

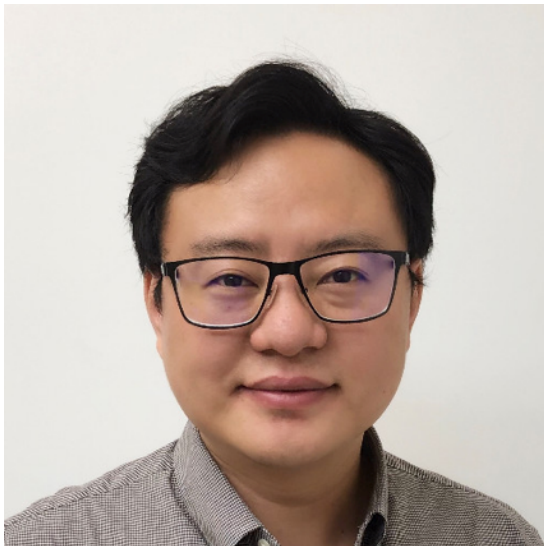


NETWORKING  
STORAGE

# Scale-Out File System Architecture Overview

Live Webcast  
February 28, 2019

# Today's Presenters



**Zhiqi Tao**  
**HPC Solution Architect**



**John Kim**  
**Mellanox**

# SNIA-At-A-Glance



**185**

industry leading  
organizations



**2,000**

active contributing  
members



**50,000**

IT end users & storage  
pros worldwide

# SNIA Legal Notice

- The material contained in this presentation is copyrighted by the SNIA unless otherwise noted.
- Member companies and individual members may use this material in presentations and literature under the following conditions:
  - ◆ Any slide or slides used must be reproduced in their entirety without modification
  - ◆ The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA.
- Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

**NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.**

# Target Audience

- This talk is intended to provide a general overview of the Scale-Out File System Architecture
- Not intended to show preference to certain technologies
- Provide some general guideline when evaluating Scale-Out File System storage solutions.

# Agenda

- General principles when architecting a storage solution based on a scale-out file system
- Use cases for scale-out file systems
- Hardware and software design considerations for different workloads
- Storage challenges and tradeoffs
- Common benchmark and performance analysis approaches
- Popular scale-out file systems in the market

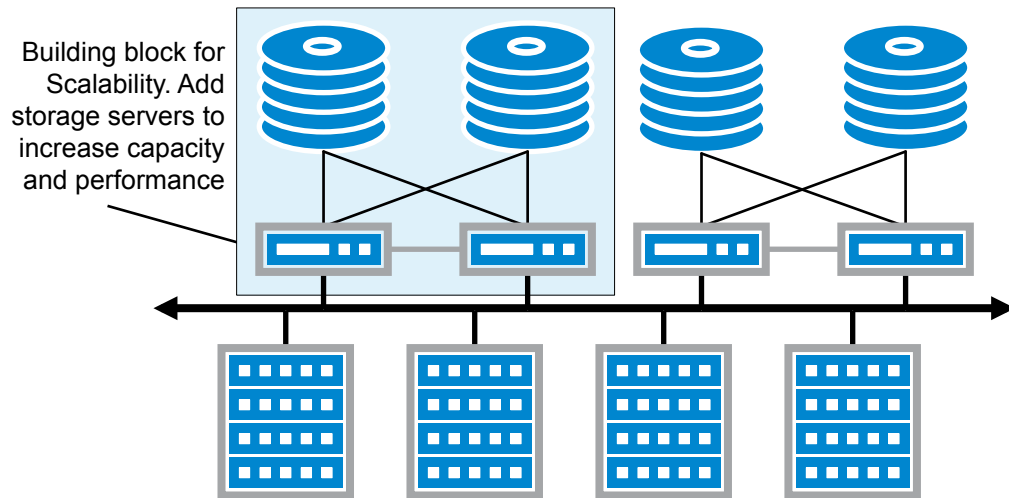
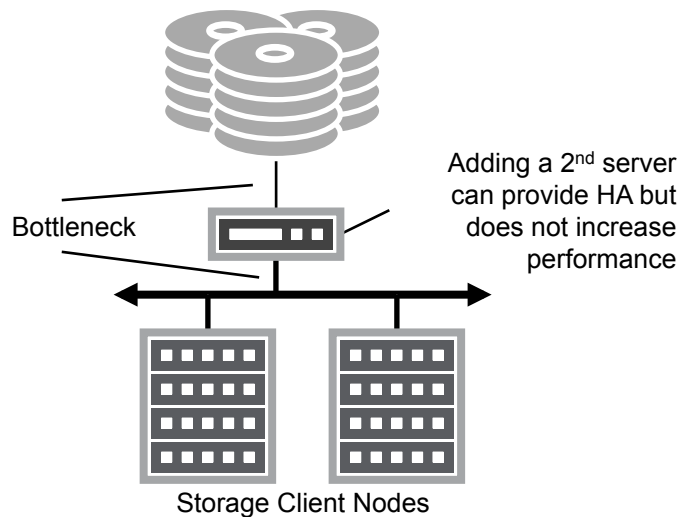
## ➤ File System

- ◆ Data management system that allows computers and applications to share read and write access to data organized as files in directories or folders.

## ➤ Scale-Out

- ◆ Solution that runs simultaneously on multiple hardware systems or nodes and utilizes the compute power, storage, and network resources of many hardware systems simultaneously.

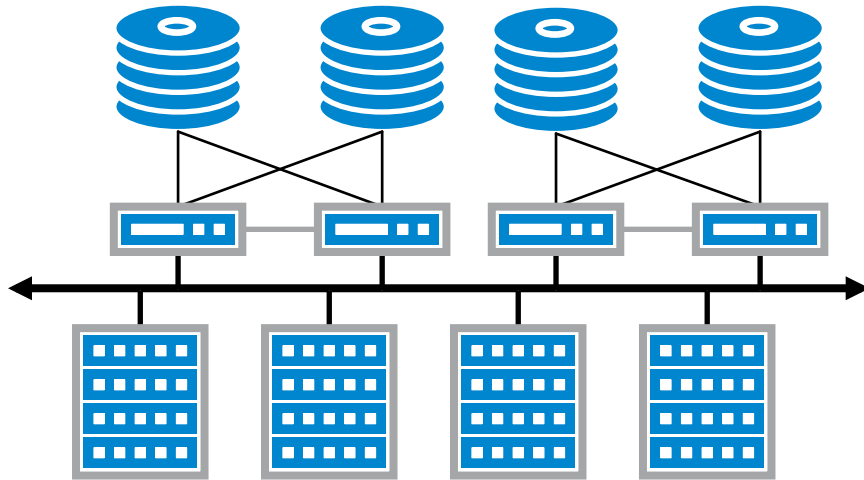
# Scale-Up vs. Scale-Out



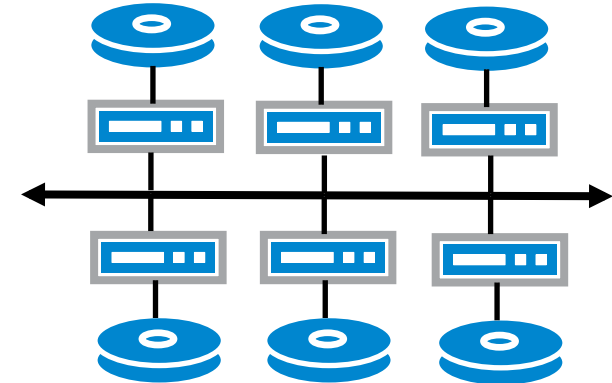


# Different Scale-Out Approaches

## Dedicated Storage and Dedicated Client Nodes



## Hybrid Storage and Client Nodes



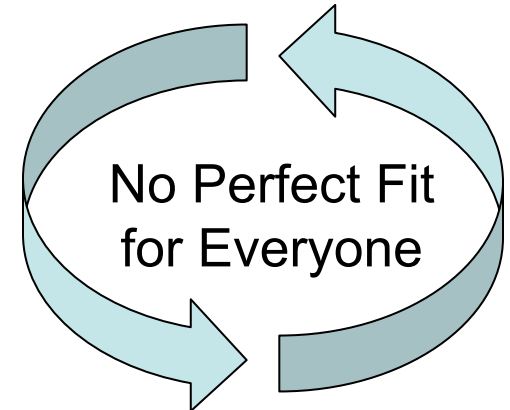
# Common Use Cases

- High-performance computing
  - ◆ Also AI and machine learning
- Technical applications
  - ◆ Oil and gas, semiconductor design (EDA)
  - ◆ Aerospace/automotive, life sciences, weather forecasting
- Media and entertainment
  - ◆ Video rendering, transcoding
- Big data

# Design Considerations

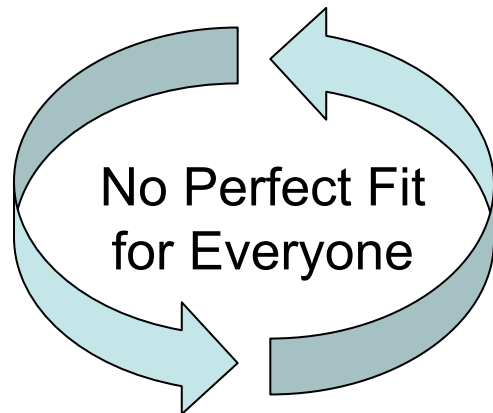
Aspect	Considerations
Capacity	Size, growth forecast, expansion strategy
Access model	Application access: protocols, networks, client software
Performance	Evaluation criteria: throughput, latency, scalability, real workloads
Reliability	Fault tolerance
Availability	Downtime vs. Operable time
Cost	Hardware, software, power/cooling, management
Manageability	Training, day-to-day burden, tools, support, renew, upgrade, expand...
Data Life Cycle	Keep data forever? Discard or tiering (move old data to archive)

- Generic servers, storage array, network, disks
  - ◆ Pros:
    - Probably lower cost
    - Might be possible to combine and mix multiple vendors' hardware
  - ◆ Cons
    - More complex to support
- Customized hardware
  - ◆ Pros
    - Normally elegant, polished, integrated design
    - Simpler to support
  - ◆ Cons
    - Vendor lock-in, possibly higher cost



# Software in Scale-Out Storage

- Software is the key differentiator in storage.
- Open source software stack
  - ◆ File System, Volume Manager, System Management, High Availability, Data migration ...
  - ◆ Many choices but often not tightly integrated.
- Proprietary
  - ◆ Often more polished but with extra costs
- Equal opportunity cost on both
- Understand your support plan



# Common Challenges

- Storage is often an after-thought.
- Scale, but only up to a certain level
- Cache and Consistency; Distributed Locking
- Metadata and Data can have different HW requirements
- Performance and Async/Sync Data Replication
- Networking requirements
- Balance
- High Availability and Timing
- Built-in tools for provisioning, management, triage

# Common Benchmark Software

- Vdbench - Swiss Army Knife
  - FIO – Generate various workloads
  - IOR, IOzone – Throughput with different IO sizes
  - Mdtest – Metadata and small files
- 
- But nothing can replace the real-life applications
  - Be creative on what could go wrong and test performance under “Rainy Day” Scenarios.

# “Rainy Day” Scenarios

- Hardware failure: drives, controllers, servers, network
- Software upgrade
- Disaster recovery
- Silent data corruption
- New or larger workloads
- Eventually there is always a tradeoff between features, capacity, safety, cost, etc.



# Popular Scale-Out File Systems

- Very scattered market. No single dominating technology.
  - ◆ Search for: “Parallel file system” or “Cluster storage”
  - ◆ Open source and proprietary options
  - ◆ Software-only and hardware+software options
- Many cloud service providers have their own scale-out storage as well.

# Summary

- Scale-out file systems run across multiple systems
- Offer some advantages in performance and scale
- Many versions available
- Many options and many tradeoffs to consider
- Test “real world” workloads and “rainy day” scenarios

# After This Webcast

- Please rate this webcast and provide us with feedback
- This webcast and a PDF of the slides will be posted to the SNIA Networking Storage Forum (NSF) website and available on-demand at [www.snia.org/library](http://www.snia.org/library)
- A full Q&A from this webcast, including answers to questions we couldn't get to today, will be posted to the SNIA-NSF blog: [sniansfblog.org](http://sniansfblog.org)
- Follow us on Twitter [@SNIAANSF](https://twitter.com/SNIAANSF)

**Thank You**

