projects in ICEPP:

東京大学
素粒子物理国際研究センター
International Center for Elementary Particle Physics
The University of Tokyo

ATLAS experiment at LHC

MEG experiment at PSI ($\mu \to e\gamma$ rare decay)



R&D for ILC

ATLAS Japan

- ✓ ATLAS–Japan group
  - 17 institutes and ~150 members (including students)

ICEPP regional analysis center

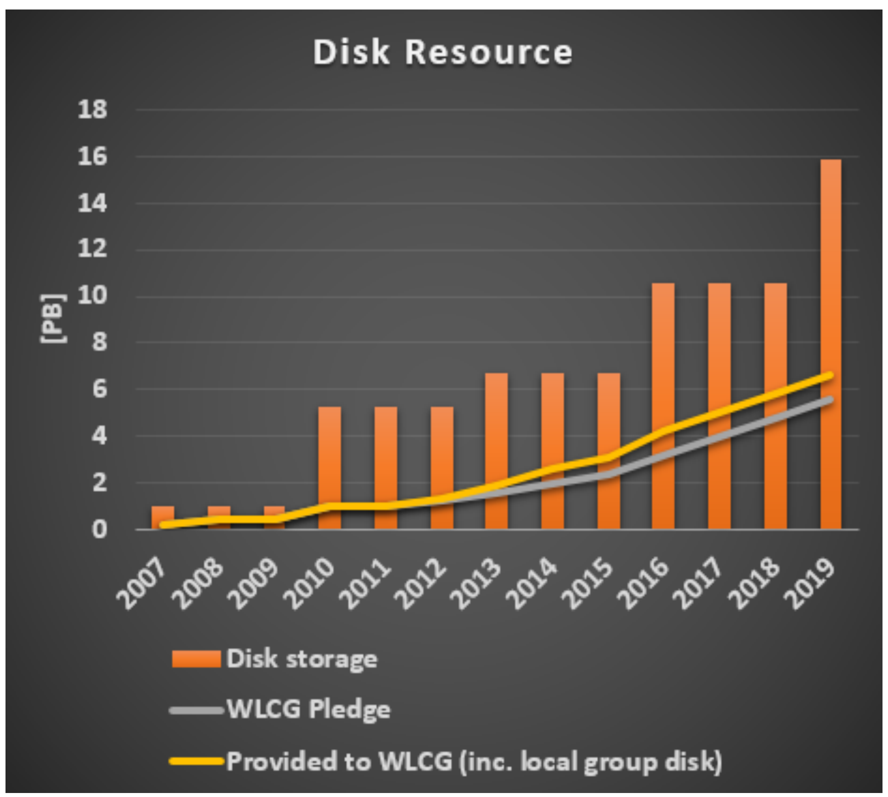
- ✓ ICEPP has provided computing resources to WLCG and ATLAS-Japan group as a representative of ATLAS-Japan group
  - No other WLCG site in ATLAS-Japan
  - "Regional analysis center" at ICEPP for the ATLAS experiment has been in operation since 2007

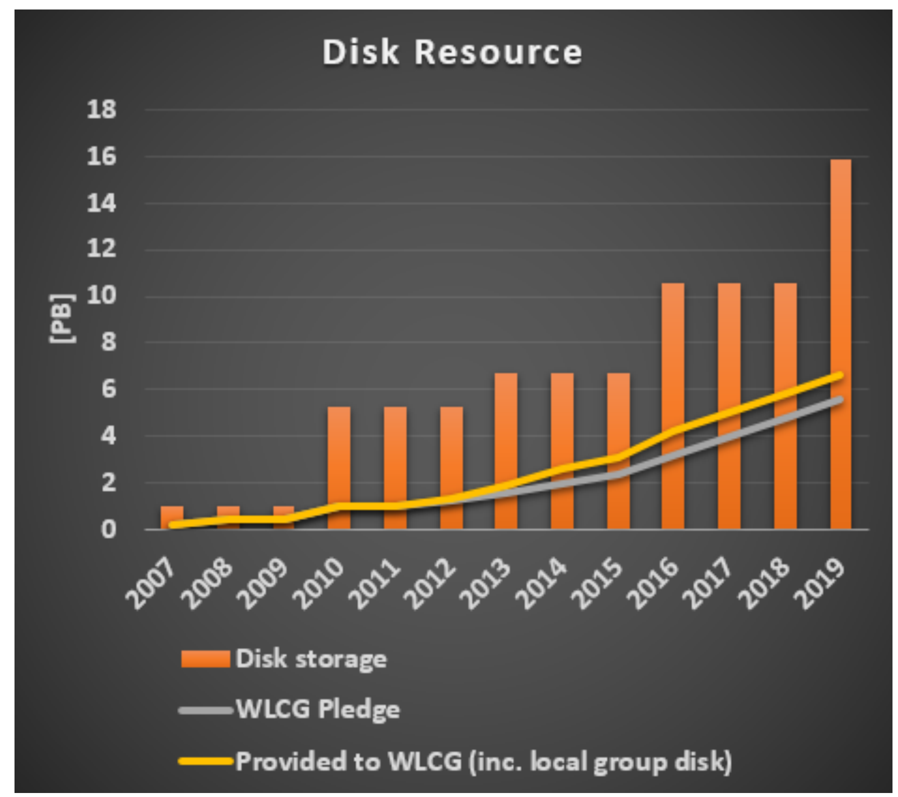
- ✓ Regional analysis center is operated for two kinds of groups:
  - WLCG (~80% of total resources) ← **Main focus of this presentation (Tokyo Tier2)**
    - Only ATLAS VO is supported
  - Non-WLCG (~20% of total resources)
    - Computing resources are provided to ATLAS-Japan members
    - They can login to interactive nodes and use batch system

System overview

✓ Hardwares in the center

- Supplied by a lease and replaced in every 3 years
  - Architecture in the system is uniform to reduce operation cost
  - The system can be kept as powerful as possible to satisfy the WLCG pledge and ATLAS-Japan requirement
- Electrical and cooling systems stay the same



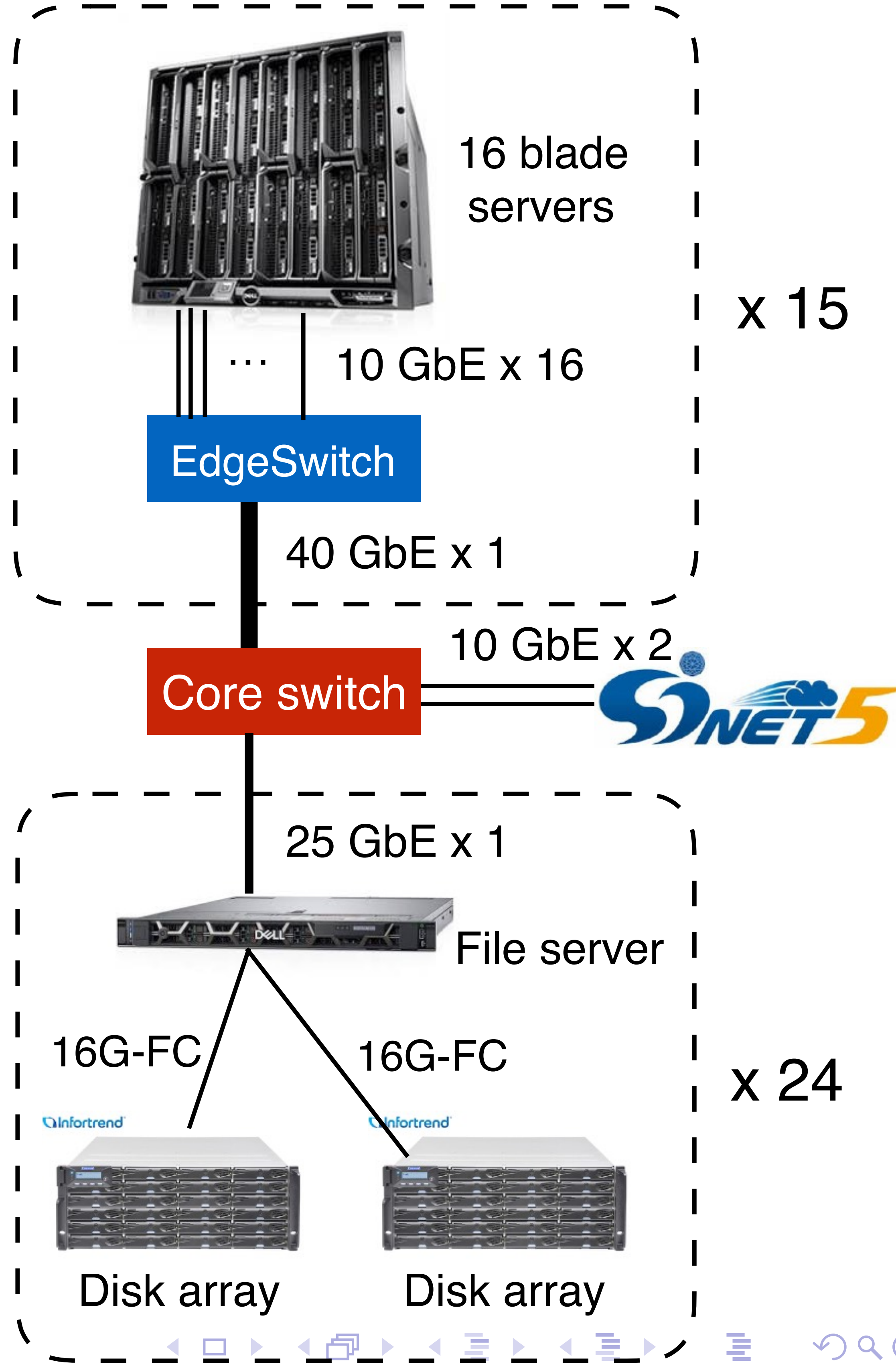


4th system (2016-2018) → ~10k CPU cores and ~10 PB disk
5th system (2019-2021) → ~10k CPU cores and ~15 PB disk

Configuration of Tokyo Tier2 (5th system)

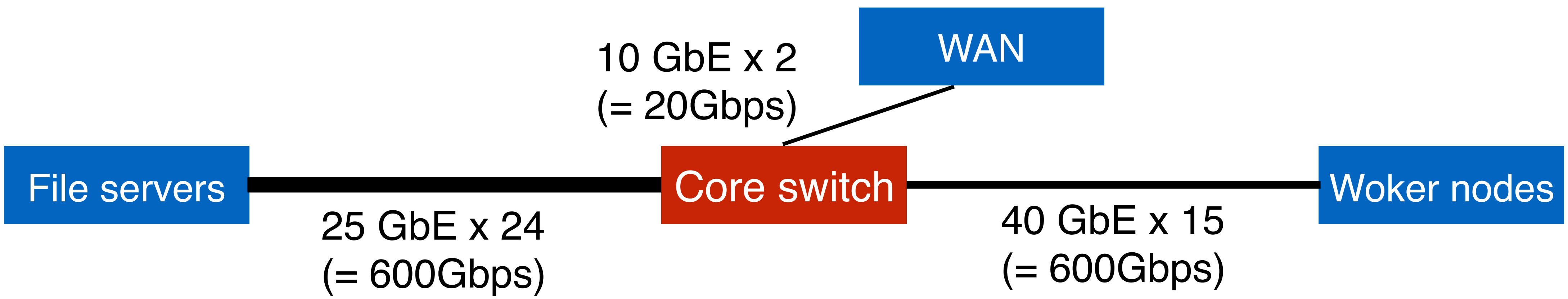
✓ Worker nodes (WN):
- Dell PowerEdge M640 (blade server) x 240
- CPU: Intel Xenon Gold 6130,
  32 CPU cores/ node
  18.97 HS06 / core
  (Hyper Threading is not enabled)
- Memory: 96 GB / node
- Disk: 1 TB HDD (RAID1)
- Managed by ARC+HTCondor

✓ File servers (FS):
- Dell PowerEdge R640 (1U server) x 24
- 2 Disk arrays = 2 x 220 TB (RAID6)
- Managed by DPM

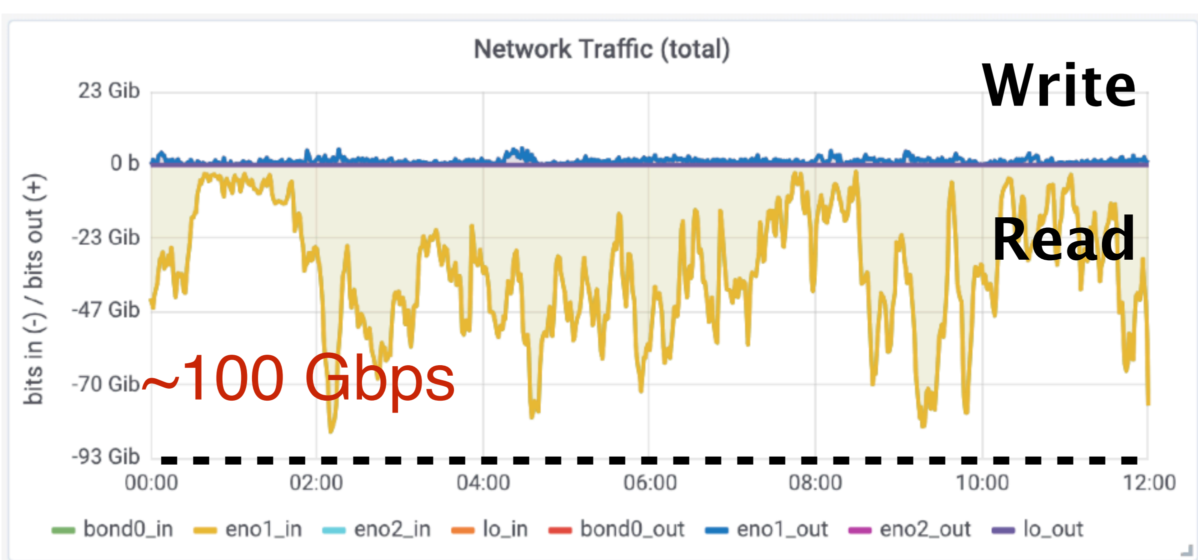


Data throughput among subsystems

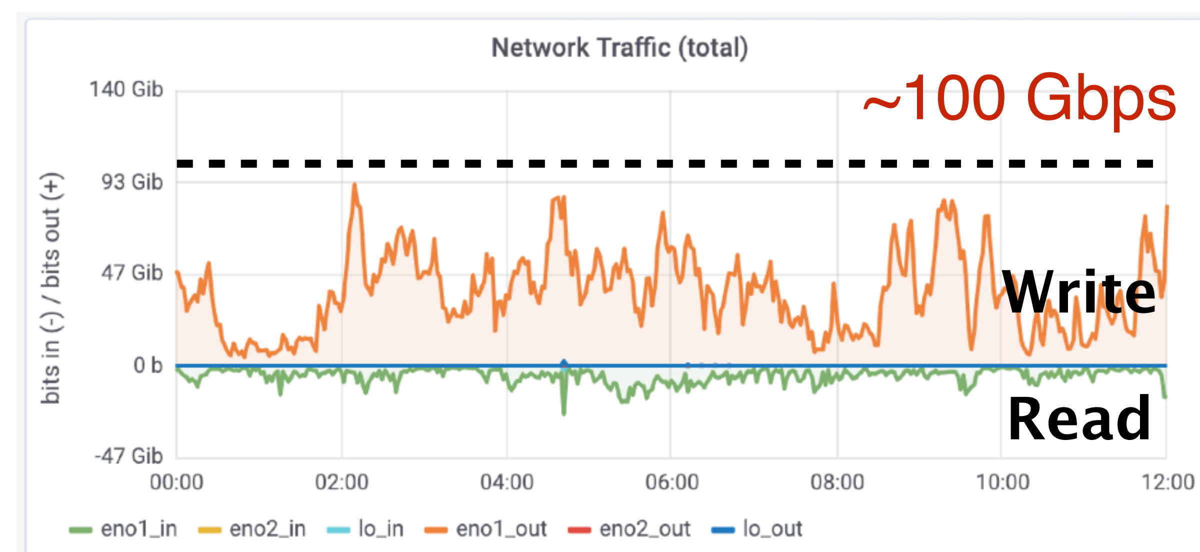
✓ Typical ATLAS Grid jobs at the Tokyo Tier-2 copy data from the file servers to its worker nodes (stage-in)



Network traffic of file servers



Network traffic of worker nodes



– Local network bandwidth is enough so far (< 15% utilization)

Performances of disk I/O

One file server | One worker node

Throughput

%utilization
(=IOPS x service time)

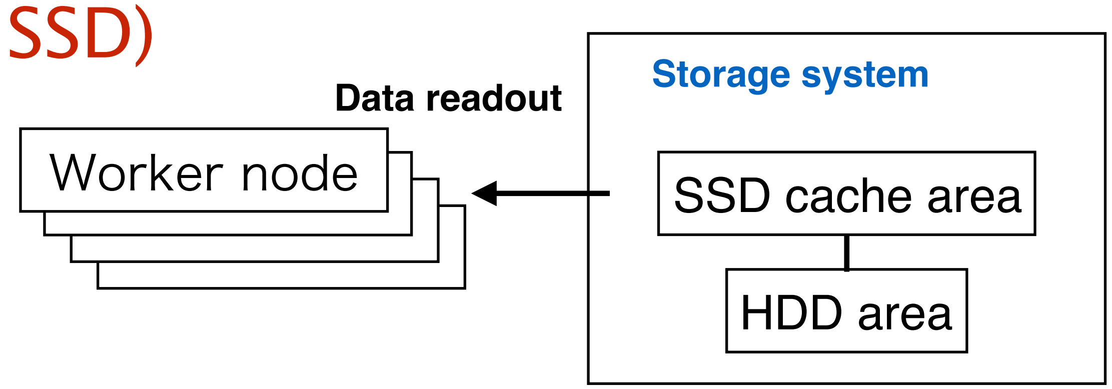
✓ % utilization of disk I/O (FSs and WNs) reaches > 75 %

Examination of cache area

- ✓ Several improvements of the disk I/O performance are under consideration for worker nodes:
  - Change the local disk configuration from RAID1 to RAID0
  - Use direct I/O (XRootD) to the file servers

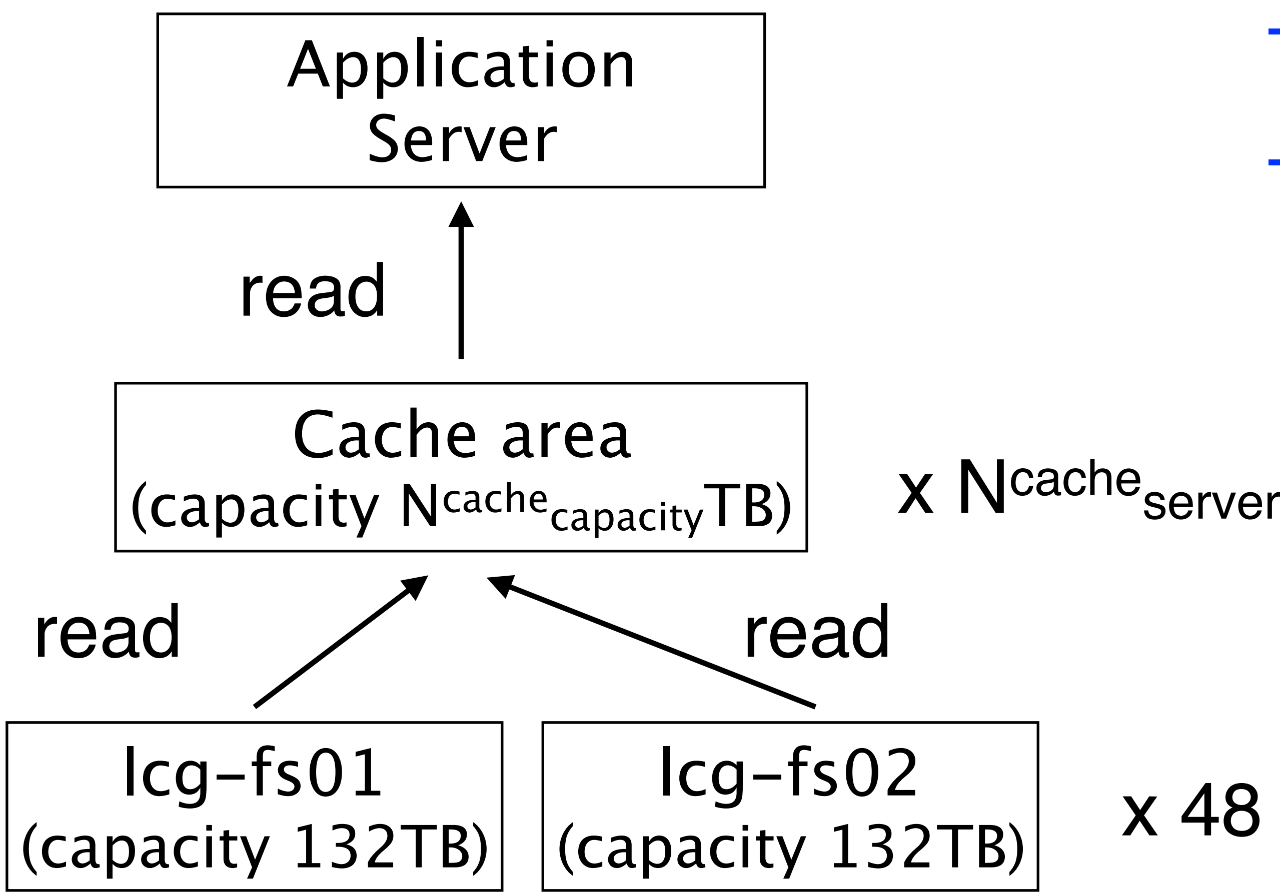
- ✓ Also, improvements of the performance and scalability of the file servers are important:
  - The number of CPU cores in the system will increase, and additional resources (cloud, HPC) are planned to integrate
  - It is difficult to simply scale the number of disk arrays due to the limited area of the computing room
  - Studying a possibility to construct a cache area to improve the performances using fast devices (e.g. SSD)

→ **Whether the cache works effectively depends on the data access pattern**



Cache simulation

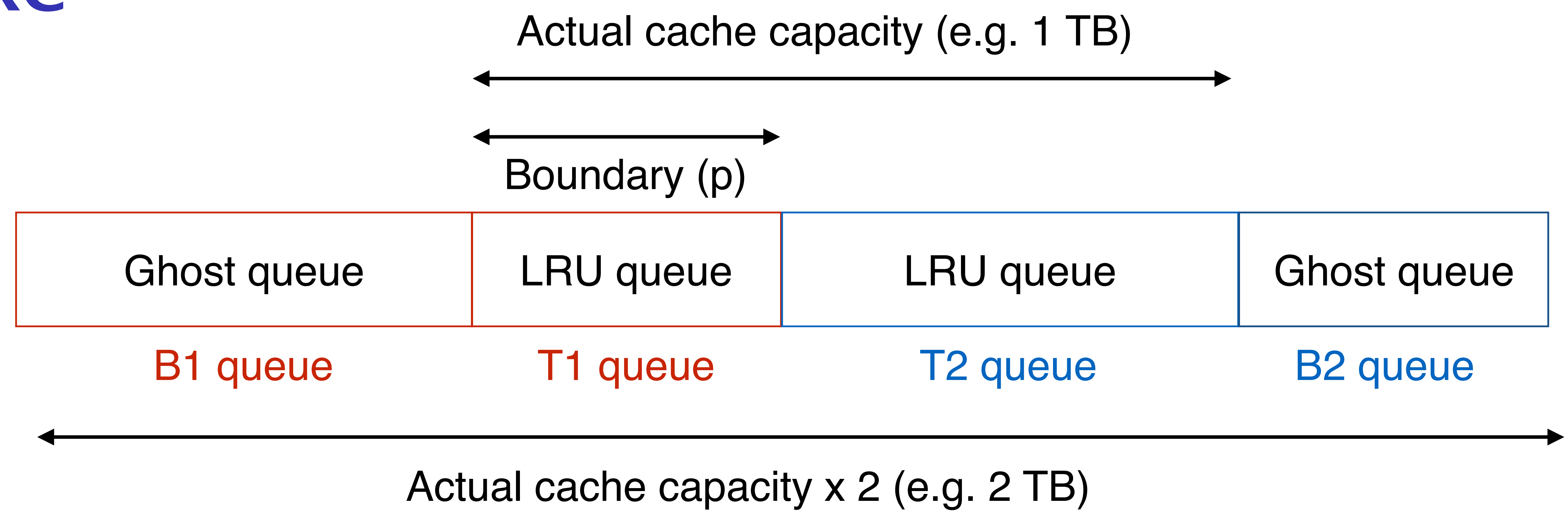
✓ Simulate the cache behavior using data transfer logs for 2017 - 2018 (4th system) at Tokyo Tier-2



*** There were 48 file servers (lcg-fs{01..48}) in the 4th system**
*** Each filer server had 132 TB capacity**

- Only read direction is considered
- Parameters of the simulation:
  - $N^{cache}_{capacity}$: Cache capacity
  - N^{cache}_{server}: # of cache areas (servers)
    (Each file server has cache area if N^{cache}_{server}=48)
  - **Type of cache algorithm**:
    - Least Recently Used (LRU)
      (Deletes the least recently used items first)
    - Least Frequently Used (LFU)
      (Delete the least often used items first)
    - Adaptive Replacement Cache (ARC)

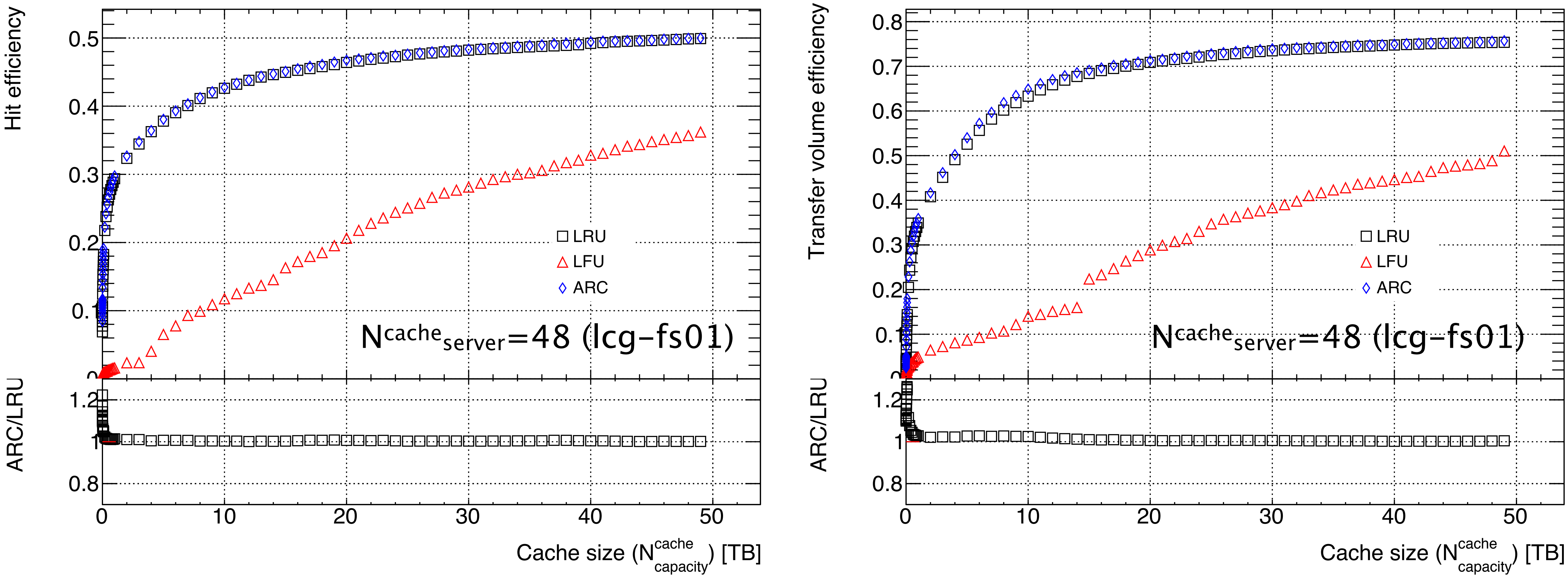
ARC



- ✓ T1, B1 queues: Contain files that have been accessed only once recently
  - Files evicted from T1 enter B1. B1 has only meta data for keeping track
- ✓ T2, B2 queues: Contain files that have been accessed at least 2 times
  - Files evicted from T2 enter B2. B2 has only meta data for keeping track
- ✓ For example:
  - Hits in B1 queue will increase the size of T1 (p), hits in B2 queue will decrease p
    → The algorithm continuously revises how to invest in T1 and T2 according to the access pattern.

Cache hit rate

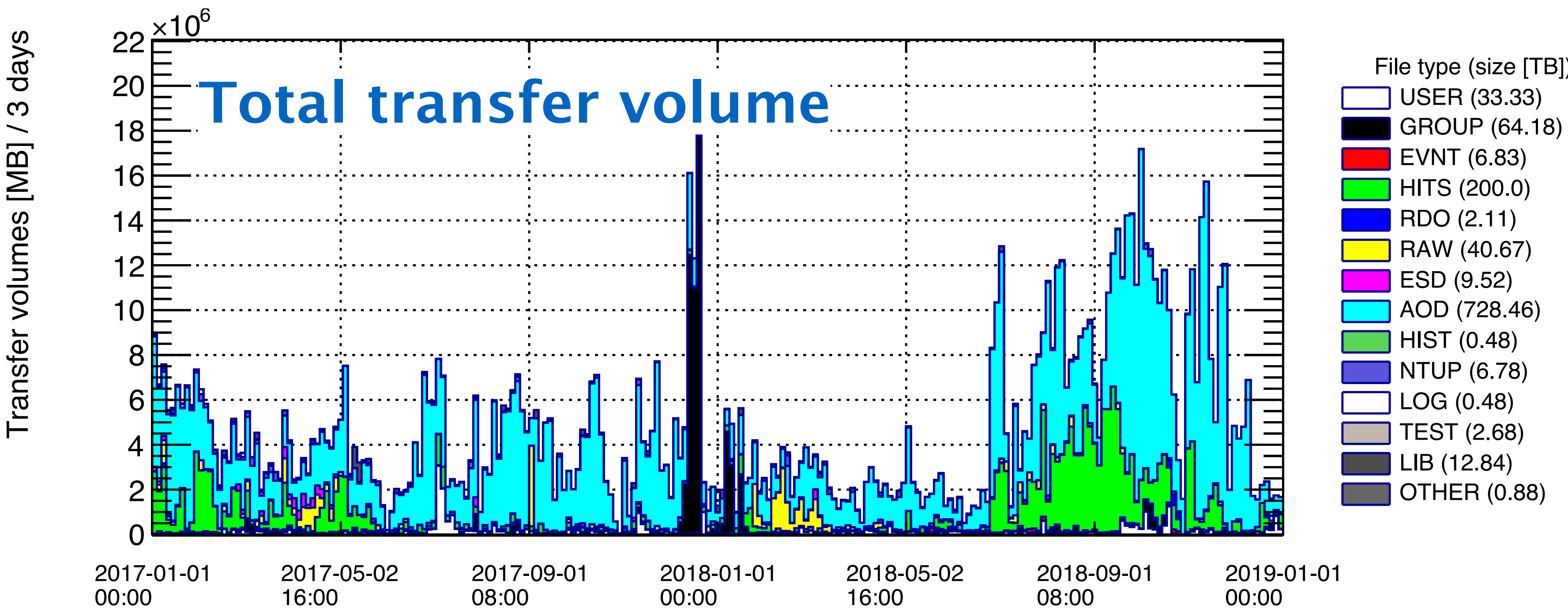
Hit efficiency = (# of data access from cache)/(Total # of data accesses)
Transfer efficiency = (Transfer volume from cache)/(Total transfer volume)

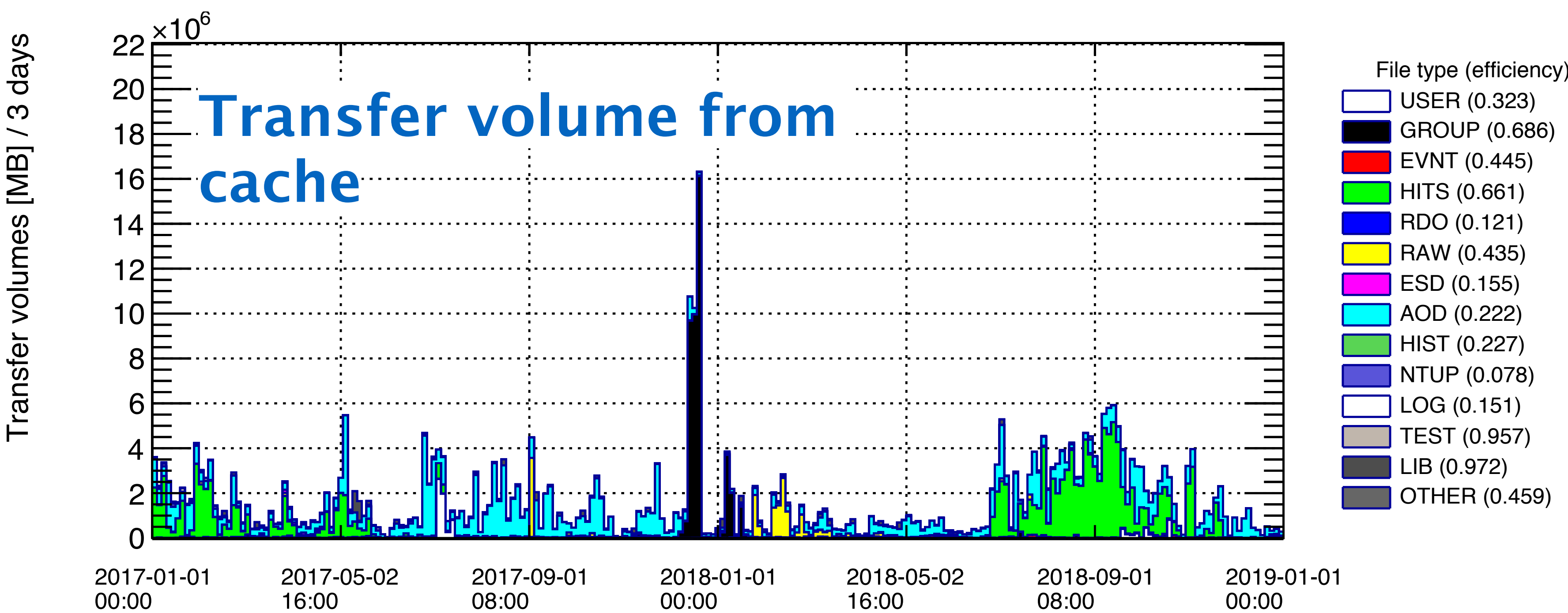


✓ LRU and ARC show similar efficiencies

- If $N^{cache}_{capacity}$ = 10 TB (≒ 10% of total capacity), almost maximum efficiencies will be archived
- If $N^{cache}_{capacity}$ =1 TB (≒ 1% of total capacity), the cache hit eff. will be ~30%
- (If the cache capacity is very small (< 0.1%), ARC gains +20% efficiency compared to LRU)

Characteristic of cached files



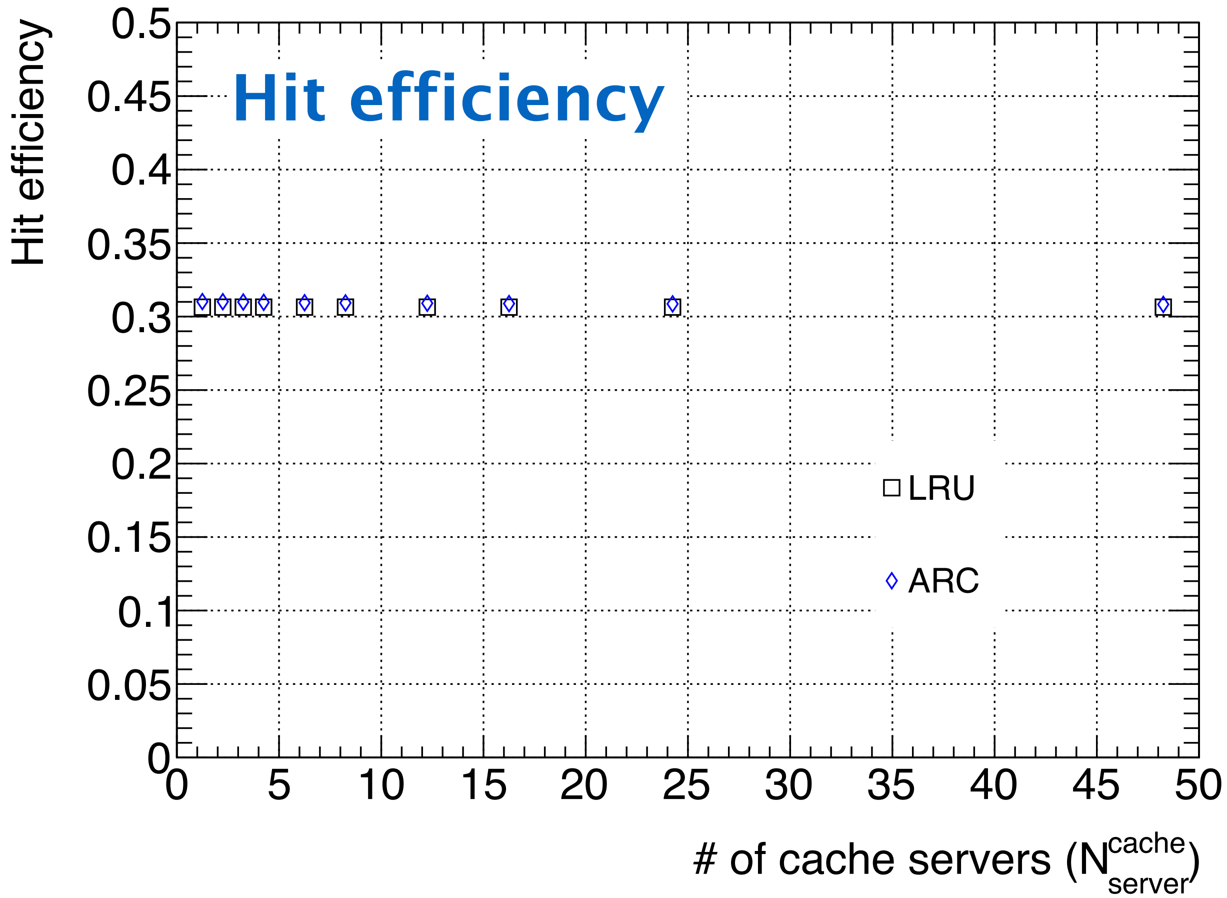


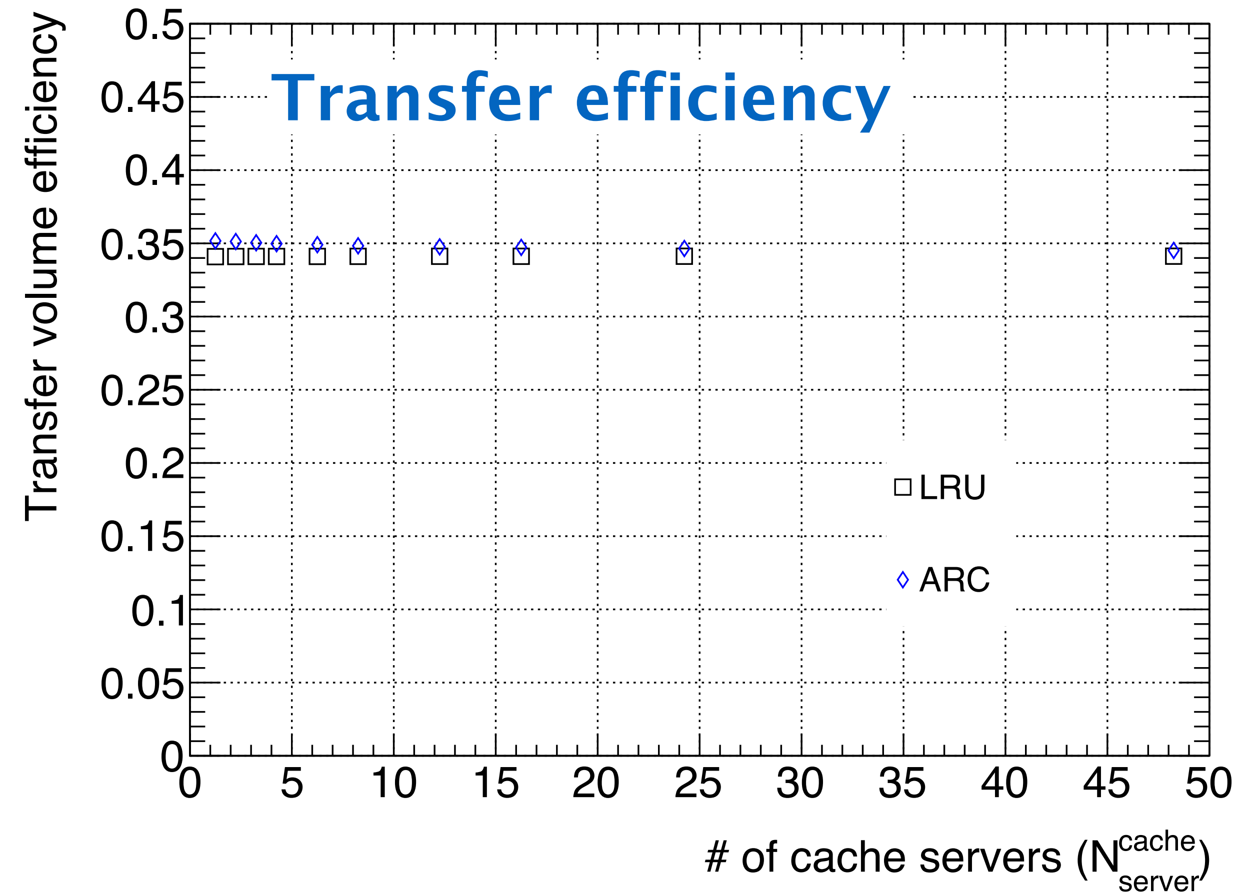
- **HITS**: Files after detector simulations (e.g. Geant4)
- **AOD**: Files after particle reconstruction, which are mainly used for physics analysis

→ The cache works effectively for the HITS files because common HITS files are used in pile-up simulations

$N^{cache}_{capacity}$=1 TB, N^{cache}_{server}=48 (lcg-fs01), LRU algorithm

Number of cache areas





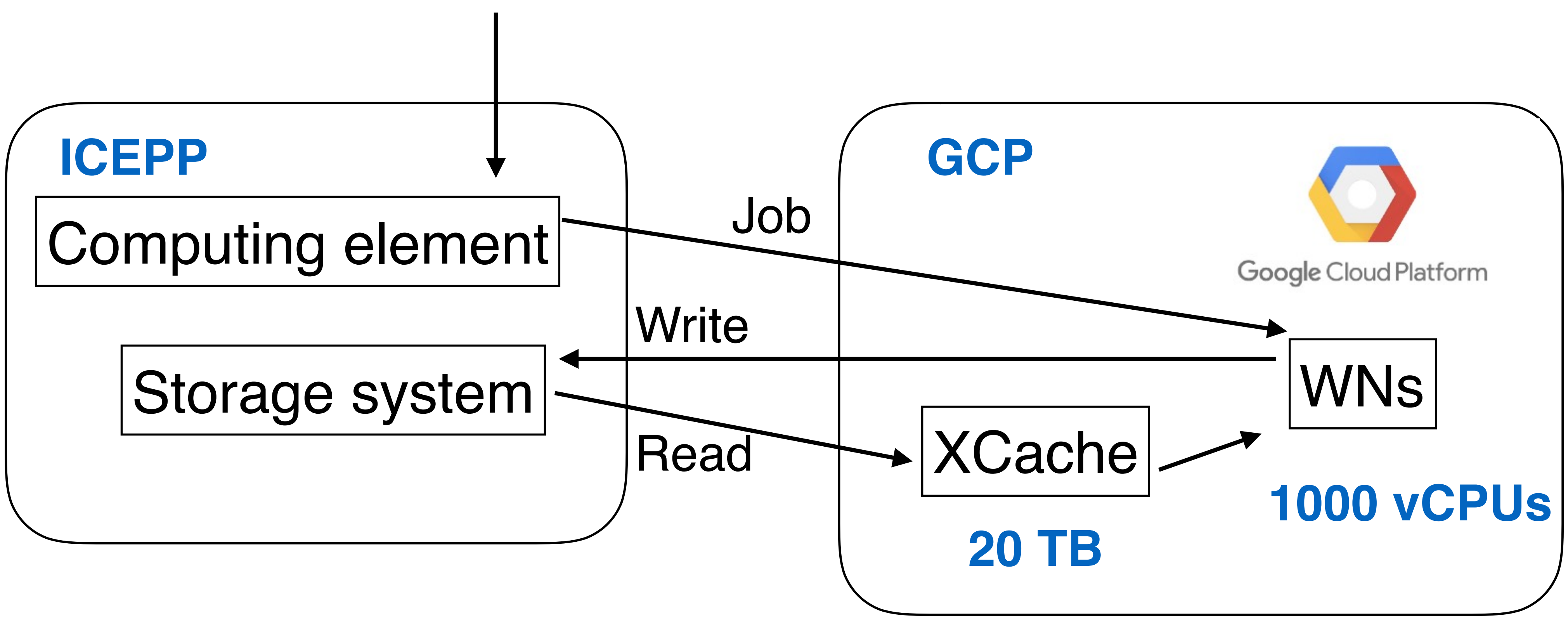
- ✓ Set total $N^{cache}_{capacity}$ to 48 TB
  - If N^{cache}_{server}= 48, each file server has 1TB cache area
  - If N^{cache}_{server}= 1, 48 TB cache area covers all file servers

- ✓ Variation in the efficiencies for N^{cache}_{server} is less than 1%

  → In terms of securing network bandwidth and I/O performances, it is required to prepare multiple cache areas (servers)

XCache test

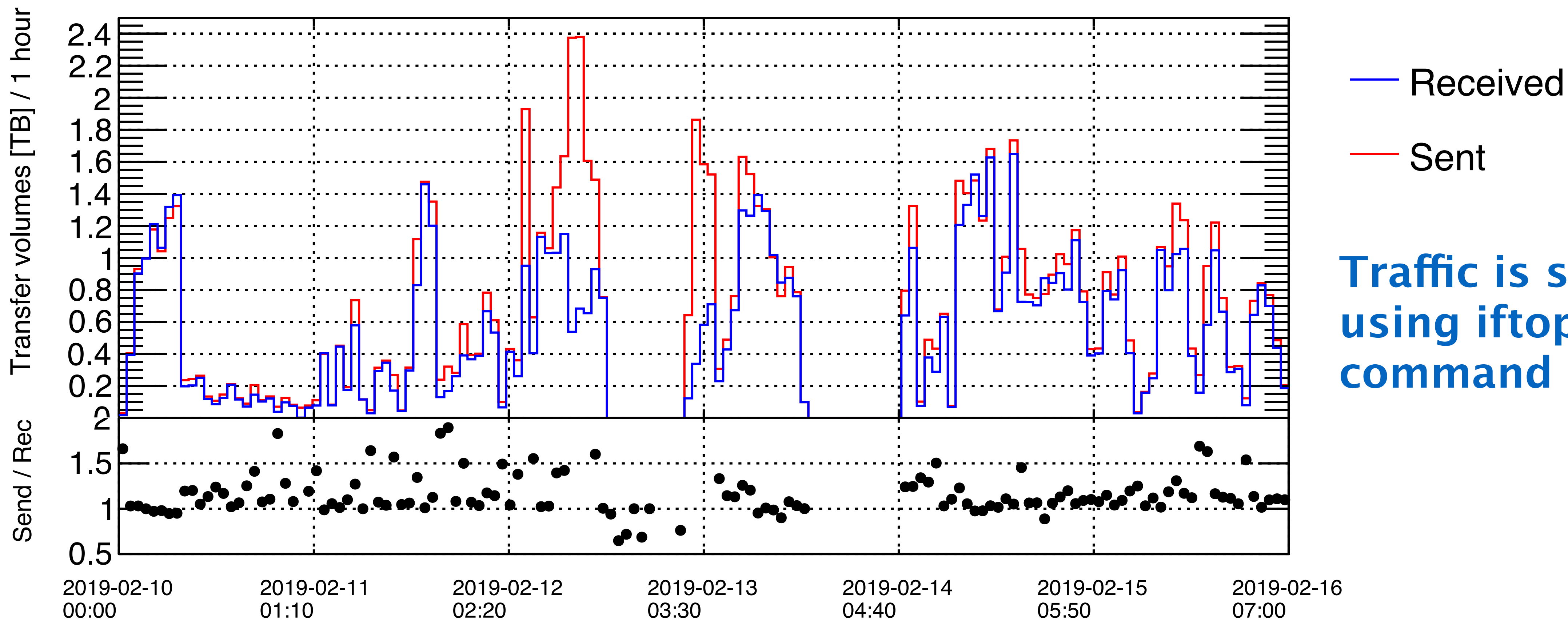
- ✓ XCache was deployed, and the cache hit rate was measured using real ATLAS Grid jobs
  - Worker nodes and XCache were deployed on commercial Cloud (Google Cloud Platform)
    →Easy to buy and scale the resources for quick tests



- * Operated XCache server for 7 days
- * XCache has 20TB (≒ 0.3% of total storage) as cache area

XCache results

Network traffic of XCache server



Traffic is sampled using iftop command

	Hit eff.	Transfer eff.
Obs. (XCache)	0.20	0.19
Simulation(*)	0.24	0.26

(*) $N^{cache}_{capacity}$ = 20 TB
N^{cache}_{server} = 1
Algorithm = LRU

✓ Reasonable agreements are observed

→ Simulation can provide reasonable information to consider the cache construction in the future

Summary

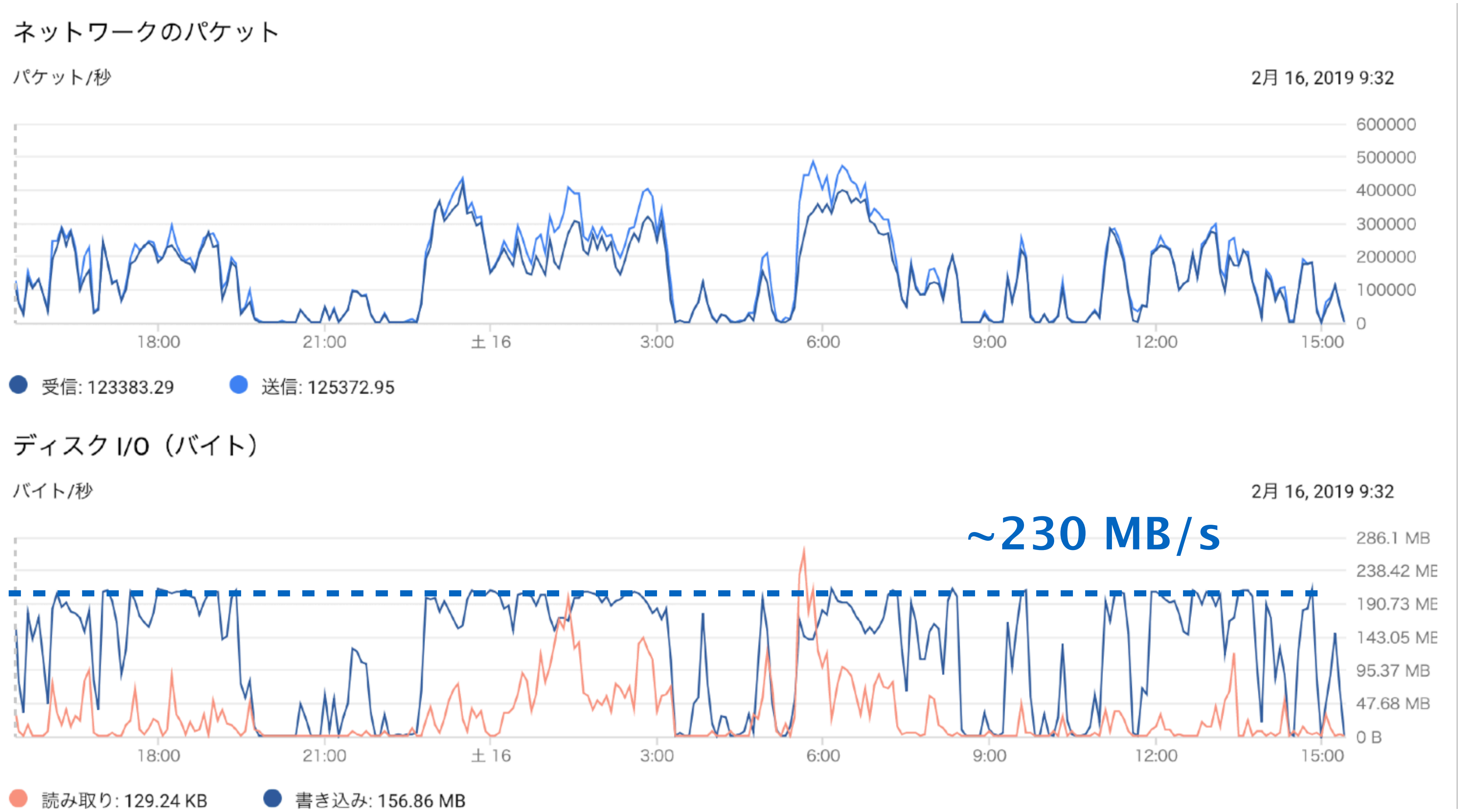
- ✓ Simulated the cache hit rate using transfer logs at Tokyo Tier2 to consider a possibility to construct the cache in the future
  - Simple LRU algorithm shows similar performances compared to the ARC algorithm
    - The cache hit rate is ~ 30% if the cache capacity is 1% of the total capacity
  - The cache works effectively especially for the pileup simulation (HITS file)
  - Variations in the efficiencies for the number of cache areas (servers) are small (<1%)
  - The simulation shows reasonable agreements with the XCache test
- ✓ Will investigate and get cache hardwares based on the simulation, and test its I/O performances

Backup

XCache experience

- ✓ 20 TB permanent SSDs were used as cache area
  - ~ $4,500 per month

- ✓ Disk I/O throughput was a bottle neck...

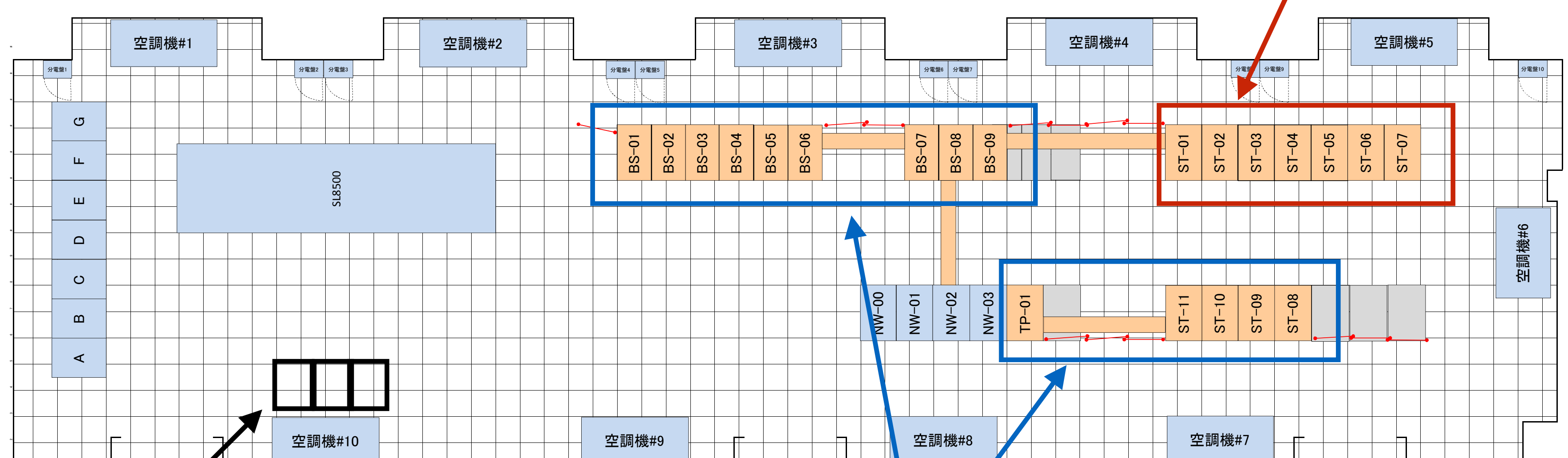


Reduced system during migration

- ✓ Late Nov. 2018
  - Old hardware devices other than WLCG storage system were removed
  - Reduced system was constructed at the edge of the room
    - 768 CPU cores and full Grid services were available
  - 16 hours site downtime

Server layout during the migration

WLCG storage system



Reduced system (3 server racks)

Removed

~300 m²