

Mining Supercomputer Jobs' I/O Behavior from System Logs

Xiaosong Ma



جامعة حمد بن خليفة

HAMAD BIN KHALIFA UNIVERSITY

OLCF Architecture Overview

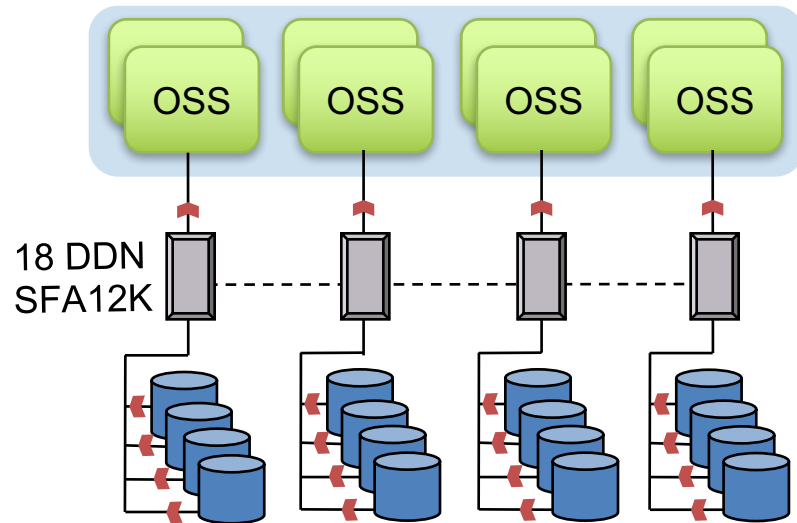


Rhea
512 node
Development
Cluster

Eos
736 Node
Cray XC30
Cluster

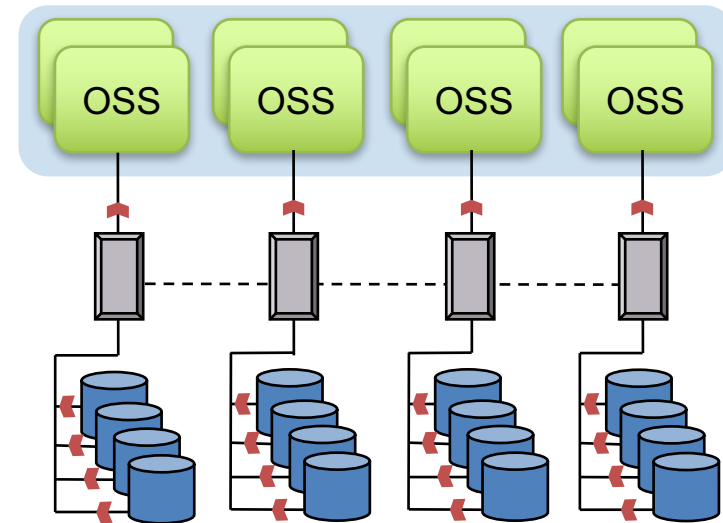
Scalable IO Network (SION) - Infiniband

144 OSS Servers



Atlas1

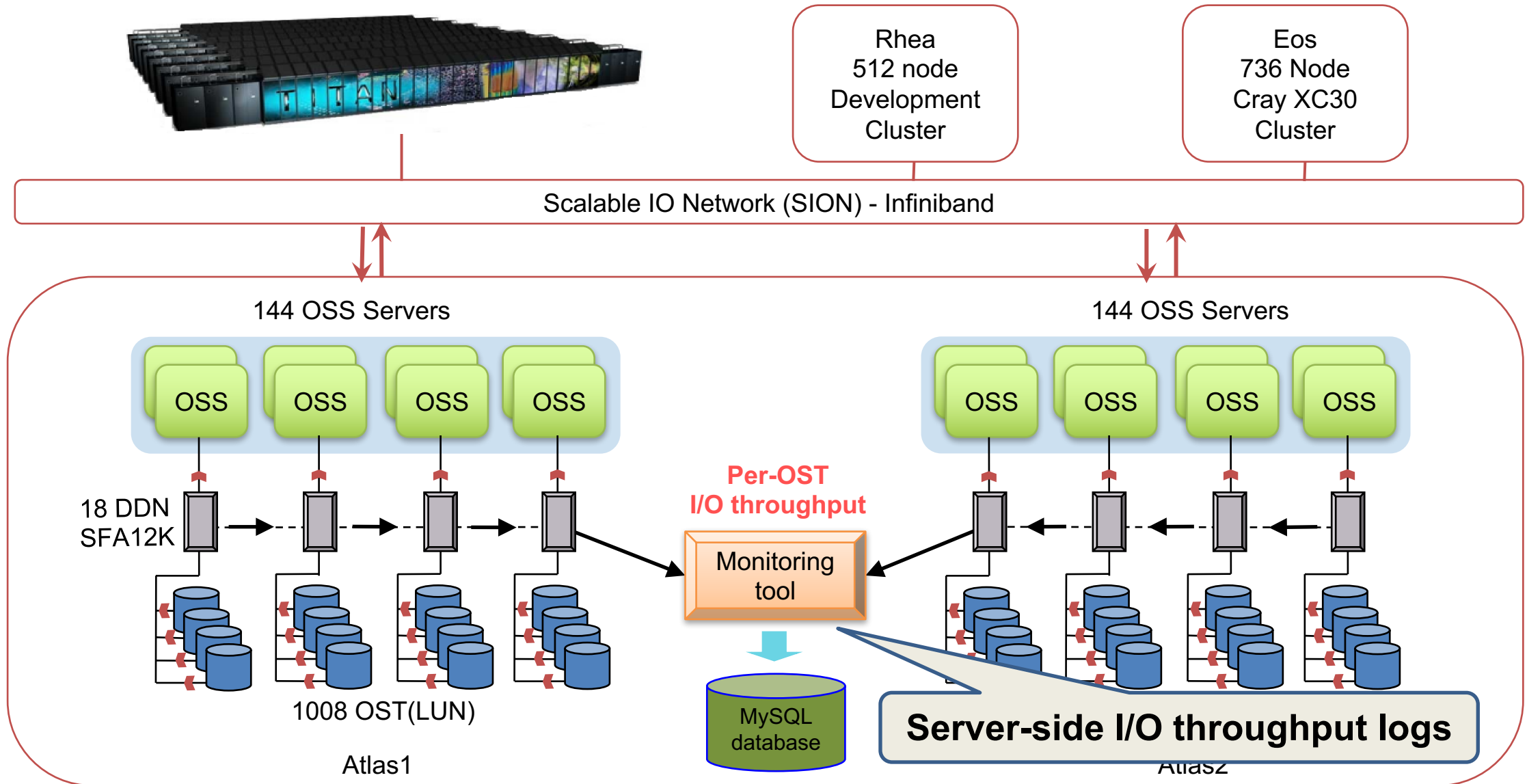
144 OSS Servers



1008 OST(LUN)

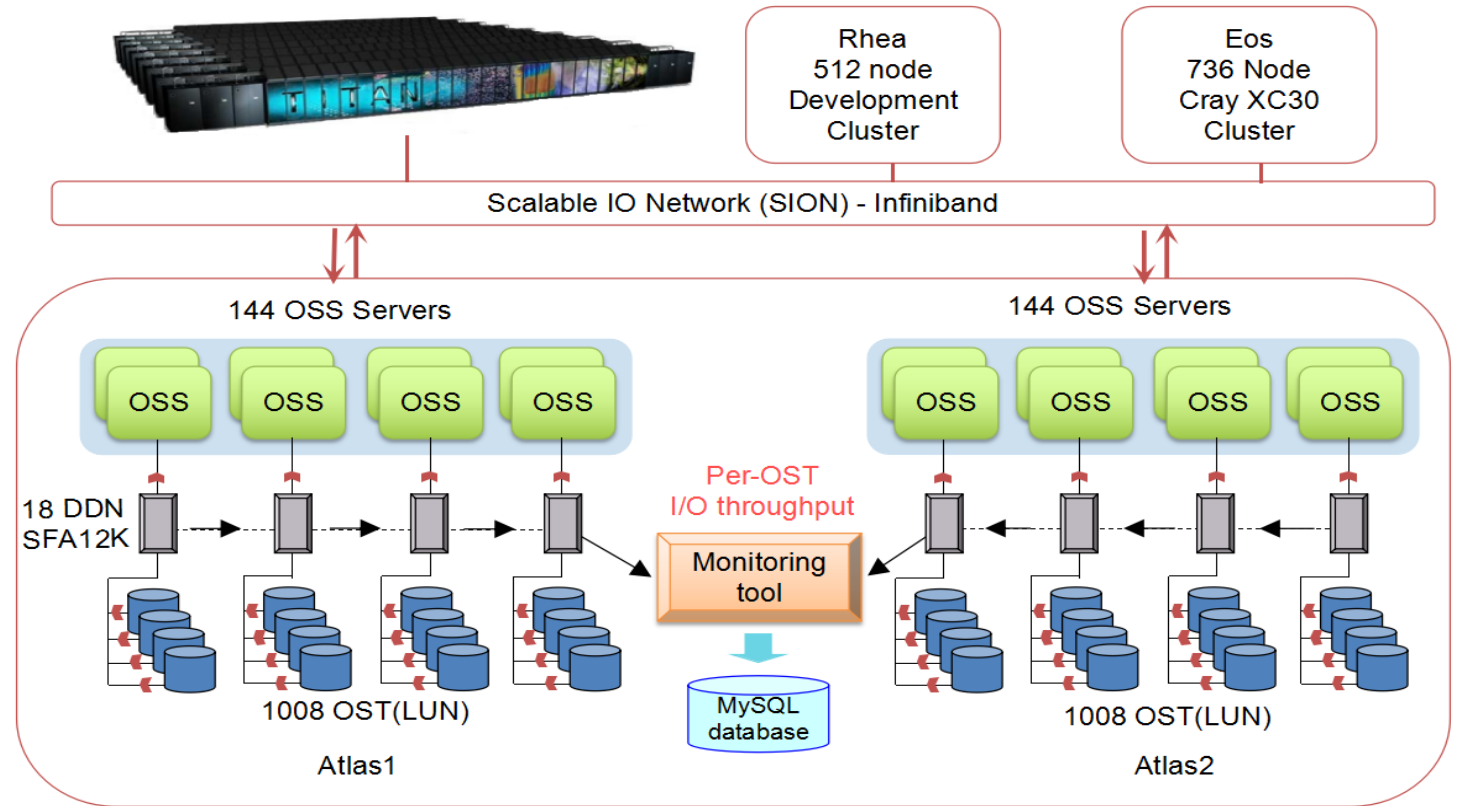
Atlas2

OLCF Architecture Overview

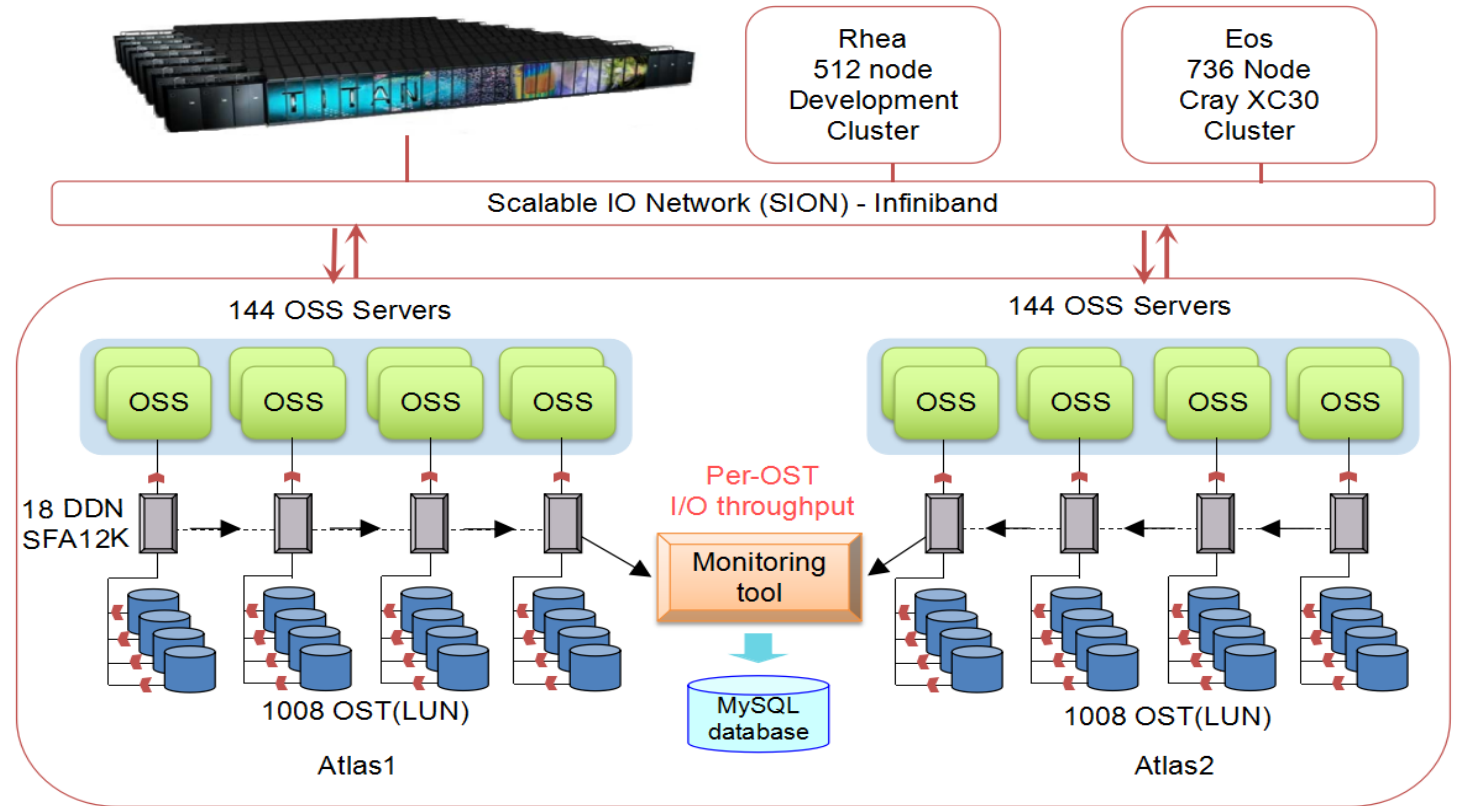


Server-side I/O Throughput Logs

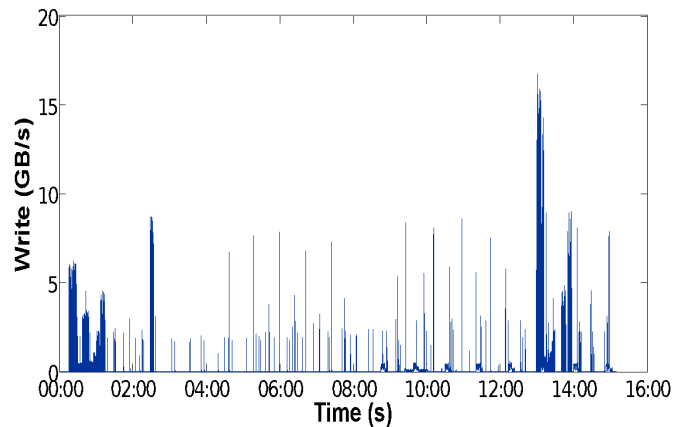
**RAID controller
Coarse-granule logging**



Server-side I/O Throughput Logs



RAID controller Coarse-granule logging



Server-side I/O Throughput Logs

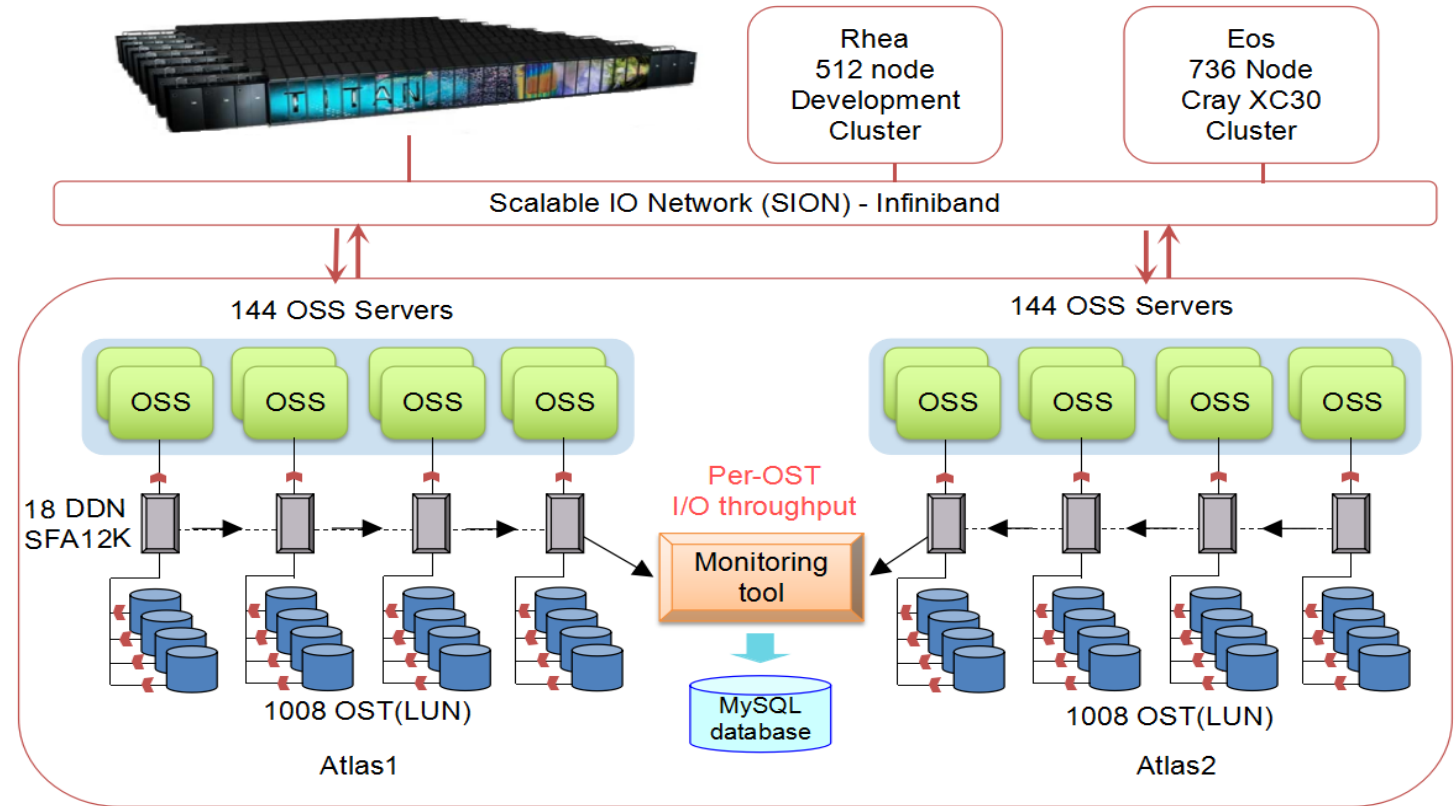
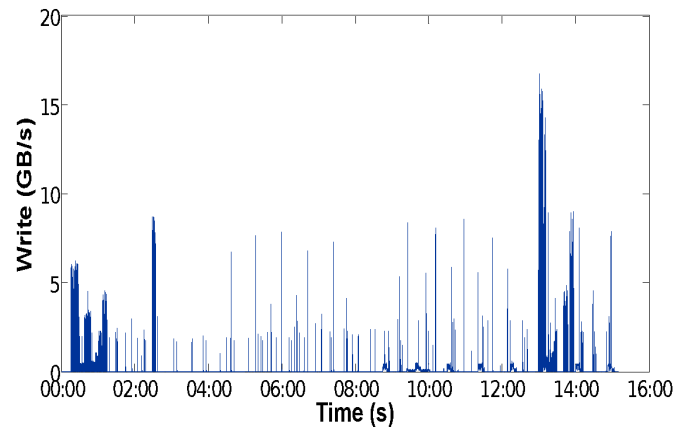
I/O throughput logs

Zero overhead

No impact on user IO

No user effort

RAID controller Coarse-granule logging



Server-side I/O Throughput Logs

I/O throughput logs

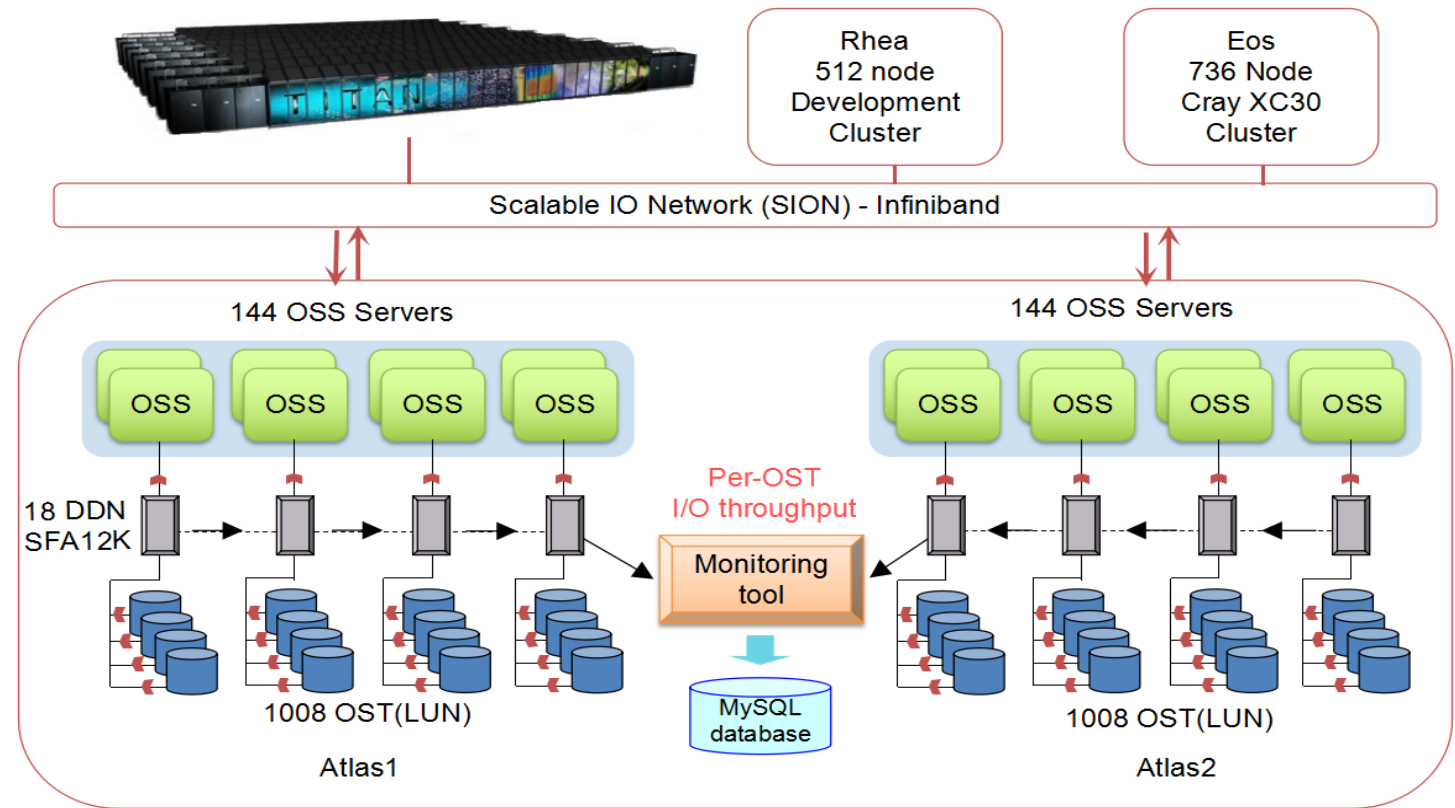
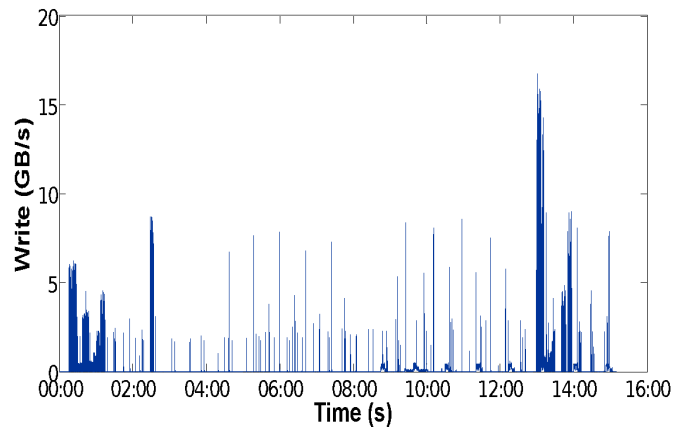
Zero overhead

No impact on user IO

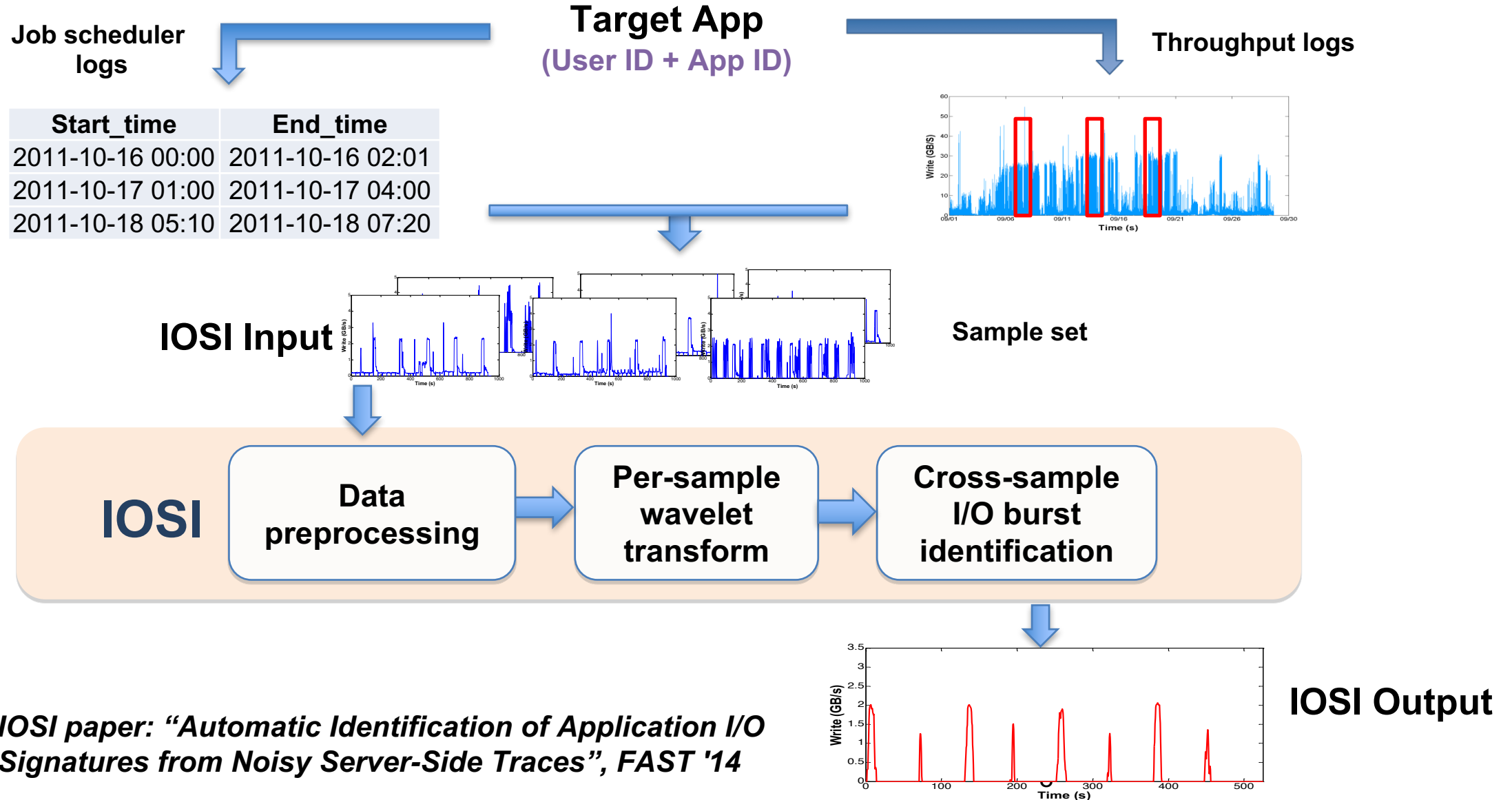
No user effort

Mixed I/O traffic

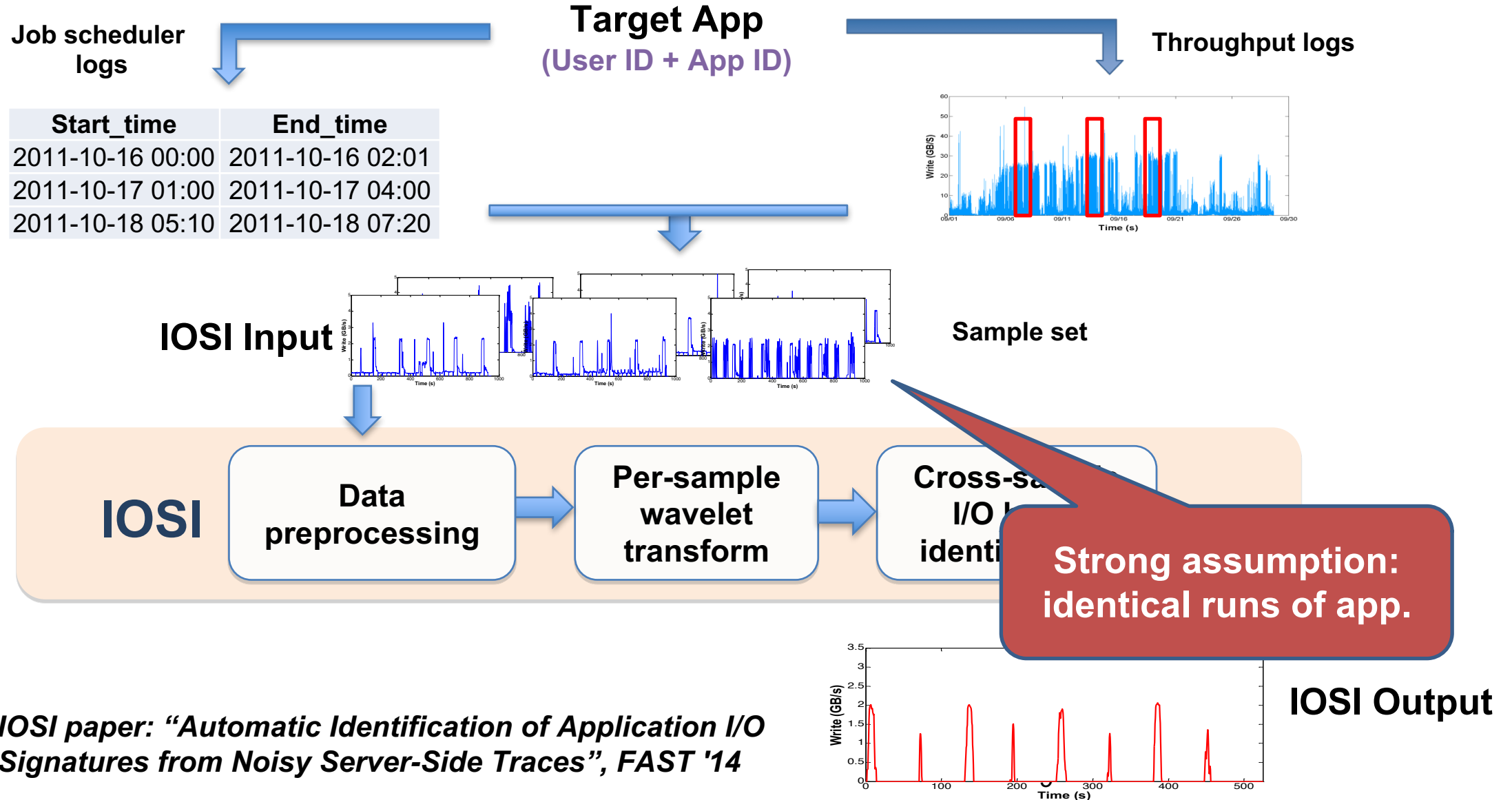
RAID controller Coarse-granule logging



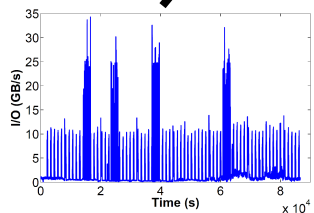
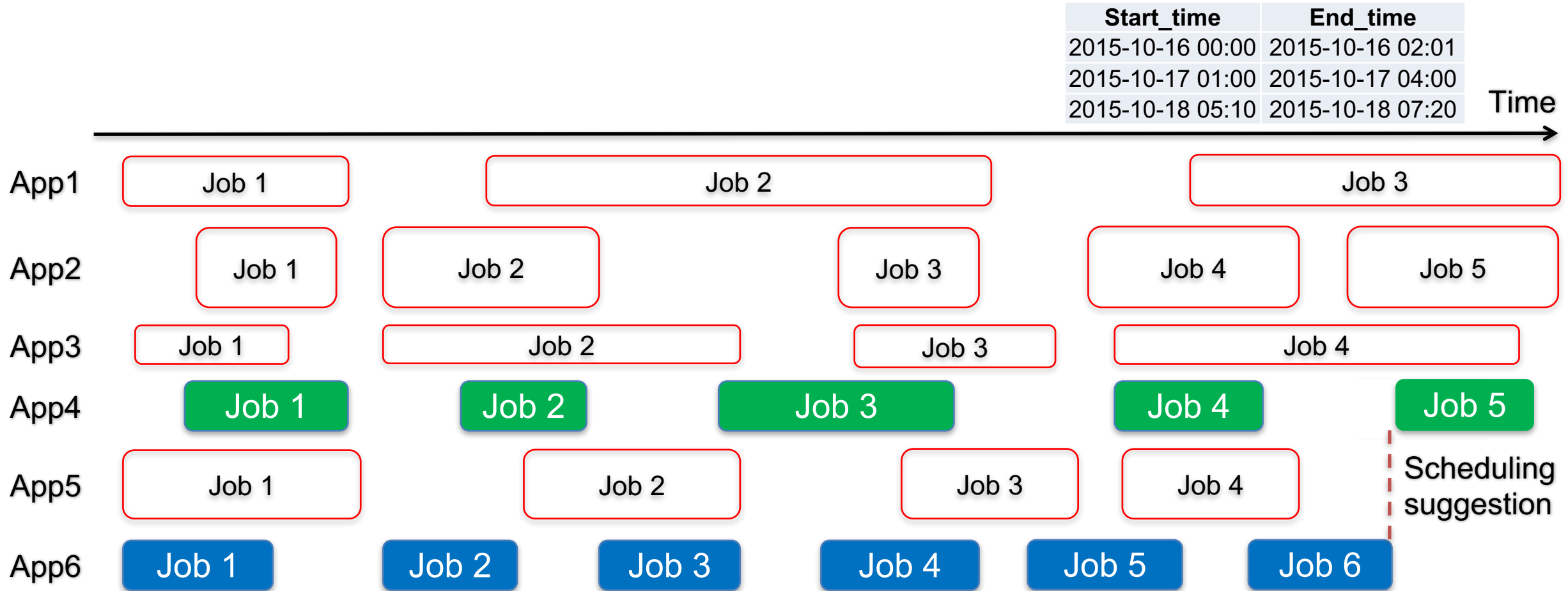
Prior Work: IOSI Workflow



Prior Work: IOSI Workflow



AID: Automatic I/O Diverter



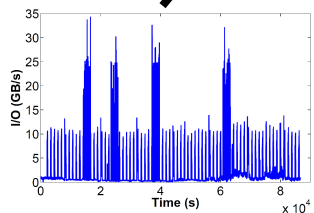
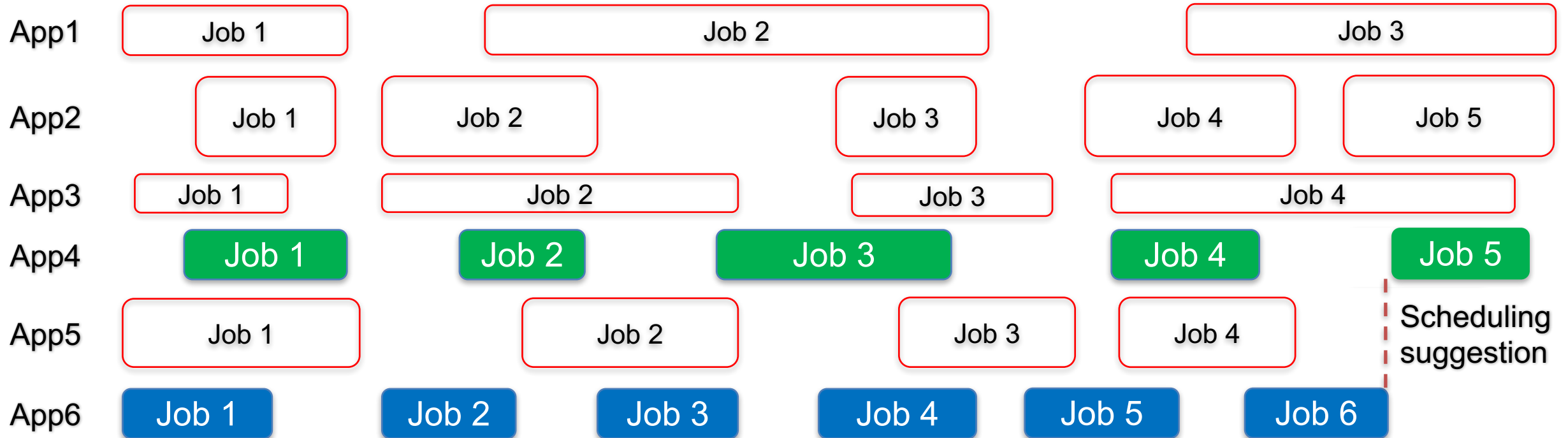
**Automatically identifying I/O-heavy apps
(No prior knowledge, no user involvement)**

AID: Automatic I/O Diverter

SC|16 Tech paper presentation:
Thursday 2pm, 355D

Start_time	End_time
2015-10-16 00:00	2015-10-16 02:01
2015-10-17 01:00	2015-10-17 04:00
2015-10-18 05:10	2015-10-18 07:20

Time →



Automatically identifying I/O-heavy apps
(No prior knowledge, no user involvement)

Application I/O Characterization Results

Name	Value
Total number of logged jobs	181,969
Unique applications identified	9,998
Initial I/O-intensive candidates	95
Candidates passing scope checking	67
Candidates passing minimum support	42
User-verified candidates	8

**Result from 5 months' Titan I/O traffic and job logs
(User verification by email)**

Application I/O Characterization Results

Name					Value
Total number of logged jobs					181,969
Unique applications	ID	Node	Time(m)	OST	App. Domain
Initial I/O-intensive c	1	8192	1440	64	Geo-sciences
Candidates passing	2	250	6-60	1008	Combustion
Candidates passing	3	2048	30-185	1008	Astrophysics
User-verified candid	4	1760	720	180	Combustion
	5	1024	110-230	1008	Systems research
	6	200	30-190	1008	Combustion
	7	1008	13-17	1008	Computer Science
	8	16388	43-310	800	Environmental

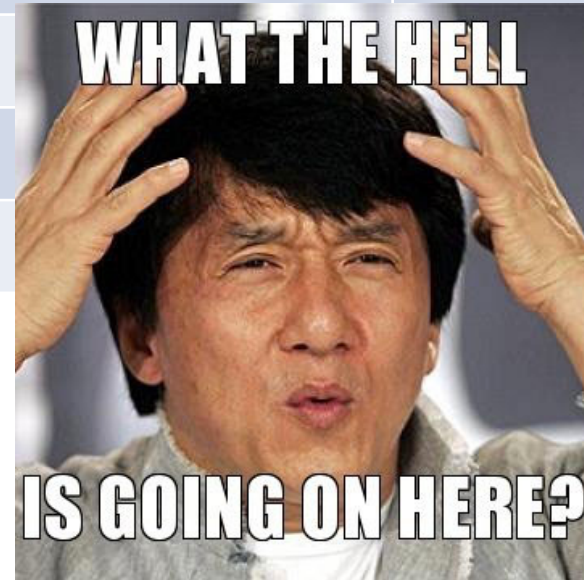
User-verified I/O-intensive applications

Application I/O Characterization Results

Name	Value
Total number of logged jobs	181,969
Unique applications identified	9,998
Initial I/O-intensive candidates	95
Candidates passing scope checking	67
Candidates passing minimum support	42
User-verified candidates	8

Application I/O Characterization Results

Name	Value
Total number of logged jobs	181,969
Unique applications identified	9,998
Initial I/O-intensive candidates	95
Candidates passing scope checking	67
Candidates passing minimum support	42
User-verified candidates	8



Application I/O Characterization Results

Name	Value
Total number of logged jobs	181,969
Unique applications identified	9,998
Initial I/O-intensive candidates	95
Candidates passing scope checking	67
Candidates passing minimum support	42
User-verified candidates	8

Applications not using parallel I/O systems well!

- Similar finding as Huong 2015 HPDC work (Darshan)
- Motivates better I/O performance data analysis
- Connecting programs to systems

Questions?

Xiaosong Ma

xma@qf.org.qa

Qatar Computing Research Institute, Hamad Bin Khalifa University

I/O Contention on Large-Scale HPC Systems

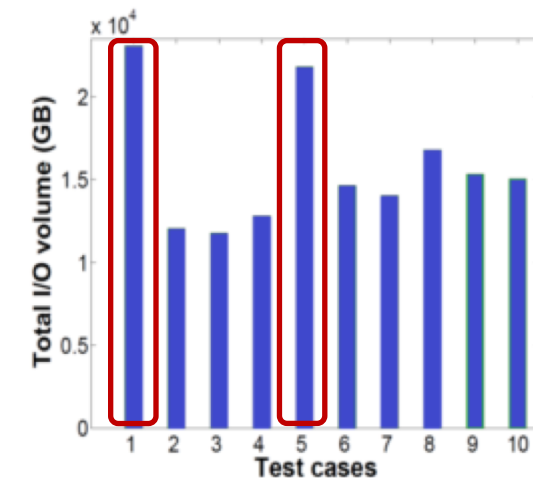
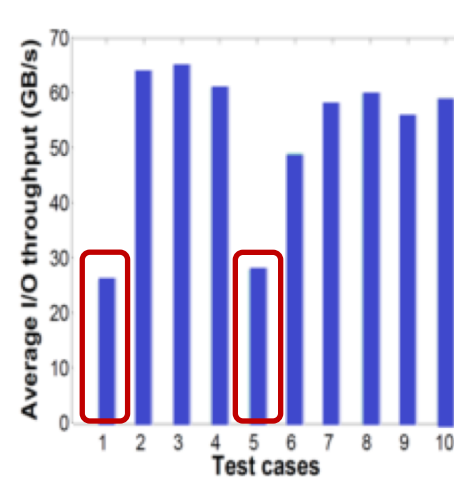


ORNL's Titan
(World's #3 Supercomputer)

- 27.1 PF Peak performance
- 18,688 compute nodes
 - 16-core AMD Opteron
 - Nvidia Tesla GPU
 - 32 + 6 GB memory
- 3-D Torus interconnect

Performance variance on HPC

- Shared parallel file system
- I/O-heavy jobs collision -> I/O performance degradation



I/O performance variance on Titan with IOR [6]

CDF of per-OST I/O throughput

