

JAXA Site Presentation

~Reliability of Data Management~

Japan Aerospace Exploration Agency

Supercomputer Division

FUJITA, Naoyuki
fujita@chofu.jaxa.jp

JAXAのサイトプレゼンテーション JAXA Site Presentation

～データ管理の信頼性～ ～Reliability of Data Management～

宇宙航空研究開発機構

Aerospace Exploration Agency

(≡ Research and Development)

スーパーコンピュータ活用課

Supercomputer Division

(≡ Utilization Division)

藤田 直行

WISTERIA field, Go straight

fujita@chofu.jaxa.jp



Family name



First name



藤田 直行
WISTERIA field, Go straight

JSS2 (JAXA Supercomputer System Generation 2)

Phase 3 Apr. 2016 ~

Apr. 2017

宙 (SORA: Supercomputer for earth Observation, Rockets, and Aeronautics) System Configuration

Chofu Aerospace Center

Main System (SORA-MA)

Peak Performance : 3.49 PFLOPS
Total Memory : 100 TiB
Total #Nodes : 3,240
Total #Cores : 103,680
(1 CPU, 32 GiB / node)



Pre-Post System (SORA-PP)

Peak Performance : 75.2 TFLOPS
Total Memory : 12 TiB
Total #Nodes : 200
Total #Cores : 2,400
(2 CPU, 64 GiB / node)



Large Memory System (SORA-LM)

Peak Performance : 2.10 TFLOPS
Total Memory : 3.5 TiB
Total #Nodes : 5
Total #Cores : 80
(2 CPU, 1.0 TiB x 2 nodes)
(2 CPU, 0.5 TiB x 3 nodes)



Login System (SORA-LI)

Peak Performance : 1.34 TFLOPS
Total Memory : 1.5 TiB
Total #Nodes : 4
Total #Cores : 48
(2 CPU, 0.37 TiB x 4 nodes)



Storage Area Network (Infini Band FDR)

Ethernet (1G, 10 G Ethernet)

Control and Management System



File System (SORA-FS)

Magnetic Disk : 7 PB



Remote Systems

LAN and SINET5

Tsukuba Space Center

Pre-Post Node (SORA-TPP)

8.40 TFLOPS 1.5 TiB 25 nodes 300 cores
(2 CPU, 64 GiB / node)

Login Node (SORA-TLI)

0.336 TFLOPS 0.37 TiB 1 node 12 cores
(2 CPU, 0.37 TiB / node)

File System (SORA-TFS)

200 TB



Kakuda Space Center

File System (SORA-KFS)

100 TB



Sagamihara Campus

File System (SORA-SFS)

100 TB



J-SPACE

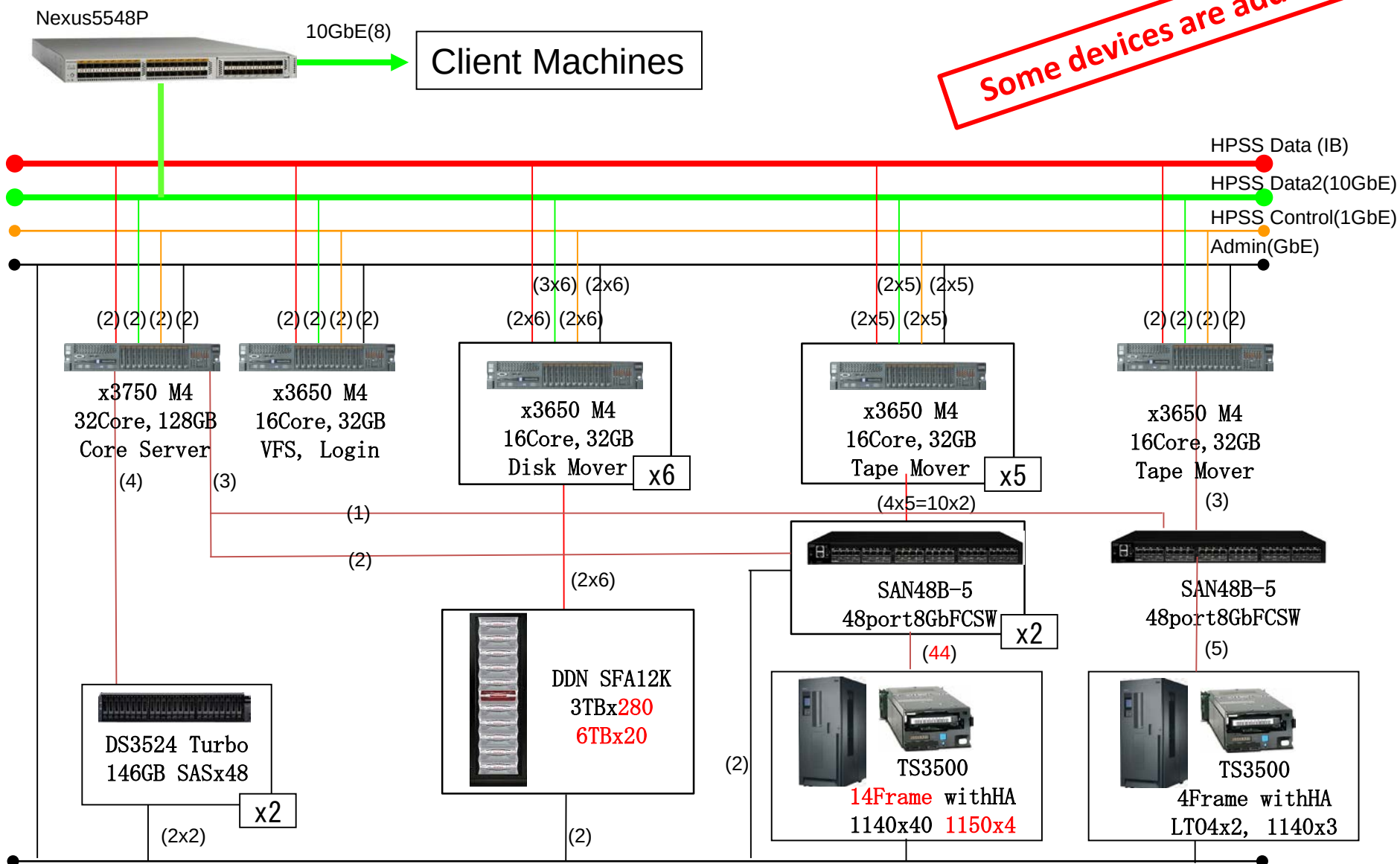
Jaxa's Storage Platform for Archiving, Computing, and Exploring

Archiver System

Disk Cache : 0.7 PB
Tape : 20 PB



J-SPACE H/W Configuration

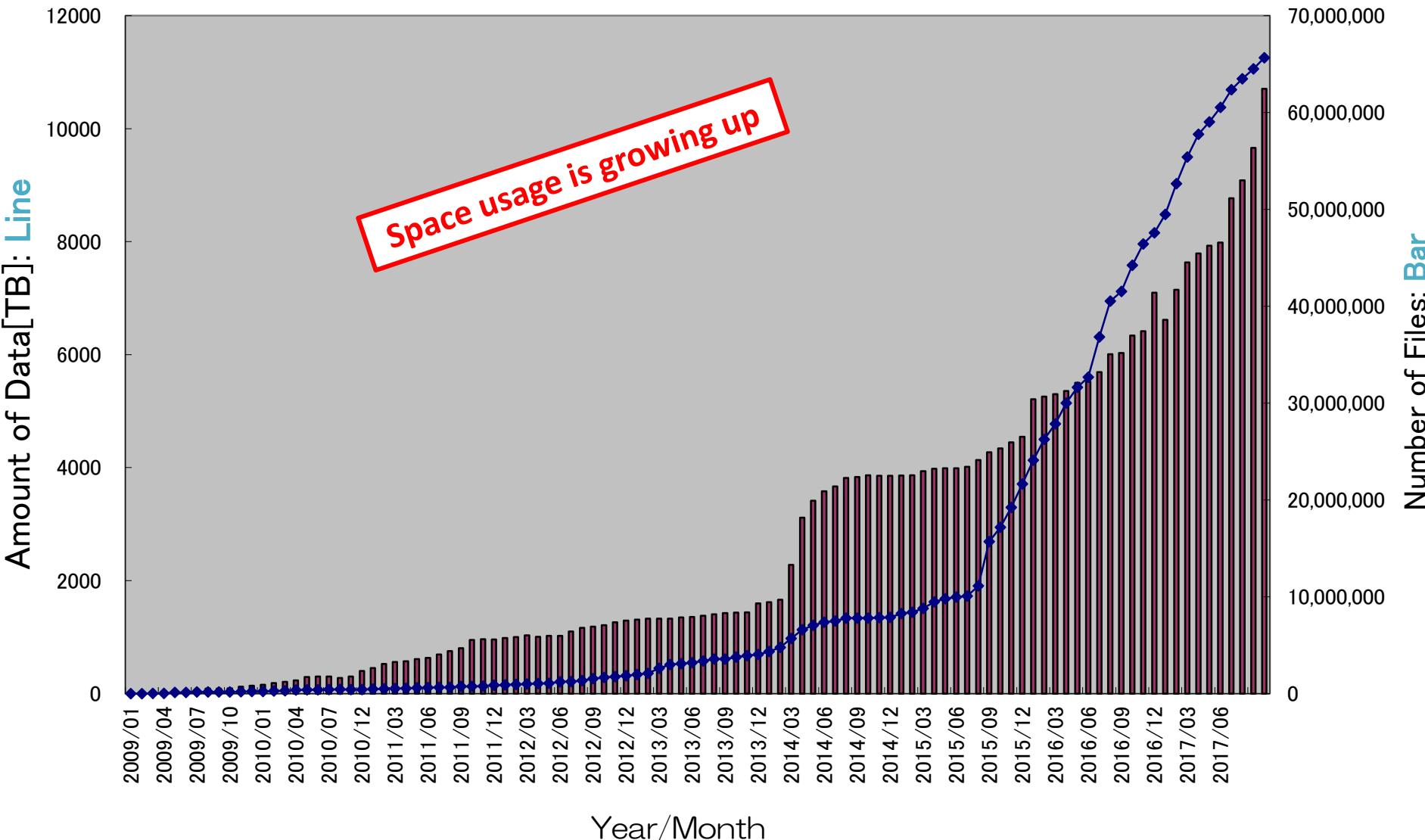


■ J-SPACE S/W version & services

Some devices are added

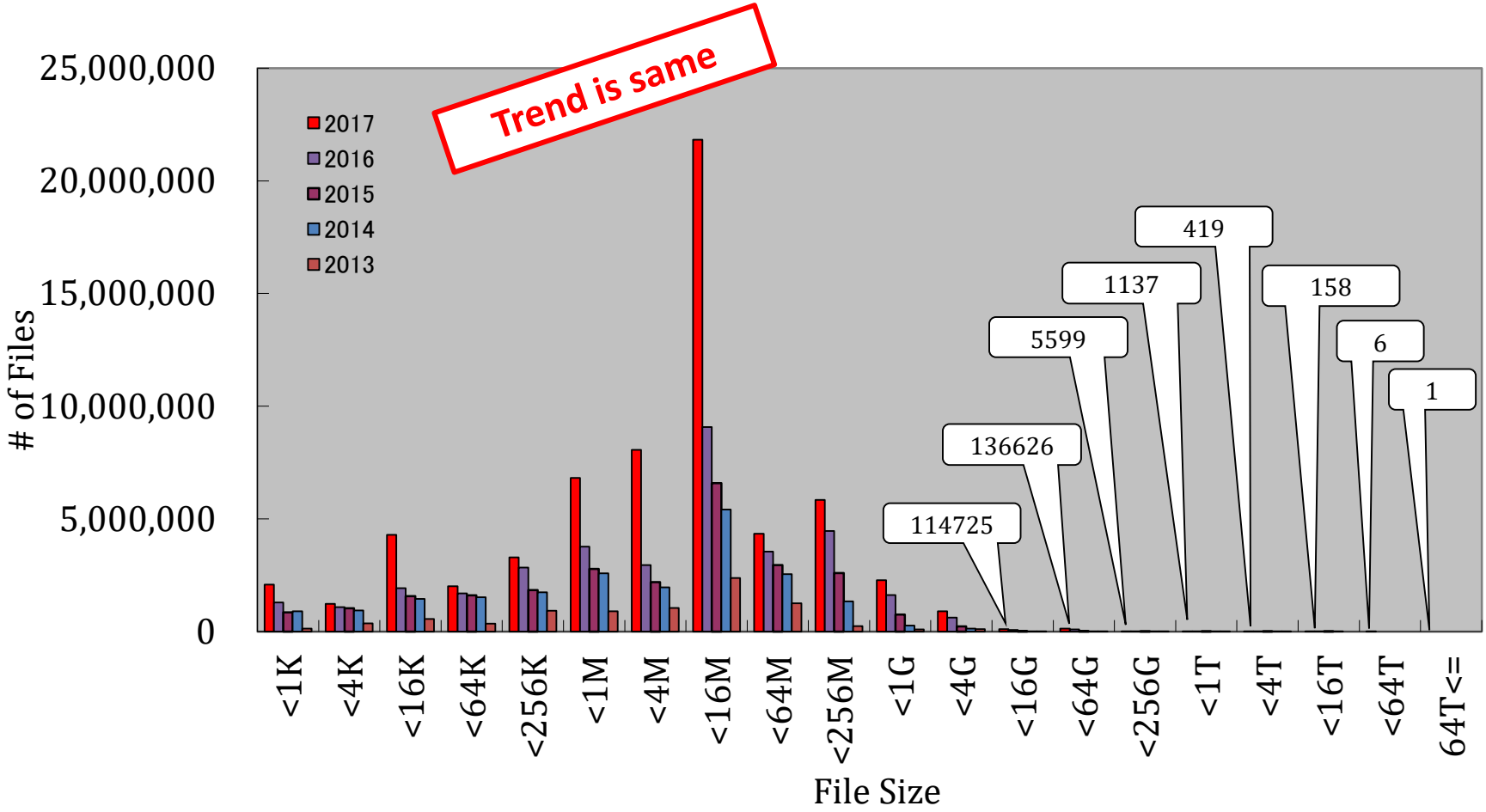
- Operation started at August 2014
- HPSS: 7.4.2.1
- DB2: 10.5 Fixpack3a
- OS: RHEL 6.4
 - AIX => N/A (Replaced with RHEL) ...orz
- Tape Drives:
 - IBM TS1140X(40+3) drives, IBM TS1150x4
 - IBM LTO4x2 drives
- Tape Cartridges:
 - 3592: 11317+1382
 - LTO4: 2884
- Cache Disk: DDN SFA12K 0.73PB
- User Data: 11.0PB(Logical) / 6.2M Files
- Number of used COS: 26
- UI
 - pftp
 - hsi/htar
 - VFS, NFS via VFS, ftp via VSF, scp via VFS

Space Usage on HPSS



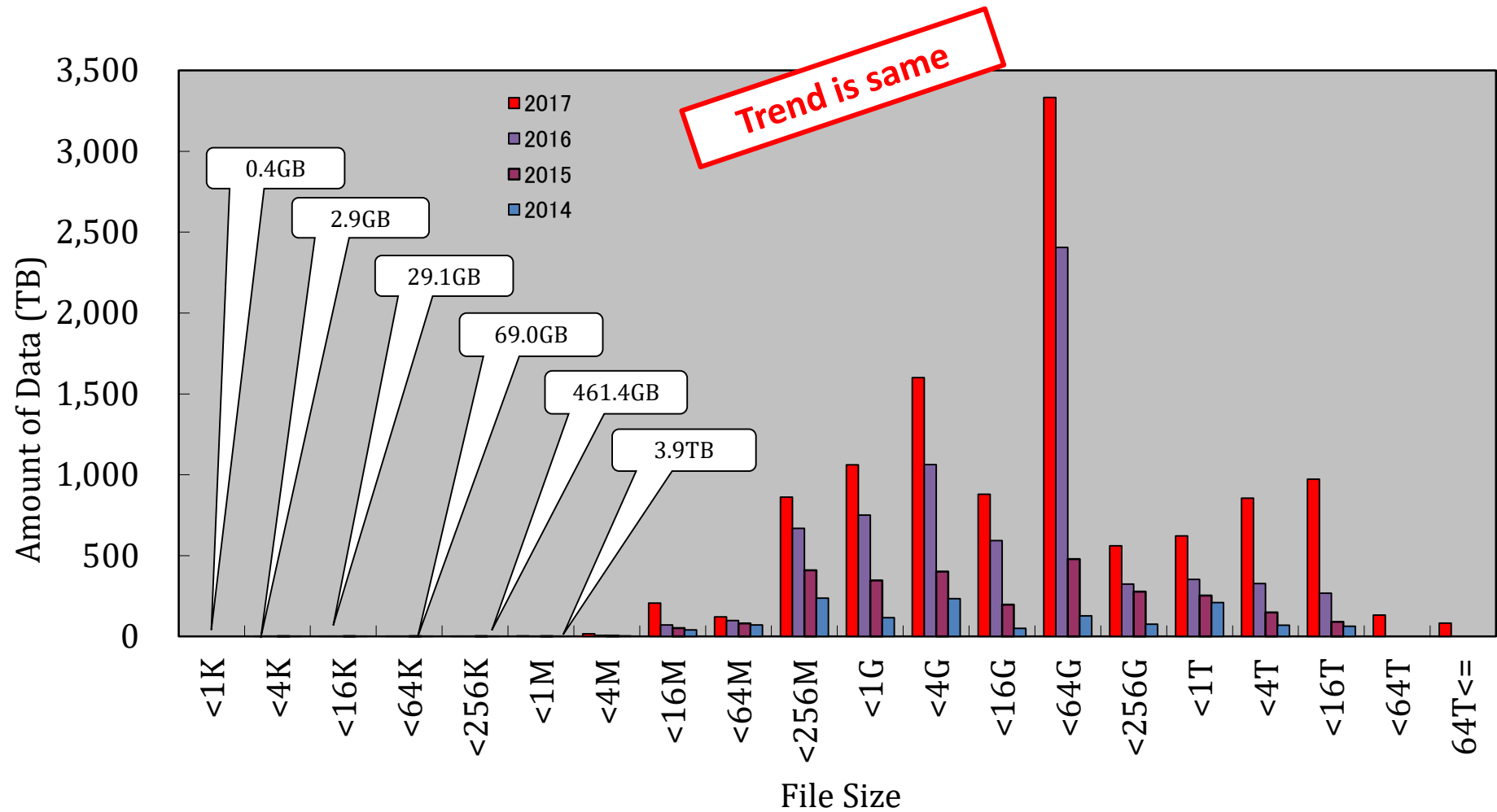
■ Number of files on HPSS (by size, by year)

Distribution of File Size #1 (2013-2017)



■ Amount of DATA on HPSS (by size, by year)

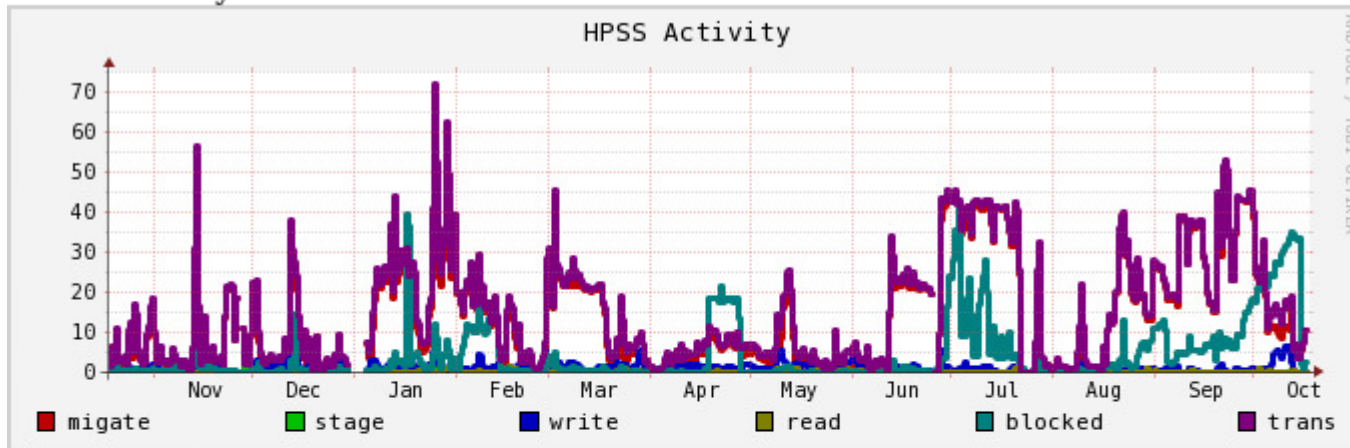
Distribution of File Size(2014-2017)



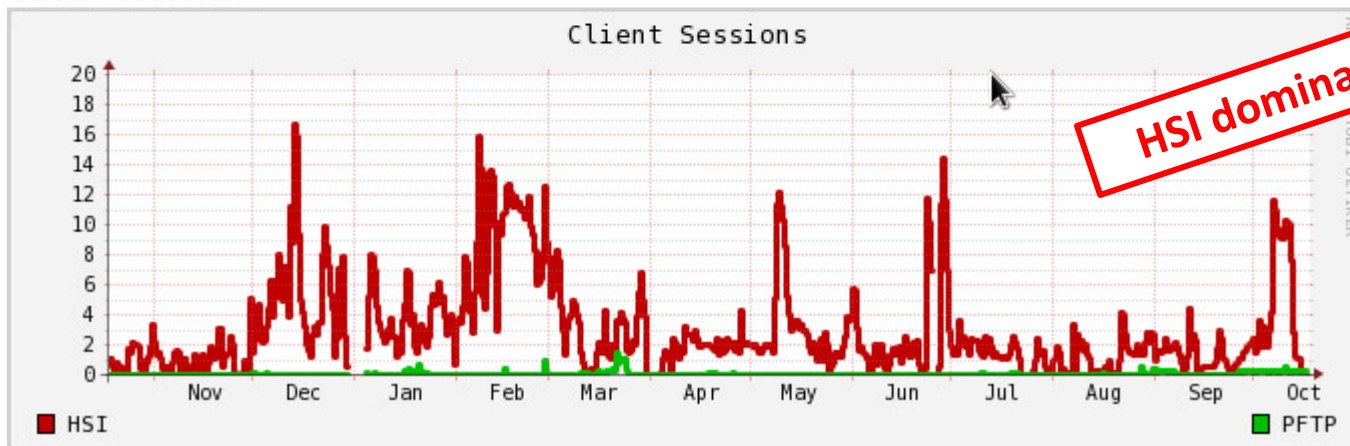
■ HPSS Activity and Client Sessions

(Usage Monitoring by Ganglia)

HPSS Activity



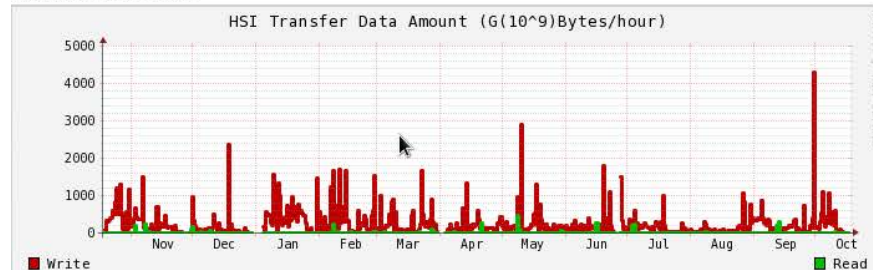
Client Sessions



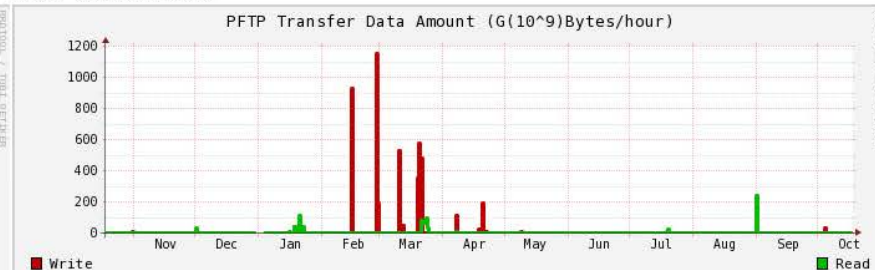
Transfer Data & Files

(Usage Monitoring by Ganglia)

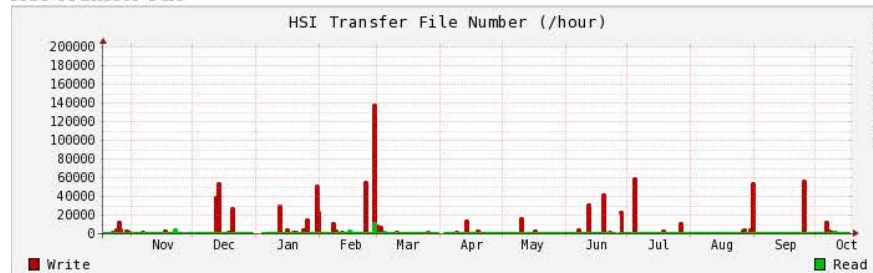
HSI Transfer Data



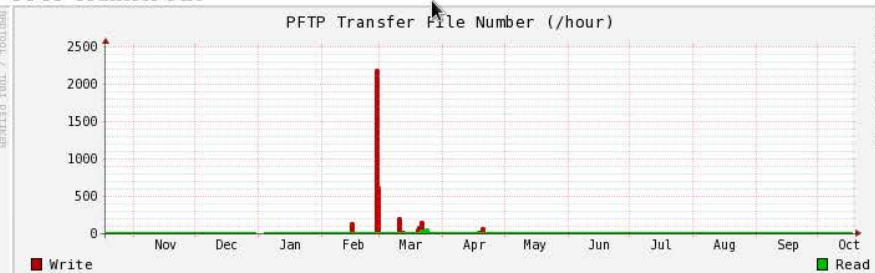
PFTP Transfer Data



HSI Transfer File



PFTP Transfer File



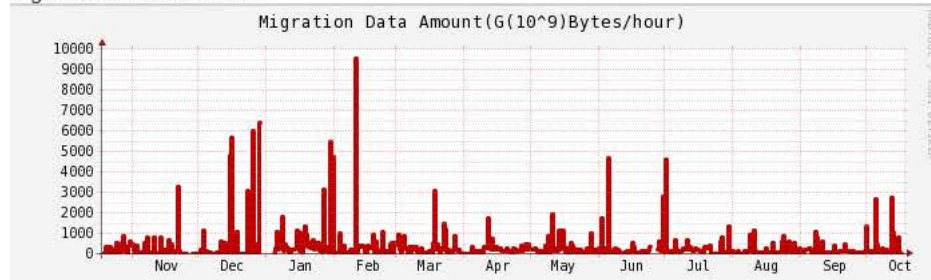
Wish Item 1

Not sure if we can monitor VFS data transfer in terms of data amount and file number but we want..

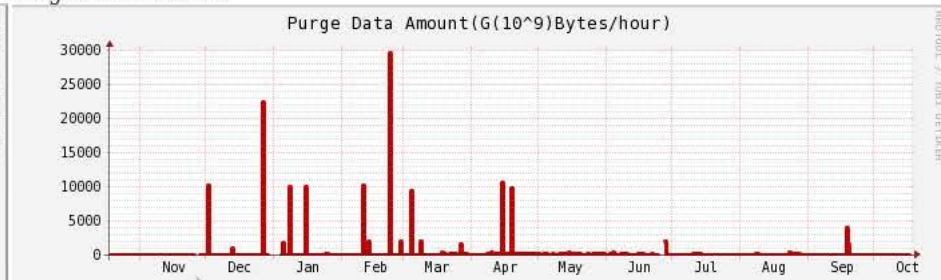
Migrate/Purge Data & Files

(Usage Monitoring by Ganglia)

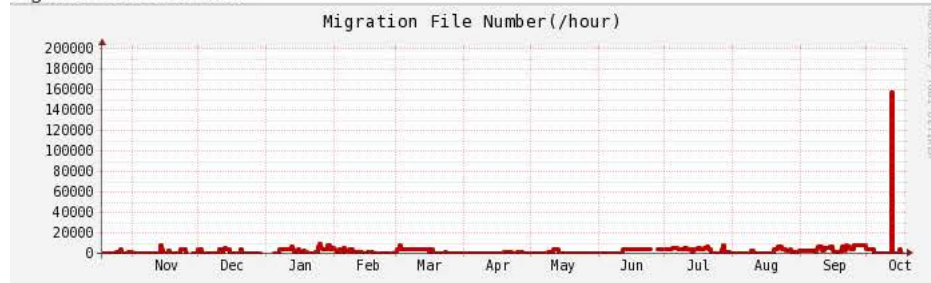
Migration Data Amount



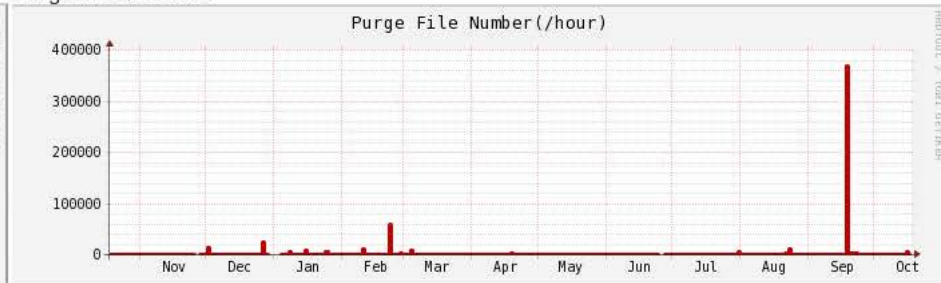
Purge Data Amount



Migration File Number



Purge File Number



Wish Item 2

We are waiting for a new HPSS version to be able to give information about staging statistics

JSS2 (JAXA Supercomputer System Generation 2)

Phase 3 Apr. 2016 ~

Apr. 2017



(SORA: Supercomputer for earth Observation, Rockets, and Aeronautics) System Configuration

Chofu Aerospace Center

Main System (SORA-MA)

Peak Performance : 3.49 PFLOPS
Total Memory : 100 TiB
Total #Nodes : 3,240
Total #Cores : 103,680
(1 CPU, 32 GiB / node)



Pre-Post System (SORA-PP)

Peak Performance : 75.2 TFLOPS
Total Memory : 12 TiB
Total #Nodes : 200
Total #Cores : 2,400
(2 CPU, 64 GiB / node)



Large Memory System (SORA-LM)

Peak Performance : 2.10 TFLOPS
Total Memory : 3.5 TiB
Total #Nodes : 5
Total #Cores : 80
(2 CPU, 1.0 TiB x 2 nodes)
(2 CPU, 0.5 TiB x 3 nodes)



Login System (SORA-LI)

Peak Performance : 1.34 TFLOPS
Total Memory : 1.5 TiB
Total #Nodes : 4
Total #Cores : 16
(2 CPU, 0.37 TiB x 4 nodes)



Storage Area Network (Infini Band FDR)

Ethernet (1G, 10 G Ethernet)

Control and Management System



File System (SORA-FS)

Magnetic Disk : 7 PB



Remote Systems



Tsukuba Space Center

Pre-Post Node
8.40 TFLOPS
(2 CPU, 64 GiB / node)

Login Node (SORA-TLI)
0.336 TFLOPS 0.37 TiB 1 node 12 cores
(2 CPU, 0.37 TiB / node)

File System (SORA-TFS) 200 TB

Kakuda Space Center

File System (SORA-KFS) 100 TB

Sagamihara Campus

File System (SORA-SFS) 100 TB

21 Sep. 2017

8 Sep. 2017

16 Jul. 2015

Banquet Day
HUF2017

16 Oct. 2017

9 May 2017
2nd copy

9 May 2017
1st copy

1 May 2017

7 May 2017

Archiver System

Disk Cache : 0.7 PB

Total : 10 PB

HPSS



Failure history

Phase 3 Apr. 2016 ~

Apr. 2017

System Configuration

Host System (SORA-PP)

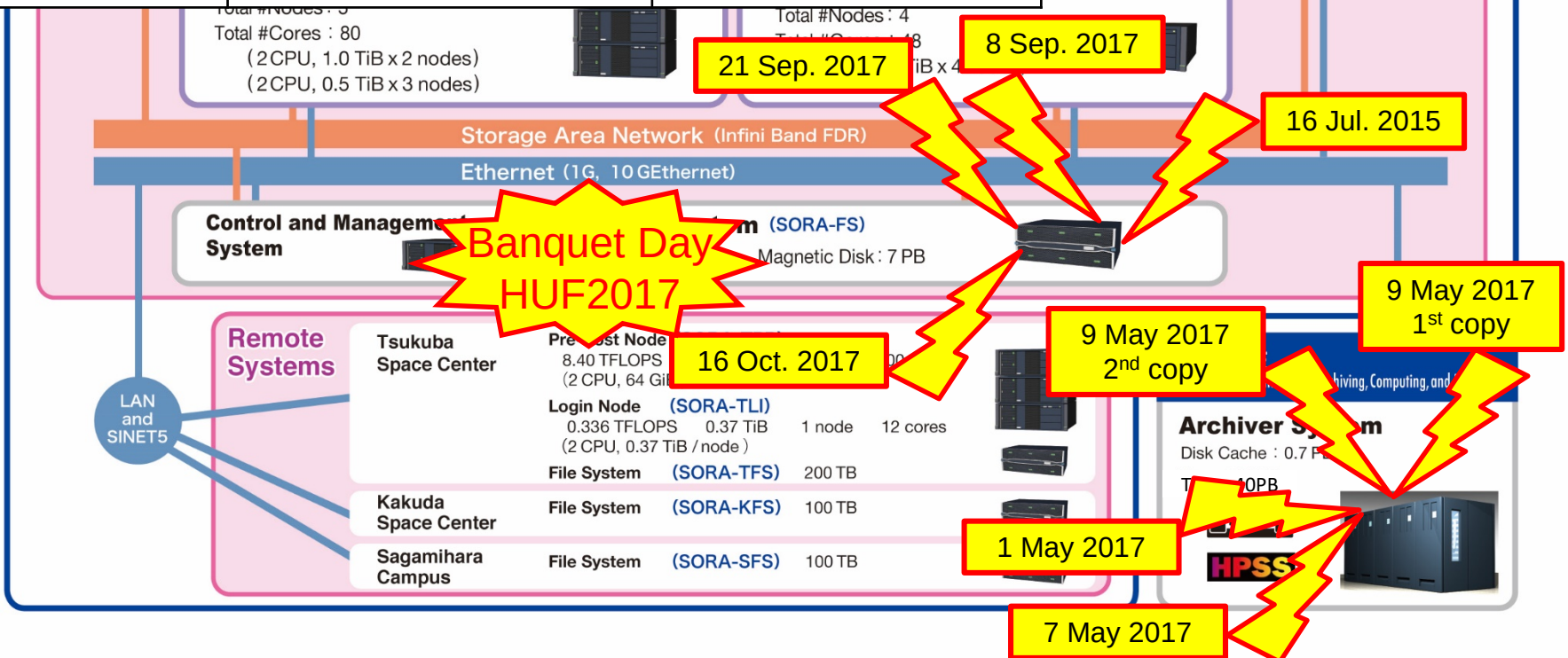
Performance : 75.2 TFLOPS
Memory : 12 TiB
Nodes : 200
Disks : 2,400
I/O : 64 GiB / node



(S-LI)

OPS

Date	Issue / Cause	Remark
16 Jul. 2015	Silent Data Corruption	
1 May 2017	Disk controller failure?	S.D.C.?
7 May 2017	Disk controller failure?	S.D.C.?
9 May 2017	1 st copy unreadable	Media damage?
9 May 2017	2 nd copy unreadable	Debris ?
8 Sep. 2017	Meta data un-match	- File System S/W bug? - H/W failure?
21 Sep. 2017	Meta data un-match	
16 Oct. 2017	Meta data un-match	



mini-Survey on HPSS Reliability

- Have you lost **any bitfile** on your HPSS?
(JAXA: Yes) ...9 raised hands
- Have you lost **user file** on your HPSS?
(JAXA: No) ...9 raised hands
- Have you experienced on Silent Data Corruption (*including suspicious one*)?
(JAXA: Maybe Yes) ...4 raised hands
- Have you experienced on unknown(unresolved) data loss?
(JAXA: Yes) ...2 raised hands

Questions!

mini-Survey on File System Reliability

- Have you lost **(a part of) file** on your file system? *(JAXA: Yes) ...8 raised hands*
- Have you lost **user file** on your file system? *(JAXA: Yes) ...8 raised hands*
- Have you experienced on Silent Data Corruption (*including suspicious one*) on your file system? *(JAXA: Yes) ...6 raised hands*
- Have you experienced on unknown data loss on your file system? *(JAXA: No) ...0 raised hand*

Questions!

What should we pay attention to?

- Media product quality?
- Tape cartridge and tape drive collision and rubbing?
- Machine room environment?
- What should we monitor media condition to avoid data-loss?
 - Mount count?
 - Mileage of tape media?
 - Retention count?
 - Error logs from media?

What are you monitoring?

What should we do for end-user?

Thank you very much.

