

Exploring Near Future Directions for Data Storage

Presented by Brad O'Neill
Senior Analyst
Taneja Group Inc.
brad@tanejagroup.com
www.tanejagroup.com

Your IT World: More Complex Than Ever

- You're managing more domains
- Your teams are growing more complex
- Your business issues are faster, tougher, wider in scope
- And guess what? Nobody is cutting you any slack...
- If you don't step back and reframe, you're simply sunk!

You need every edge you can get!

What We'll Do Today...

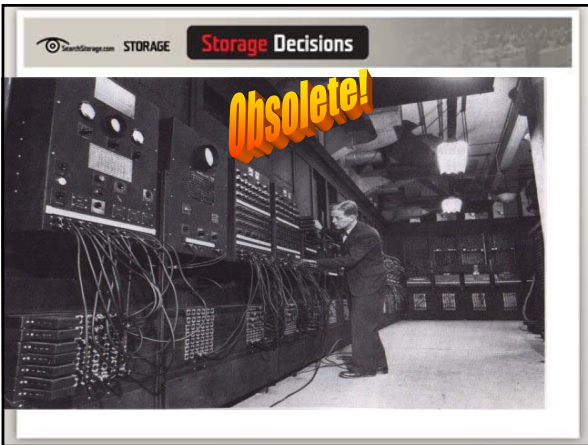
- Brainstorm three future scenarios
- Explore five "dials" we can turn
- Discuss the probable futures
- Q&A

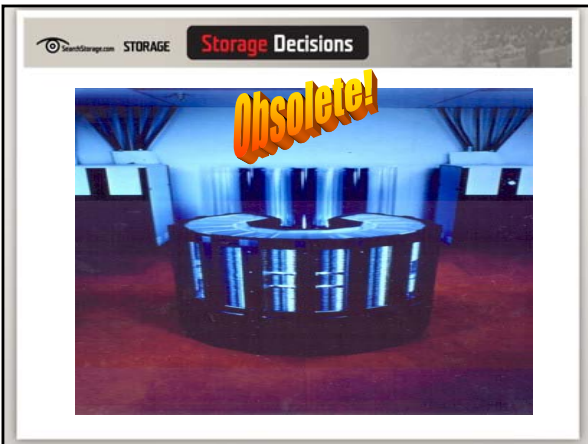
Before the future, let's look at the past...












STORAGE **Storage Decisions**



1996- ~You Decide

STORAGE **Storage Decisions**

Your Storage Future 2008-2012: Three Scenarios

Scenario #1: Aggressive transformation
Scenario #2: Moderated advancement
Scenario #3: Conservative change

STORAGE **Storage Decisions**

Five "Dials" Tuning In These Futures

1. Virtualization adoption
2. Data protection innovations
3. I/O-specific array architectures
4. Remote-branch office tech
5. File management strategies

External Factors That Impact the Near Future of Storage

- The competitive landscape
 - M&A can be both good and bad
- Vendor profit margins
 - A tech market downturn will slow us
- X Factors
 - Think Enron, 9/11, Katrina...

Storage Scenario #1: "Aggressive Transformation"

- Shift to clustered and modular storage
- Automatic workload management
- Mostly network-resident management
- Integrated end-to-end virtualization
- Totally application driven protection
- Automatic multiple-site capabilities
- Converged interconnects/networks

Storage Scenario #2: "Moderated Advancement"

- Mix of monolithic, modular, clustered
- Vendor-specific workload management
- Some network-resident controls
- Vendor-driven virtualization schema
- Key apps have integrated protection
- Vendor-specific multiple-site capabilities
- Case-based converged network fabrics

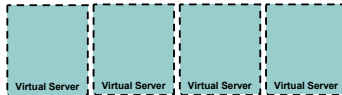
Storage Scenario #3: "Conservative Change"

- Still led by monolithic, modular mindframe
- Vendor-silos for workload automation
- Still limited network-resident controls
- Poor end-to-end virtualization schema
- Key apps have integrated protection
- Limited integrated multiple-site features
- Little progress on network convergence

The 1st "Dial"

Virtualization Adoption

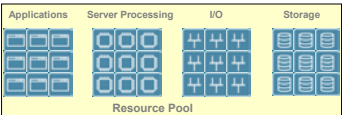
Virtualization: Think Beyond Storage



Virtual Server Virtual Server Virtual Server Virtual Server



Intelligent Fabric



Resource Pool

Think in 3D

1. Compute resources
2. Network resources
3. Storage resources

“Virtualized Data Center” This Will Take Some Effort!

- Server virtualization
 - Virtualizes the physical CPU, Memory, I/O of servers
- Server edge / Network virtualization / IO virtualization
 - Virtualizes the physical network topology and network identifiers
- Storage virtualization
 - Virtualizes physical block storage devices
- File virtualization
 - Virtualizes files and namespaces across file serving resources

Example Issues To Resolve

Issue: Current server virtualization stresses fabric

- **Much higher SAN attach rate than traditional applications**
- **Shared storage & CLVM is common deployment for VM mobility**
- **HBAs must be specifically qualified to run at hypervisor level for VMWare**
- **Single HBA shared across all virtualized guest OSes**

Example Issues To Resolve

Issue: Must invest in cutting edge management

- **NPIV: Present a virtual n_port to guest OSes in virtual machines**
- **Allow storage administrators to use standard tools to meter and bill storage**
- **Auto-confirm virtualization compatibility in key management tools**

Potential Futures from Virtualization

- Advanced services capabilities
- Painless migrations
- Much better SLA capabilities
- If not "end-to-end", at least federated capabilities...

The 2nd "Dial"

Data Protection Innovations

Key Shift: All About Recovery

- Recovery-based innovation sets pace
- D2D2T: Create multiple disk tier environments
 - CDP: Recovering right data, right time
 - Emulation: Getting from tape to disk
 - DPM: Automating and managing recovery

The Three Major Goals of Recovery Management

#1. Ensure recovery

- Improve success rates and reliability

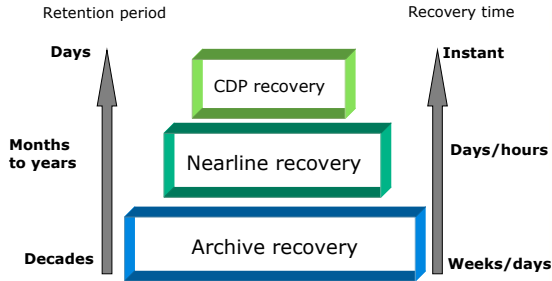
#2. Speed up recovery

- Decrease risk exposure

#3. Integrate recovery

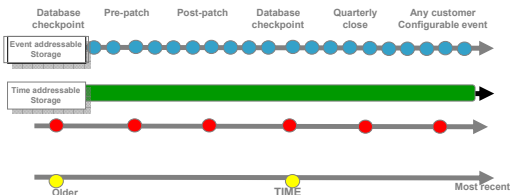
- Connect top-level applications to recovery sets

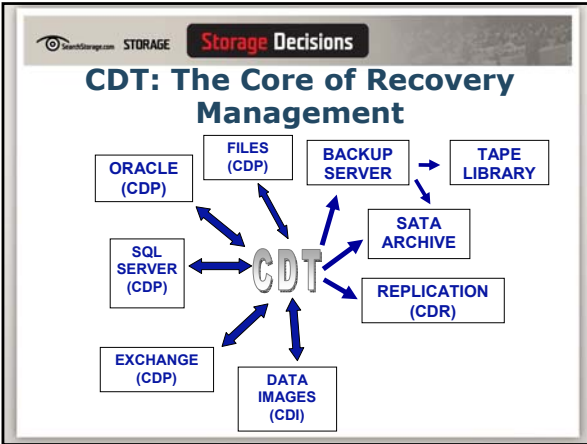
Recovery Is a Continuum



Key Enabler: Continuous Data

- Daily backup Snapshots → Point-in-time (PIT) based
- Any point in time (APIT) → Vanilla "CDT" → App-Aware APIT recovery
- Annotated business/application processes →

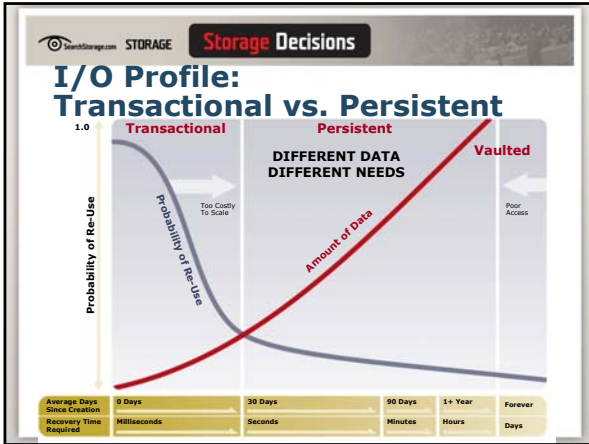




-
- Net-Net on Data Protection Innovations**
- BU/R Is Now "recovery management"
 - Requires new tools investment
 - Co-ordination of elements matters now
 - 3rd Party vendor innovation is key

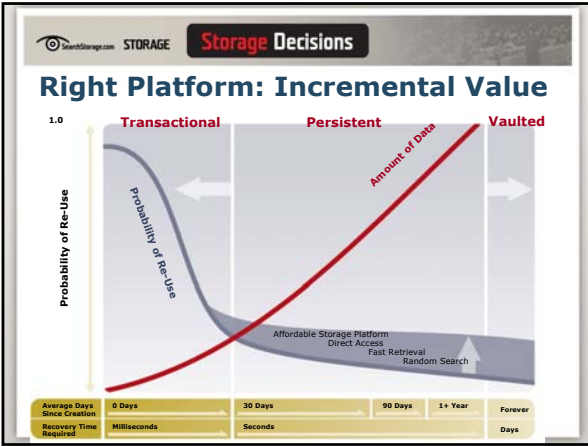
The 3rd "Dial"

I/O-Specific Array Architectures



Transactional Data	Persistent data
<p><u>Origins and traits</u></p> <ul style="list-style-type: none"> • DB, OLTP, ERP, email • Highly dynamic • Short shelf life • High IOPS • Random read/write • Information capture & creation • Structured data (mostly) • Consistency restrictions 	<p><u>Origins and traits</u></p> <ul style="list-style-type: none"> • BU/R, archives, records • Immutable • Long-term retention • Data integrity • Bandwidth centric • Event-driven • Reference content • Data accumulation

-
- STORAGE Storage Decisions**
- ### So, What Makes A Persistent Platform?
- Cost-effective, modular disk
 - Hyper-density
 - Multi-modal access
 - Pluggable services (de-dup, search, VTL, etc.)
 - Energy efficiency
 - Long-term retention viability



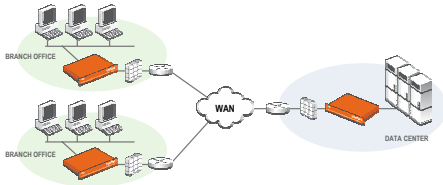
Key Insights on I/O Profile

- Insight: Without I/O profiling, we don't break the "one-size" monolithic array mindset
- Transactional platforms not optimal for persistent duty. No excuse anymore!
- Don't ignore Persistent I/O! It's not a stepchild
- You will never optimize storage ROI if you don't manage information by I/O profile

The 4th "Dial"

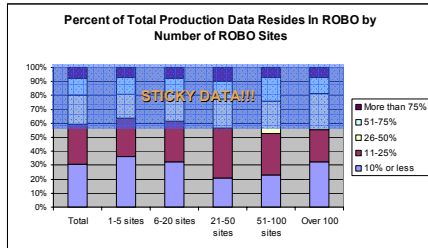
Remote and Branch Technologies

ROBO: More Than Shiny Boxes



70% Say: "Edge strategy is core!"

Why ROBO Will Shape Storage: Data at Edges Will Stay There!



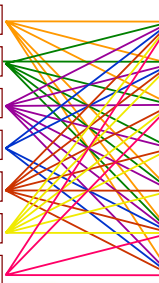
ROBO Is Distributed Computing

IT Business Drivers

- Productivity
- Cost reduction
- Security best practices
- Business continuity
- Globalization
- IT consolidation
- Regulatory compliance

ROBO IT Initiatives

- Server consolidation
- Web apps
- Server-based Computing
- VoIP
- Backup consolidation
- Disaster recovery
- Collaboration



Some Top ROBO Issues To Solve

1. Security-related issues
 - Servers, apps, networks, users
2. Distributed collaboration
 - Product development, product workflow issues
3. Data Protection and disaster recovery
 - Backups, site fail-over, remote replication
4. Infrastructure consolidation and optimization
 - Applications, servers, storage, networks

Hot ROBO Innovation Areas

- WAN optimization
- Capacity optimization
- Application acceleration
- Data coherency controls

Potential Returns From Today's ROBO Investments

- Storage parked at edges, but still managed
- Massive data and network reduction
- Advanced app virtualization

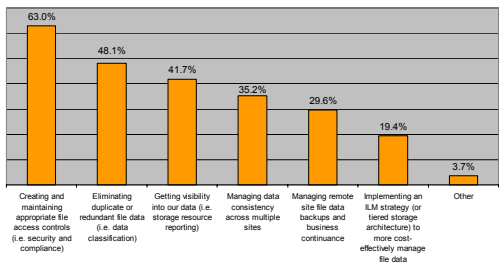
The 5th "Dial"

File Management Strategies

Unstructured Data: Tail That Wags The Data Center

- Massive data growth in files (3x-7x annually)
- Very poor data visibility
- Poor utilization/lowered ROI
- Poor performance for file-based data
- File and NAS management/scaling complexity
- Remote/branch office collaboration issues
- Data consolidation challenges
- Legal and/or compliance pressures

Issues: Controls, Redundancy, Visibility



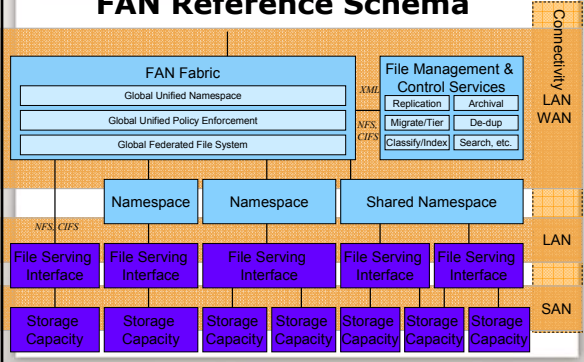
Taneja Group File Management Survey, 2006

A Way Out: File Area Networking

A FAN is comprised of the following six elements:

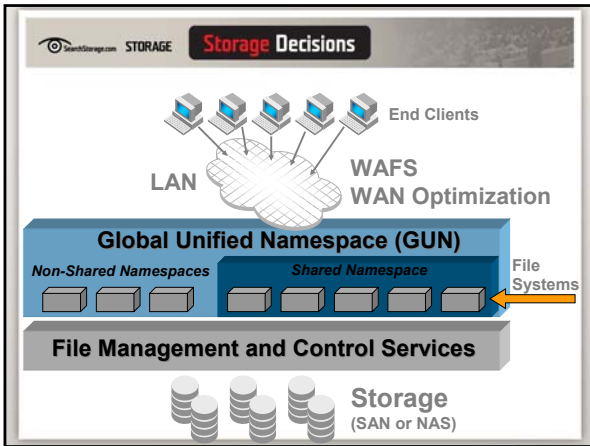
1. Storage devices
2. File serving devices and/or interfaces
3. FAN fabric: Namespaces, policies, and advanced file system semantics
4. The end client machines in an enterprise
5. Connectivity between end client machines and file namespaces
6. File management services that augment the namespaces and/or behalf of both end clients or administrators

FAN Reference Schema



What a Coherent FAN Provides...

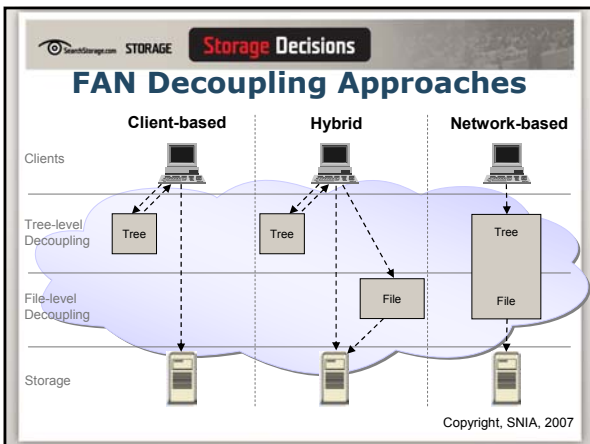
- **CONTROL:** Enterprise-wide, pervasive controls of file data.
- **VISIBILITY:** File visibility and access based on business values.
- **TRANSPARENCY:** Seamless access across geographies.
- **SERVICE LEVERAGE:** Ability to deploy software as a true "service" to the entire infrastructure, not app-specific silos.
- **ROI PLATFORM:** Measurable ROI for file data



Storage Decisions

The FAN Mindset: Decoupling Approach Drives It All

- Client-based (out-of-band)
 - OS service or agent loaded on client system
 - Tree-level granularity with asynchronous updates
- Hybrid (dual-band)
 - Combines client-based and network-based
- Network-based (in-band)
 - Continuous network-resident decoupling
 - File-level granularity and synchronous updates



Potential File Management Futures

- “Virtual” file access capabilities
- Enterprise-wide controls
- Easily deployed new services
- Network-resident controls

Pulling It Together...

The “Dials” For Our Future

1. Virtualization adoption
2. Data protection innovations
3. I/O-specific array architectures
4. Remote-branch office tech
5. File management strategies

How Likely Are Our Scenarios?

- Scenario: "Aggressive transformation"
~15% probability
- Scenario: "Moderated advancement"
~75% probability
- Scenario: "Conservative change"
~10% probability

Summary

- What you buy TODAY will determine our collective future
- Explore, support and reward true innovation
- Create a mix of 3rd party and start-up vendors
- Put vendor vision through a reality test

Questions!

brad@tanejagroup.com
