SDC 20
SNIA INDIA
Storage Developer Conference
December 4-5, 2020
*BY Developers FOR Developers*

# Optimized Storage and Network layer for Next Gen Industry workloads

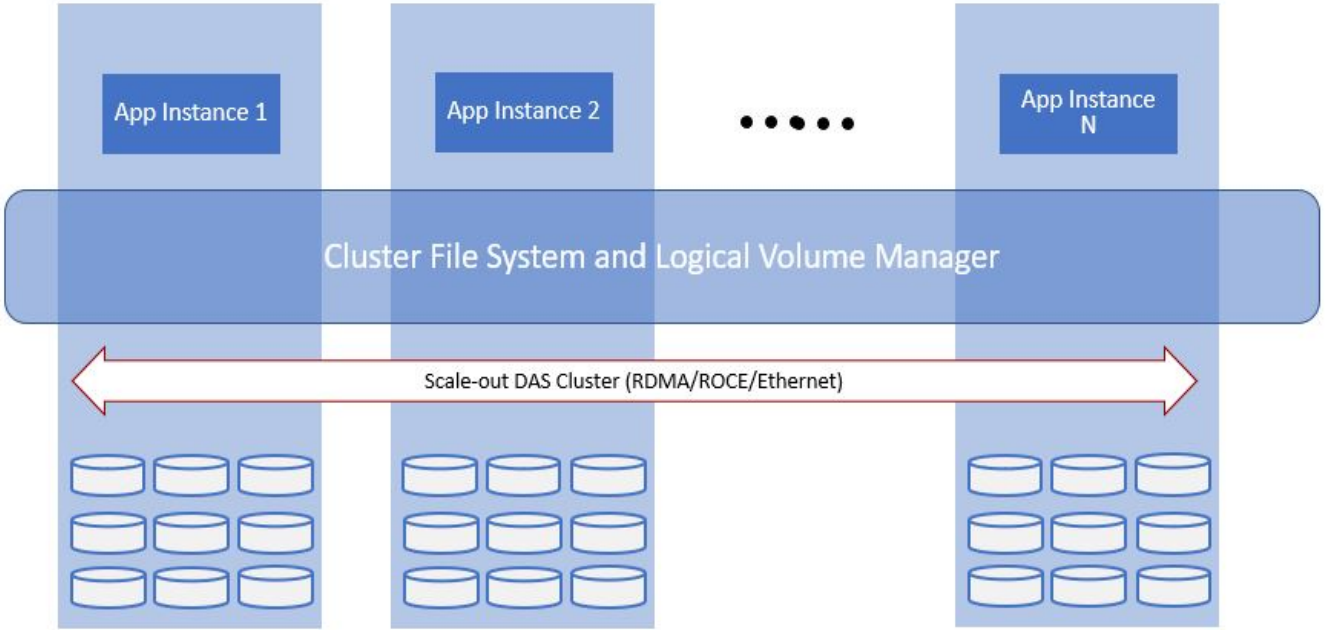
Ajinkya Nakave
Prajakta Zagade
Sumit Dighe
**Veritas**

# Agenda

- Brief overview
- Challenges involved
- Solutions
- Performance study
- Conclusion

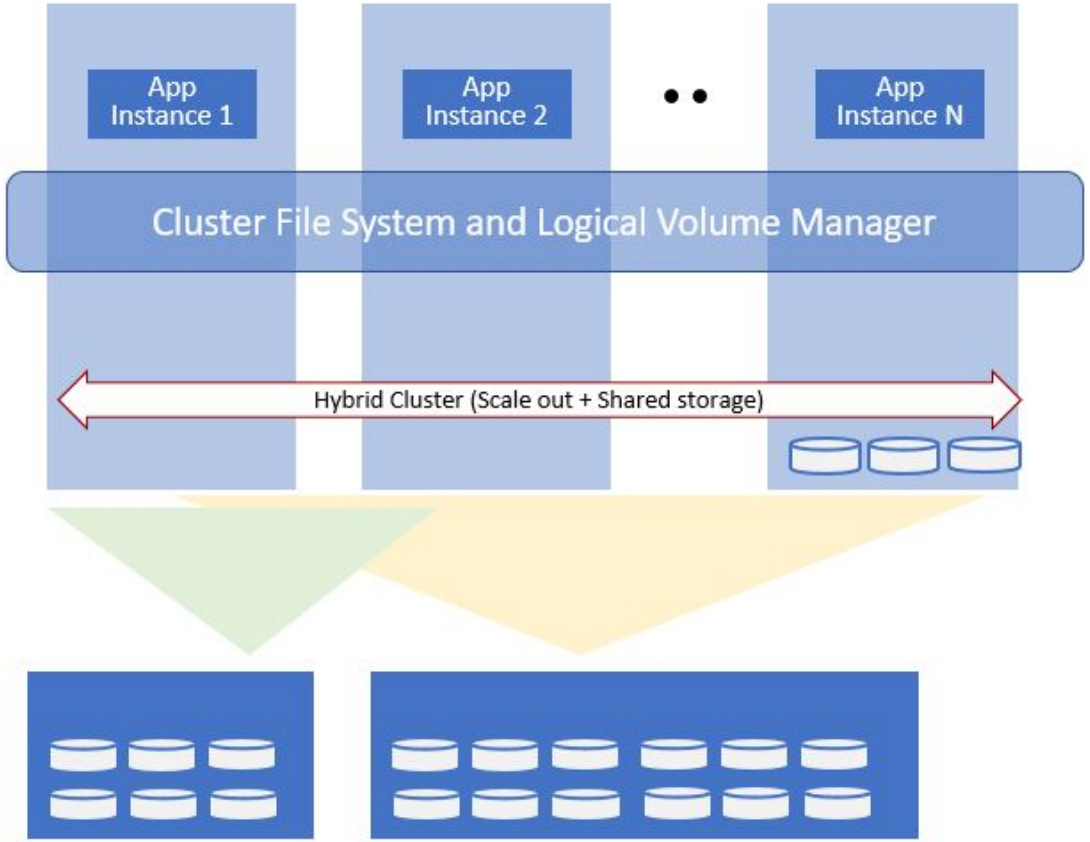
2020 Storage Developer Conference India. © Insert Your Company Name.  All Rights Reserved.

# Hyper-Converged Architecture




2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# Dis-aggregated Architecture




2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

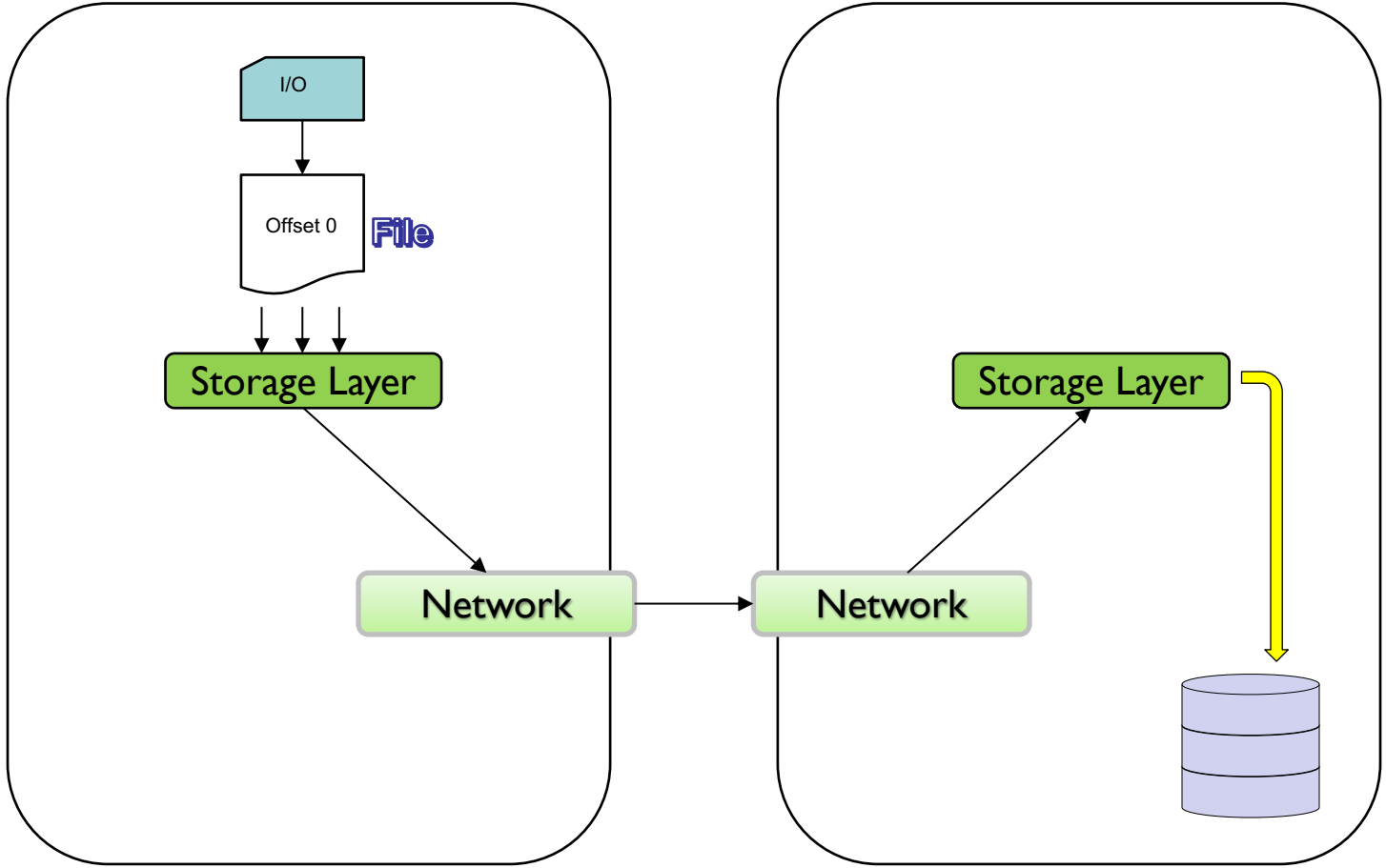
# Possible Optimizations

- Multithreading
- Batching
- Reduced context switching
- Use of multiple Q pairs/connections
- Load balancing across ports
- Is strict flow control at network layer needed ?

2020 Storage Developer Conference. © Insert Your Company Name. All Rights Reserved.

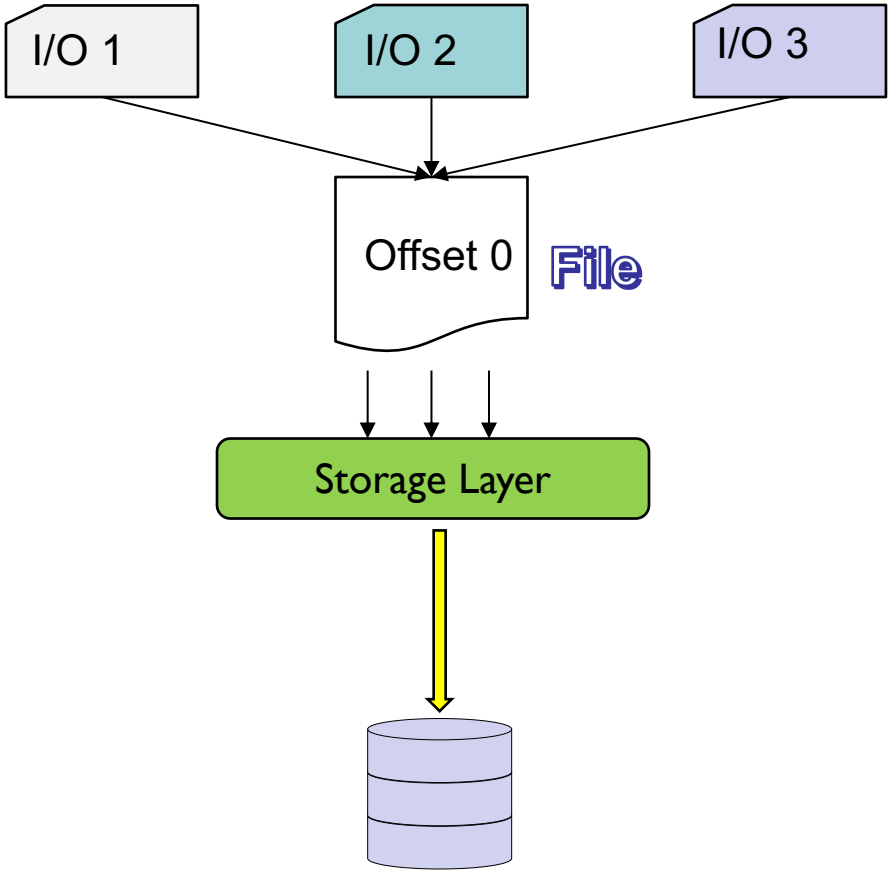
# Deep dive

# How IO Flows




2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# Ordering of I/Os



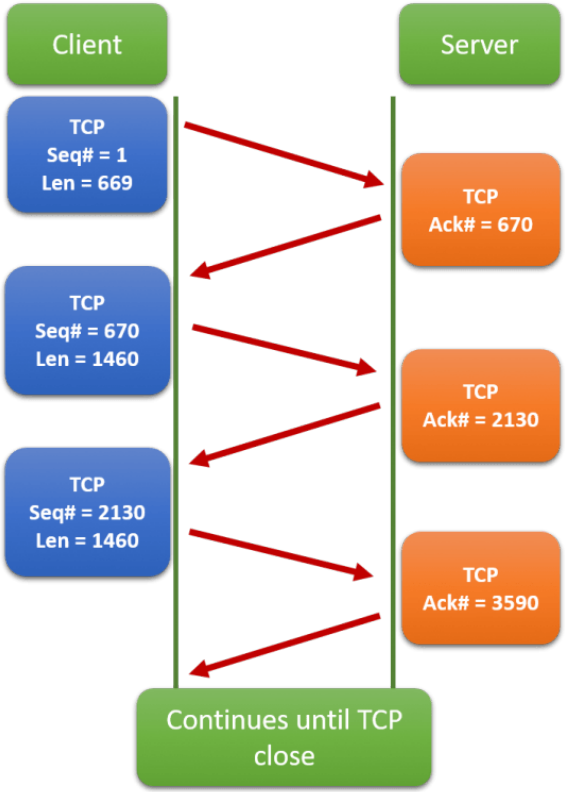

2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# Packet Handling

- Network Layer
  - Ordering of packets
  - Flow control
  - Guaranteed Delivery (Data Resends-Duplicate detection)

2020 Storage Developer Conference. © Insert Your Company Name. All Rights Reserved.

# Ordering of packets




2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# Ordering-Sequencing

- Ordering requires
  - Complex protocols: consumes CPU and increased latency
- Out of order delivery may cause
  - corruption to data
  - Some protocols cause retransmissions
  - No of acks on network increases due to smaller window
- Found duplication in handling IO Ordering:
  - File System layer: Using range locking
  - Network layer: Using Sliding Window Protocol

2020 Storage Developer Conference. © Insert Your Company Name. All Rights Reserved.

# Flow control

- To cope up with fast Sender
- Found duplication in handling Flow control:
  - Application IO:
    - Most of the applications issue finite IOs.
    - Next set of IOs wait till previous comes back
  - Network layer: Using Sliding Window Flow control Protocol
  - Receiver window field can be adapted to higher values to avoid unnecessary flow control

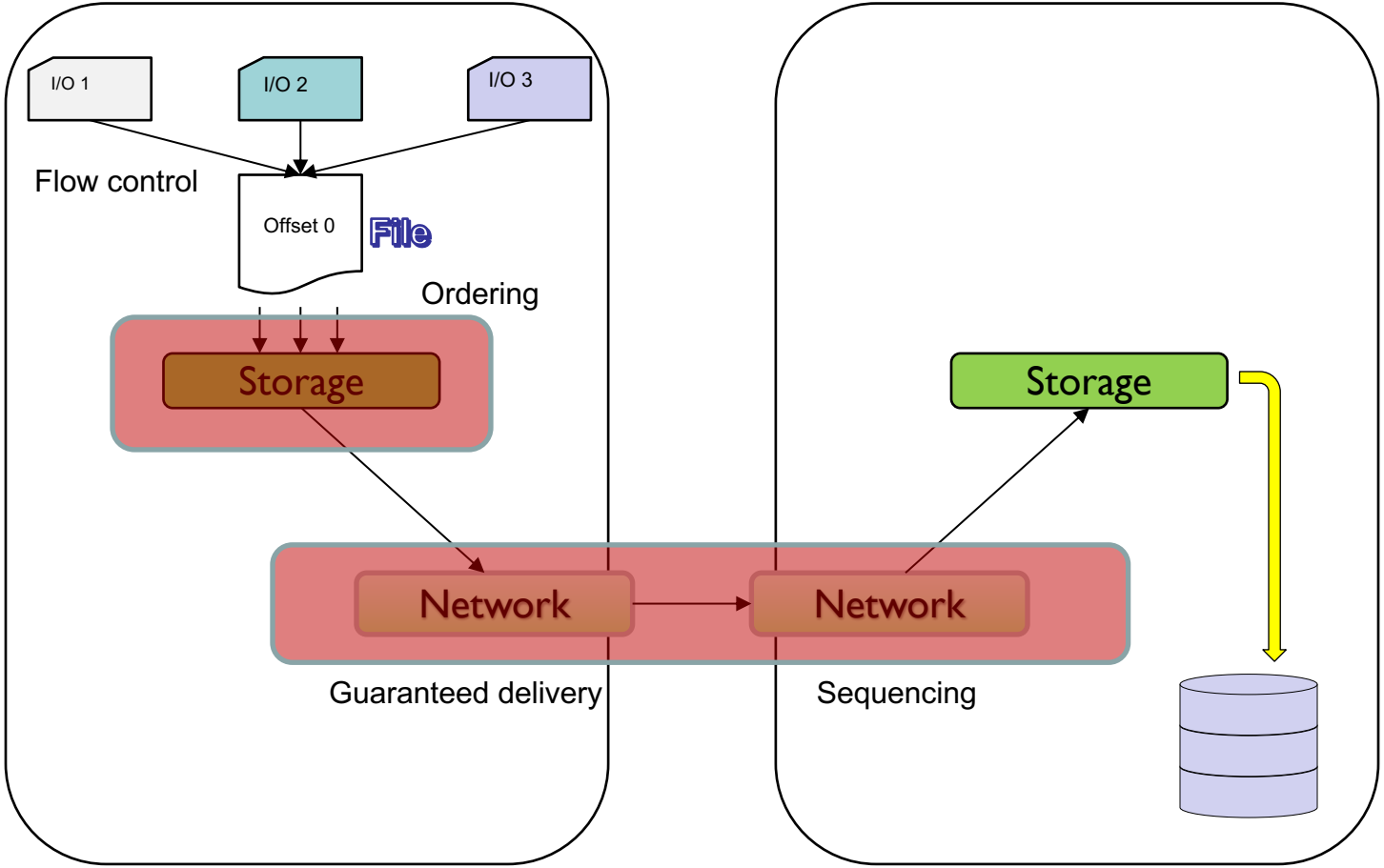
2020 Storage Developer Conference. © Insert Your Company Name. All Rights Reserved.

# Guaranteed Delivery

- Packets may get dropped
- Guaranteed delivery has to ensure data to be delivered
- Sequence numbers allow receivers to discard duplicates
- Network layer protocol is best place to handle this

2020 Storage Developer Conference. © Insert Your Company Name. All Rights Reserved.

# How IO Flows



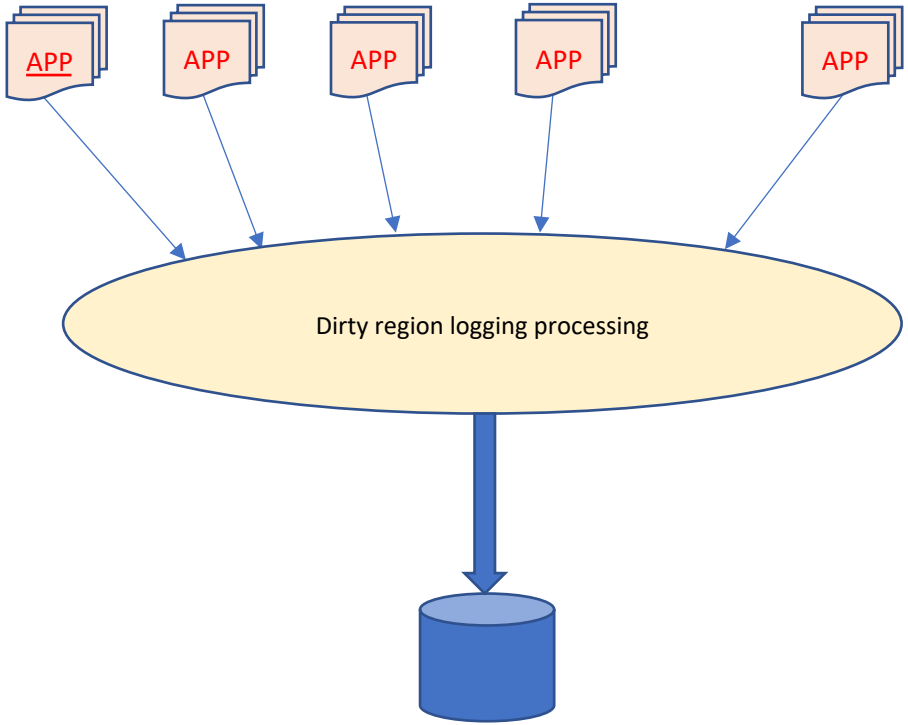

2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# How to optimize I/O Paths

- Sequencing-Ordering
- Flow control
- Guaranteed Delivery (Data Resends-Duplicate detection)
- Resiliency / Storage Redundancy (RAID levels, Erasure coding)

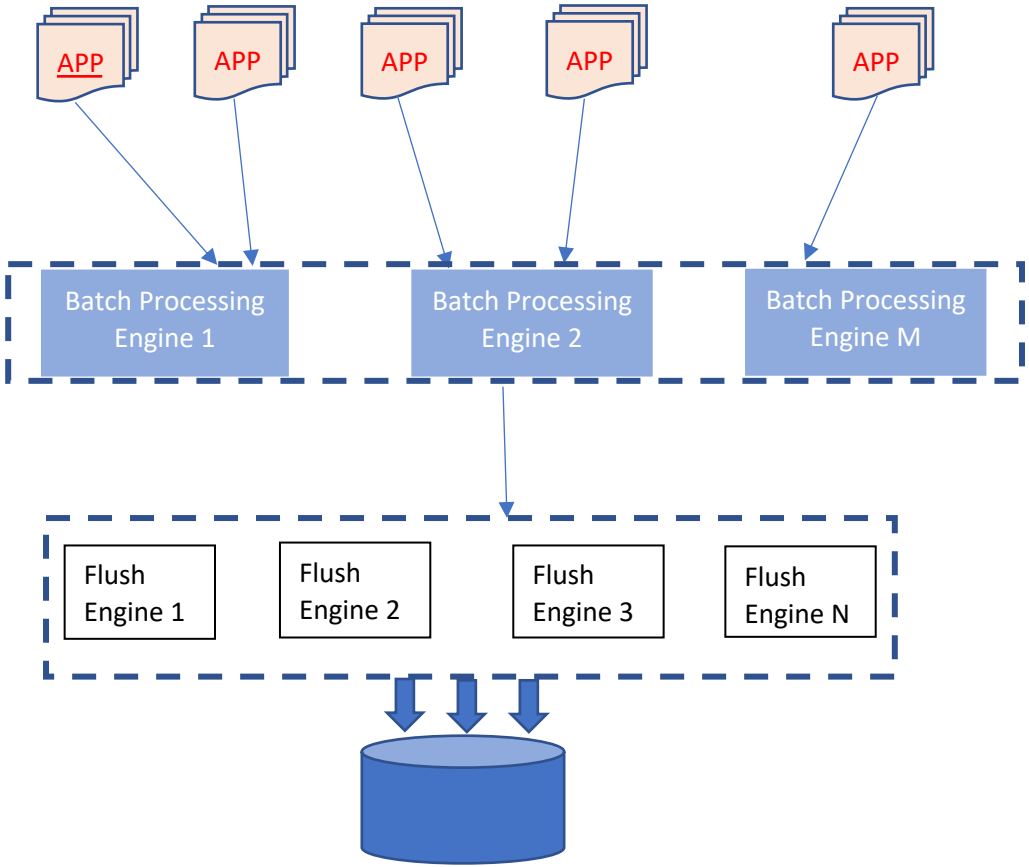
2020 Storage Developer Conference. © Insert Your Company Name. All Rights Reserved.

# Mirroring Overhead: Dirty region tracking




2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# Mirroring Overhead: Optimizations




2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

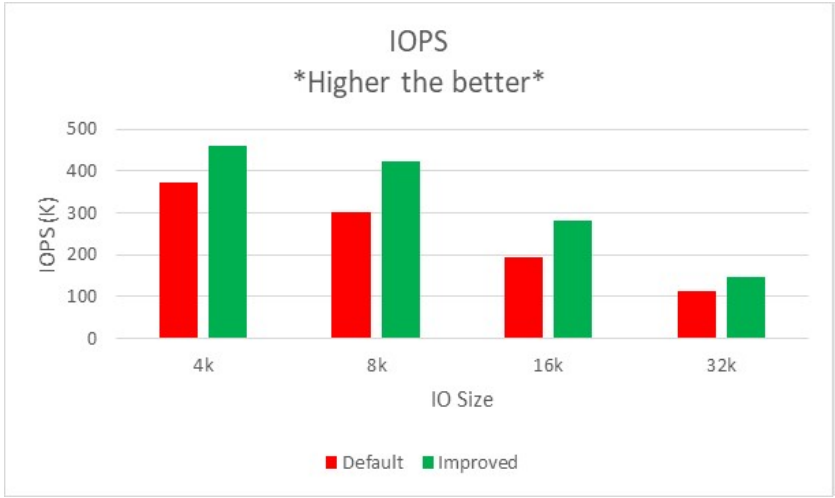
# Additional considerations

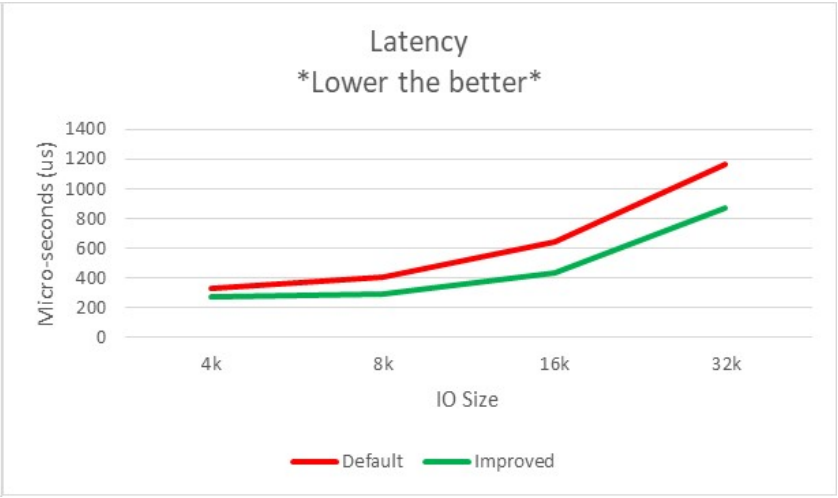
- Receive Side Scaling (RSS)
- Zero copy
- blk-mq for SSD devices
- TCP – UDP offloading
- Multiple socket / queue-pair (RDMA)
- NUMA aware handling for CPU load balancing

2020 Storage Developer Conference. © Insert Your Company Name. All Rights Reserved.

# Performance Study

# Performance Study: Results with no sequencing at N/W layer



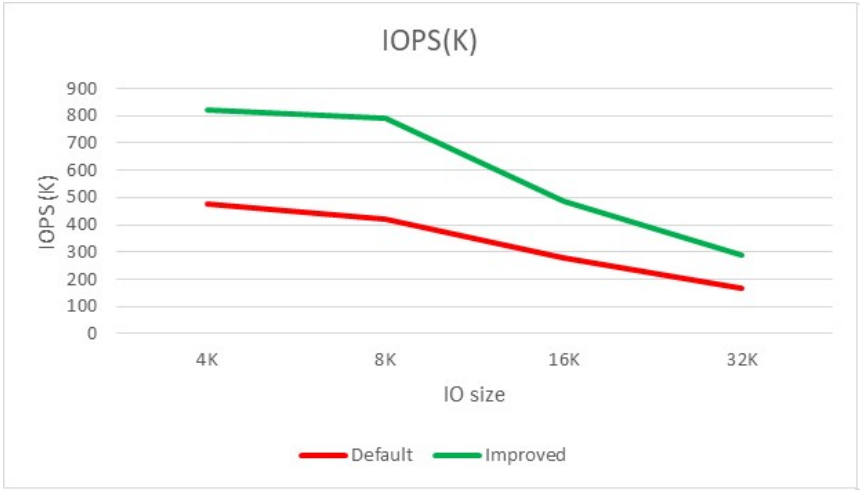

- Default: Sequencing enabled at N/W layer
- Improved: No sequencing at N/W layer

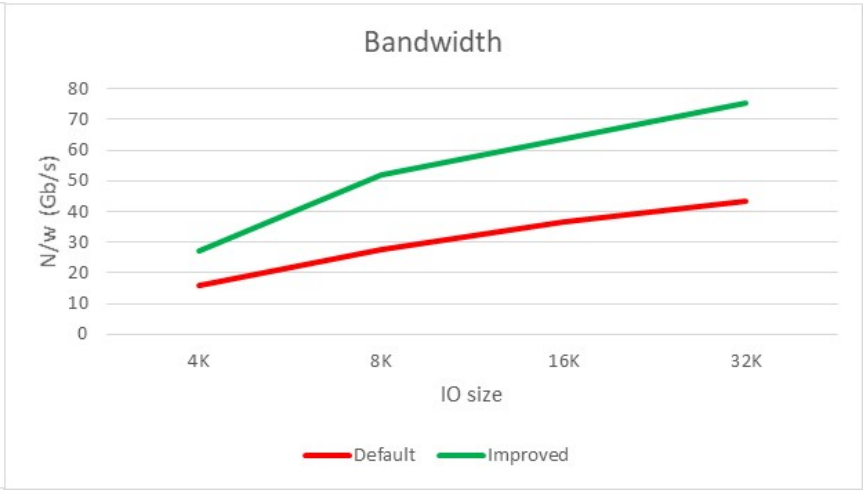
Workload Details:

- Fio – libaio ioengine, iodepth=8, numjobs=8
- **Storage: loopback device** (to avoid storage bottleneck)

2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# Performance Study: Results with all optimizations



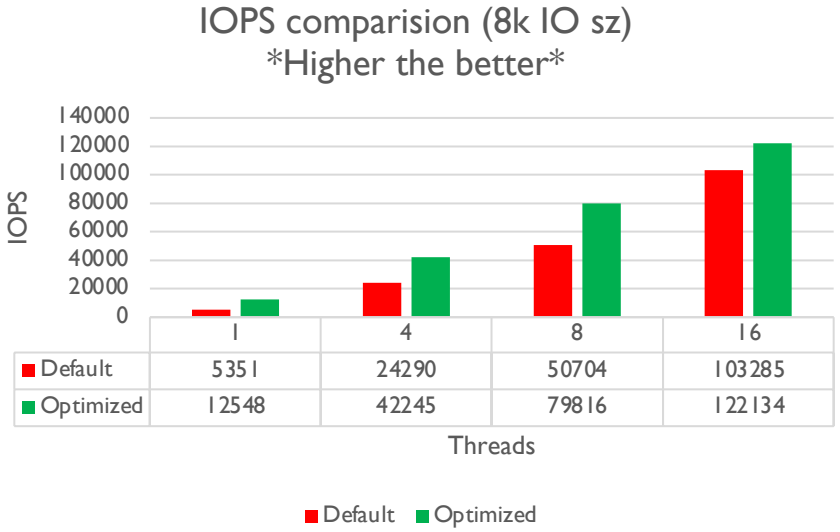

- Default: No optimizations
- Improved: Consists all optimizations discussed so far

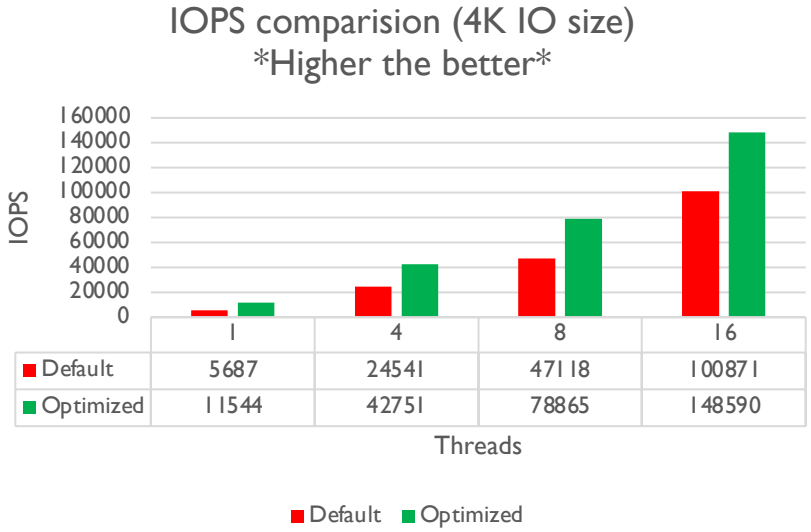
Workload Details:

- Fio – libaio ioengine, iodepth=8, numjobs=8
- **Storage: loopback device** (avoid storage bottleneck)
- Same workload ran from 4 nodes

2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# Performance Study: Mirror layout




- Default: No optimizations
- Improved: All optimizations discussed so far

NOTE:

- Storage layout is mirror across nodes
- Each write updates dirty region bitmap on disk

Storage: NVME.
network link: 2 10G
Benchmark: FIO-libaio
IO sizes (4k/8k).
Iodepth = 8 (Random write)

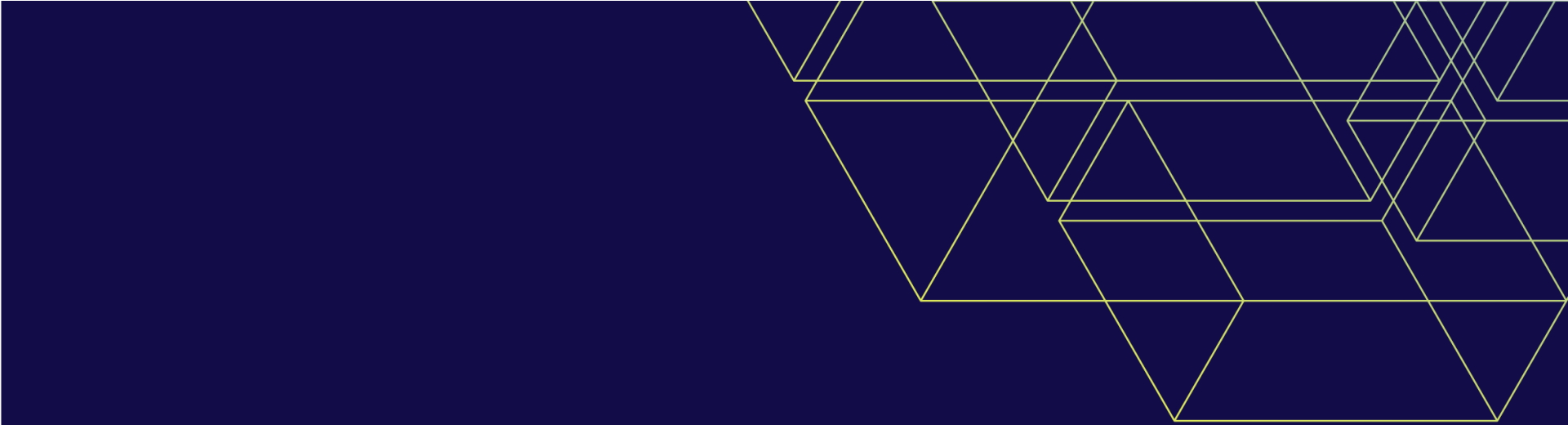
2020 Storage Developer Conference India. © Insert Your Company Name. All Rights Reserved.

# Summary

# Summary

- Discussed various techniques to optimize storage and network layer
- Performance statistics: Up to 1.5X to 2X performance improvement

2020 Storage Developer Conference. © Insert Your Company Name. All Rights Reserved.



**Please take a moment**
**to rate this session.**

**Your feedback matters to us.**