Reliability of NVM devices for I/O Acceleration on Supercomputing Systems

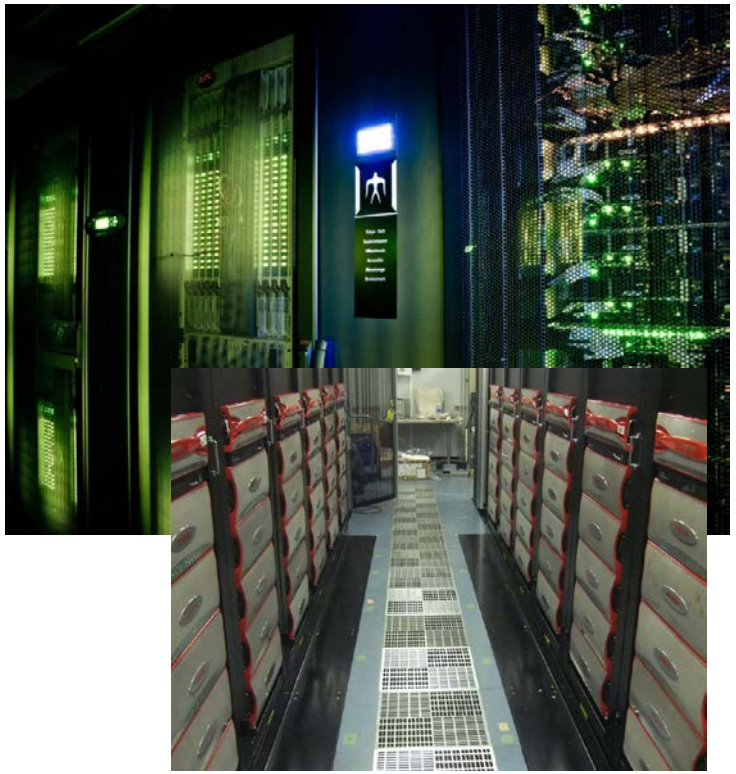
Hitoshi Sato*[1], Shuichi Ihara[2)], Satoshi Matsuoka[1)]

*[1] Tokyo Institute of Technology

*[2] Data Direct Networks Japan

Requirements for Next-gen I/O Subsystem

- TSUBAME2
  - Pioneer in the use of local flash storage
    - 200TB of SSDs for productive operation since 2010
    - Tired and hybrid storage environments, combining local flash with external Lustre FSs
- Industry Status
  - Various flash devices emerging
    - Available at reasonable cost
  - New Approaches
    - Flat Buffers
    - Burst Buffers etc.,



Issues for Introducing Flash Devices on Next-gen Supercomputers

- Flash Devices have various performance characteristics

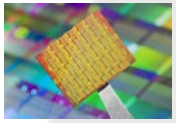
- How much reliability of flash devices is required for burst buffers for next-gen supercomputers?
  - Maximize
    - Throughput, IOPS, Capacity, Reliability, etc.
  - Minimize
    - Cost for introducing flash devices on large-scale systems

Today's Talk

- Preliminary Evaluation of the endurance for SSD devices to support supercomputing I/O workloads
- Lustre I/O Monitoring for I/O workload analysis to evaluate the endurance

Towards Extreme-scale Supercomputing

- Memory/Storage Architecture
  - NVM(Non-Volatile Memory, Flash) is a key device for I/O subsystems
    - Capacity, Non-Volatility, Low energy consumption (vs. DRAM)
    - High throughput, Low latency, Low energy consumption (vs. HDD)
  - Various Flash Devices
    - Throughput: 200MB/s ～ over 1GB/s
    - IOPS: 10,000 ～ over 100,000
    - Interfaces : PCI-e attched, SATA3, mSATA, m.2, etc.





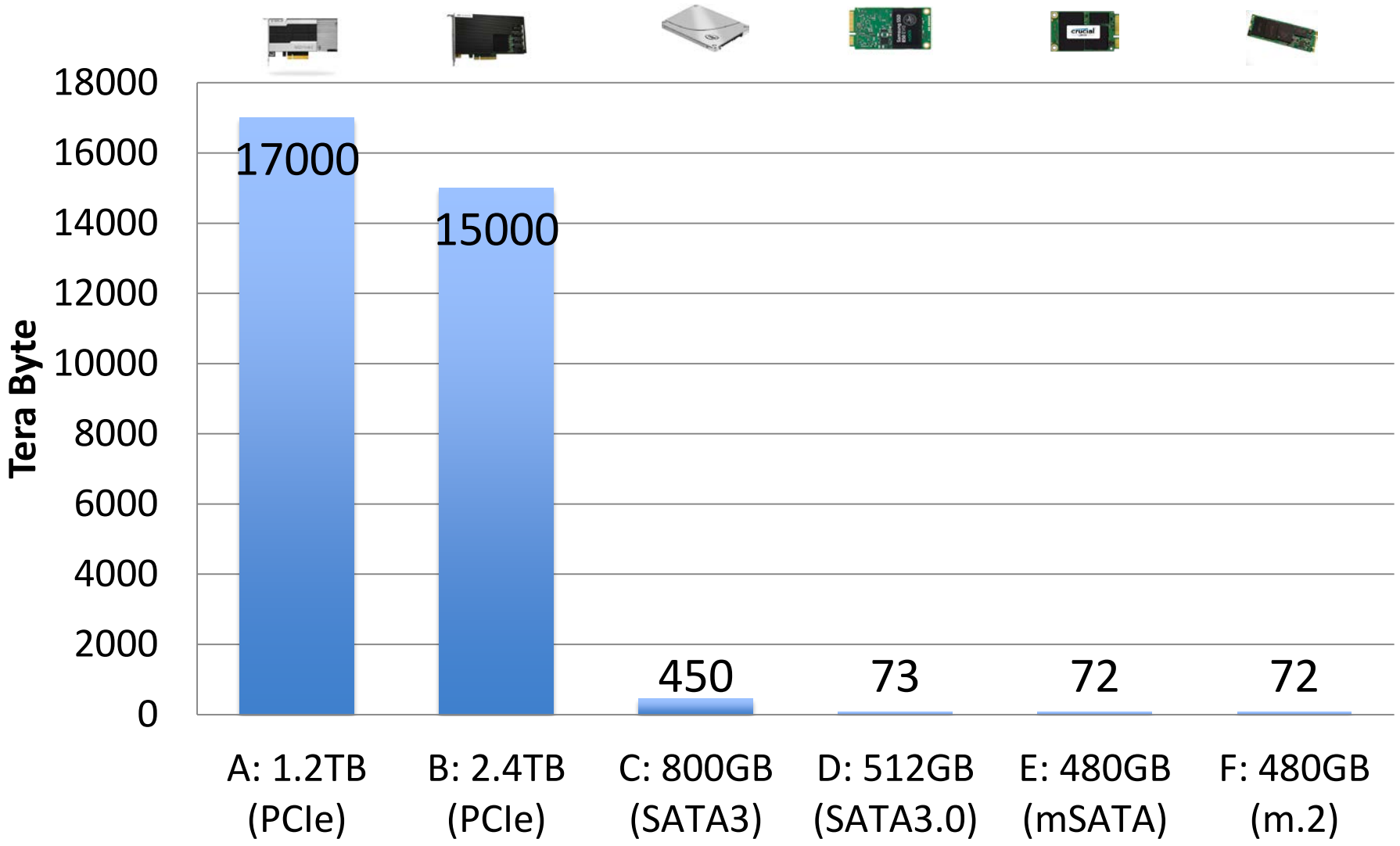

Reliability Evaluation of Flash Devices

- Power on Hours
  - Should support for operation duration
  - 4～5 years
- Endurance
  - Peta Byte Written (PBW), Tera Byte Written (TBW)
  - Deterioration Progress (%) = Total Write / PBW or TBW
- Average Erase Count (AEC)
  - Upper limit of erasure count (EC) is determined for a flash device
  - Deterioration Progress (%) = AEC / Upper Limit of EC
- Incidence of Bad Blocks
  - Impact for wear-leveling, GC

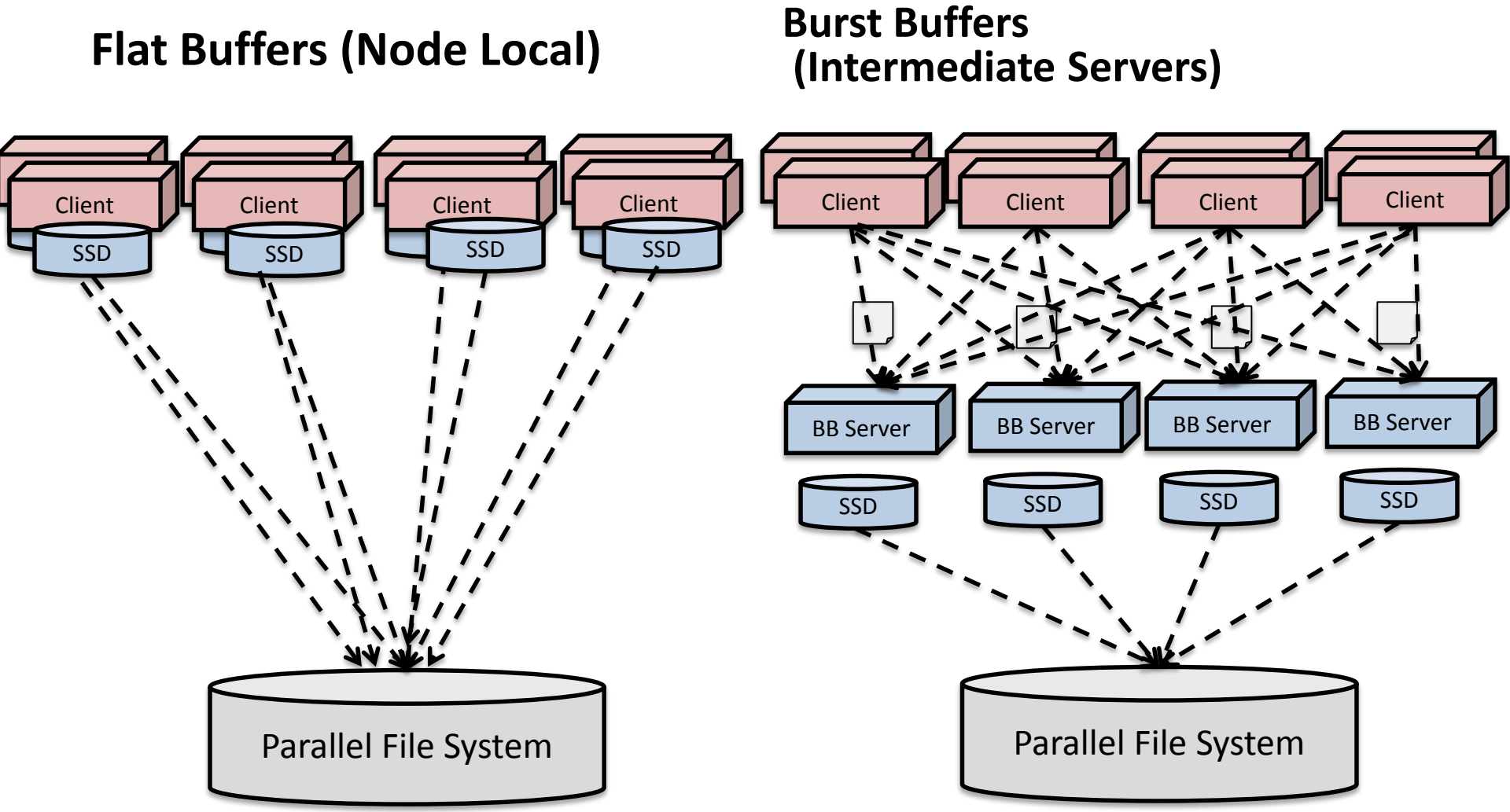
Reliability Evaluation of Flash Devices

- Power on Hours
  - Should support for operation duration
  - 4～5 years
- Endurance
  - Peta Byte Written (PBW), Tera Byte Written (TBW)
  - Deterioration Progress (%) = Total Write / PBW or TBW
- Average Erase Count (AEC)
  - Upper limit of erasure count (EC) is determined for a flash device
  - Deterioration Progress (%) = AEC / Upper Limit of EC
- Incidence of Bad Blocks
  - Impact for wear-leveling, GC

Endurance of Existing Flash Devices



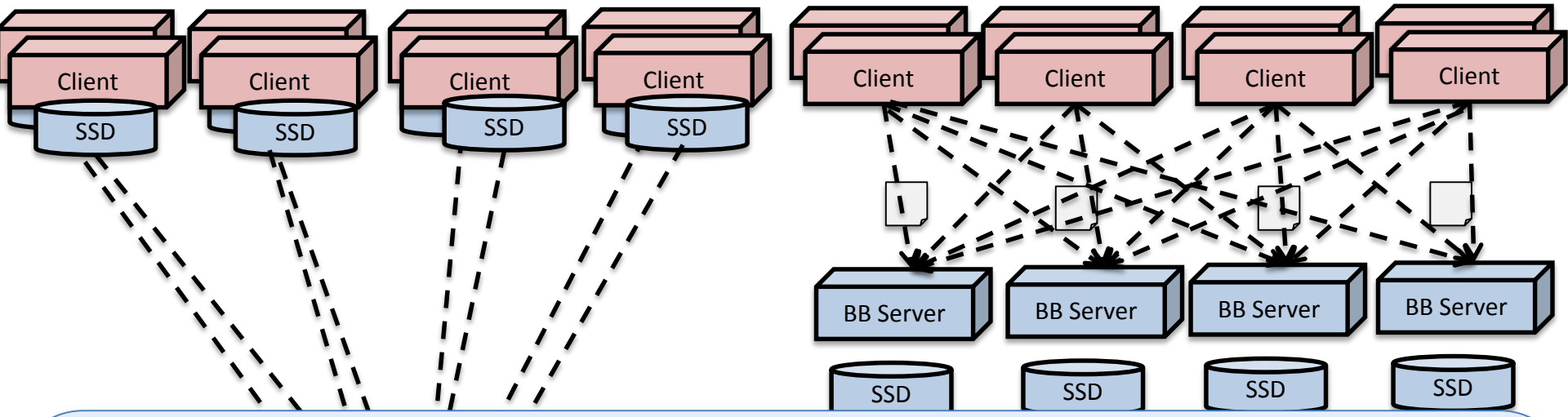
SSD Configuration on Supercomputers



SSD Configuration on Supercomputers

Flat Buffers (Node Local)

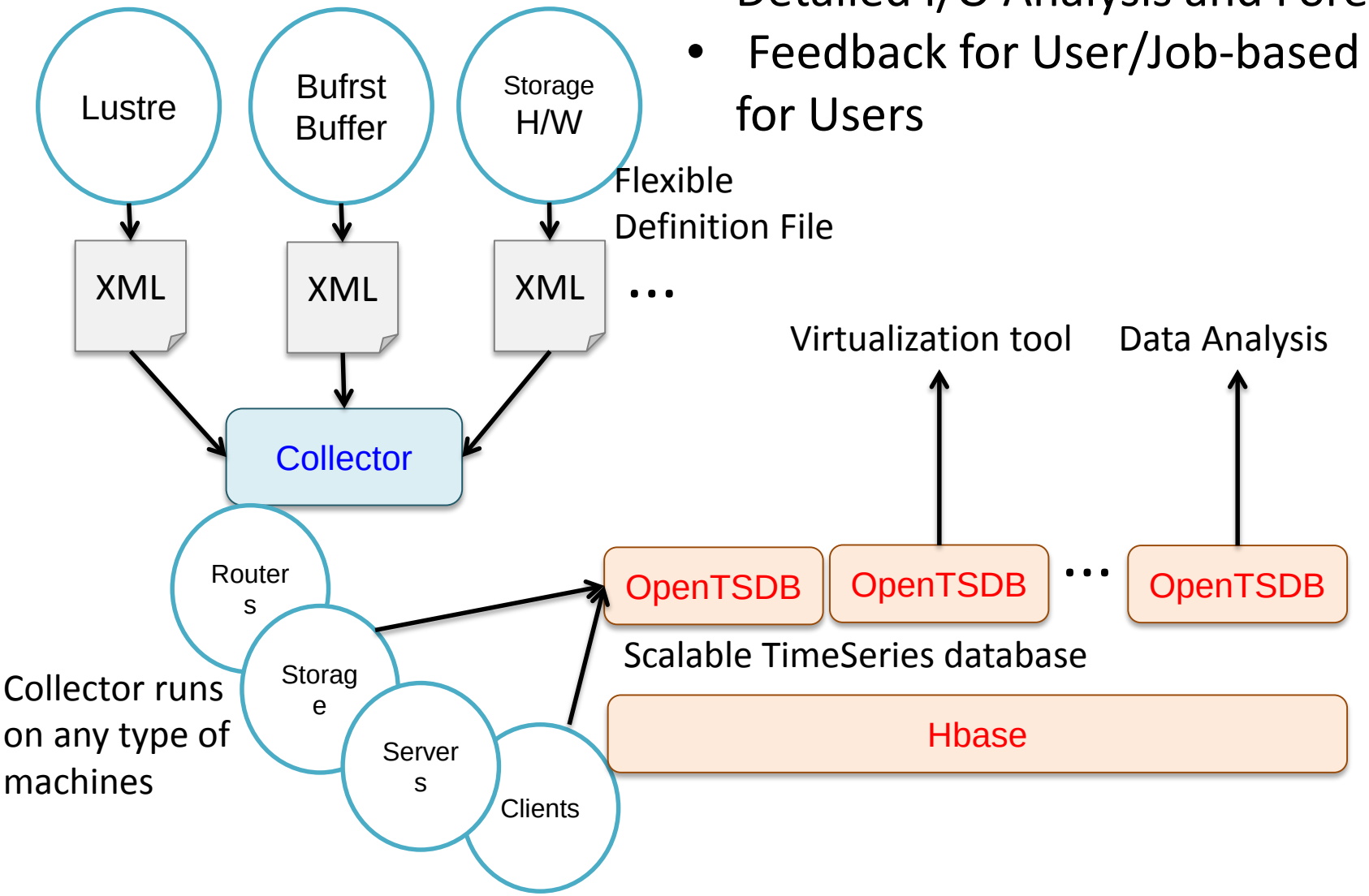
Burst Buffers (Intermediate Servers)



- How many SSD devices can we use to support the endurance of the volumes during the operation duration?
- We analyze the required endurance based on Lustre I/O monitoring

Lustre I/O Monitoring

- Detailed I/O Analysis and Forecast
- Feedback for User/Job-based Stats for Users



Lustre I/O Monitoring

- Scalable and Flexible
  - Many Lustre stats, more than, 10M, 100M... stats
    - #OST * #client * stats * sampling
    - Future : #OST * #client * stats * JOBID(UID, etc) * sampling
  - Today, collected client's exported stats from Lustre servers
    - Lustre-2.1 is still running on server, no jobstats
  - Store data into scalable backend database
    - OpenTSDB and HBase
  - Selectable OpenSource frontend
  - All Lustre version support
    - /proc/fs/lustre structure changes on several lustre versions
    - A shadow definition XML of each version's /proc/fs/lustre

Endurance Evaluation

- Scheme
  - Estimate Total Write Size [PB] during operation years (1～5 years) based on Lustre Write Data Rate [GB/s] (Avg, Max)
  - Map the Total Write Size to aggregated SSD volumes
  - Evaluate the endurance of the aggregated SSD volumes based on the Total Write Size

TSUBAME2 System Overview

11PB (7PB HDD, 4PB Tape, 200TB SSD)

Computing Nodes: **17.1PFlops(SFP), 5.76PFlops(DFP), 224.69TFlops(CPU), ~100TB MEM, ~200TB SSD**

Thin nodes 1408nodes (32nodes x44 Racks)

HP Proliant SL390s G7 1408nodes
CPU: Intel Westmere-EP 2.93GHz
6cores × 2 = 12cores/node
GPU: NVIDIA Tesla K20X, 3GPUs/node
Mem: 54GB (96GB)
SSD: 60GB x 2 = 120GB (120GB x 2 = 240GB)

Local SSDs

Medium nodes

HP Proliant DL580 G7 24nodes
CPU: Intel Nehalem-EX 2.0GHz
8cores × 2 = 32cores/node
GPU: NVIDIA Tesla S1070,
NextIO vCORE Express 2070
Mem:128GB
SSD: 120GB x 4 = 480GB

Fat nodes

HP Proliant DL580 G7 10nodes
CPU: Intel Nehalem-EX 2.0GHz
8cores × 2 = 32cores/node
GPU: NVIDIA Tesla S1070
Mem: 256GB (512GB)
SSD: 120GB x 4 = 480GB

Interconnets: Full-bisection Optical QDR Infiniband Network

Core Switches

12switches

Voltaire Grid Director 4700 ×12
IB QDR: 324 ports

Edge Switches

179switches

Voltaire Grid Director 4036
×179
IB QDR : 36 ports

Edge Switches /w 10GbE ports

6switches

Voltaire Grid Director 4036E ×6
IB QDR:34ports
10GbE: 2port

QDR IB(×4) × 20

QDR IB (×4) × 8

10GbE × 2

SFA10k #1 SFA10k #2 SFA10k #3 SFA10k #4 SFA10k #5 SFA10k #6

GPFS#1 GPFS#2 GPFS#3 GPFS#4

GPFS+Tape

Lustre

Home

/gscr

HOME HOME

iSCSI

"Global Work Space" #2

"Global Work Space" #3

"cNFS/Clusterd Samba w/ GPFS"

"NFS/CIFS/iSCSI by BlueARC"

2.4 PB HDD + ~4PB Tape

3.6 PB

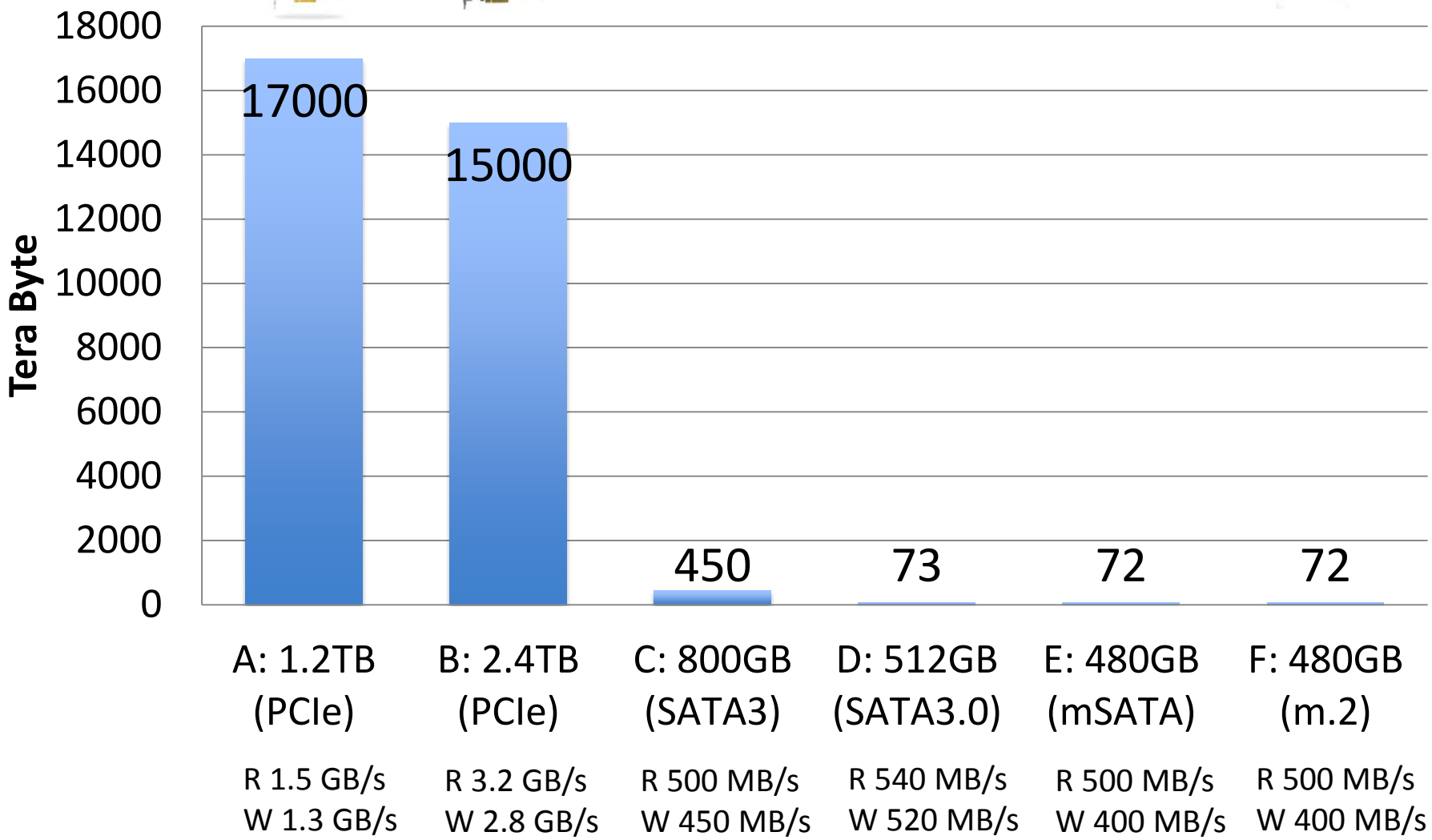
Parallel File System Volumes

Home Volumes **1.2PB**

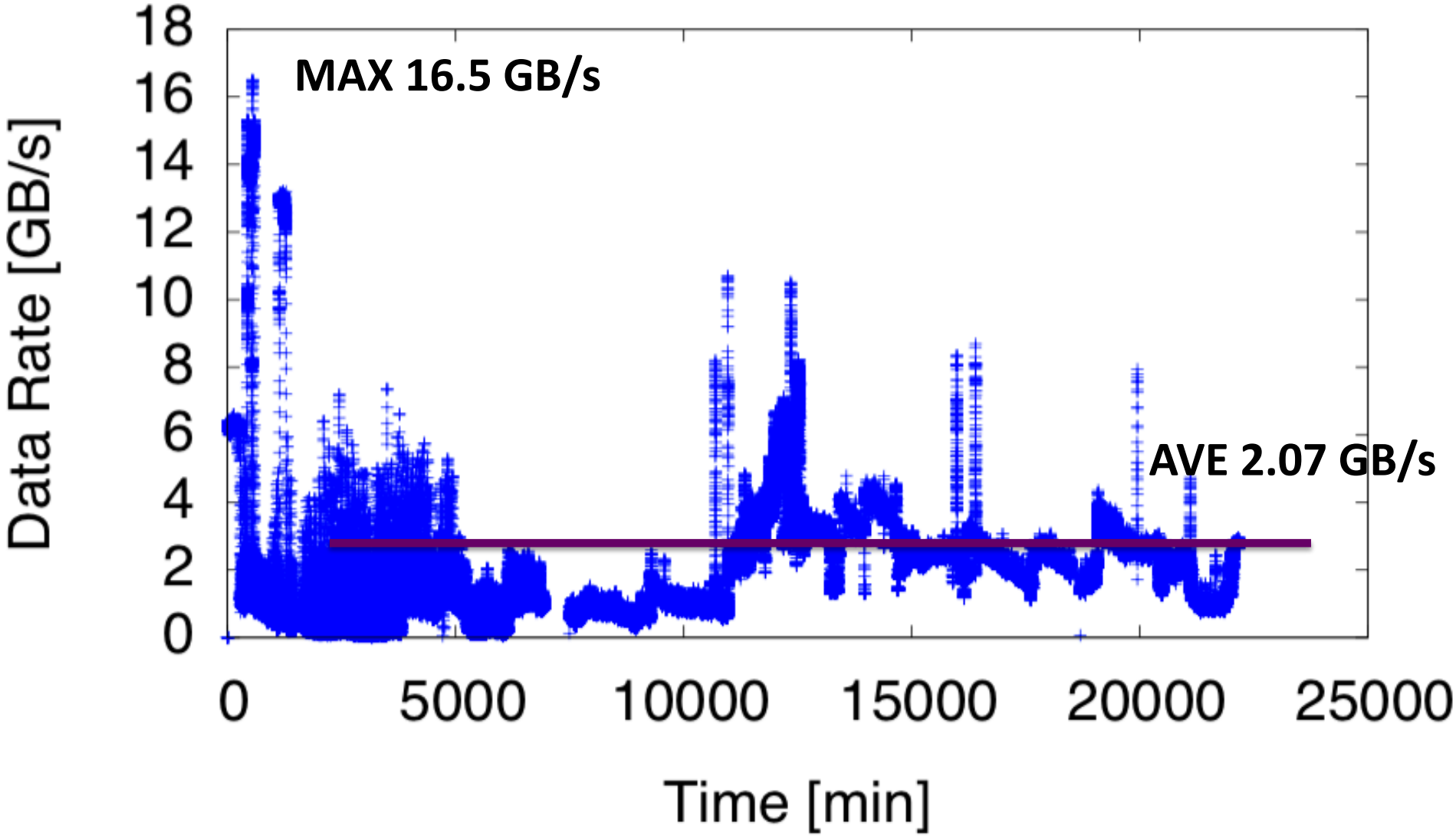
Configuration of SSD Volumes

- Configuration
  - Same as TSUBAME2
  - 2 devices per node * 1408 nodes = 2816 devices
- Assumption
  - Written Data are equally distributed to the aggregated SSD volumes
  - The endurance (TBW/PBW) of the aggregated SSD volumes is equal to the aggregation of the endurance of each SSD device

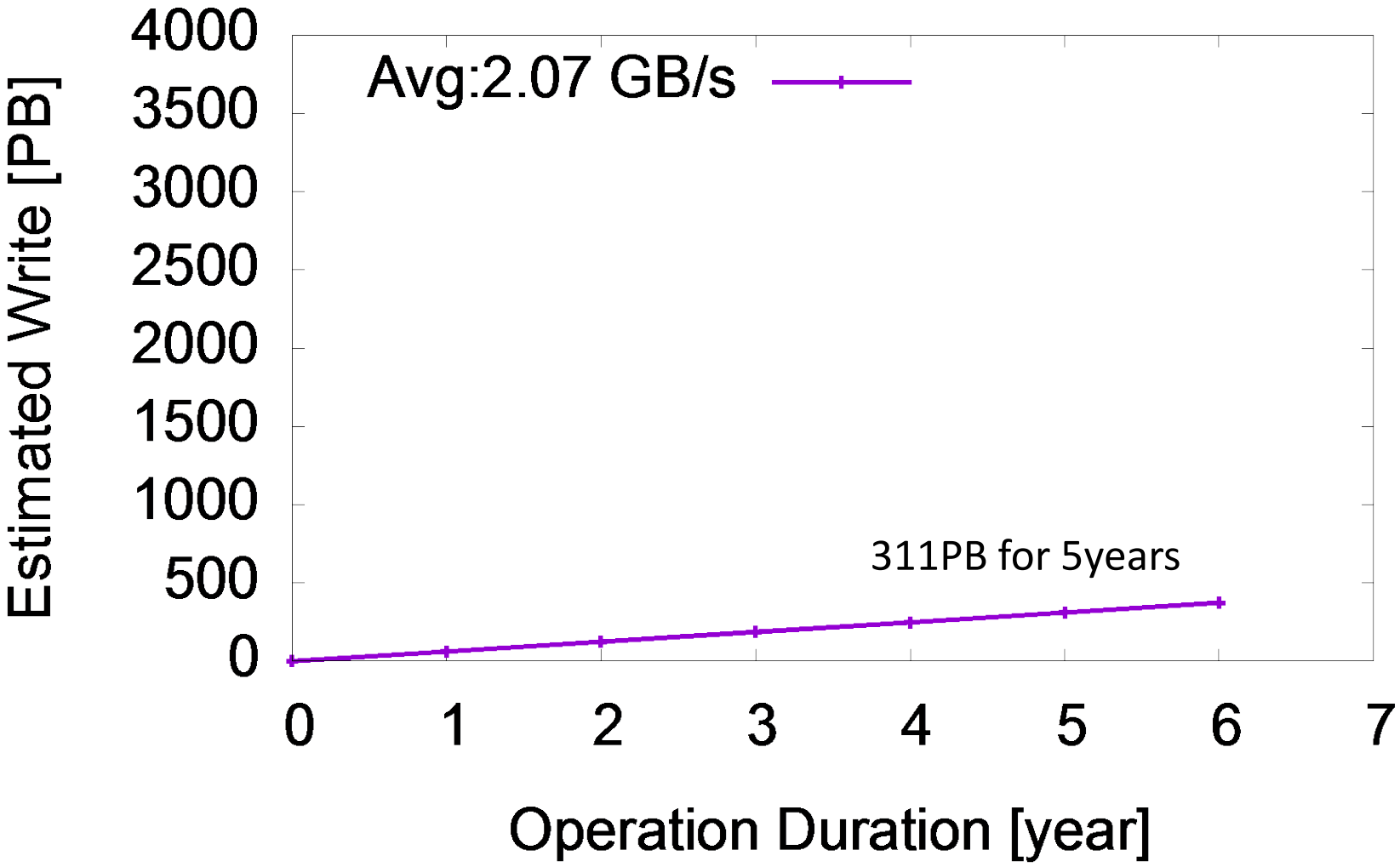
Target Devices



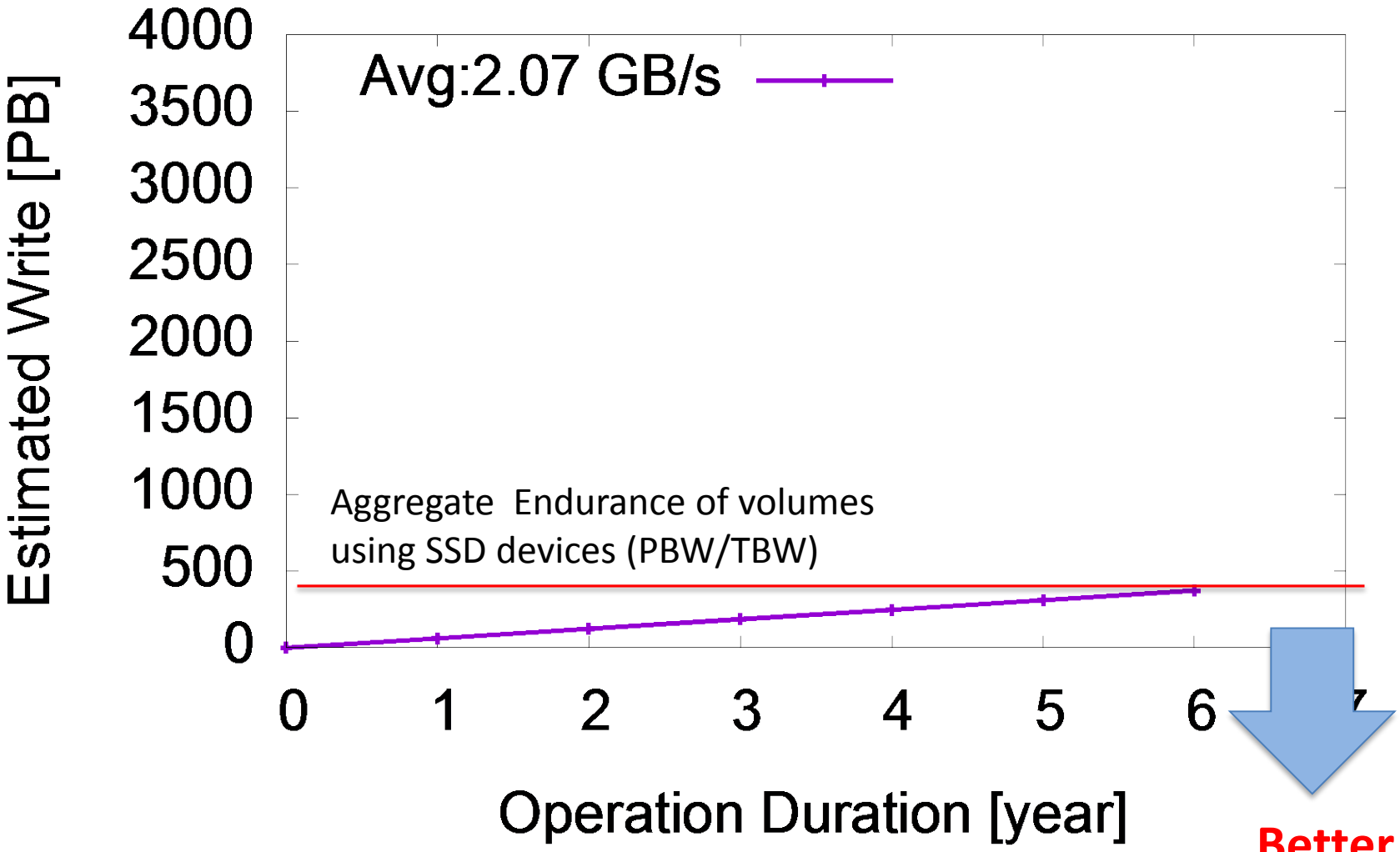
Aggregate Data Rate for TSUBAME2's Lustre Volumes



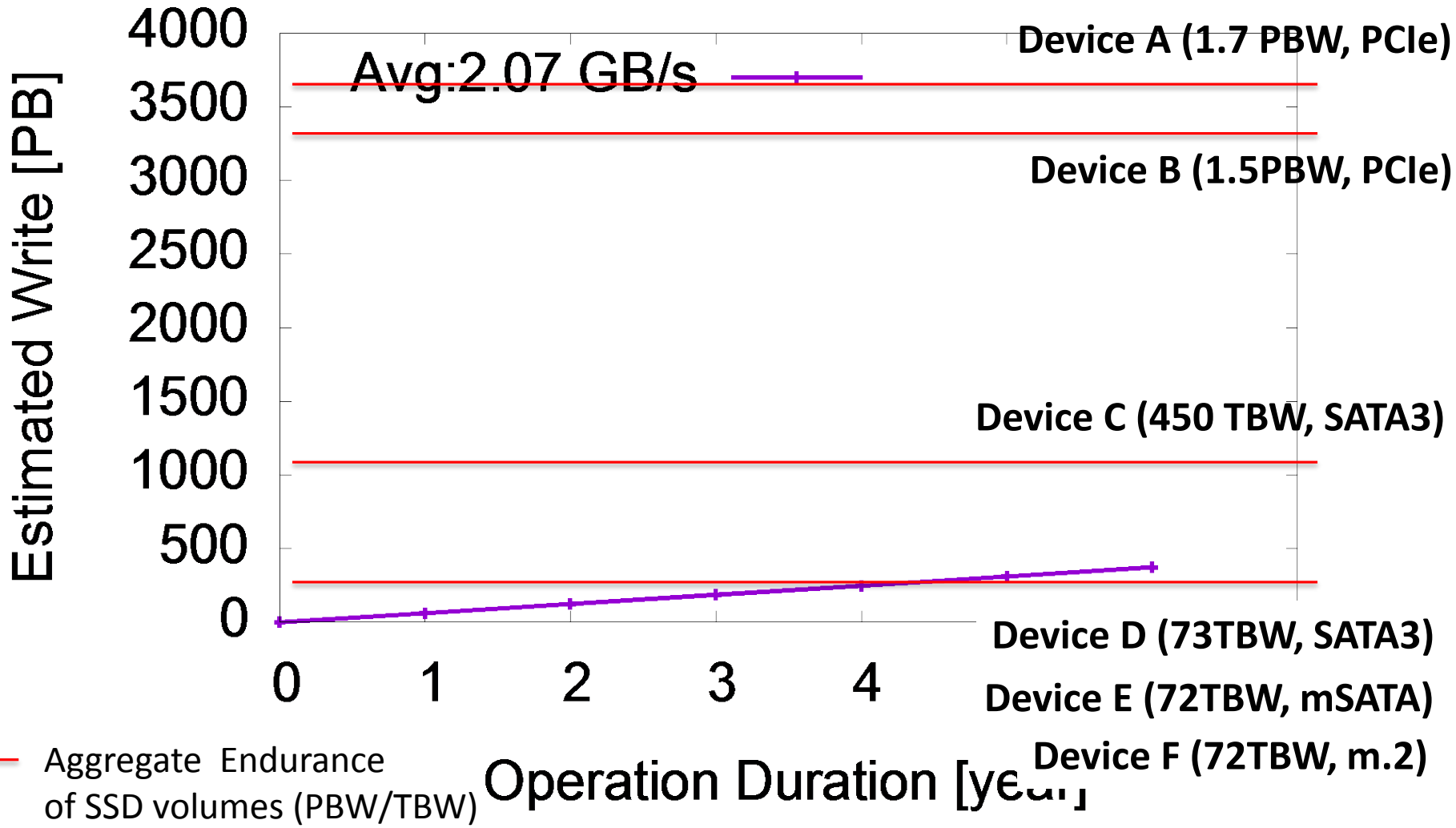
Estimated Write Based on Lustre I/O Workload on TSUBAME2



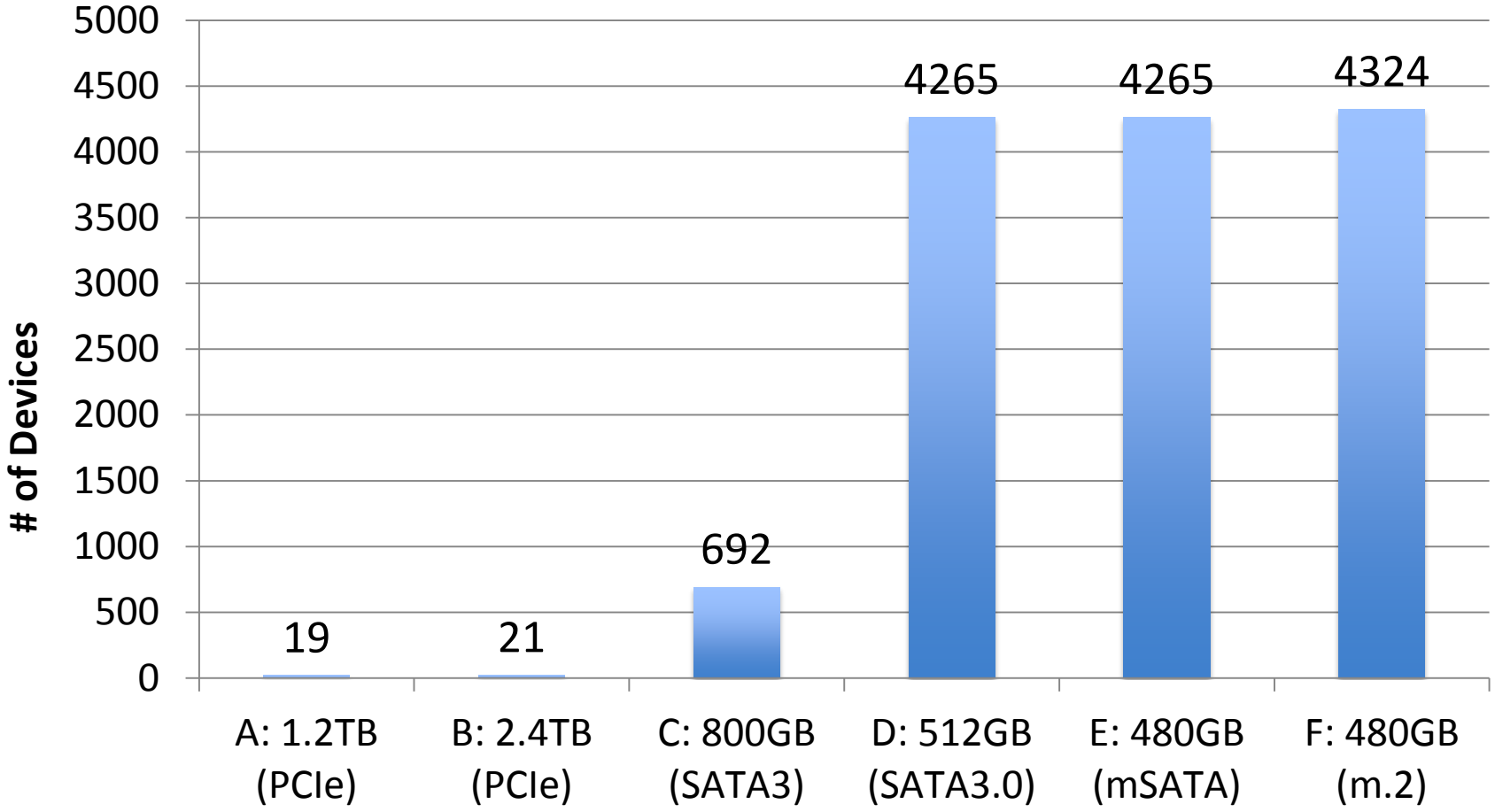
Endurance Based on TSUBAE2 SSD Configuration (Using 2816 devices)



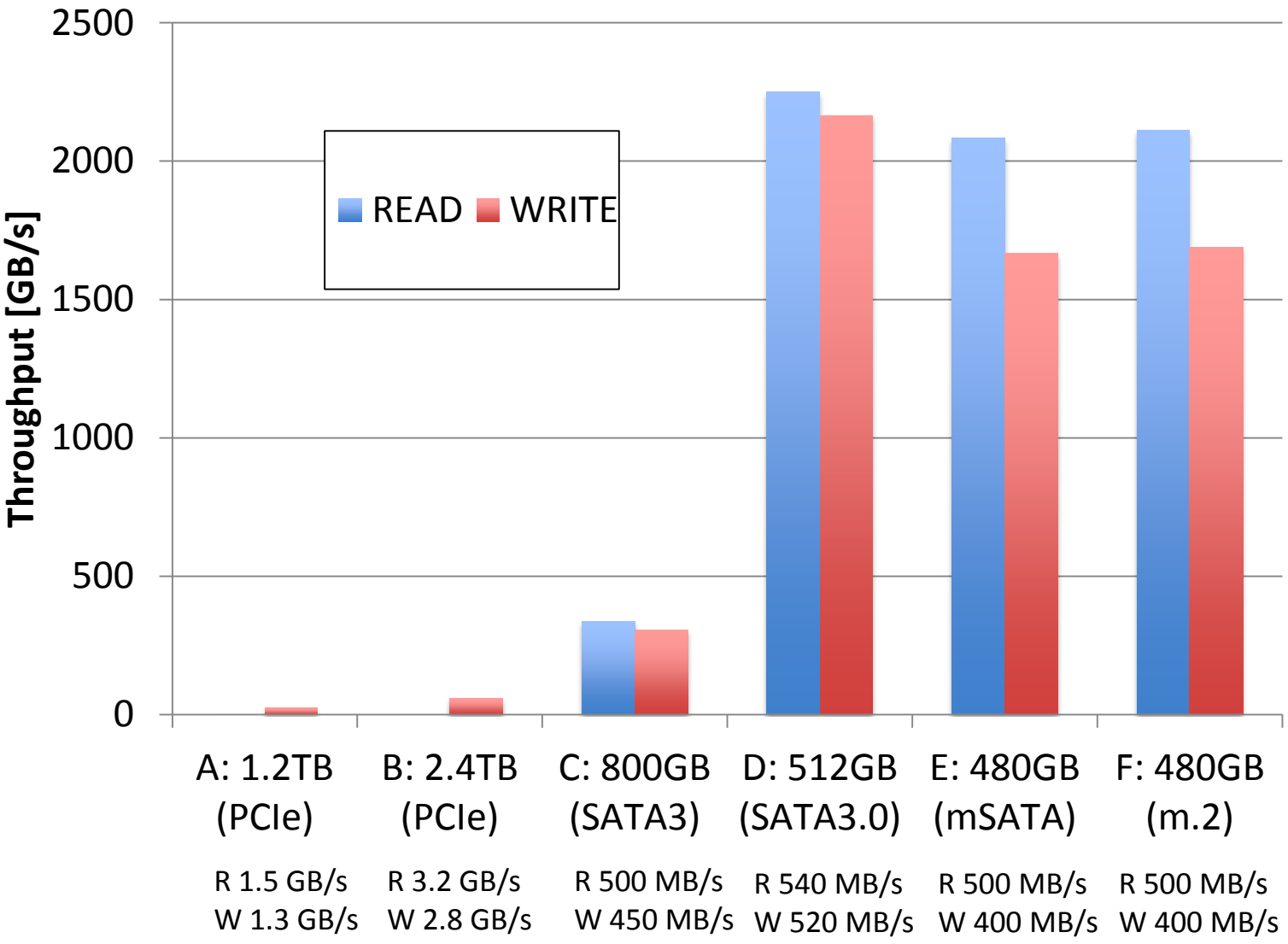
Endurance Based on TSUBAE2 SSD Configuration (Using 2816 devices)



of Required Devices to Support Estimated Total Write Size during Operation Duration (311PB for 5 years)



Aggregate Throughput



Discussion

- We can configure SSD-based buffer volumes using commodity-based SSD devices
  - Based on Lustre Write I/O workload
  - Several thousand of devices needed
- If we use high end SSD devices (PCI-e attached Flash), we can consolidate the SSD devices
  - Required Performance under the limited endurance of the SSD volumes
- Further evaluation by using more detailed I/O behavior is needed
  - Emulation of Burst Buffers etc.

Summary

- Preliminary Evaluation of Endurance for SSD devices to support supercomputing I/O workloads
  - Based on Lustre I/O monitoring

- High-performance Lustre I/O Monitoring for detailed I/O workload analysis