SNIA

STORAGE NETWORKING INDUSTRY ASSOCIATION

EDUCATION

Storage Grid: Relevance and Overview

Abbott Schindler,

Hewlett-Packard Company

SNIA Legal Notice

- The material contained in this tutorial is copyrighted by the SNIA.
- Member companies and individuals may use this material in presentations and literature under the following conditions:
  - Any slide or slides used must be reproduced without modification
  - The SNIA must be acknowledged as source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA Education Committee.

Abstract

Storage Grid: Relevance and Overview

Storage infrastructures based on new architectures are emerging in the marketplace. The new architectures are built around Storage Grid. This session will explore what Storage Grid are, their basic elements, and how the industry is implementing them. Also, in an emerging world of Grid Computing, the tutorial will aid in understanding how storage, data, compute, and application can work together. Storage Grid will be compared to conventional storage in terms of business benefits and compatibility with existing storage environments.

Highlights

- Understand what Storage Grid offer
- Overview of possible implementations
- Understand Storage Grid benefits
- Impact Storage Grid may have to the evolution of storage environments

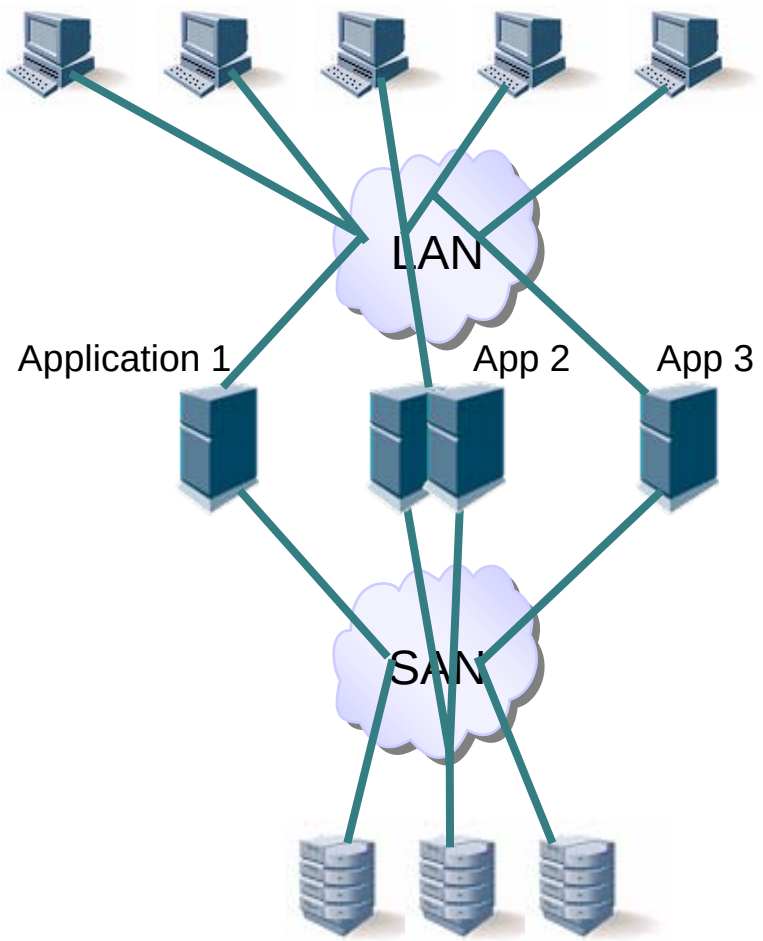
Agenda

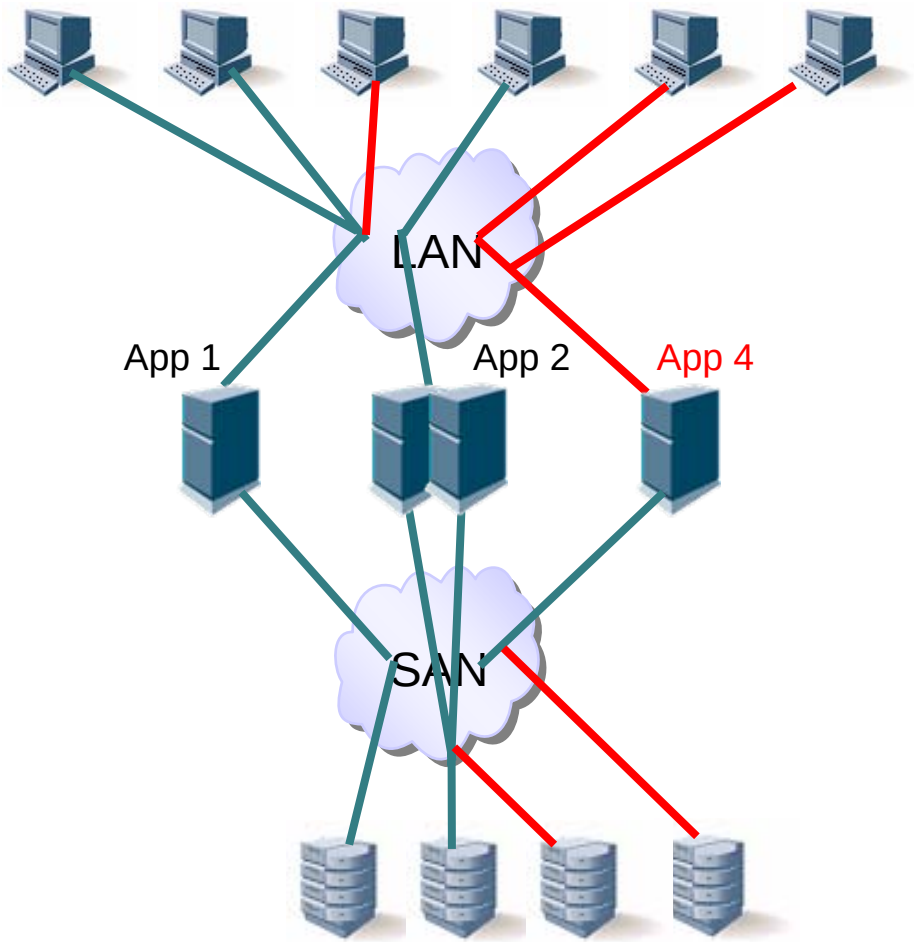
- Tutorial goal: understand what a Storage Grid is and its relevant to you
- Storage Grid rationale
  - Grid Computing and storage
- Storage Grid attributes
- Proposed Storage Grid Model
- Compare Storage Grid with other approaches

Challenges

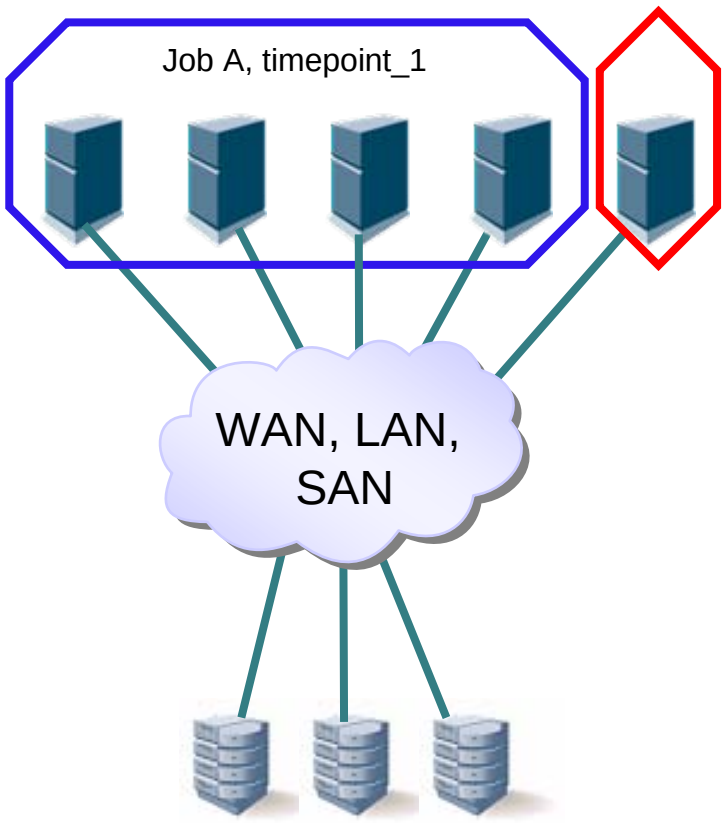
- Growing amounts of information
  - Focus is information and managing it, not just data
- Increasingly dynamic and fluid environments
  - Organizations, applications, and the encompassing business environment
- Economic constraints
  - Efficient utilization of assets
  - Financials (ROI, TCO, etc.)
- New types of workloads

Dynamic Environments

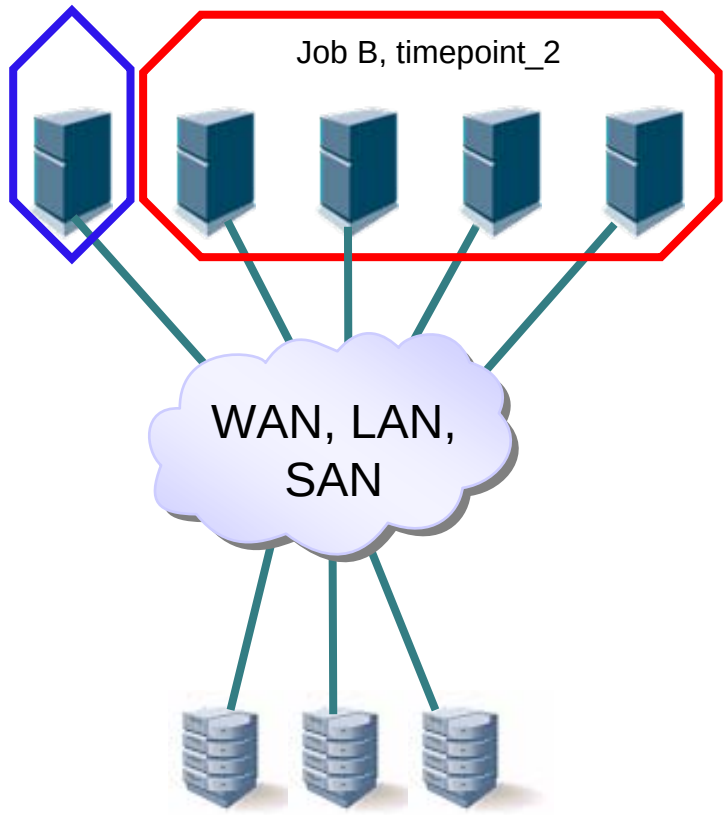




Grid Computing adds a Dimension



Networked storage resources



Networked storage resources

Resulting Storage Needs

- Deliver storage services easily and flexibly
  - Capacity, performance, availability, object types, etc.
- Easy to deploy and grow
  - Compatibility/interoperability
- Easy to manage
  - Complexity can't increase as the environment scales
- Adaptable
  - Able to accommodate changing demands (workloads, SLOs, etc.)
- Cost-effective
  - Acquisition, maintenance, management, etc.

Compute Grids and Storage

- Can use any kind of storage
  - Today: DAS, NAS, SAN
- Future: accommodate special qualities of Grid Computing environments
  - Batch-oriented resource allocation
  - Distributed compute resources come and go
  - Changing locality of data reference
  - Primarily file, database orientated
  - High storage performance
  - Resources are delivered and consumed as services

Terminology

- Consumer
  - Any application or entity that utilizes storage resources
- Storage for Grid
  - Storage capabilities/interfaces that specifically address Grid Computing environments
  - Concerned with storage access connections, not infrastructure
- Storage Grid
  - A unified, managed infrastructure that leverages grid-computing concepts to present a single utility-like storage system
- Delivered service
  - A storage capability that is provided as a fungible resource to consumers

Storage for Grid

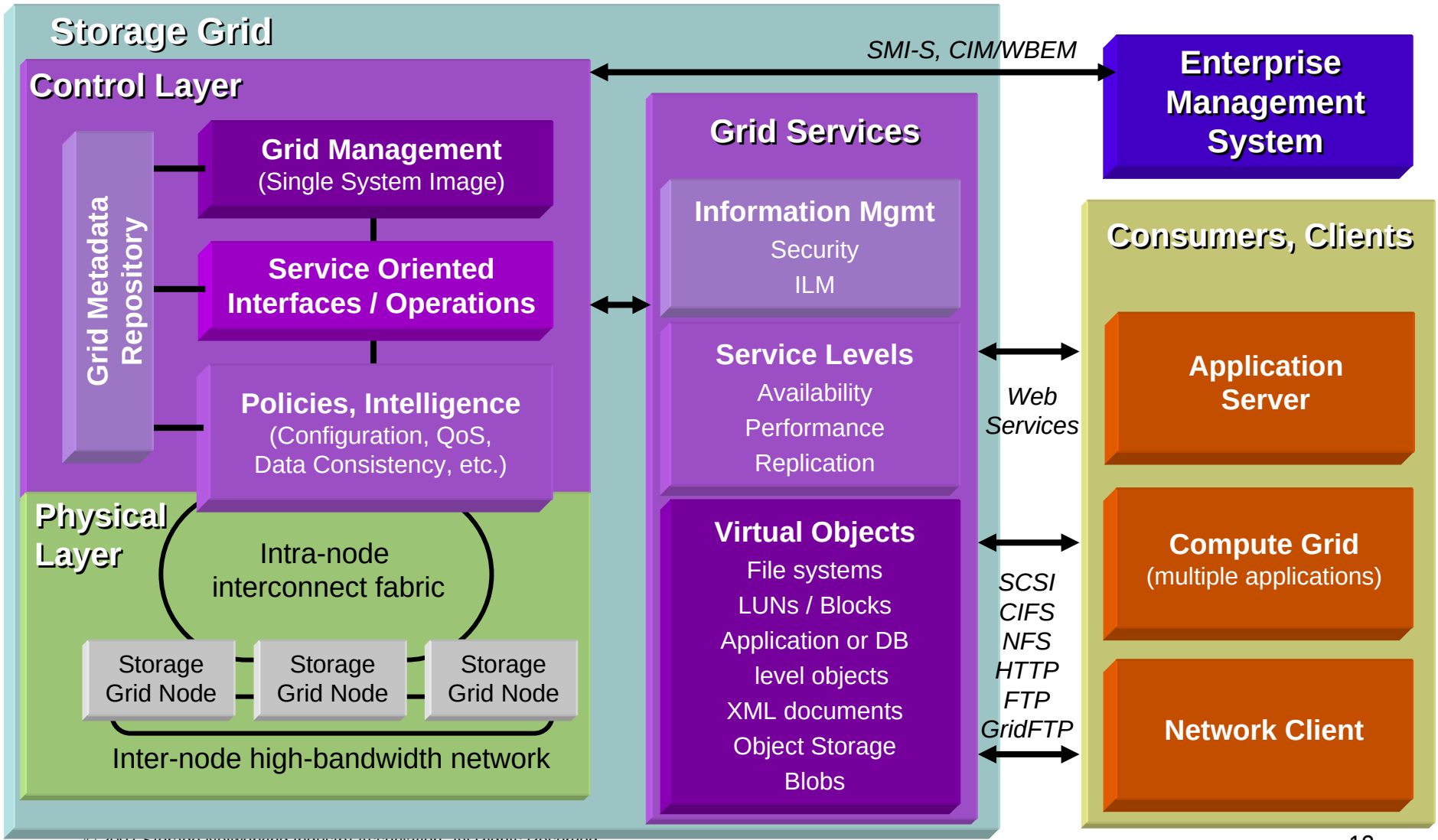
- Interface layer between Grid Computing and storage
- Provides compute grid computing-specific storage services
  - Services added onto conventional storage capabilities
- Long term promise for storage
  - Data transport
  - Supports simultaneous distributed data access
  - Able to accommodate changing data access patterns
  - Quality of Service specifiers, feedback mechanisms
  - Unified end-to-end security model

Storage Grid Qualities

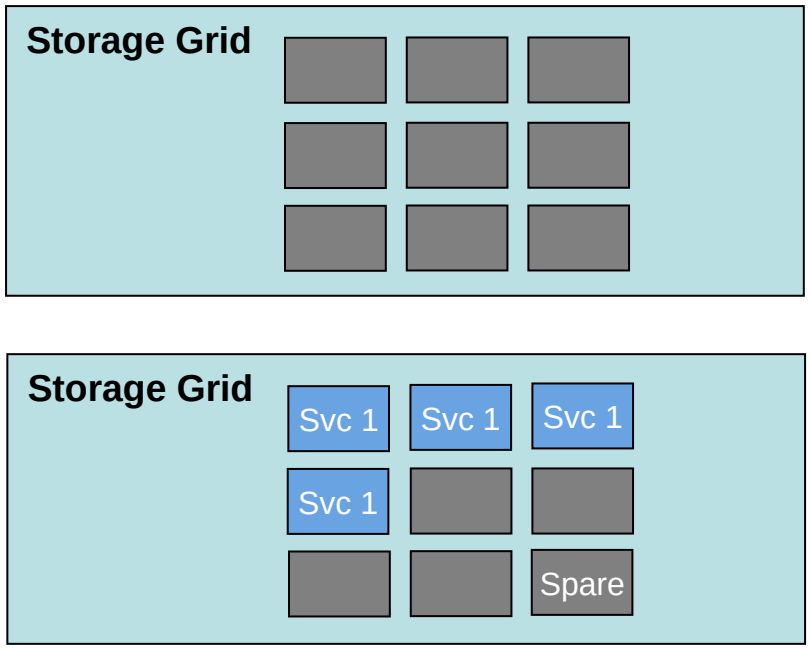
SNIA
Storage Networking Industry Association
EDUCATION

- Scale-out architecture
  - Scales with predictable results
  - Allows new capabilities to be incorporated easily
- Managed and accessed as a single system
  - Administratively, a single management point
  - Elemental details managed by the system itself
- Service delivery model
  - Storage capabilities are delivered as services (eg. capacity, performance, availability, security)
  - Able to respond to consumer SLO with minimal administrator intervention
- Consumption-oriented provisioning
  - Automated on-demand storage services delivery
  - Allows for more accurate accounting/chargeback

Enterprise View of a Storage Grid Model (a *proposal*)



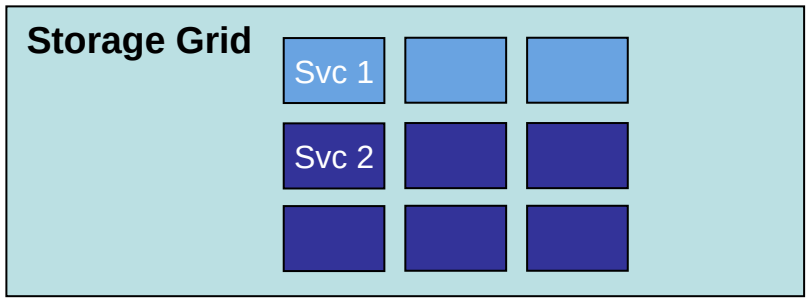
Creating Storage Grid Capabilities

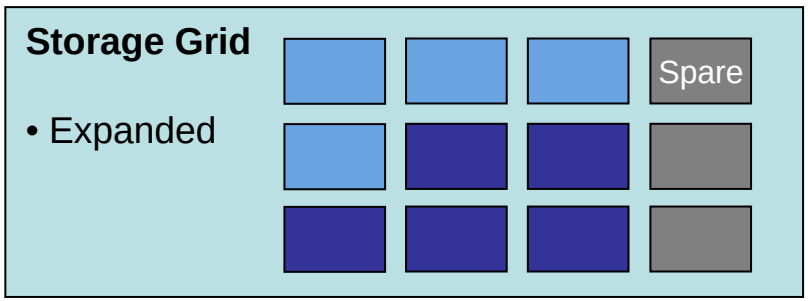


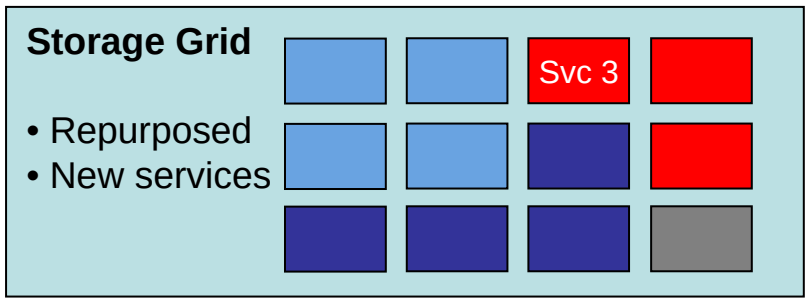
Storage Grid infrastructure
- Nodes + raw storage capacity

"Purpose" the nodes
- Connect raw storage resources to appropriate nodes
- Load storage application software onto appropriate nodes

Advanced Capabilities

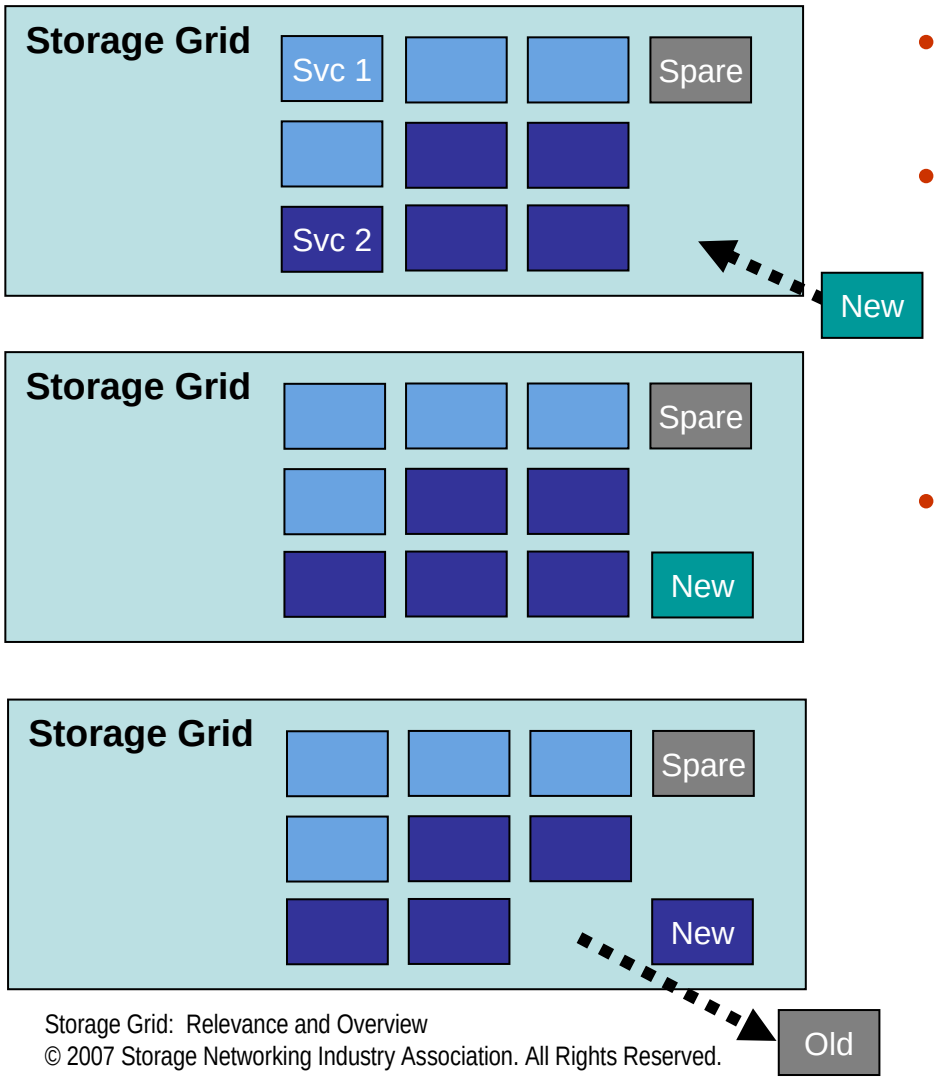






- Storage Grid can provide multiple services concurrently
- Dynamic expansion
  - Responsive to applications
  - Automatic load optimization
- Dynamic repurposing
  - Automatic load optimization
  - SLO adherence

Resource Lifecycle Management



- Inherently straightforward
  - Minimal application impact
- Expand a Storage Grid
  - Install an element
  - Auto-discover, bind
  - Provision
  - Redistribute workload
- Decommission elements
  - Drain data
  - De-present
  - Unbind
  - Retire and de-install
  - Also applies to replacement/upgrade

Storage Clusters and Storage Grids

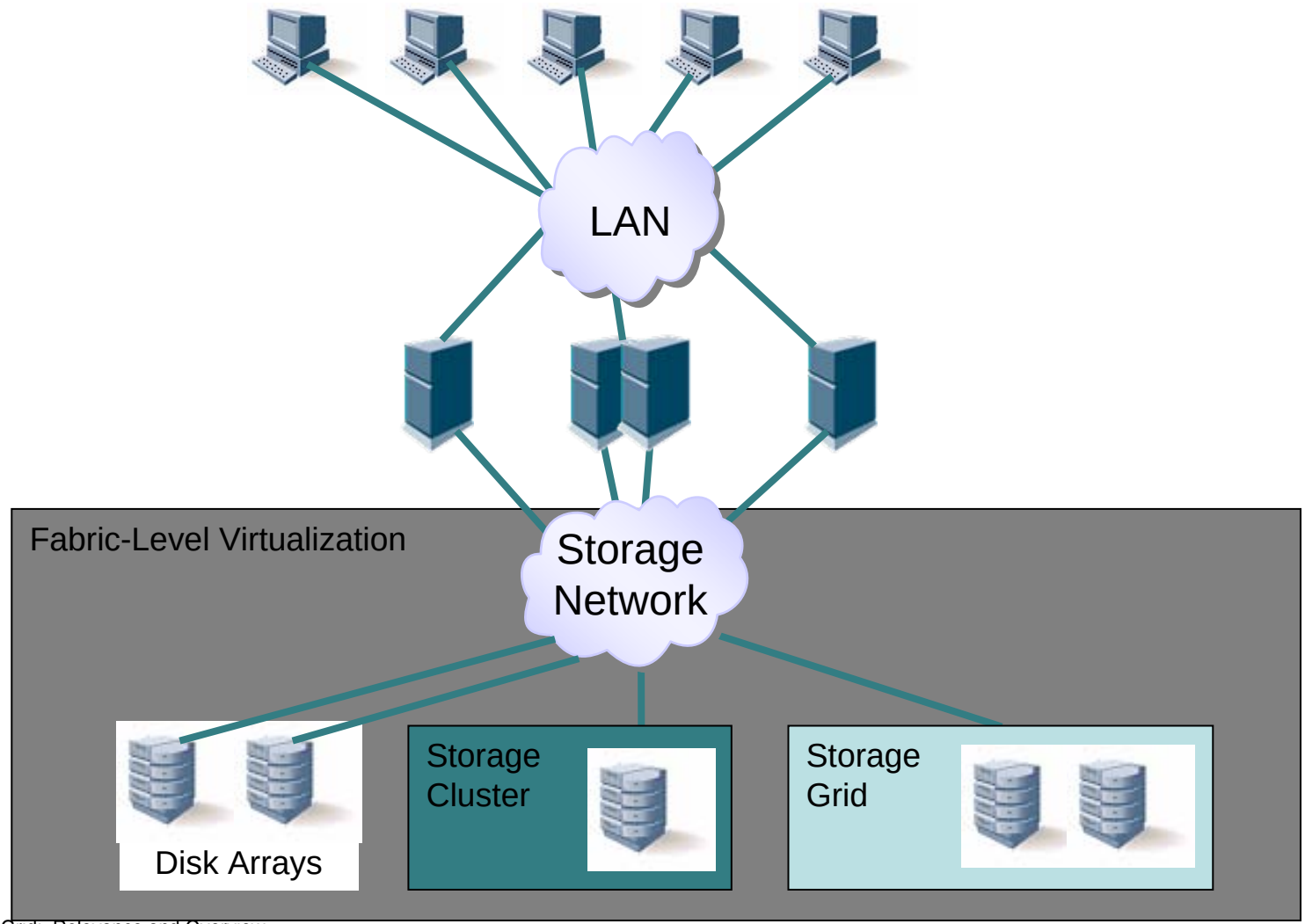
SNIA
Storage Networking Industry Association
EDUCATION

- Two approaches for aggregating resources
- Storage Cluster
  - General focus is high availability, scalability, or both
  - Not focused on higher levels of intelligence
  - Provide a fixed set of capabilities
- Storage Grid
  - Present a broad and extensible set of capabilities
  - Can embed higher levels of control, automation
  - Resources can be reassigned dynamically

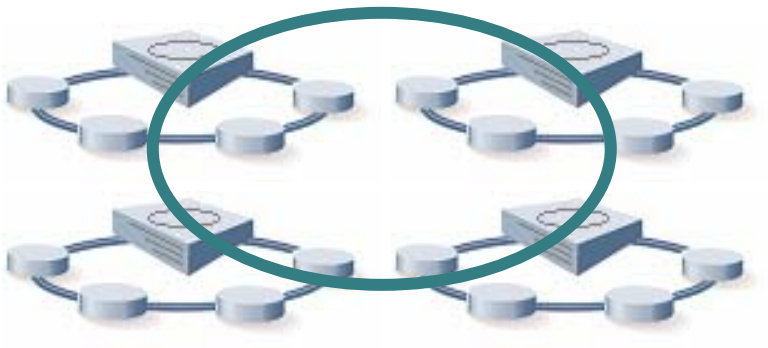
Storage Virtualization and Storage Grids

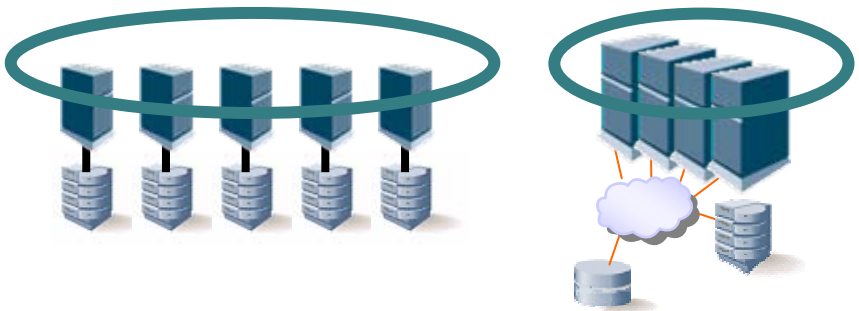
- Two key approaches for aggregating heterogeneous resources
- Virtualization in the fabric
  - Aggregates heterogeneous storage systems
  - Typically applied to homogeneous types of resources, but is emerging for heterogeneous resources
- Storage Grid, over time:
  - Aggregate heterogeneous and disparate systems
  - Dynamically configure to optimize asset utilization
  - Automatically manage data placement among the incorporated assets

Putting It Together



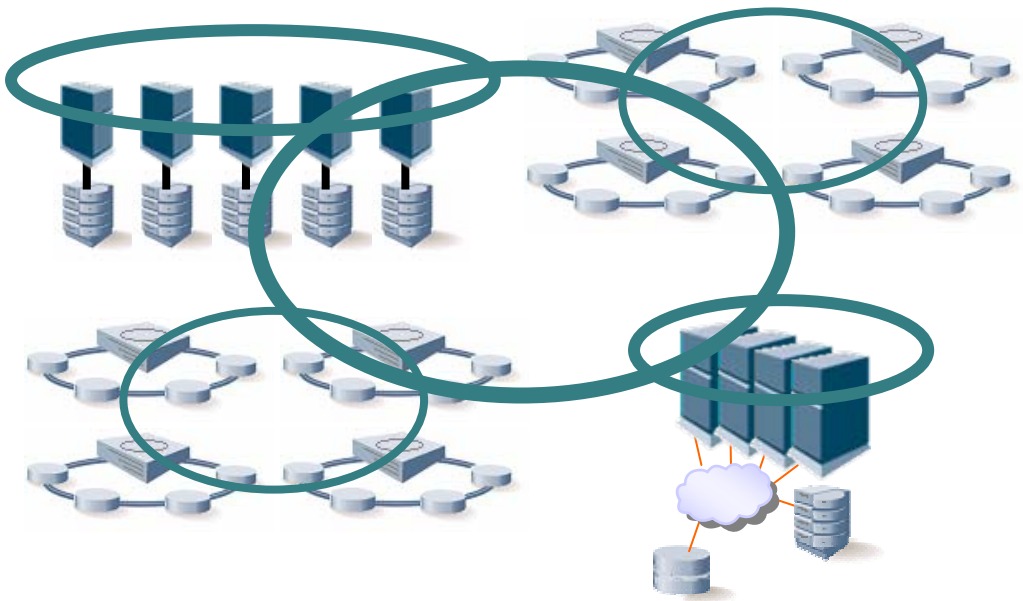
Storage Clusters





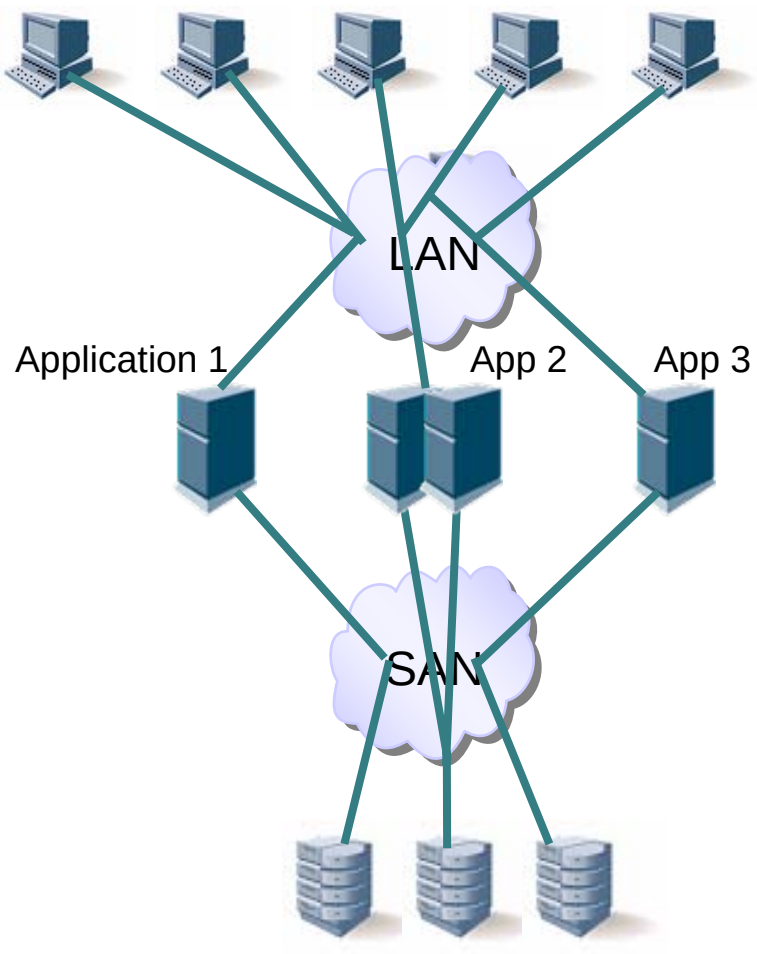
- Generally support a single function
- Basic building blocks are joined together
- Can be seamlessly expanded dynamically by adding addition nodes (scale-out architecture)
- Examples
  - Block-serving array service
  - File-serving NAS service
  - Tape library service

Storage Grids



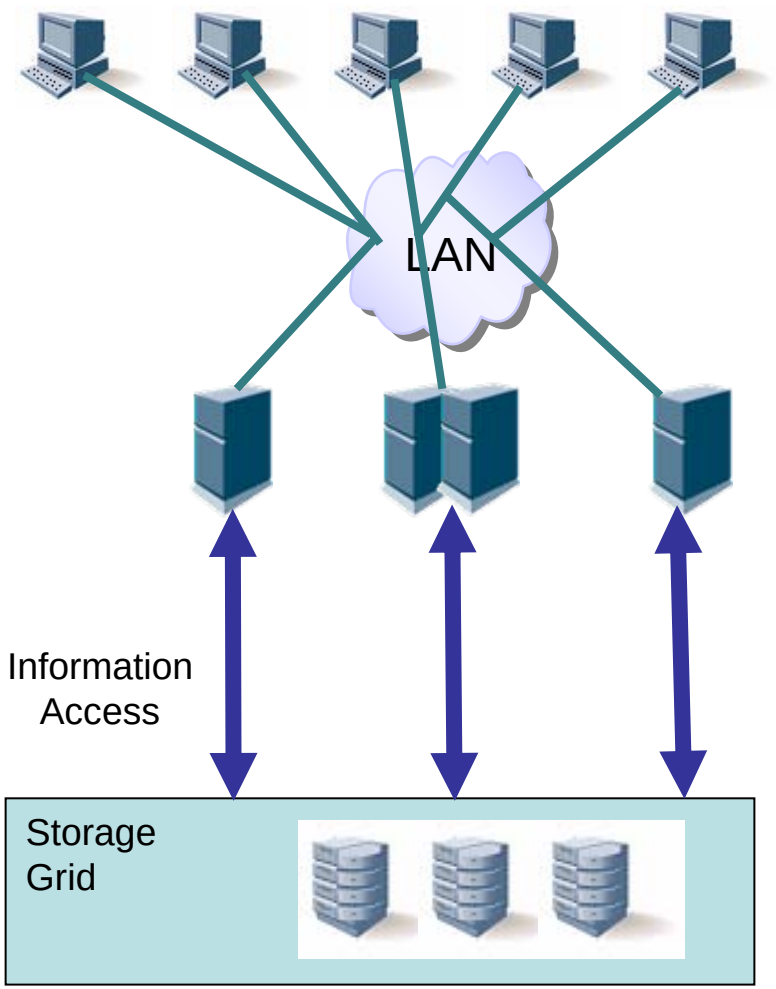
- Service-Oriented Architecture
- Combines a broad set of functions into a single entity
- Multiple functions can be presented independently or as composite services
- Management software dynamically adjusts the grid to accommodate varying demands

Traditional Storage



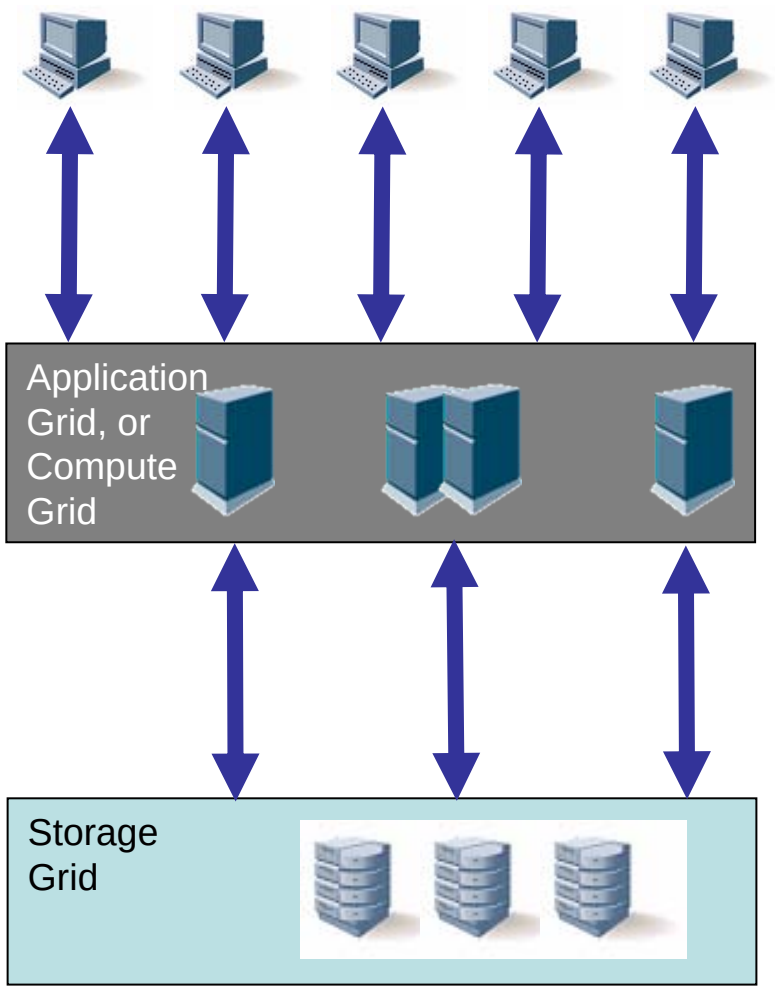
- Data flow is among a large number of independent entities
  - Servers, clients, storage devices, etc.
- Storage services depend on infrastructure
  - Redundant components, paths
  - Manual resource allocation and provisioning
  - Independent resource management

Applying Storage Grids



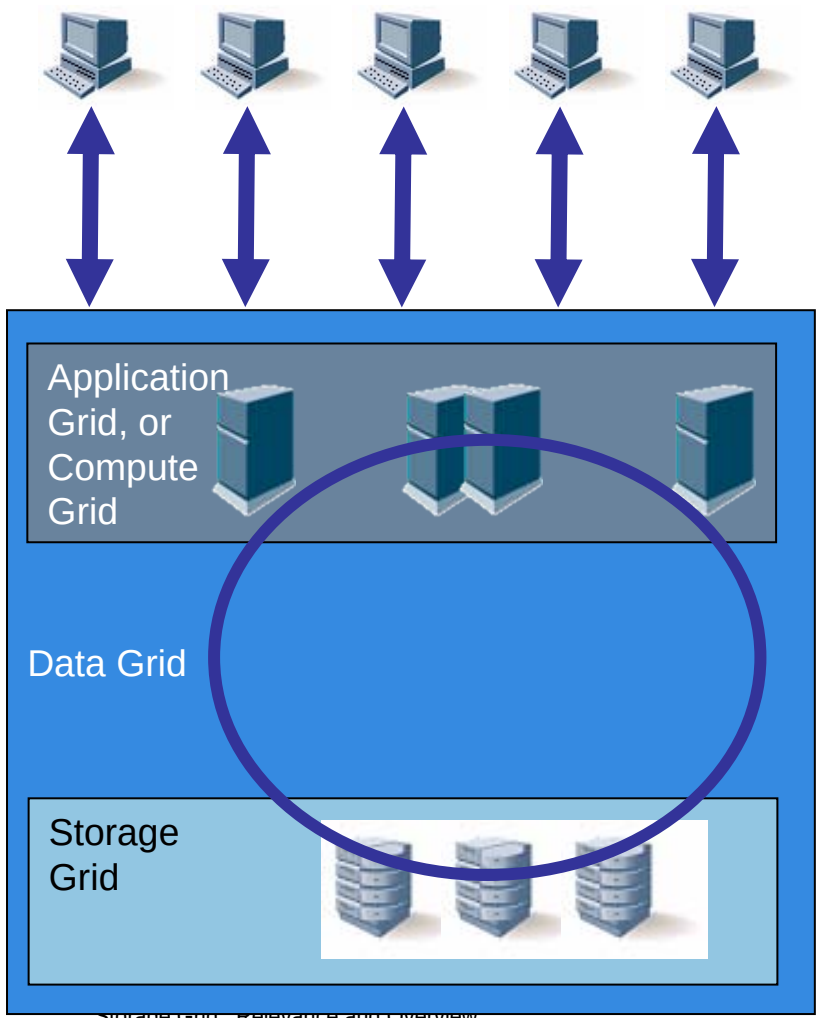
- Data flow is among a smaller number of virtualized entities
  - Servers, clients, Storage Grid
- Storage services are more reliable, flexible, scalable
  - Embedded storage grid management ensures SLOs

Complementary Grids



- Data flow is among an even smaller number of virtualized entities
  - Clients, various grids
- IT services are more reliable, flexible, scalable
  - Grid capabilities extend throughout information processing stack

In the Future, Grids May Converge



- Data flow is among an even smaller number of virtualized entities
  - Clients, a heterogeneous "Data Grid"
- IT services are more reliable, flexible, scalable
  - Grid capabilities are unified throughout information processing stack

Storage Grid Standards

Next Steps and Going Forward

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION
EDUCATION

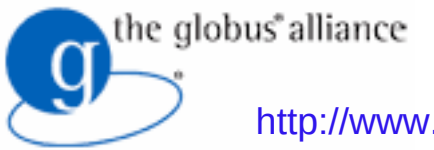
- New standards-based capabilities are needed
- Service Oriented Interfaces
- Management integration, automation
  - Look for areas to extend or expand SMI-S
  - Common APIs, toolkits, frameworks
- Data transport protocols
- Standard interfaces for data services
- Security
  - End-to-end security model

Existing Grid Standards and Alliances Bodies

http://www.gridforum.org/

http://www.oasis-open.org/

http://www.dmtf.org/



http://www.globus.org/

- Open Grid Forum
  - Overall architecture for grid computing
  - Definition of overall service taxonomy
  - Enterprise grid requirements
- OASIS
  - Middleware / Web services focused
  - Protocols (including management)
  - WS-RF, WS-Notification, WSDM, …
- Distributed Management Taskforce (DMTF)
  - Management and Information models
  - Utility Computing Working Group
- Globus Alliance
  - Enterprise harden the Globus Toolkit
    - De facto standard open source grid middleware

Summary

- Storage Grid is an emerging architecture built with tenets derived from grid computing
- Storage Grid are different from storage clustering and other multi-node architectures
- Key attributes include scale-out, single system image (to consumers and to manage), dynamic scalability and automated resource delivery
- Long-term promise includes:
  - Higher level storage functionality
  - Higher level of integration with applications
  - Improved adherence to SLOs through greater automation
  - Simpler growth and incorporation of new technologies

Q&A / Feedback

- Please send any questions or comments on this presentation to SNIA: trackvirtualization@snia.org

Many thanks to the following individuals for their contributions to this tutorial.

SNIA Education Committee

Abbott Schindler
Ken Wood
Curt Kolovson
Rob Peglar
John Easton
Sorin Faibish
David Black