

EMC Backup Recovery Systems

Next Generation Backup Solutions

Erwin Freisleben
Advisory Technology Consultant

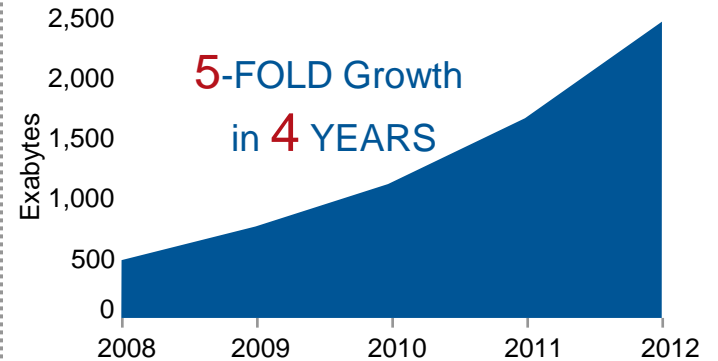
The Backdrop for Data Protection



Major trends influencing the transformation of backup environments

- **Data growth**
 - Typically represents a factor of 4 to 30 times production capacity
 - Daily, weekly, and monthly full backups kept for months or years
 - New requirements to keep more data for longer periods
- **Server virtualization**
 - Dynamic environment causes increased complexity
 - VM sprawl creates protection challenges
 - Server consolidation creates high server utilization leaving little bandwidth for backup

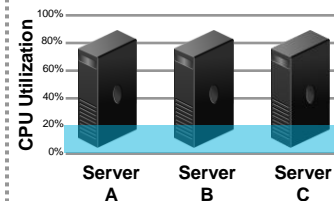
Digital Information Created and Replicated Worldwide



Source: IDC Digital Universe white paper, sponsored by EMC, May 2009

Old Paradigm

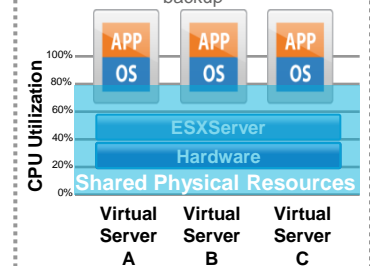
Physical Environment: Low overall server utilization and plenty of bandwidth for backup



20 percent resource utilization

New Paradigm

Virtual Environment: High overall server utilization and little bandwidth for backup



80 percent resource utilization

Data Deduplication is the Enabler

Insights from Gartner



Key findings:

“Production deployments of deduplication for backups have progressed at an unusually high rate for such a recent technology.”

Market implications:

“Gartner views this technology as transformational because it radically decreases the economics of disk-based backup and recovery.”

“...too compelling to ignore.”

Recommendations:

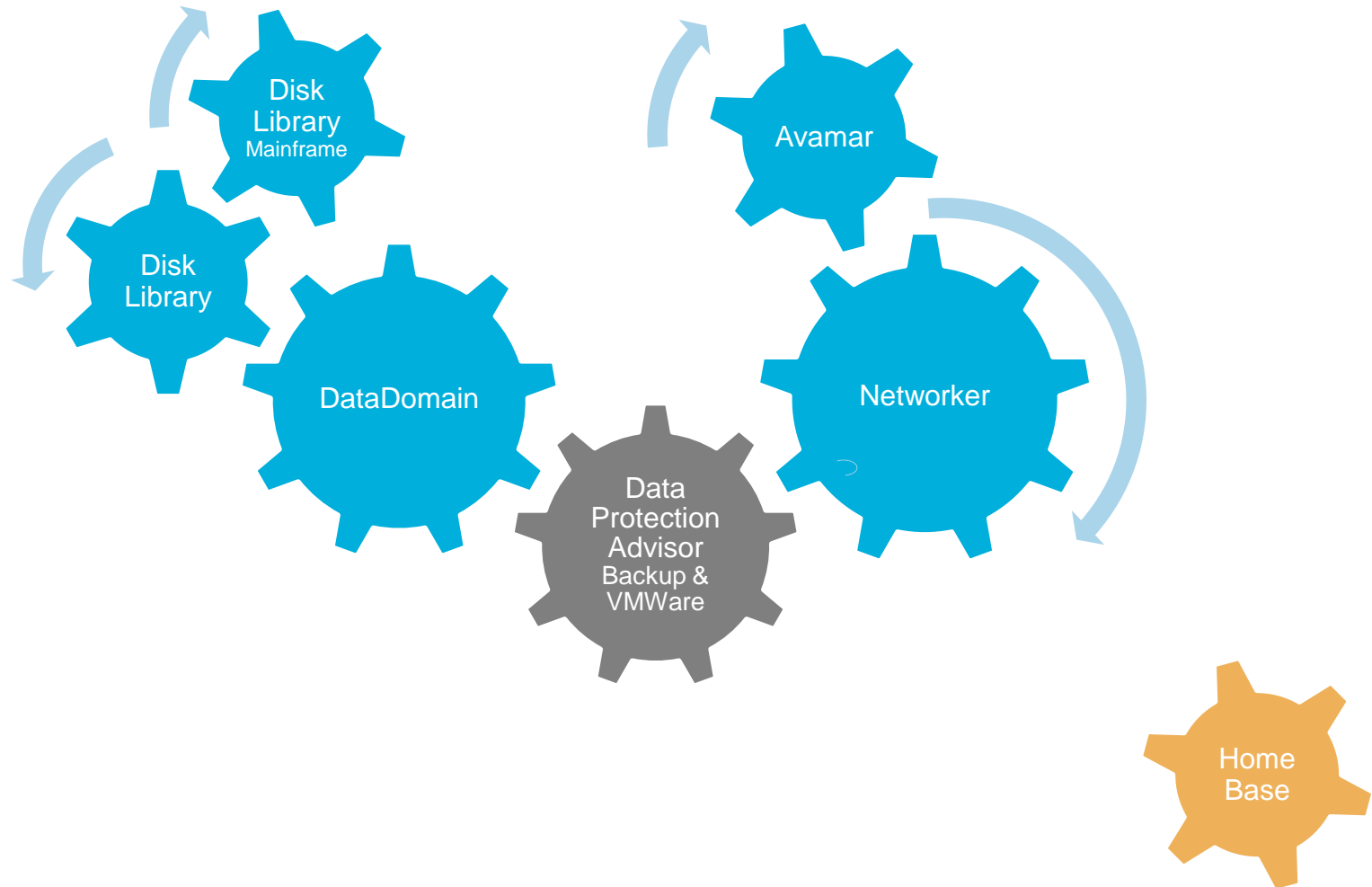
“...ensure that your organization is comfortable with the robustness and maturity of the vendor's approach.”

“By 2012, deduplication will be applied to 75% of all backups.”

Gartner.

Source: Dave Russell, Gartner.

BRS Product Portfolio



Data Protection Advisor

Unified monitoring, analysis, and reporting



Single view for data protection management, including backup applications, replication, and continuous data protection (CDP)

Unify—See—Act

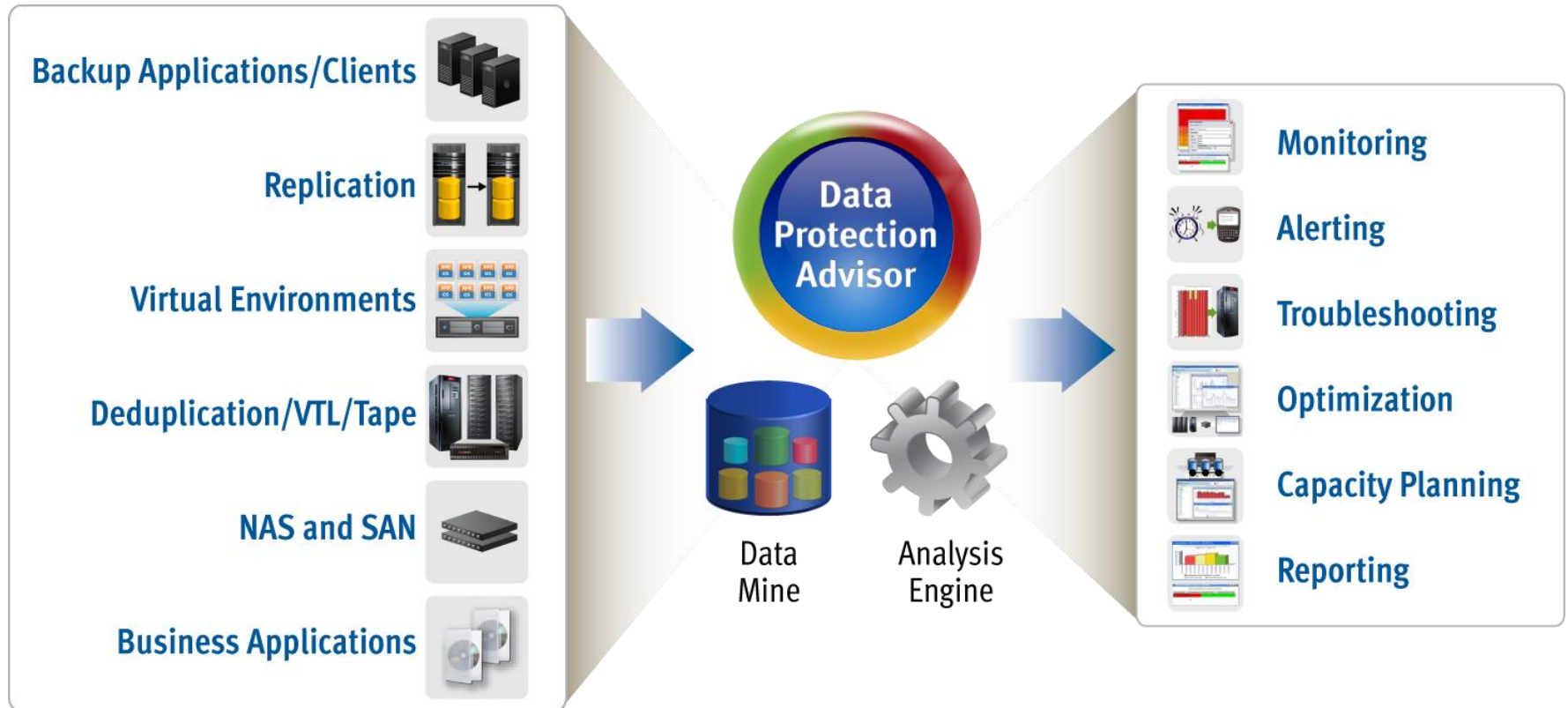
- Automatically collect data from across data protection infrastructure
- Transform disparate data into new insights, actionable information, and service-level agreement statistics
- Make better, faster decisions to meet service level agreements, correct issues, and drive efficiency

Payback typically 12 months or less

- Reduces costs with automation and new insight that drives efficient use of people and infrastructure
- Improves compliance, lowers risk with greater visibility into service-level agreement recoverability statistics
- Minimizes complexity with a single, integrated view across the environment

With EMC Data Protection Advisor...

Automated, single view eliminates manual effort, data silos



Payback typical in 12 months

- ✓ Reduces costs—insight drives efficiency
- ✓ Lowers risk—manage to service level agreements
- ✓ Minimizes complexity—single, integrated view

Data de-duplication: WHERE

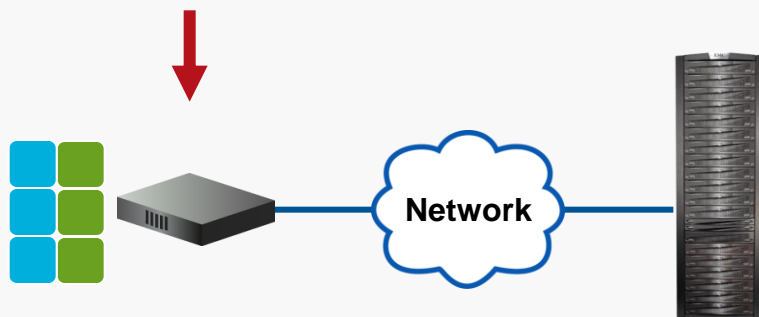
Source

- Client software agents identify repeated sub-file data segments at the source
- Only new, unique segments are transferred across the network and stored to disk during backup operations
- Benefits include shorter backup window and lower bandwidth requirements

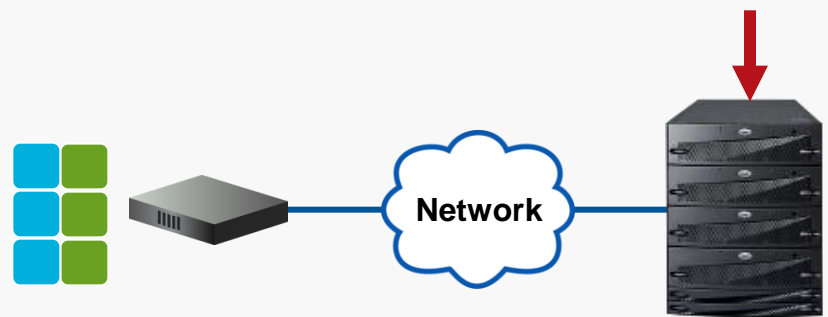
Target

- Backup application sends native data to a target storage device
- Data is de-duplicated once it reaches the target
- De-duplication can happen during or after backup
- Found in VTLs or LAN B2D appliances
- Transparency to backup application offers users a “plug and play” experience

DE-DUPLICATION AT SOURCE

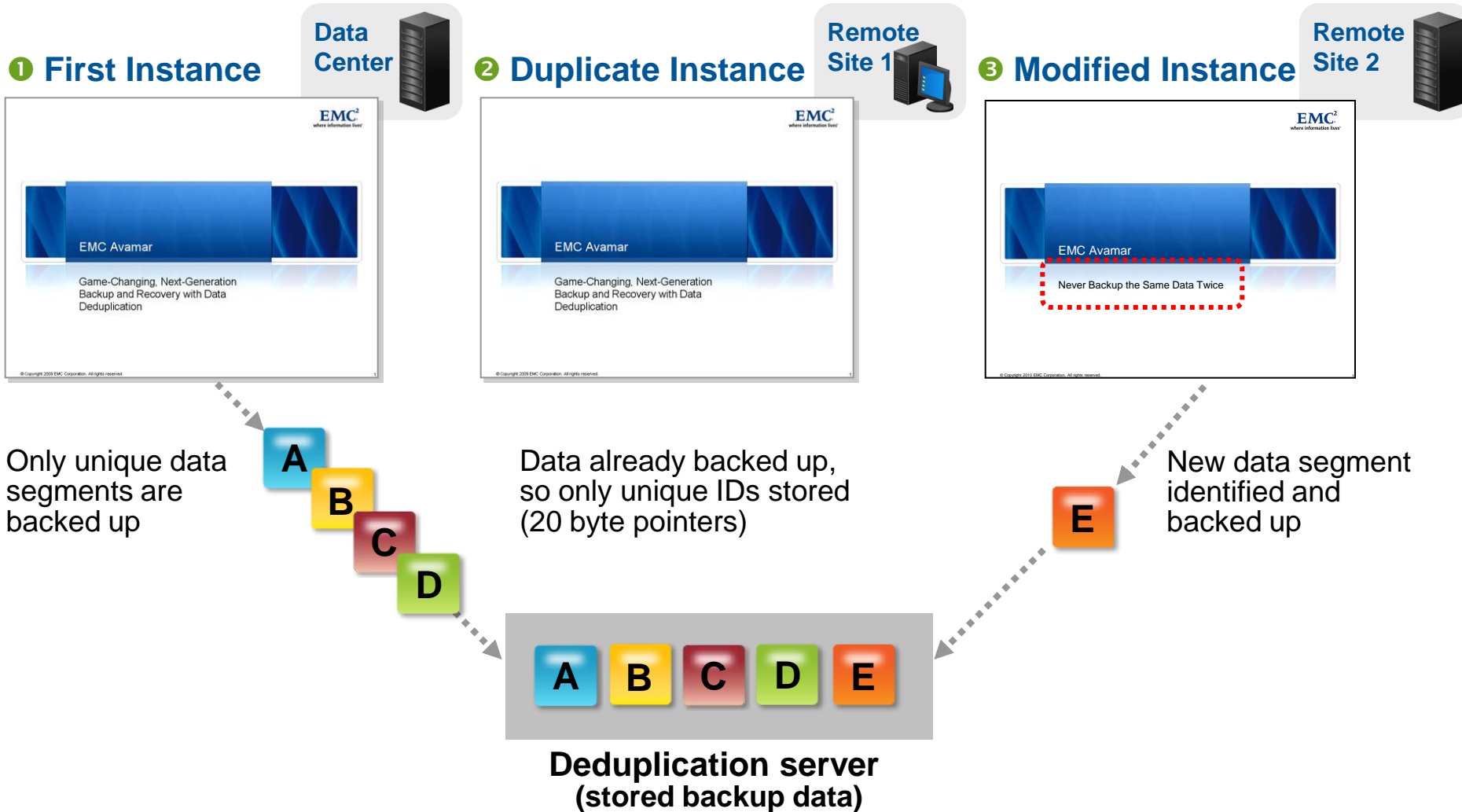


DE-DUPLICATION AT TARGET



Avamar: How it Works

Simple example of global, sub-file deduplication



Avamar: Real-World Results

Avamar daily full backups versus traditional daily full backups



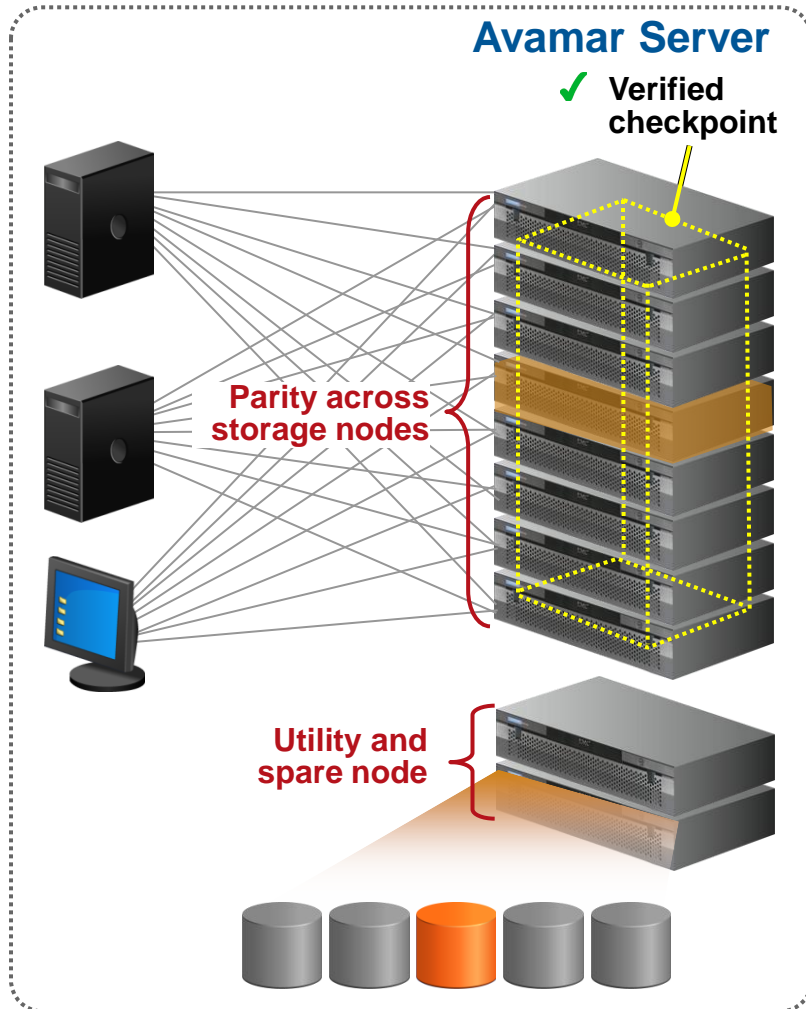
Data Type	Amount of Primary Data Backed Up	Amount of Data Moved Daily
Windows file systems	3,573 GB	6.1 GB
Mix of Windows, Linux, and UNIX file systems	5,097 GB	11.7 GB
Engineering files on NAS (NDMP backups)	3,265 GB	24.2 GB
Mix of 20% databases, 80% file systems (Windows and UNIX)	9,583 GB	80.0 GB
Mix of Linux file systems and databases	7,831 GB	104.2 GB

Source: EMC

While results will vary by data type and mix, Avamar can dramatically improve backup performance and efficiency

Avamar's Innovative Architecture

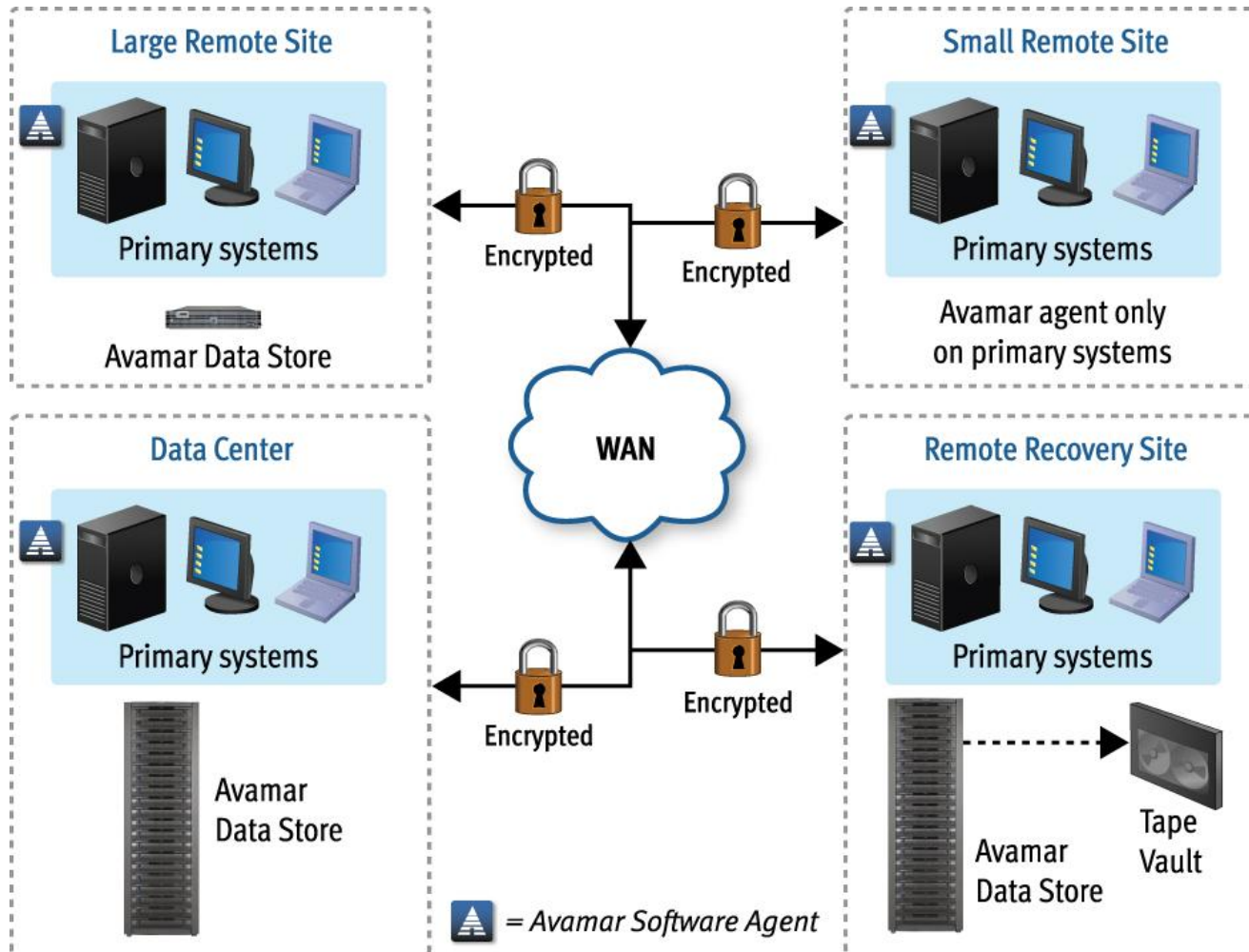
EMC²
where information lives[®]



U.S. Patent No. 6,826,711

- Redundant Array of Independent Nodes (RAIN) architecture
 - Each server node includes CPU, memory, and internal disk storage
 - Provides high availability and fault tolerance across nodes
- Grid architecture for online scalability and performance
- Daily Avamar server integrity checks
- Data recoverability verified daily
- RAID protection from disk failures
- Available for Avamar software and Avamar Data Store

Avamar's Flexible Deployment





Fully integrated software/hardware solution

- Complete EMC backup, recovery, and disaster recovery offering
 - Avamar backup and recovery software with integrated source/global data deduplication
 - EMC-certified hardware built to order—fully configured and delivered
- Built-in high availability
 - Avamar RAIN technology
 - Spare node, RAID
 - Redundant power distribution
 - Redundant switches, NICs for automatic failover on user's network
- Simplifies
 - Purchase (single vendor, certified hardware)
 - Deployment (minimizes onsite setup)
 - Service (single vendor support)

Avamar Supported Operating Systems and Applications



Client operating systems supported

- Microsoft Windows 7 Professional, Enterprise, Ultimate
- Microsoft Windows Server 2003 Standard and Enterprise
- Microsoft Windows Server 2008 R2
- Microsoft Windows Vista and XP
- Red Hat Enterprise Linux (RHEL) 3.0, 4.0, 5.0
- Red Hat Enterprise Linux (RHEL) 5.4 on zLinux
- Red Hat Linux 7.2, 7.3, 9.0
- Solaris 8, 9, 10, 10 x86
- SUSE Linux Enterprise Server (SLES) 8, 9, 10, 11
- SUSE Linux Enterprise Server (SLES) 10SP3 on zLinux
- Apple Macintosh OS X 10.4.x, 10.5.x, 10.6.x
- CentOS 4.x, 5.x
- Debian 4, 5
- Free BSD 6.2
- HP-UX 11.0, 11iV1, 11iV2, 11iV3
- IBM AIX 5.2, 5.3, 6.1
- NetWare 6.5
- Novell Open Enterprise Server OES SP2
- Oracle Enterprise Linux R5
- SCO OpenServer 5.0.5, 5.0.6, 5.0.7
- SCO UnixWare 7.1.3

Application modules

- IBM DB2 9.1, 9.5, 9.7
- IBM Lotus Domino 6, 7, 8
- Microsoft Exchange 2003, 2007, 2010
- Microsoft Office SharePoint Server 2007
- Microsoft SQL Server 7.0, 2000, 2005, 2008
- NDMP (EMC Celerra with DART 5.5, 5.6)
- NDMP (NetApp Data ONTAP 6.5, 7.0.4, 7.0.5, 7.0.6, 7.1x. 7.2, 7.3.x)
- Oracle 9i, 9.2, 10, 10g, 10gR2, 11g, 11gR2

VMware infrastructure

- ESX 3.0.x, 3.5, 3i
- vSphere 4

• = ESX client operating system-certified using RHEL 4 client (support within virtual machines depends on VMware support of client operating system)

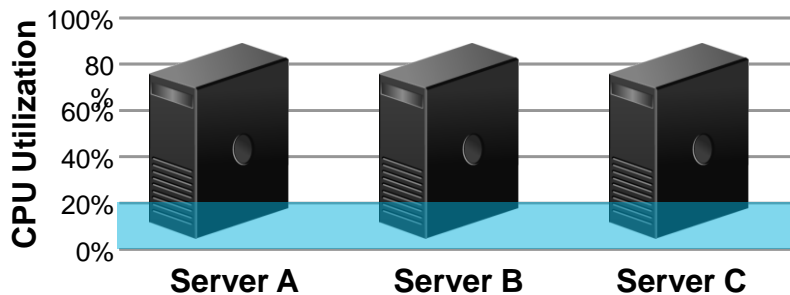
No charge for client agents or application modules

VMware Backup Challenges

Virtualization changes server, application, and information paradigm...
backup must evolve to deliver even greater consolidation and value

Old Paradigm

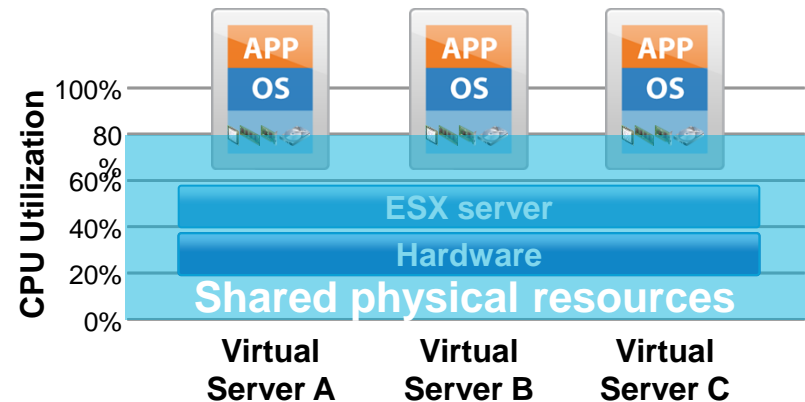
Physical environment: Low overall server utilization and plenty of bandwidth for backup



20% resource utilization

New Paradigm

Virtual environment: High overall server utilization and little bandwidth for backup



80% resource utilization

Avamar Solutions for VMware Infrastructure

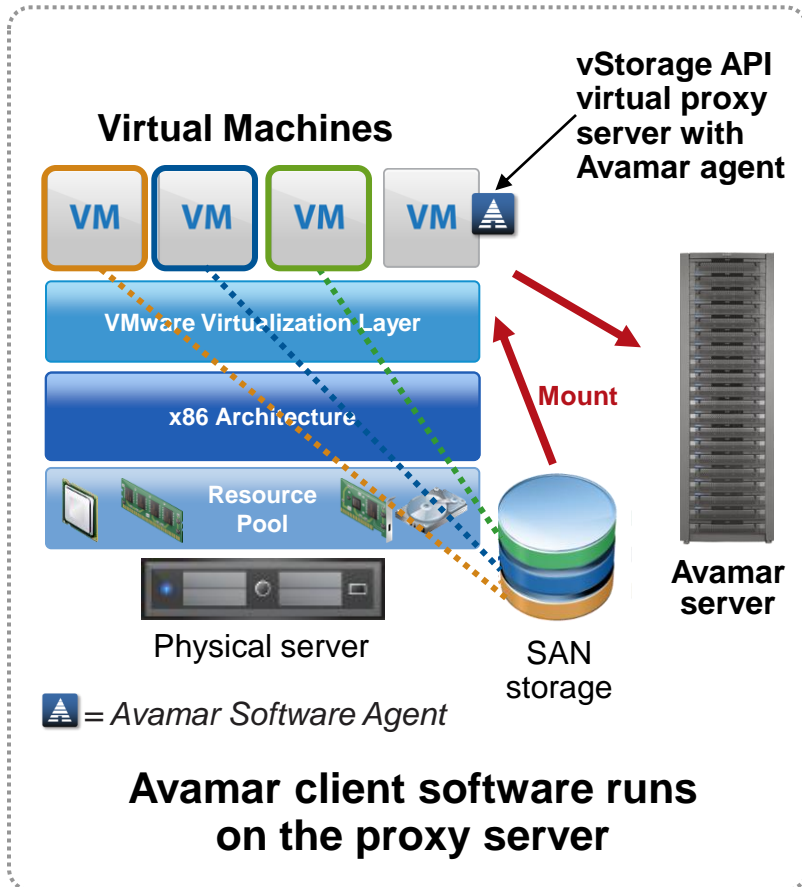
Flexible, fast, efficient, and reliable backup and recovery



Avamar and VMware vSphere 4 integration

- Universal support for all vSphere 4 backup options
 - Daily full guest- and image-level backups
- vStorage API integration
 - Greater backup and restore flexibility
- VMware vCenter Server integration
 - Single point of management within Avamar for all VMware backup options

Avamar Image-Level Backups through Storage APIs



- Avamar agent resides on the proxy server
- Fast, daily full backups
- Deduplicates within and across VMDK files
- Supports both file-level (Windows only) and image-level recovery
- Avamar replication provides disaster recovery for VMDKs
- CBT Support

Avamar 5 Changed Block Tracking (CBT) Effect

With Change-Block

Session									
Status	Error Code	Start Time (GMT)	Elapsed	End Time (GMT)	Type	Progress Bytes	New Bytes	Client	
✓ Completed		2009-11-10 21:47	00h:00m:56s	2009-11-10 21:48	On-Demand Backup	60.0 GB	<0.05%	AVAMAR079	
✓ Completed		2009-11-10 21:47	00h:00m:56s	2009-11-10 21:48	On-Demand Backup	60.0 GB	<0.05%	AVAMAR028	
✓ Completed		2009-11-10 21:48	00h:01m:04s	2009-11-10 21:49	On-Demand Backup	60.0 GB	<0.05%	AVAMAR041	
✓ Completed		2009-11-10 22:13	00h:01m:10s	2009-11-10 22:14	On-Demand Backup	100.0 GB	<0.05%	AVAMARX136	
✓ Completed		2009-11-10 21:45	00h:01m:15s	2009-11-10 21:46	On-Demand Backup	8.0 GB	0.1%	AVAMARX060	
✓ Completed		2009-11-10 21:45	00h:01m:15s	2009-11-10 21:46	On-Demand Backup	8.0 GB	0.1%	AVAMARX062	
✓ Completed		2009-11-10 21:43	00h:01m:59s	2009-11-10 21:45	On-Demand Backup	8.0 GB	0.1%	AVAMARX063	
✓ Completed		2009-11-10 21:43	00h:01m:59s	2009-11-10 21:45	On-Demand Backup	8.0 GB	0.2%	AVAMARX061	
✓ Completed		2009-11-10 21:43	00h:01m:59s	2009-11-10 21:45	On-Demand Backup	8.0 GB	0.2%	AVAMARX061	
✓ Completed		2009-11-10 21:43	00h:01m:59s	2009-11-10 21:45	On-Demand Backup	8.0 GB	0.2%	AVAMARX061	
✓ Completed		2009-11-10 22:09	00h:25m:58s	2009-11-10 22:35	On-Demand Backup	40.0 GB	0.2%	AVAMAR253	
✓ Completed		2009-11-10 22:35	00h:26m:15s	2009-11-10 23:01	On-Demand Backup	40.0 GB	0.1%	AVAMAR243	
✓ Completed		2009-11-10 21:43	00h:29m:52s	2009-11-10 22:13	On-Demand Backup	40.0 GB	0.5%	AVAMAR252	
✓ Completed		2009-11-10 22:15	00h:36m:08s	2009-11-10 22:51	On-Demand Backup	20.0 GB	0.1%	AVAMAR101	
✓ Completed		2009-11-10 21:47	01h:13m:45s	2009-11-10 23:00	On-Demand Backup	250.0 GB	<0.05%	MACHINE11	

100GB VM takes 1:10

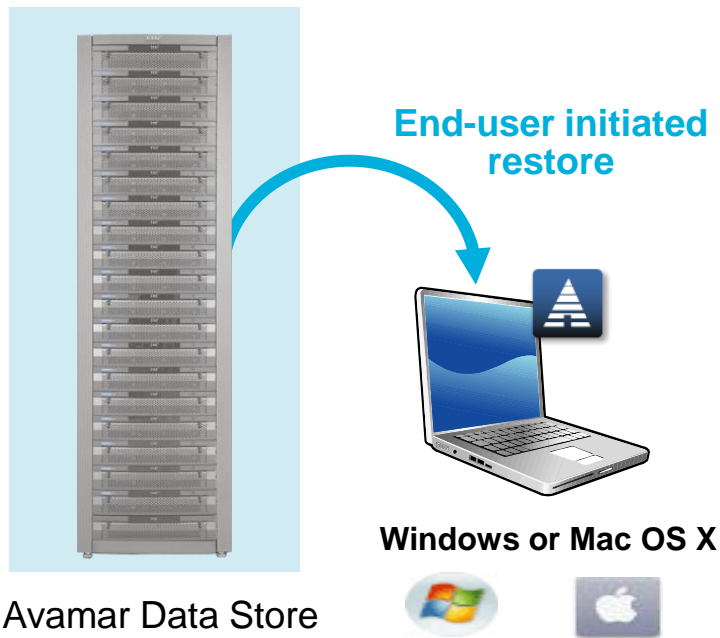
40GB VM takes 26:15

Without Change-Block

Desktop/Laptop Backup Challenges and Avamar Solution



Avamar Desktop/Laptop



- Puts power in end-users hands
 - Eliminates Help Desk for restoring data to desktop or laptop
 - End users restore their data anytime, anywhere
 - Self-managed restores
 - Document-level search engine
 - Multiple language support
- All the benefits of Avamar backup
 - Data deduplicated at the source
 - Daily full backups—up to 10-times faster
 - Minimal impact on PC and network resources during backup

Avamar's Unique Value



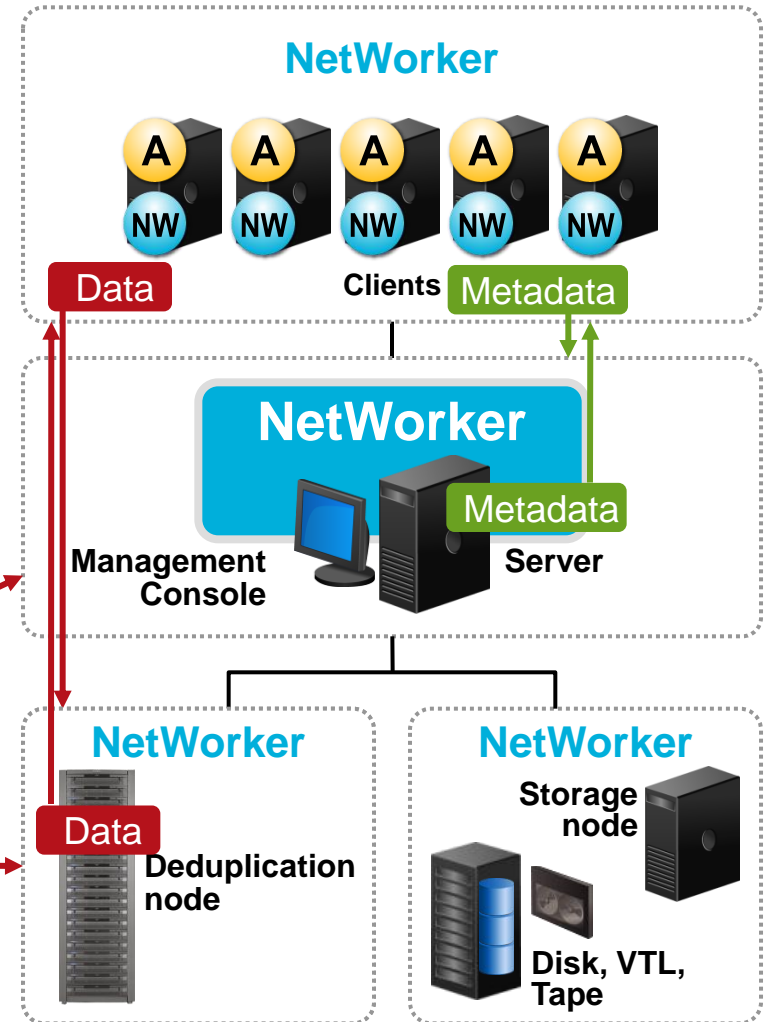
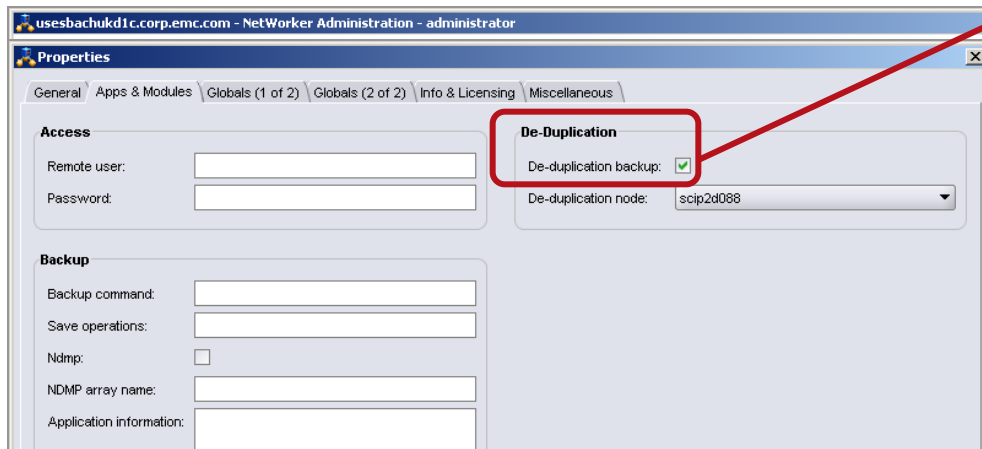
- **Deduplicates data at the client (source)**
 - Fast, daily full backups through existing infrastructure
- **Deduplicates data globally**
 - Reduces required backup disk storage by up to 50 times
- **Fast, single-step recovery**
 - Eliminates hassle of recovering last good full and subsequent incrementals to reach desired recovery point
- **Recoverability verified daily**
 - Daily server integrity checks; data recoverability verified daily
- **High availability**
 - RAIN architecture and replication—no single points of failure
- **Centralized management**
 - Streamlines administration and management of remote backups
- **Flexible deployment options**
- **Complete solution from EMC**
 - Backup, recovery, disaster recovery, and data deduplication

NetWorker and Avamar Integration



How it works

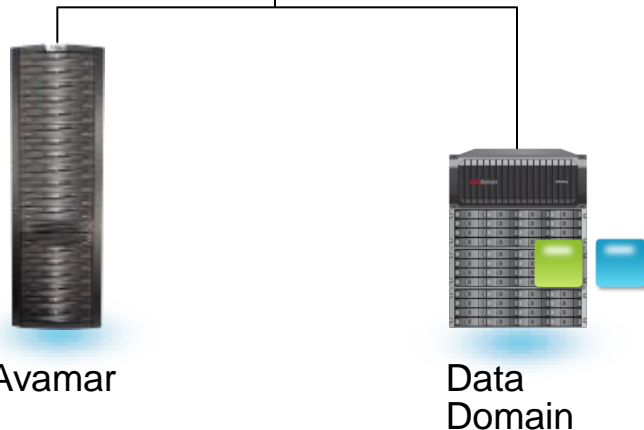
- Integrated client is a standard feature of NetWorker
- NetWorker represents the Avamar Data Store as a deduplication node
- Deduplication is enabled through NetWorker Management Console
- Save process gathers metadata and sends it to the NetWorker server and data to the deduplication node
- On recovery, required deduplication data is requested, and full dataset is restored at client



Networker: Complete Deduplication Management



File Systems and Applications



- Manage both source and target deduplication
 - Choose based on best fit for use case, cost and time to deployment
- Source using NetWorker
 - Integrated with Avamar
 - Single client
 - Common management
- Target using NetWorker
 - Support for Data Domain
 - SAN: VTL software option
 - NAS: NFS, CIFS
 - Compatible with third party solutions

Benefits of Disk-Focused Backup and Recovery Solutions



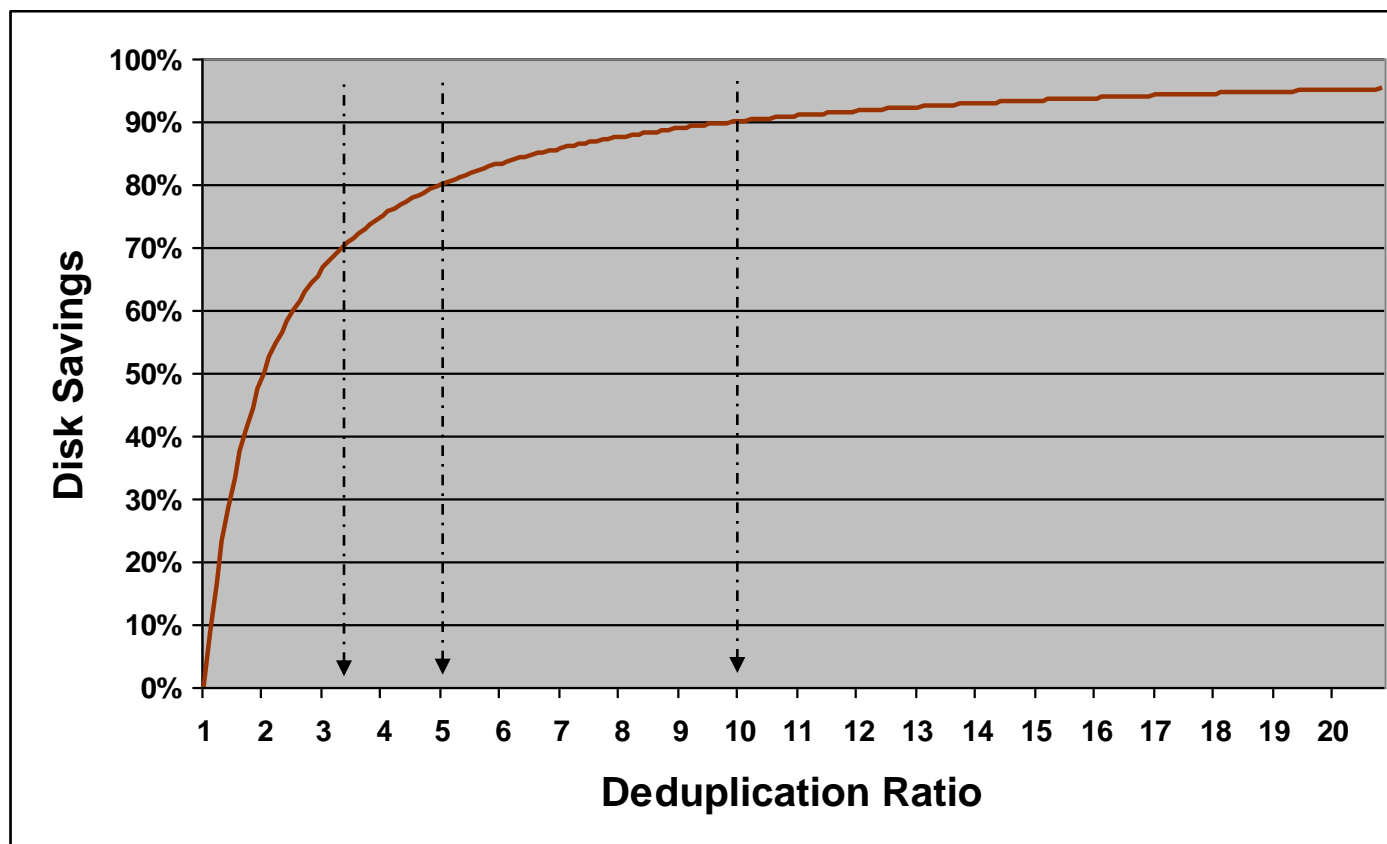
- Performance
 - Up to five-times performance increase for restore
 - Up to three-times performance increase for backup
- Availability
 - Reliable backup and restore operations
- Management efficiency
 - Save hours of staff time
 - Improve productivity
- Data retention
 - Data on disk can be remotely copied—no reliance on shipping
 - Data easily retrieved

Better performing and predictable backup and recovery

De-Duplication Ratio vs Disk Space Savings



It doesn't take a very high de-duplication ratio to save a lot of disk



Dedup Ratio	Disk Savings
1:1	0%
2:1	50%
3:1	67%
4:1	75%
5:1	80%
6:1	83%
7:1	86%
8:1	87%
9:1	89%
10:1	90%
20:1	95%
50:1	98%
100:1	99.0%
500:1	99.8%

Data Domain Deduplication Storage



Control Tier

Backup & Archive Applications



CIFS, NFS,
NDMP, OpenStorage

Ethernet

VTL over FC

Target Tier



DD880 Appliance

- 4U
- 2 - 6 ports
- 10 and 1 Gb Ethernet; 4 Gb Fibre Channel
- RAID-6
- 5.4 to 71 TB usable capacity with shelves
- 1 TB or 500 GB 7.2k rpm SATA HDD in shelf
- File system
- NVRAM
- N+1 fans and redundant, hot-plug power supplies

DR Tier

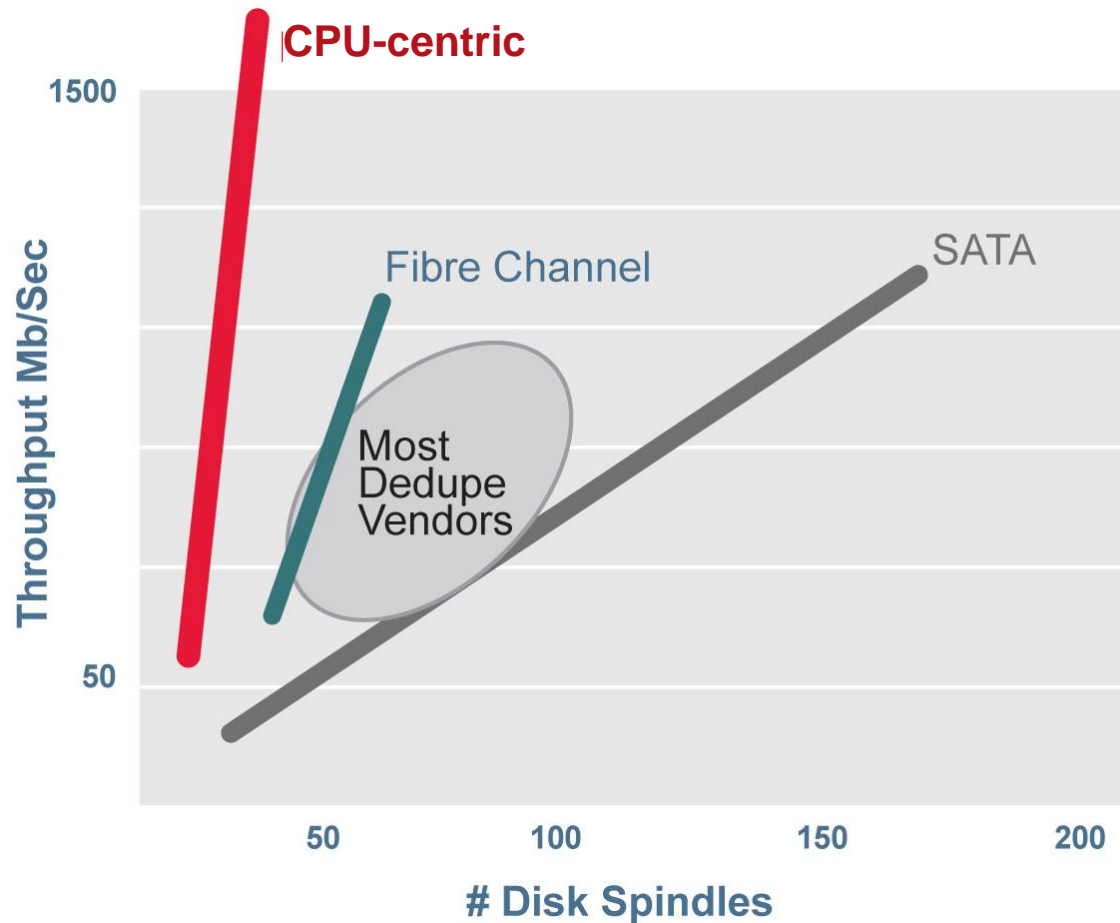


DD880 Appliance

Replication

Easy integration makes a conventional environment transformational

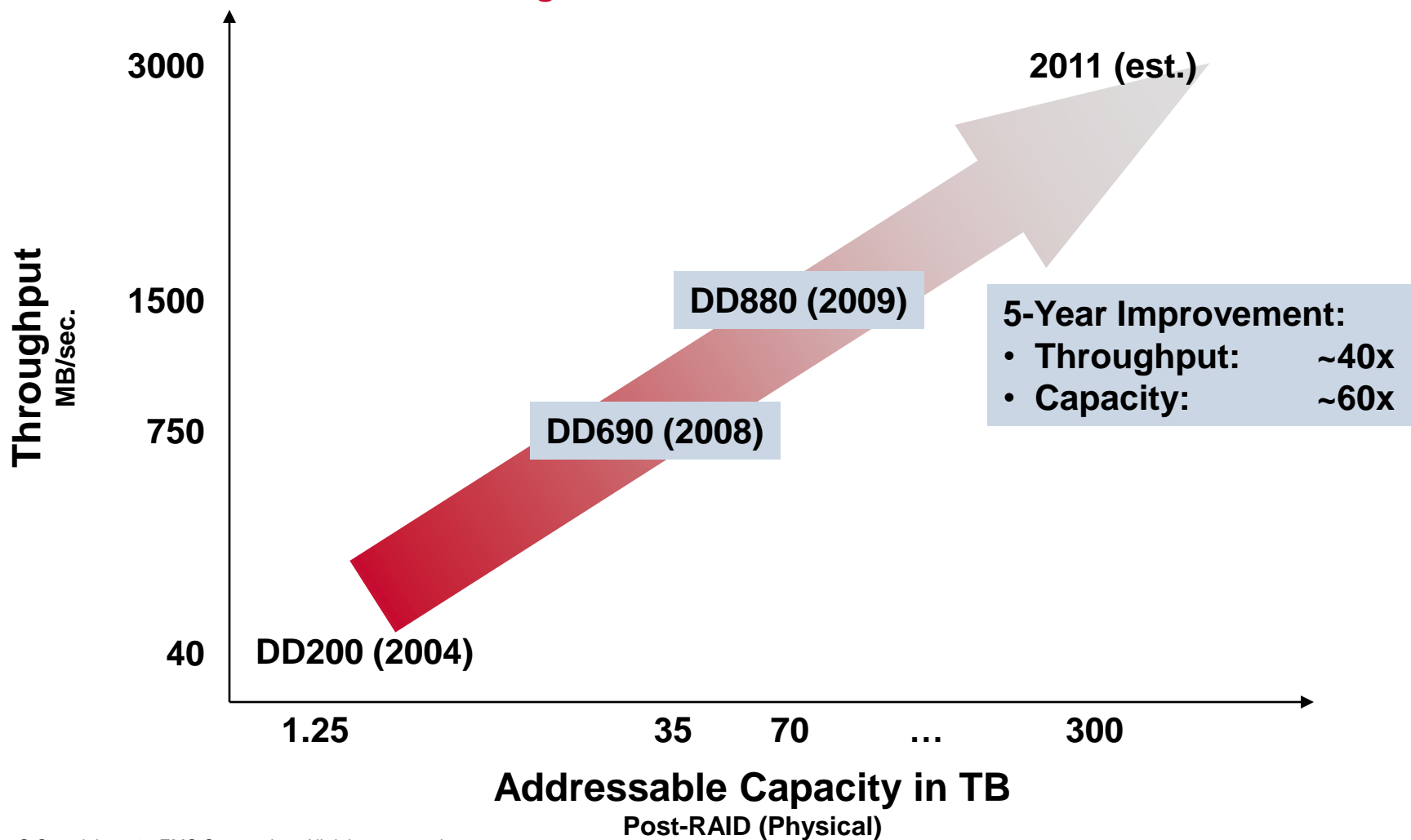
Performance: CPU-centric Dedupe vs. Disk-bound



Scalability: Data Domain Systems Trajectory

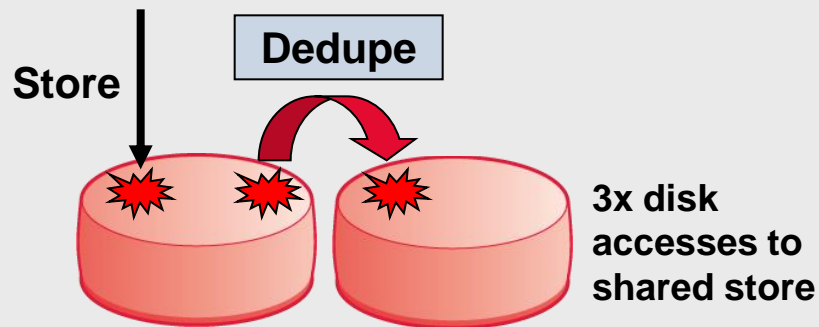


Data Domain SISL™ Scaling Architecture



Inline vs Post-Process Deduplication: Provisioning & Admin

Post Process: Dedupe After Storing

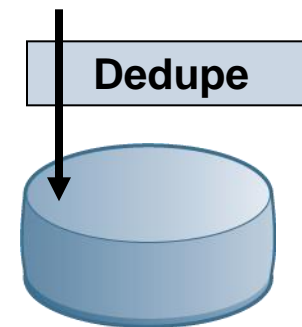


I/O contention if simultaneous actions:

- **Copy to tape:** Too slow to stream tape
- **Recovery:** SLA predictability
- **Replication:** Poor time-to-DR
- **Dedupe itself** if interleaved with backup or restore

More admin to fight these issues

Inline: Dedupe Before Storing



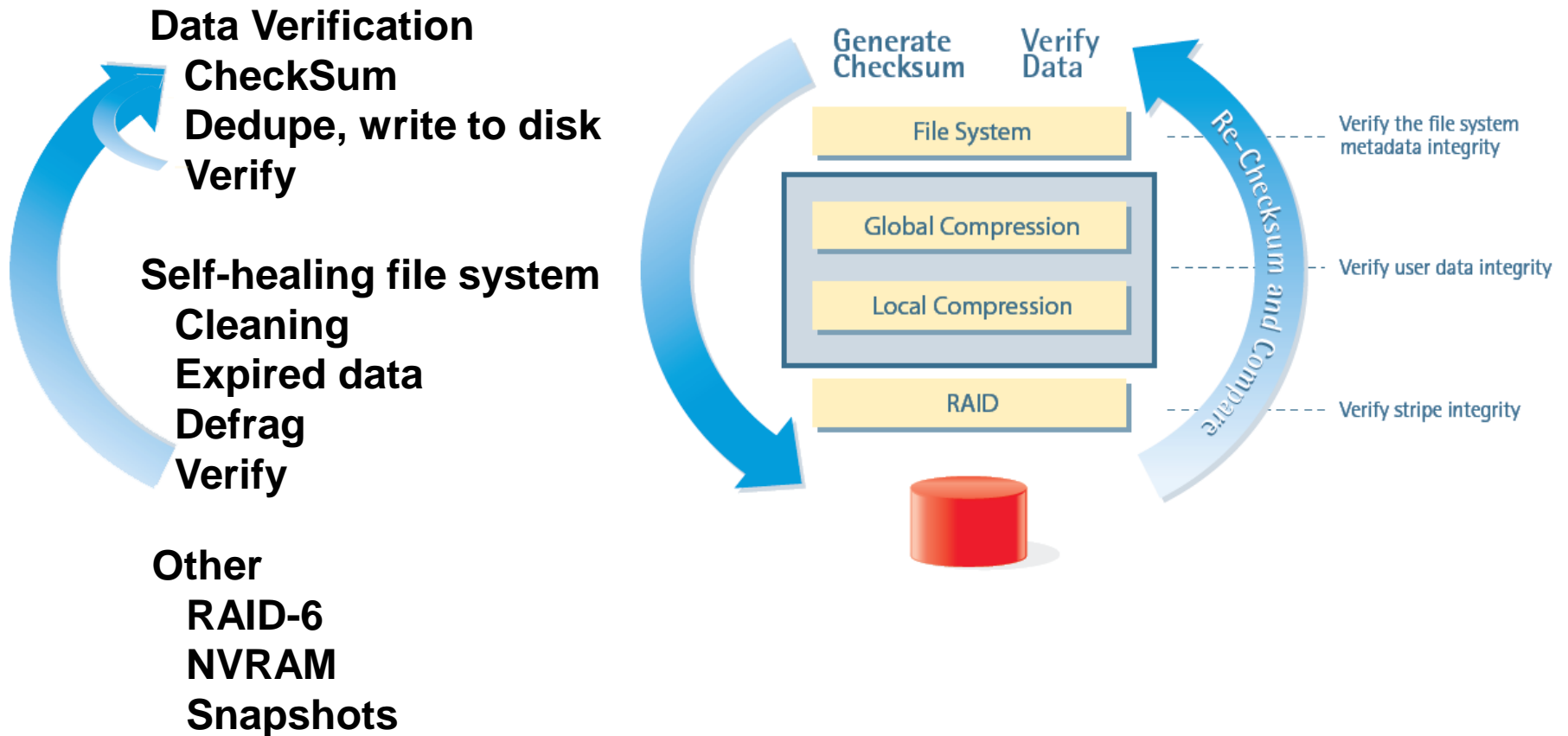
Other activities unimpeded

- **Predictable**
- **Less to manage**
- **Less to provision**

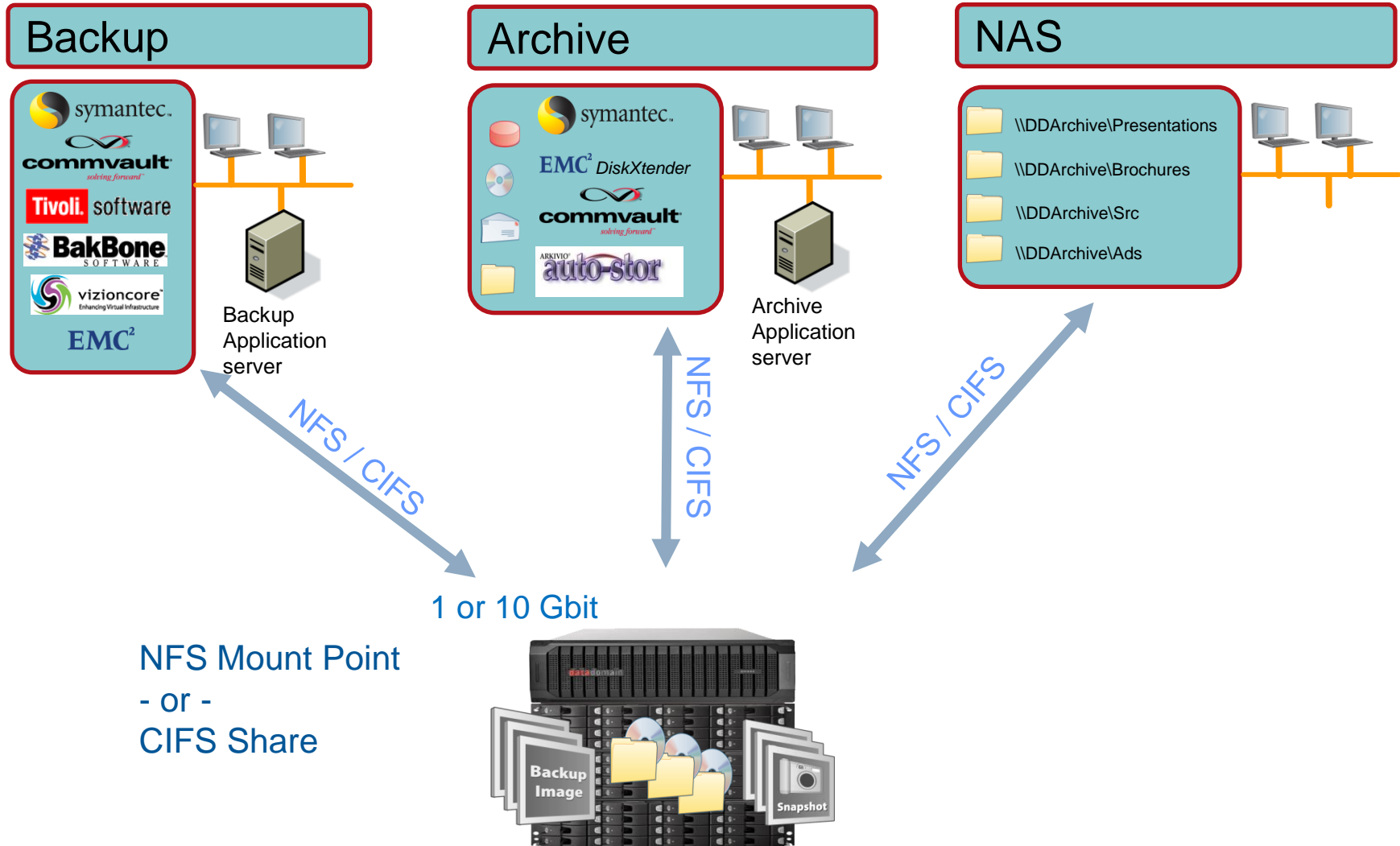
Data Integrity: Data Invulnerability Architecture



Trust but verify – hope is not a strategy

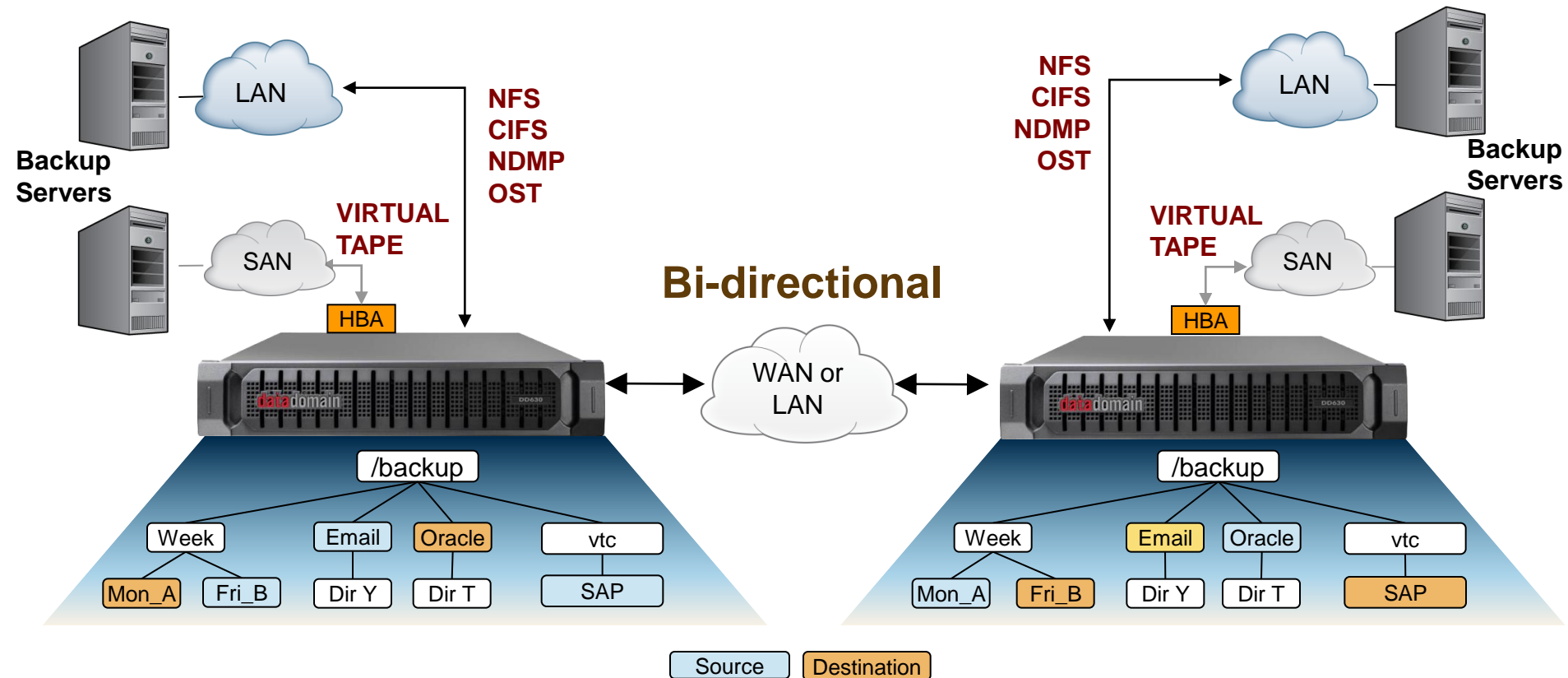


NFS / CIFS – Use Cases



Directory Replication – Cross-site Protection

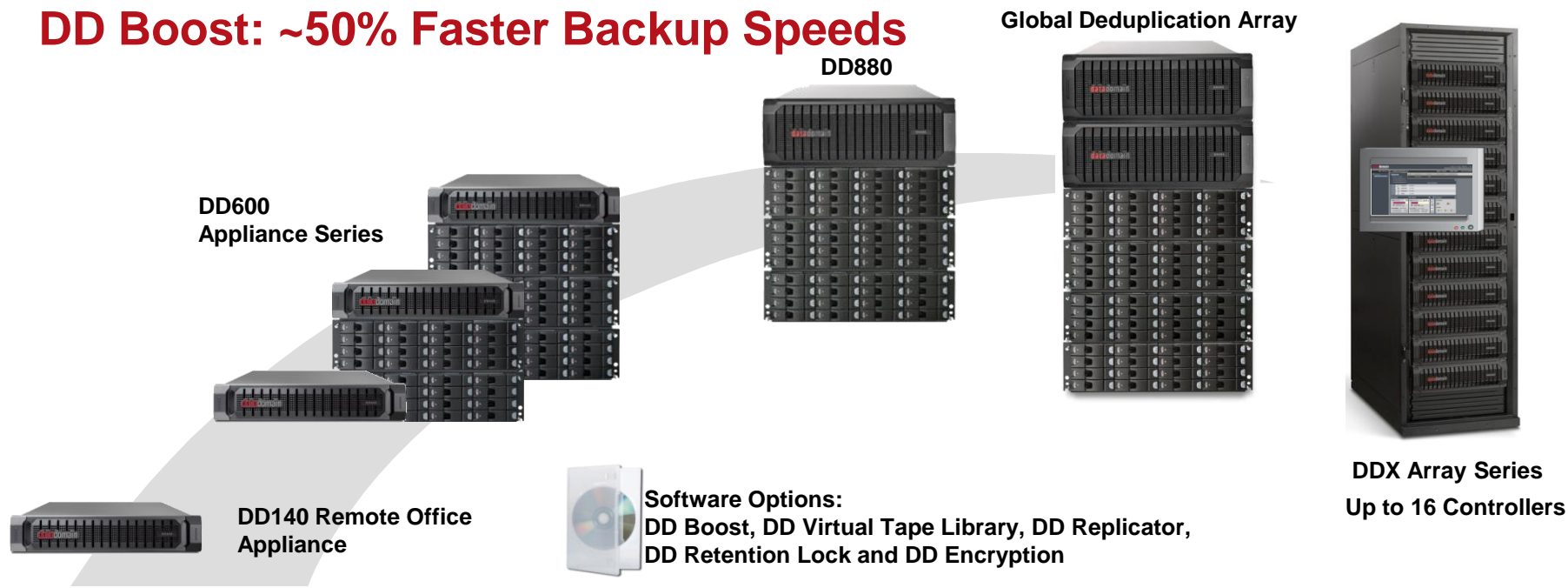
- Replicate critical directories only (weekly copy)
- Flexibility to replicate data bi-directionally
- Tapes can be organized in user defined VTL pool



Industry's Most Scalable Inline Deduplication Systems



DD Boost: ~50% Faster Backup Speeds



	DD140	DD610	DD630	DD660	DD690	DD880	GDA	DDX Array
Speed (Other)	450 GB/hr	675 GB/hr	1.1 TB/hr	2.0 TB/hr	2.7 TB/hr	5.4 TB/hr		86.4 TB/hr
Speed (DD Boost)								
Logical Capacity	17-43 TB	75-195 TB	165-420 TB	.520-1.31 PB	.710-1.7 PB	2.8-7.1 PB	5.7-14 .2 PB	45.6-114 PB
Raw Capacity	1.5 TB	Up to 6 TB	Up to 12 TB	Up to 36 TB	Up to 48 TB	Up to 192 TB	Up to 384 TB	Up to 3.07 PB
Usable Capacity	0.86 TB	Up to 3.98 TB	Up to 8.4 TB	Up to 26.1 TB	Up to 35.3 TB	Up to 142.5 TB	Up to 285 TB	Up to 2.28 PB

Confidential – NDA ONLY
Subject to Change Without Notice

Data Domain Differentiation



- **Maturity**
 - Simple
 - Consistent
 - Robust, e.g., policy-driven dedupe replication
- **Product concept: purpose-built storage**
 - Inline and simple appliance
 - System infrastructure
 - Application independent: backup, archive, and more
- **Architecture: fast, small, storage of last resort**
 - CPU-centric for price/performance
 - Data protection from the ground up



EMC²

where information lives[®]

deduped