

Storage Virtualization Strategies

Steve Norall

Senior Analyst, Taneja Group

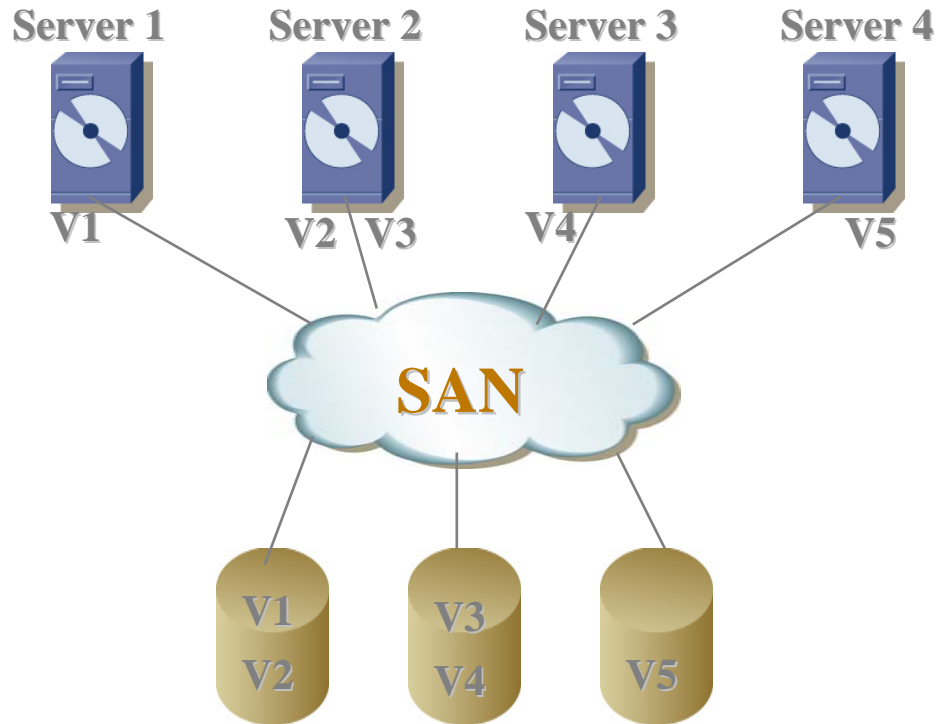
steve@tanejagroup.com

February 12, 2007

Agenda

- Storage Virtualization Approaches, Architecture, & Players
- File Virtualization Approaches, Architecture, & Players
- Recommendations to End Users

Typical SAN Today



Ultimate Vision of SAN is Incomplete

Why Storage Virtualization is En Vogue Again...

- Heterogeneous Storage is a Reality
- Data Migration, Data Consolidation are Disruptive Processes Today
- V has Passed the “Hype” and “Low Point” Stages and into “Reality” Stage
- Big guys have Ratified the Concept (IBM, EMC, HDS)



What is Storage Virtualization?

- Separation of Physical and Logical
- Enables One Common Set of Applications
- Physical can be Heterogeneous or Homogeneous
- Underwear

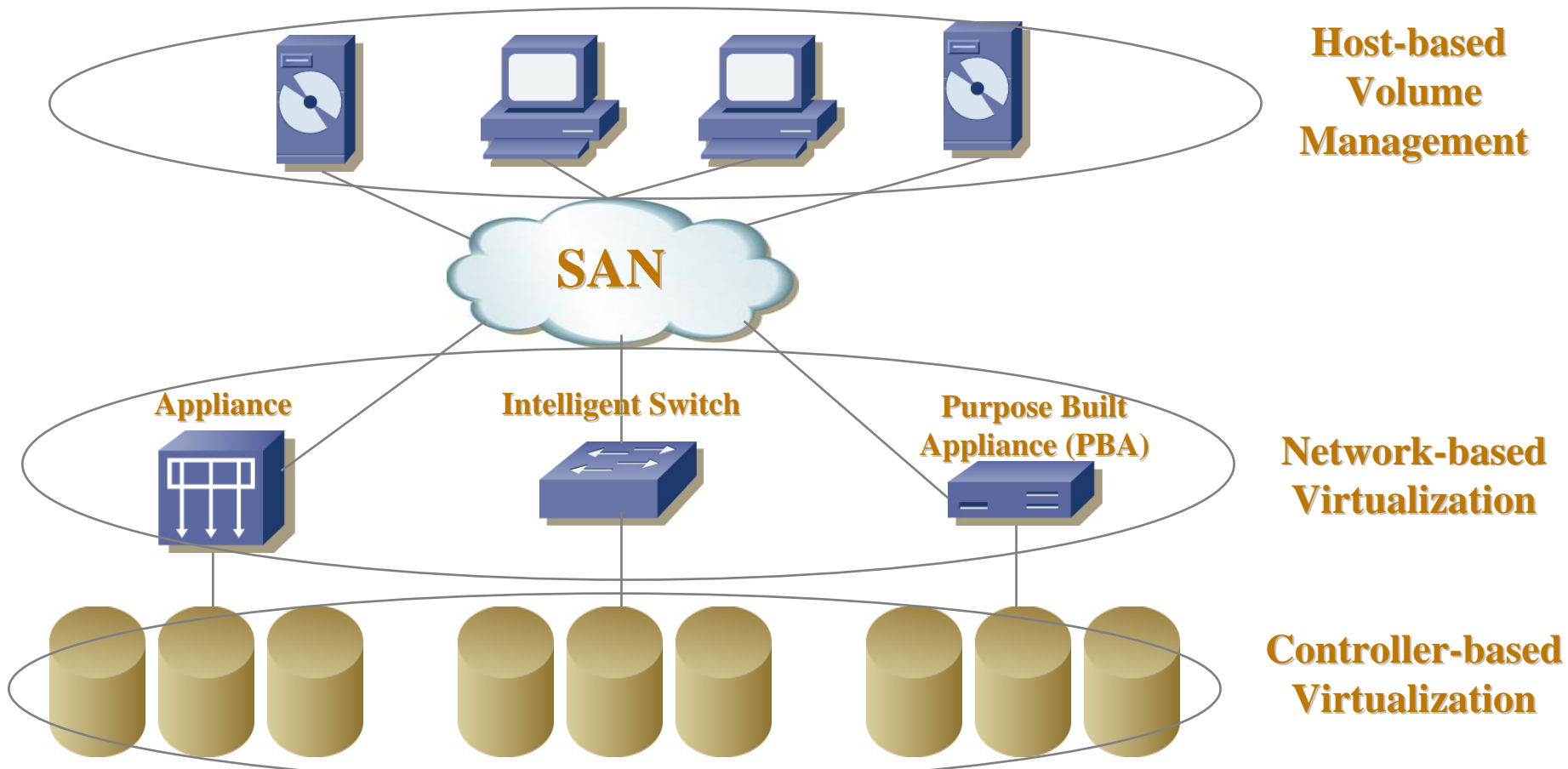
RAID equals elemental storage virtualization

What Does Storage Virtualization Enable?

- Consistent Interface for All Storage Applications
- Non-disruptive LUN Creation, Expansion, Deletion
- Non-disruptive Data Consolidation, Data Migration
- Non-disruptive RAID Level Changes
- Tiering of Storage Using Hetero Storage
- Replication Across Hetero Storage
- Backup/Restore, Archival, Thin Provisioning, etc.

Storage Virtualization Completes SAN Vision...

Where Can It Be Done?



Fundamental Concepts

- Metadata Server
 - Maintains Configuration Database
 - Virtualization Mapping Tables
- Control Path Software
 - Provides Interface Between Metadata Server and Data Path Server
 - Performs Background I/O Tasks
- Data Path Software
 - Moves Data from Server to Storage (and vice versa)
 - Performs Translation from Virtual to Physical Addresses

Note: Datapath Software is Always in the Data Path

Virtualization Architectures: In-Band vs. Out-of-Band

- **In-Band**

- Data and Control Data are managed by same compute elements
- Initial appliance implementations lacked scalability and performance
- Newer implementations much more scalable and “enterprise-class”
- Most widely deployed approach to date
- Rich set of advanced storage applications available
- Examples: DataCore (appliance), FalconStor (appliance), IBM SVC (appliance), HDS (controller)

- **Out-of-Band**

- Separates data and control paths
- Requires host agents creating management complexity and potential for security breaches
- Example: LSI Logic/StoreAge

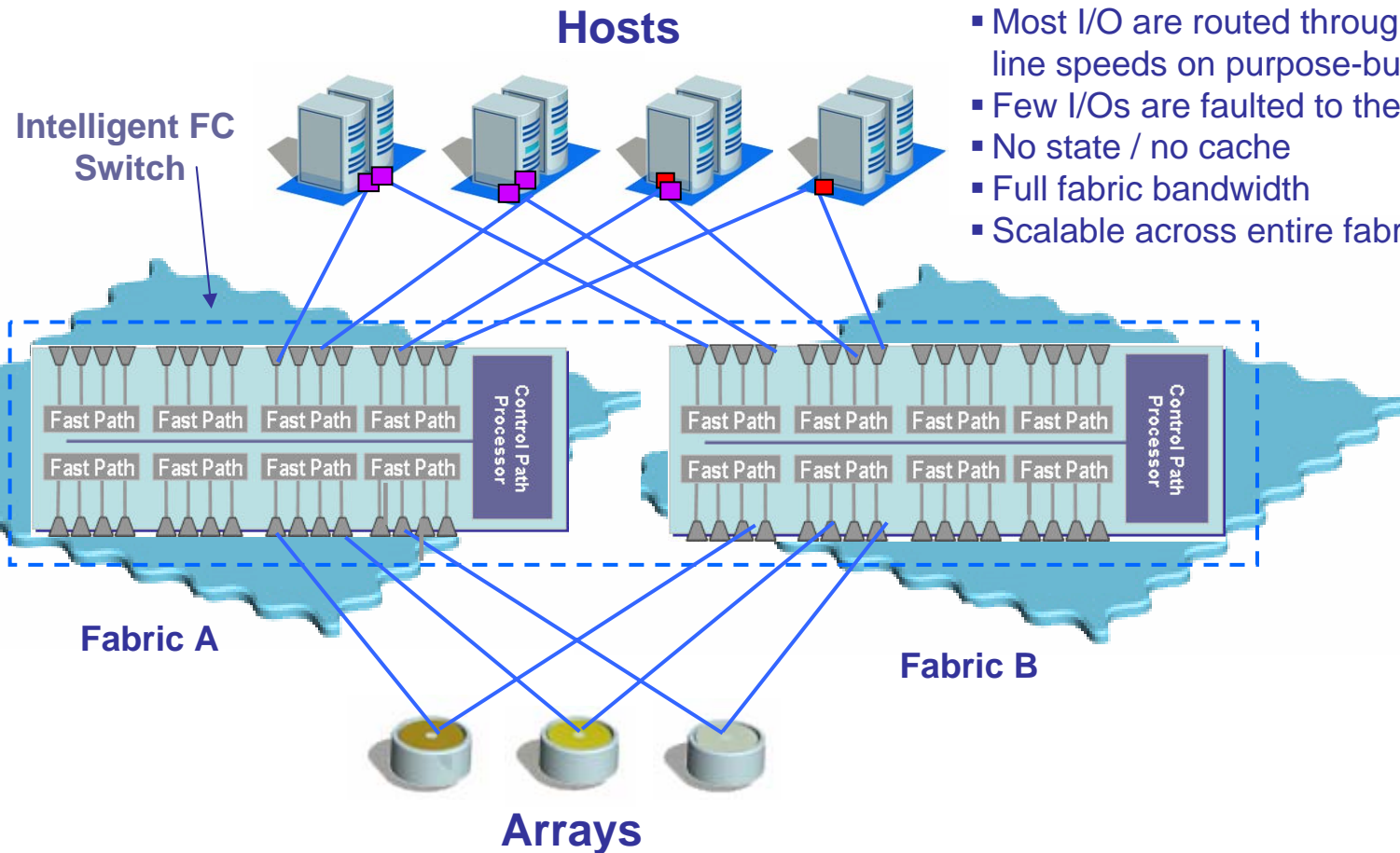
Virtualization Architectures: Split Path Architecture for Intelligent Devices (SPAID)

- **Hybrid of In-Band & Out-of-Band Approaches**
 - Separates control path from the data path
 - Leverages port-level processing capabilities of intelligent switches or purpose-built controllers
 - Promises highest levels of scalability, availability, and performance
 - Virtualization controller software unproven & lack key storage applications
 - Widespread adoption of intelligent switches has been gate to adoption of SPAID
 - Example: EMC InVista, Incipient, LSI Logic/StoreAge

SPAID Architecture

Split Path Architecture

- Most I/O are routed through the Fast Path at near line speeds on purpose-built ASICs
- Few I/Os are faulted to the Control Path (CP)
- No state / no cache
- Full fabric bandwidth
- Scalable across entire fabric



File Virtualization

- Abstraction of logical relationships above individual file & file system images
- Separates user and group relationships from specific devices
- Enables creation of exports and shares

Creates a global namespace

3 Types of File Virtualization

1. Platform-Integrated Namespaces
2. Clustered Storage Derived Namespaces
3. Network-resident Virtualized Namespaces

Platform-Enabled Namespace

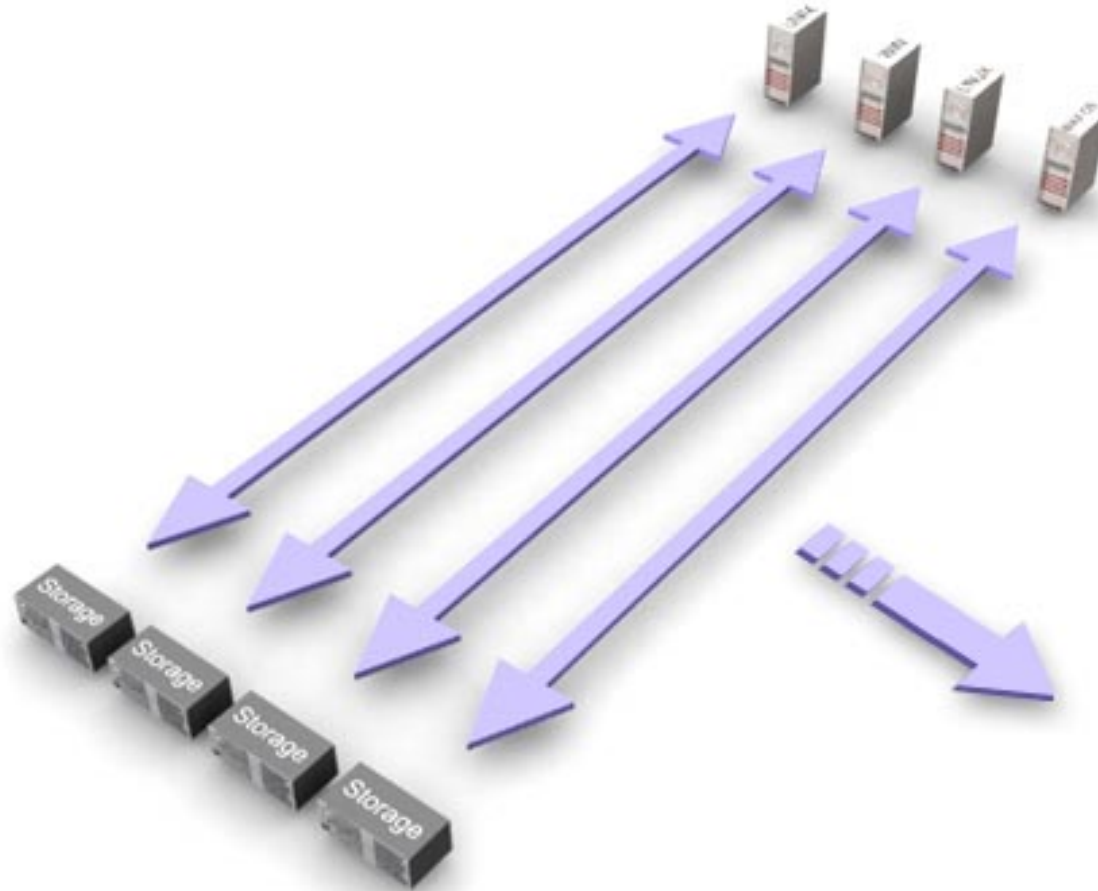
- An extension of the server's file system
- Often called a “Distributed” File System
- Provides a platform-specific means of abstracting file relationships across machines on that server platform

- Examples: Microsoft, NFS v4.0, Brocade/NuView (distributed s/w)

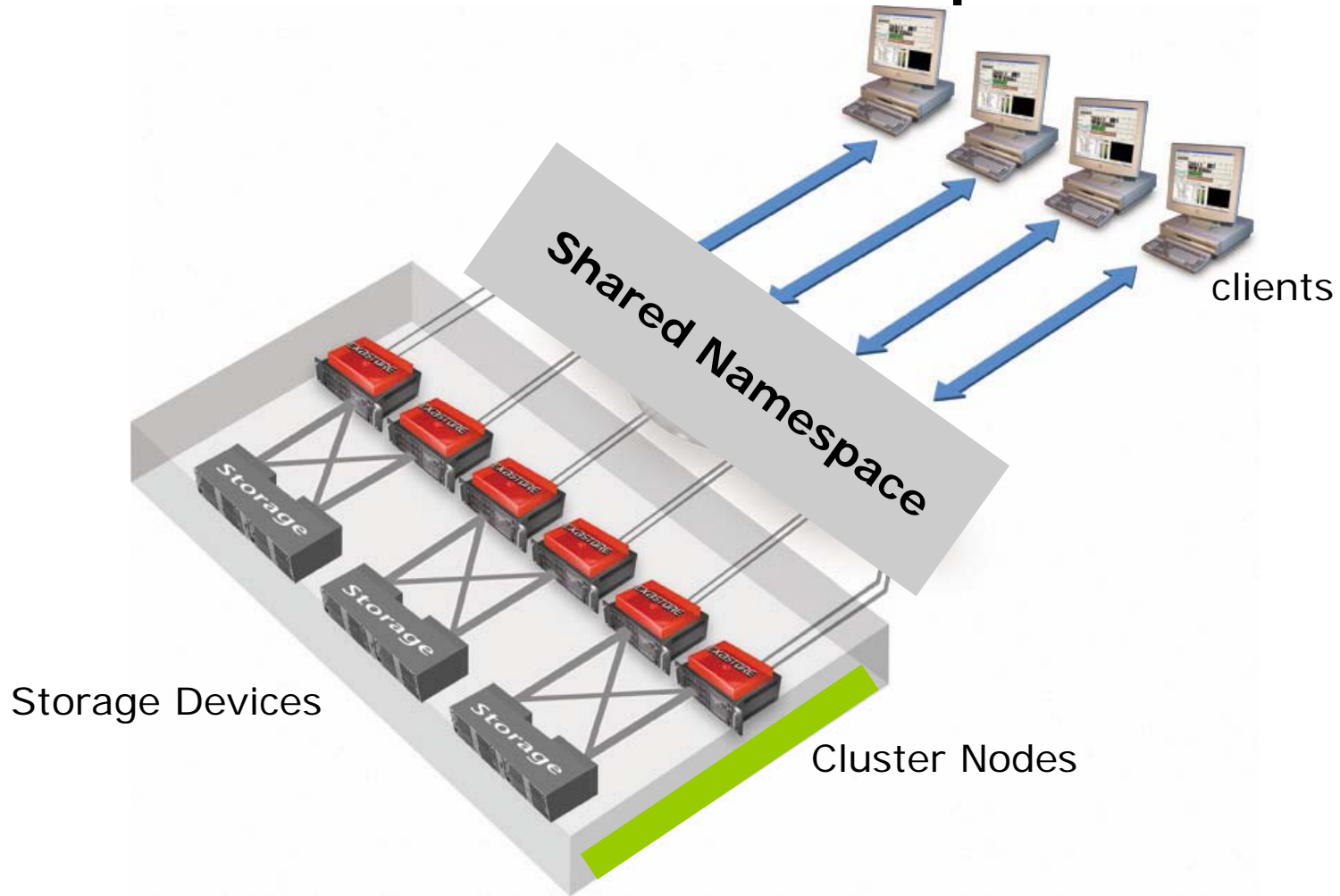
Clustered Storage Derived Namespaces

- Leverages the federated FS architecture of a cluster
- A “gift” of the NAS or DFS cluster is a common namespace
- But, that namespace is restricted to the devices on which it resides
- Examples: Exanet, Isilon, Network Appliance GX, PolyServe

Parochial File Systems Work Like This

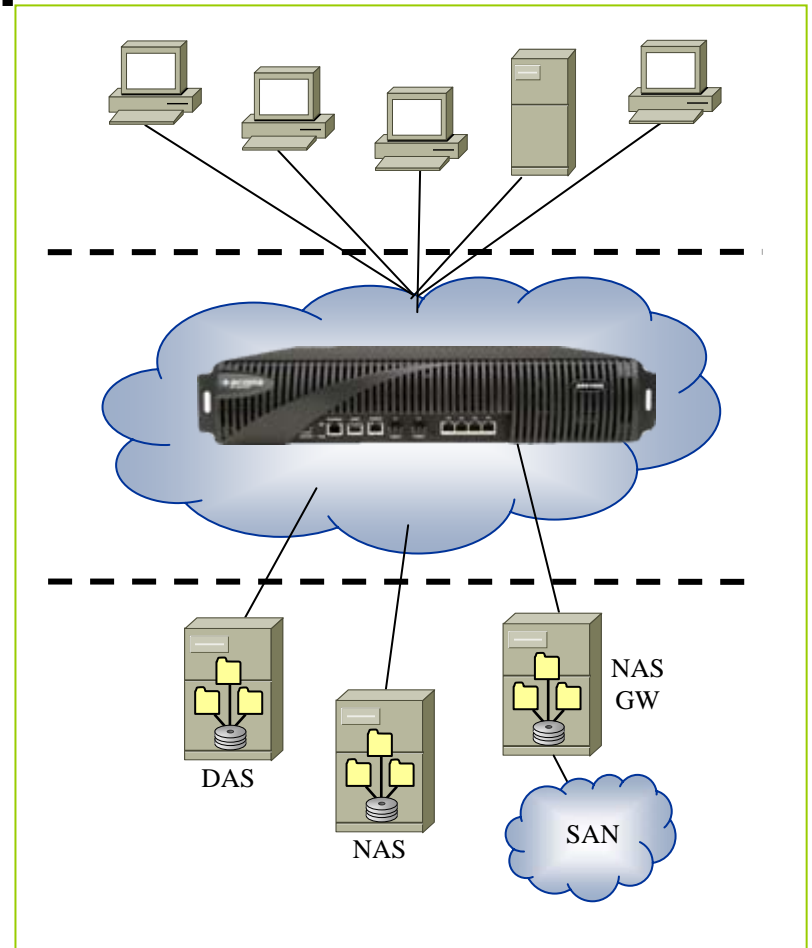


Clusters Federate File Systems: A Shared Namespace



Network-Resident Virtualized Namespaces

- Approach utilizes some manner of in-band or data path information extraction
- Deploys as a network mounted device
- Called “Network File Management”
- Examples: Acopia, Attune, EMC-Rainfinity, Neopath





Recommendations to Users

- If Mostly Homogeneous, virtualization is still important.
 - Stay with existing vendor's virtualization solution.
- If Highly Heterogeneous, go with intelligent switch or appliance
- Ensure the Applications you want are Available
- Ensure Solution is Not “All or Nothing”

Questions



Steve Norall

steve@tanejagroup.com