IBM VTL Solutions for Open Systems

Mladen Portak, IBM Storage Specialist SEE

Agenda

- Market Drivers for Data Deduplication
- ProtecTIER Overview
- Native Replication
- Product Options
- Customer Case Studies
- Summary

Why Do Invest In ProtecTIER?

Performance

2,500 MB/s ingest
3,200 MB/s Restore

Delivering superior competitive throughput for the Enterprise

Capacity

Up to 1 PB physical capacity per node

Nearly 2x competitors capability

Open Gateway Design

Disk selection freedom

Provides best of breed disk now and with future upgrades

Non-Disruption

Daily Operations
Inline de-duplication eliminates need for significant secondary processing

Implementation
Integrates well with existing backup environment and infrastructure

Continuous Availability

Dual node design for high availability

The only HA solution available

Global Deduplication

Deduplicates ALL data

Provides superior deduplication ratios

Protect More. Store Less.®

Market Drivers for Data Deduplication

What's Driving Deduplication Demand?

1. Need to reduce backup and recovery times
2. Keeping pace with capacity of data to protect
3. Remote site backup
4. Need to improve backup and recovery reliability
5. Lack of disaster recovery plan or process
6. Difficult to validate backup/recovery success
7. Tape media costs and management
8. Meeting service level agreements (SLAs)

Research Report: 2010 Data Protection Trends
*** NOTE – Drivers are the same in 2013 as they were in 2010.*

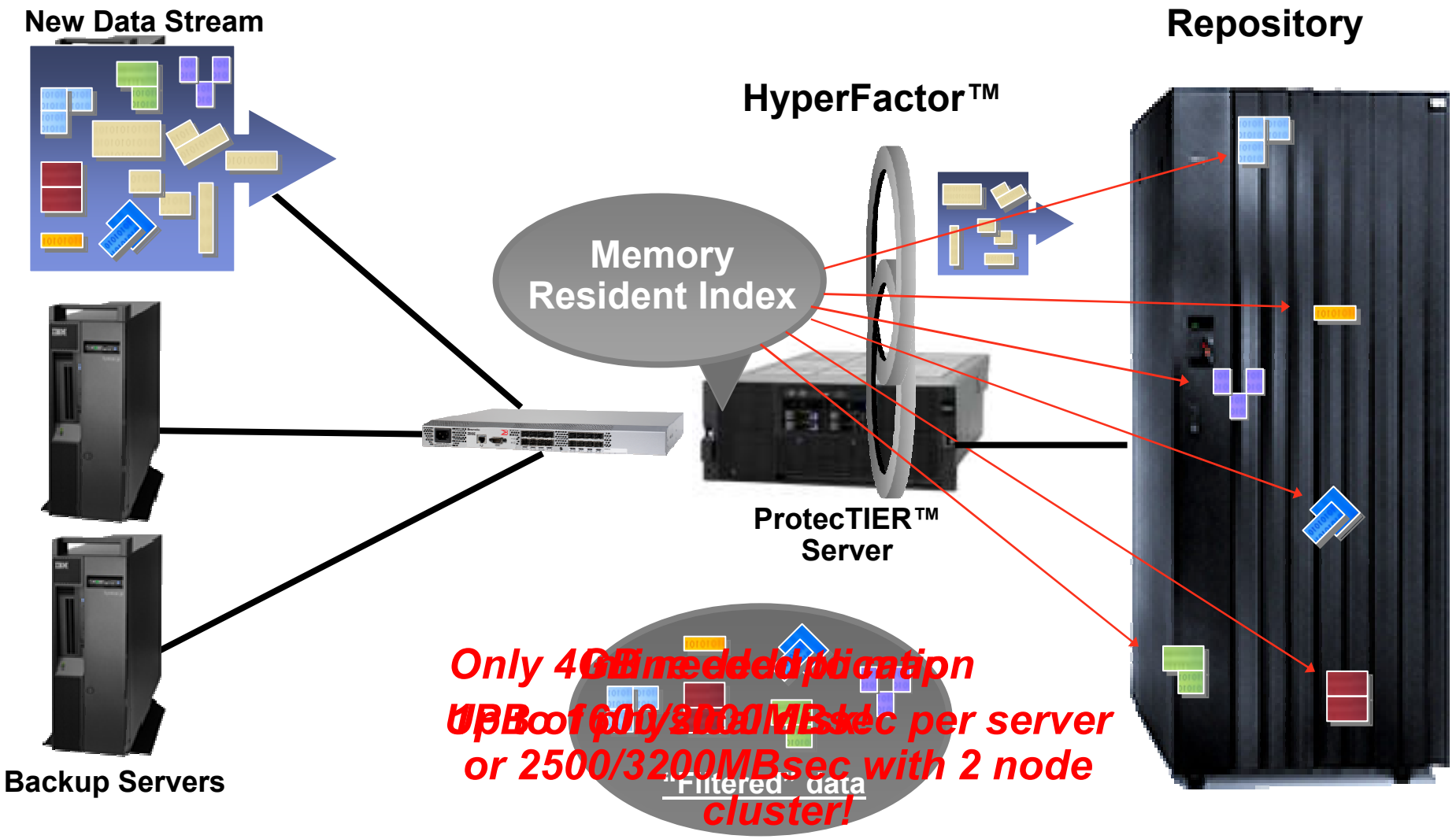
ProtecTIER Overview

What is ProtecTIER?

ProtecTIER is the first **virtual tape** product to contain patent-pending data factoring technology that IBM calls *HyperFactor*™ for **data deduplication**. ProtecTIER also supports Symantec Netbackup **OST** as well as **CIFS/NFS connectivity**…

1. In-line deduplication vs. Post Processing
2. UP TO 25:1 deduplication factoring
3. High Performance (2500/3200 MBs; backup/restore)
4. Highly Scalable (UP TO 1PB physical storage per system)
5. 100% Data Integrity
6. Clustering (Two (2) Node Cluster for Higher Availability)
7. Global Dedupe
8. Native replication (M – 1; M – M: 1 – M)
9. VTL, OST , CIFS & ***NFS*** Capable

How Does ProtecTIER Work?



Protect More. Store Less.®

Sample Customer Results

- **Mileage will vary based on the type of data, the data change rate, the data retention period required and the Back-up Application, i.e. TSM factoring avg's 6-7:1**

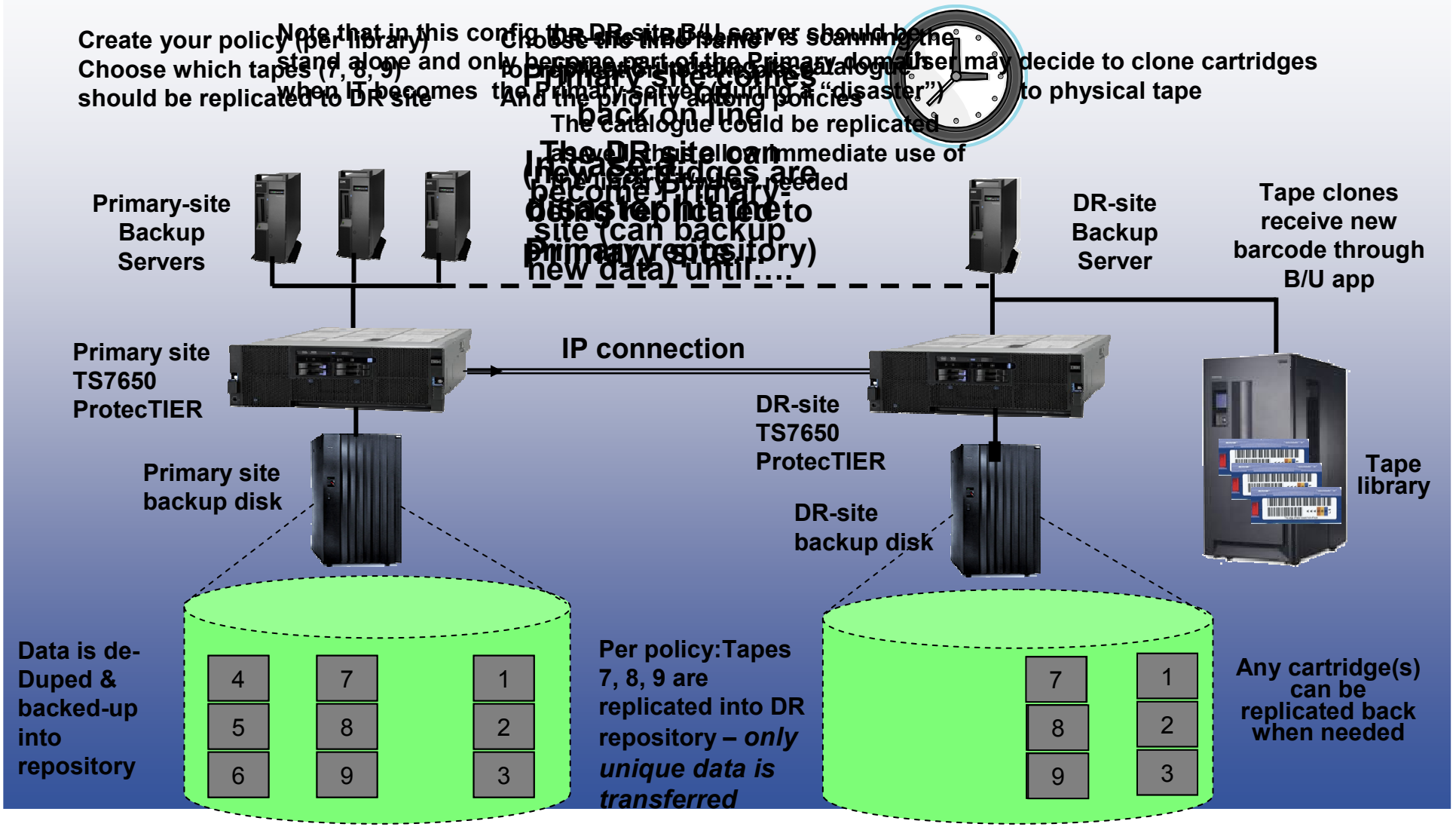
Customer Scenarios	Nominal Capacity	Physical Capacity	Factoring Ratio (Average)
Leading US Cancer Research Hospital	250TBs	22TBs	~12:1
Largest US Wireless Carrier	300TBs	30TBs	~10-30:1
Large International M&E Company	350TBs	36TBs	~15:1
F100 Worldwide Oil & Gas Corporation	900TBs	60TBs	~40-52:1

Protect More. Store Less.®

Why is Deduplication such a HOT Technology?

Native Replication

Native Replication Functionality (example)



Protect More. Store Less.®

ProtecTIER Native Replication Solutions

Powerful Deduplication and Many-to-Many Replication technologies combine to reduce required WAN bandwidth by up to 95% or more!

For customers . . .

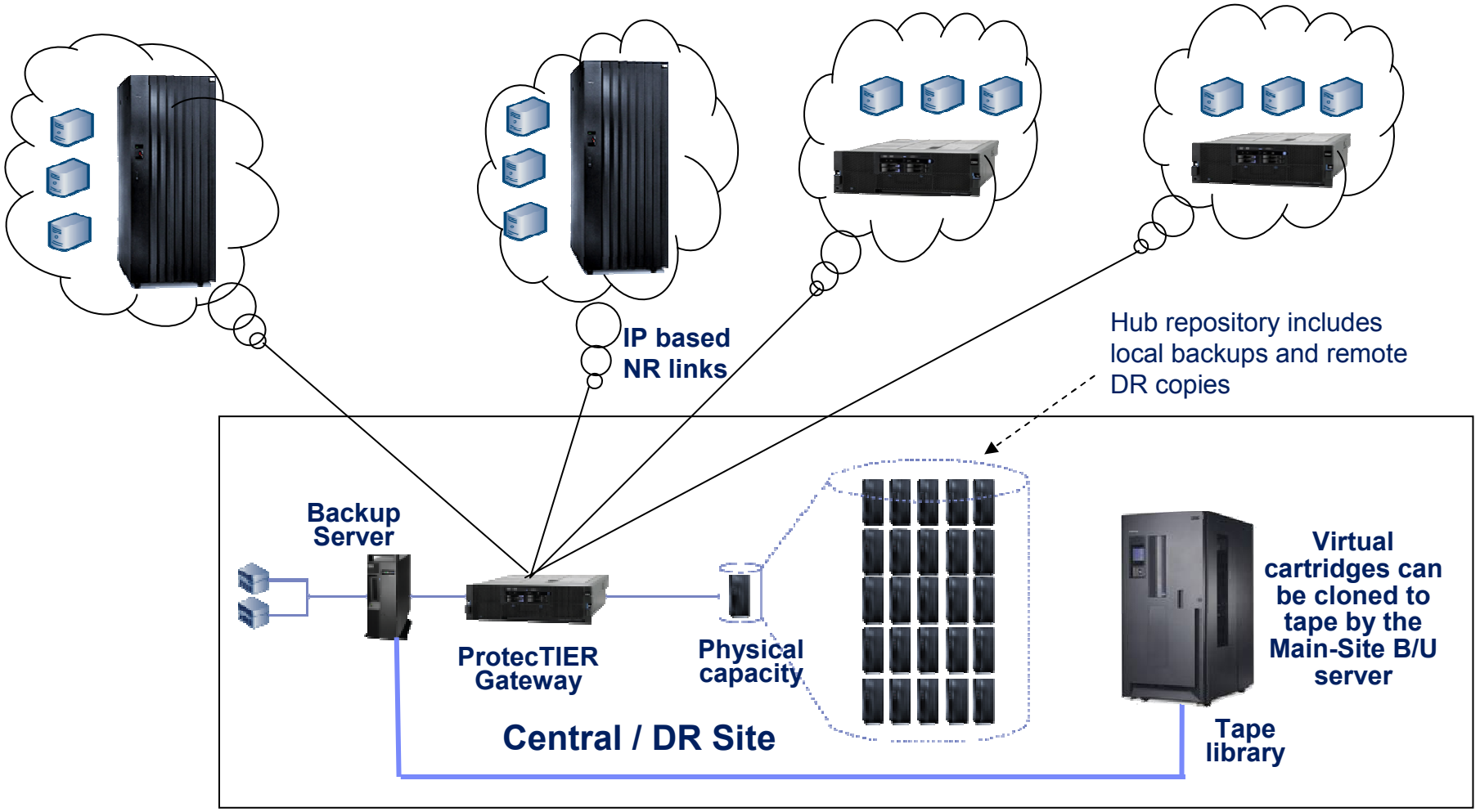
- with backup data at multiple data centers, remote offices or multiple office locations struggling with protecting mission critical data stored at those sites

ProtecTIER Deduplication with Many-to-one Replication provides . . .

- Disaster Recovery ability for organizations with multiple Data Centers or Remote Offices
- Up to 12 source repositories (spokes) replicating to a single target (hub)
- Spokes and hub can be either an Appliance or Gateway, single or dual-node cluster configuration
- Hub acts as a Disaster Recovery site for all spokes, can become a “primary” for any number of the spokes under disaster while still allowing local backups and replication from other active spokes

Many to One Replication

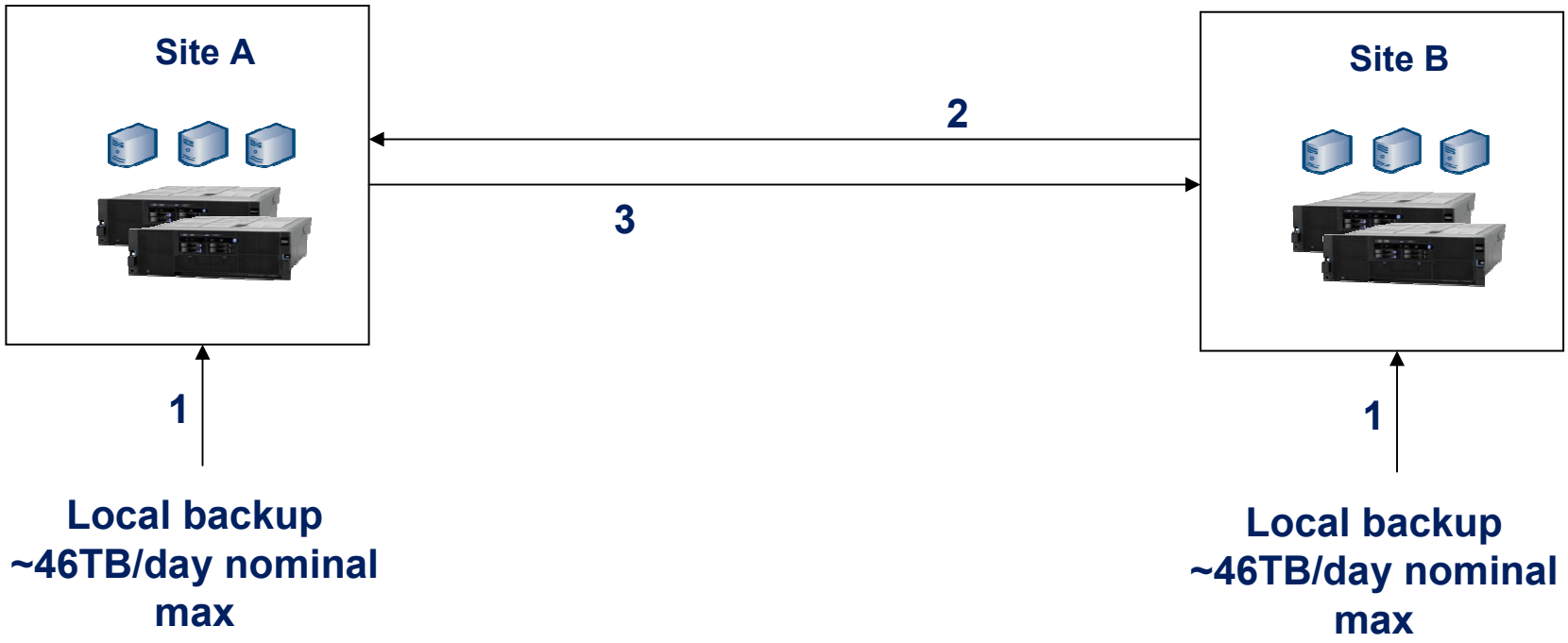
Up to 12 Branch Offices (spokes): Gateways and/or Appliances
1 target (hub): Appliance, Gateway, single or two-node cluster



Protect More. Store Less.®

ProtecTIER Bi-Directional Replication Pair

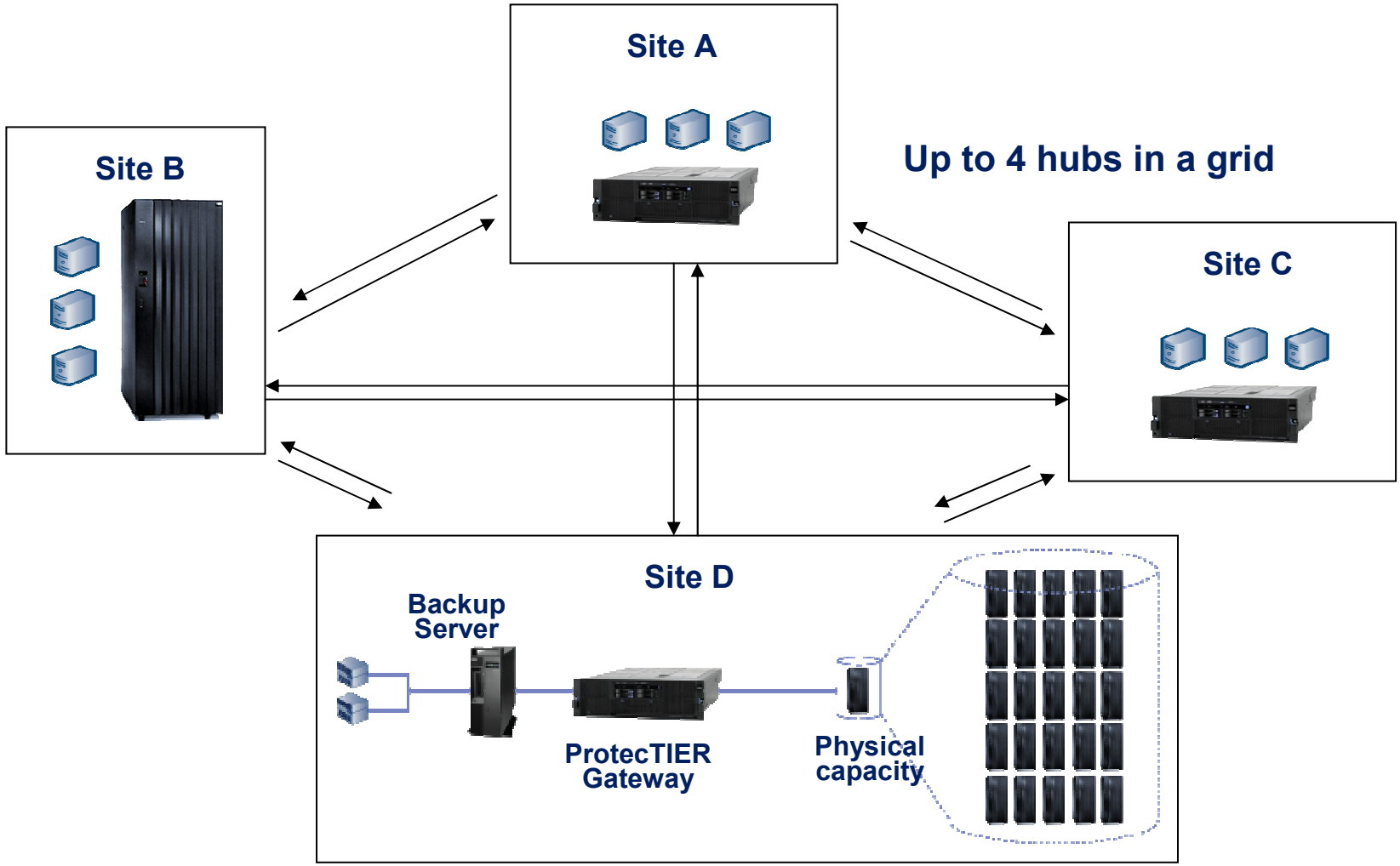
Example: 2 dual-node cluster hubs in a grid



1-3 numbers represent separate & equal-length time frames

Protect More. Store Less.®

Many to Many Replication Grid



Supports any combination of Gateways, Appliances, single or two-node clusters

Protect More. Store Less.®

Replication Benefits

1. **Dramatic improvements in Disaster Recovery operations**
   1. Automates electronic transfer of backup data to a remote site
   2. Leverages deduplication to only send unique data across the network
   3. Eliminates the risk associated with physically transporting tapes

2. **Radical cost reduction to DR operations through deduplication-enabled replication**
   1. Bandwidth, one of the most expensive costs for replication, is greatly reduced
   2. ProtecTIER requires less infrastructure at both the primary and secondary sites
   3. Saves money by eliminating the cost of transporting and storing physical tapes

3. **“Democratization" of replication – Replication for the masses**
   1. Replication no longer reserved for Tier one applications only
   2. Deduplication enables ALL applications to be replicated cost effectively
   3. Reduces risk of data loss and speeds recovery for most other applications

4. **It's all about Recovery Time Objectives (RTO)**
   1. Replication gets data to the remote site faster and safer
   2. Enables rapid restores of critical backup at backup site
   3. Applications can get back online quicker with fast disk-based recovery
   4. Deduplication enables more data to be protected with low RTO solution

ProtecTIER Product Options

IBM TS7600 ProtecTIER® Family

Support for VTL, OST, FSI CIFS & NFS

TS7620 ProtecTIER Deduplication Appliance Express

Good Performance
Entry Capacity
Low Investment
Scalable From 6 – 35TB




Single Node
Up to 300 MB/sec
6TB, 12TB, 23TB, 35TB Useable

TS7650G ProtecTIER Deduplication Gateway

High Performance
High Capacity
Flexible Storage

Single Node
Up to 1600/2000 MB/sec
1 PB useable

Highest Performance
Largest Capacity
High Availability


Active-Active Cluster
Up to 2500/3200 MB/sec
1 PB useable

Capacity and Performance

TS7650G Gateway

TS7650G ProtecTIER Deduplication Gateway

- Powerful and Flexible solution, featuring:
  - IBM ProtecTIER software with patented HyperFactor™ deduplication technology
  - IBM System x Server – multi-core server for enterprise-level performance
- Supports both IBM & Non-IBM disk
  - IBM V7000 Storwize, DS8000, XIV
  - HDS, EMC and others
- And delivers:
  - **Up to 1600/2000 MB/sec backup/restore**
  - Up to 25 times or more data reduction
  - Scalable to 1PB physical capacity
  - Enterprise-class data integrity

Protect More. Store Less.®

TS7650G ProtecTIER Data Deduplication Gateway

- Fast Backups - up to 1600MB/sec or more
- Faster Restores - up to 2000MB/sec or more
- Scalable to 1PB of useable capacity
- Up to 25x or more Data Reduction
- Enterprise Class Data Integrity
- Supports IBM & Non-IBM disk

Protect More. Store Less.®

TS7650G Gateway Dual-Node (2) Cluster

TS7650G ProtecTIER Deduplication Cluster

- Our most Powerful and Flexible solution:
  - IBM ProtecTIER software with patented HyperFactor™ deduplication technology
  - 2 IBM System x Servers – multi-core servers for maximum performance & availability
- Supports both IBM & Non-IBM disk
  - IBM V7000 Storwize, DS8000, XIV
  - HDS, EMC and others
- And delivers:
  - **Up to 2500/*3200 MB/sec backup/restore***
  - True active-active cluster technology
  - Two nodes working together as one repository
  - Easily manageable yet highly scalable

Protect More. Store Less.®

TS7650G ProtecTIER Data Deduplication 2 Node Cluster

- Fast Backups - up to 2500MB/sec or more
- Faster Restores - up to 3200MB/sec or more
- True Cluster with Global Deduplication
- Highly available
- Highly scalable
- Easily manageable

Protect More. Store Less.®

TS7650G Gateway Cluster Features

- Active-Active 2 nodes cluster (architecture will allow for increasing node count over time)
- Full repository sharing among nodes
  - Writing data to the repository
  - Reading data from the repository (restore and read reference)
  - Access to all virtual devices
- No degradation on HyperFactor efficiency (regardless of the node through which the data is received)
- Minimum cluster down-time

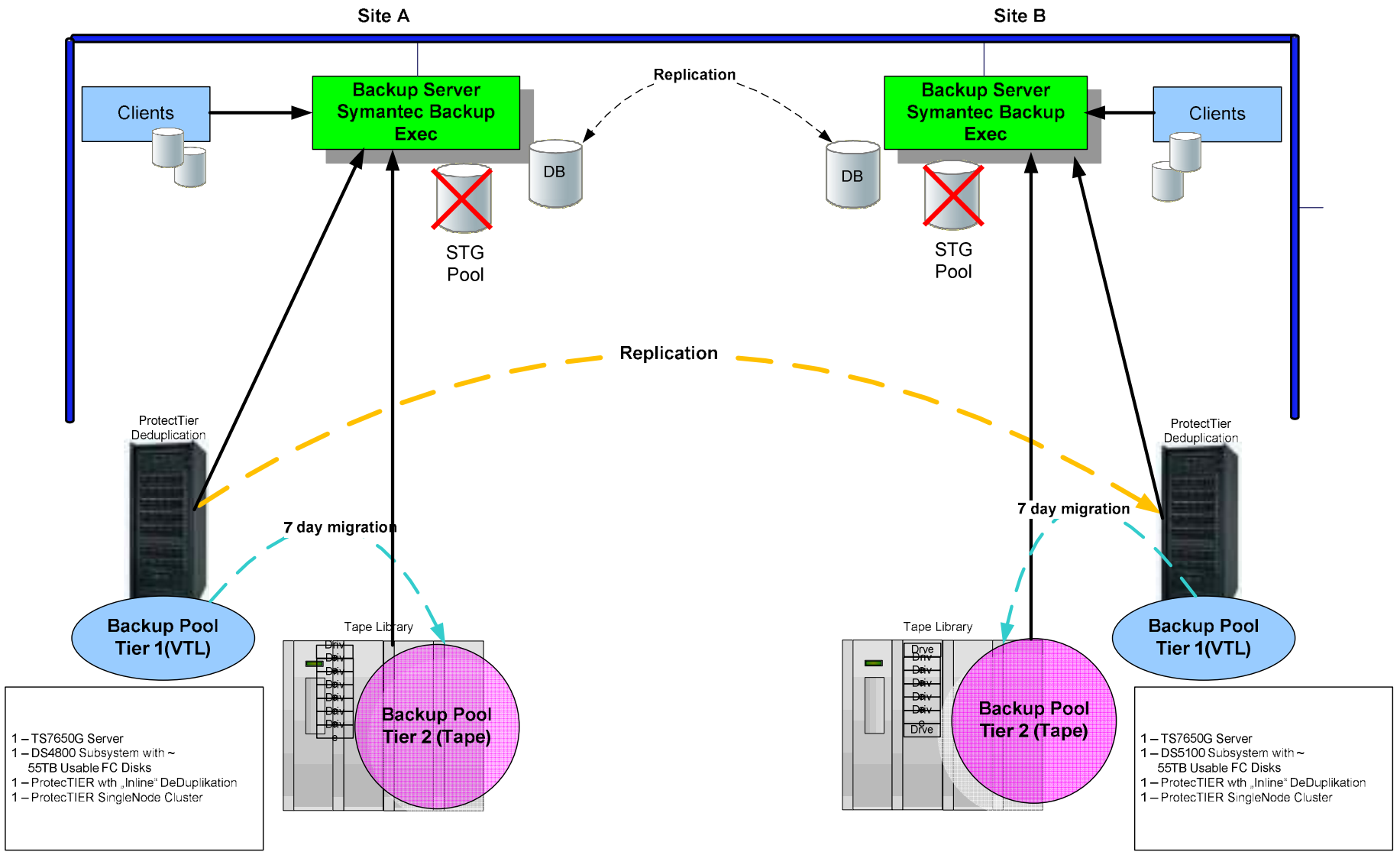
IBM's Fastest and Most Scalable Deduplication Solution!

Customer Case Studies

ProtecTIER: WW Success in 2012

- **Highlights**
  - Over 1,500 ProtecTIER customers WW
    - Approximate 30% increase in new installations
  - Over 3,000 ProtecTIER sytsems shipped.
  - Continued double digit growth for IBM deduplication
  - ProtecTIER deduping 300 PB of physical disk
  - Significant competitive wins in 2012
- **Key Wins**
  - *American Express*
  - *Virgin Media*
  - *BMW*
  - *Charles Schwab*
  - *Principal Insurance Group*
  - *Norfolk Southern*
  - *Bell Aliant*

Protect More. Store Less.®



Protect More. Store Less.®

Key elements for customer decision

Performance	Up to 2000 MBps per node, performing inline de-duplication. Solution based on FC disks.
Capacity	55TB per Site and future dynamic scalability.
100% Data integrity guarantee	Only IBM solution can provide it.
No impact to existing daily operations using deduplication.	Inline de-duplication eliminates need for significant secondary processing.
Non-disruptive implementation	Transparent installation in existing environment and integration with backup application.

Summary

Why Do Co's Invest In ProtecTIER?

Performance

2,500 MB/s ingest
3,200 MB/s Restore

Delivering superior competitive throughput for the Enterprise

Capacity

Up to 1 PB physical capacity per node

Nearly 2x other vendors capability

Open Gateway Design

Disk selection freedom

Provides best of breed disk now and with future upgrades

Non-Disruption

Daily Operations
Inline de-duplication eliminates need for significant secondary processing

Implementation
Integrates well with existing backup environment and infrastructure

Continuous Availability

Dual node design for high availability

The only HA solution available

Global Deduplication

Deduplicates ALL data

Provides superior deduplication ratios

Protect More. Store Less.®

Why ProtecTIER WINS

- Two (2) Node H/A Clustering
- Performance
  - UP TO 3200MB/sec Restore
  - UP TO 2500MB/sec Ingest
- Scalability
  - TS7620 6, 12, 23 upgradeable to 35 TB
  - TS7650G Gateway UP TO one (1) PB of physical capacity
- Flexibility - support for IBM AND 3rd party disk
  - EMC, HDS, Others

IBM ProtecTIER Solution provides….

- Store up to 25 times or more data on disk
  - 250TB reduced to only 10TB with enterprise class data integrity
- Reduce backup and restore times
  - High speed restore at UP TO 3200MB/sec or more
- Improve the reliability of backup operations
  - Eliminates mechanical & handling failures
- Drive the cost of disk based backup down
  - Reduces energy, cooling, and space required
- Increase data retention
  - Store more backup data on disk for a longer time with very little additional cost

Protect More. Store Less.®

Protect More. Store Less.®

Trademarks and Disclaimers

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Trademarks of International Business Machines Corporation in the United States, other countries, or both can be found on the World Wide Web at http://www.ibm.com/legal/copytrade.shtml.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Other company, product, or service names may be trademarks or service marks of others.

Information is provided "AS IS" without warranty of any kind.

The customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in IBM product announcements. The information is presented here to communicate IBM's current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown may be engineering prototypes. Changes may be incorporated in production models.