

# The Future of Deduplication Storage for Mainframes

Vendor Session #14337



**Scott James**  
*VP Global Alliances*  
Luminex Software, Inc.

**Randy Fleenor**  
*IBM® ProtecTIER® Group*  
IBM Corporation



# Discussion Topics

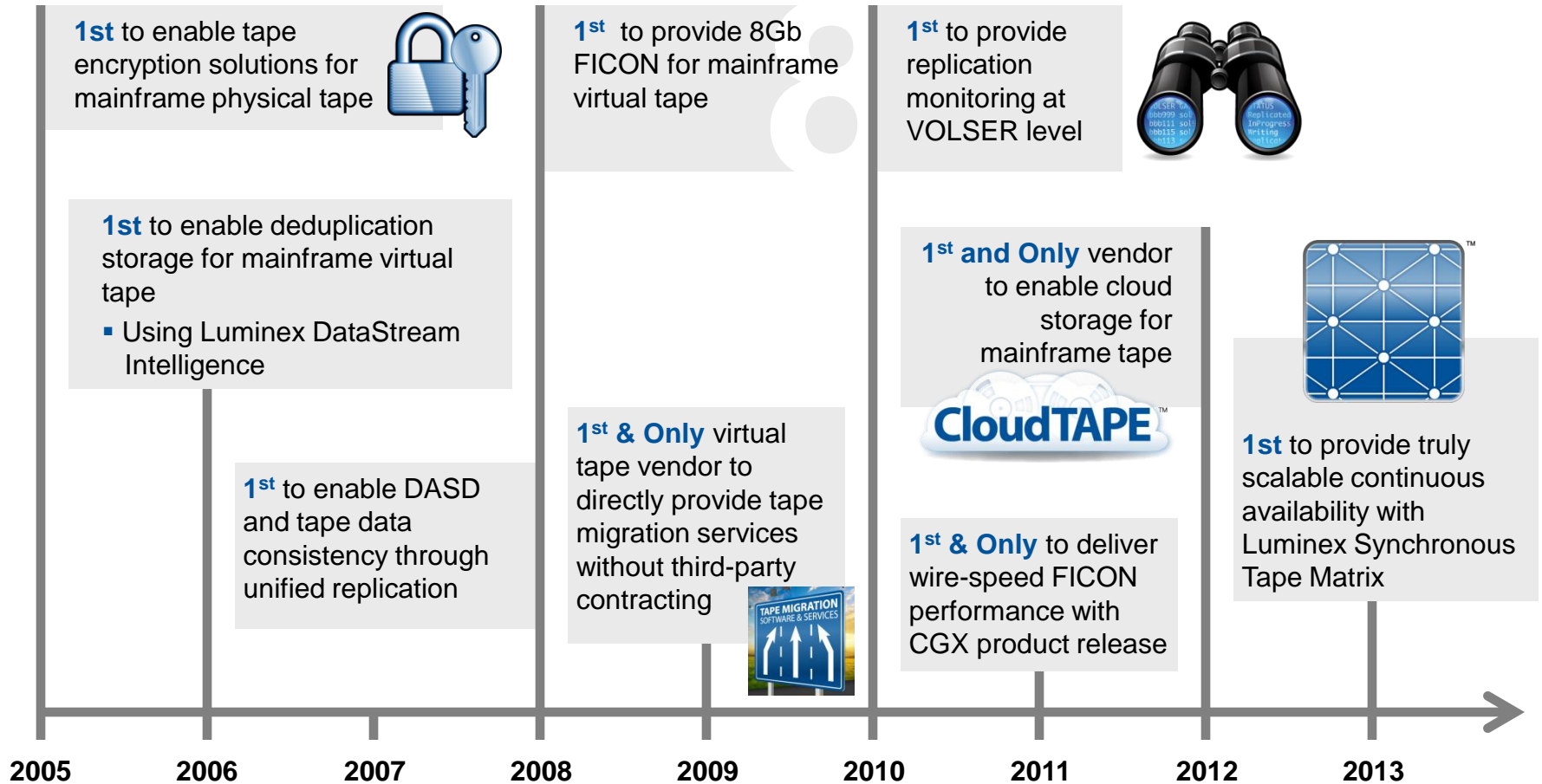
- What Happened to Mainframe Deduplication?
- Introducing IBM ProtecTIER Mainframe Edition (ME)
- Sizing Storage & Network Requirements for Deduplication Storage
- What Can Dedupe Do For You?
- Trends and Directions
- IBM® Systems Storage® ProtecTIER
  - TS7650G ProtecTIER Deduplication Gateway
  - TS7620 ProtecTIER Deduplication Appliance

# What Happened to Mainframe Deduplication?

- Primary benefit of deduplication storage for mainframe environments is to minimize WAN costs for DR
  - How many dedupe suppliers realize this?
- Luminex led the industry beginning in 2006 and maintains majority market share of production installations
  - Luminex DataStream Intelligence optimizes deduplication storage
  - Luminex CGX is used with a wide variety of deduplication storage systems and vendors
- EMC positions deduplication and designed products to work only for backups and not hsm or other tape applications
- Oracle – VSM6 / VLE provides zfs based deduplication, but does this provide any benefit for replication?



# Luminex's Heritage of Innovation



# Luminex At-A-Glance

- Luminex is an enterprise-class solution vendor focused on mainframe virtual tape
- 25+ years delivering mainframe virtual tape solutions worldwide
- Installed in all major industries on 6 continents
- Solutions scale seamlessly from entry level to most demanding multi-PB datacenters
- Luminex continues to lead the industry's Mainframe Tape-less Revolution
  - No physical tape is required



# IBM ProtecTIER Mainframe Edition (ME)

POWERED BY  
**LUMINEX**



- Luminex and IBM have raised the bar... by enabling IBM System Storage ProtecTIER Deduplication Storage Products for Use in Mainframe Data Centers

# IBM ProtecTIER Mainframe Edition

Enterprise class Deduplication powered by Luminex

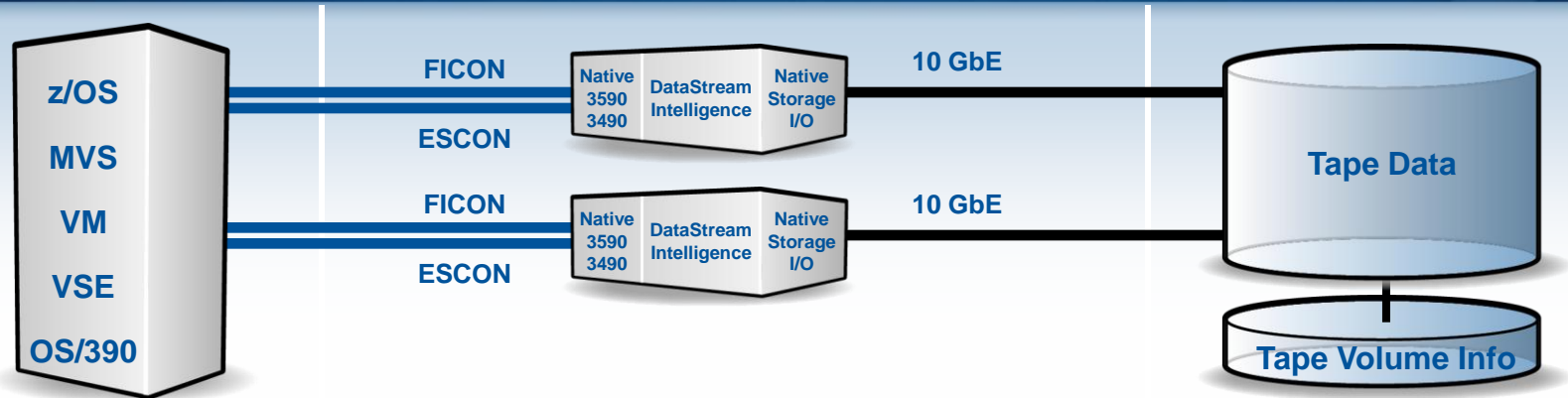


Specifications per cluster	ProtecTIER ME
Number of Virtual Devices	Up to 8192
Usable Physical Cache	Up to 1 PB
Effective Cache Capacity (15:1)	up to 15 PB
Number of Virtual Volumes	unlimited
FICON channels	2 or 4
Logical paths per FICON channel	256
8 GB FICON support	Yes



# How the Solution Works...

## IBM ProtecTIER ME



### Mainframe

- Application transparent – non-intrusive
- No MIPS required
- z/OS, MVS, VM, VSE and OS/390 supported
- Works well with all major tape management systems
- SMS via MTL or Esoterics can be used
- Optional host-based software for management

### Luminex Channel Gateways

- Emulates 3490 or 3590 mainframe tape drives
- “Wire Speed” up to 8 Gb FICON
- DataStream Intelligence™ optimizes compression and deduplication
- No dependencies on separate code to modify a backup VOLSER
- Active – Active with NSPOF
- Encryption and Key Management Option
- Modular design makes scaling throughput and/or capacity easy and cost effective

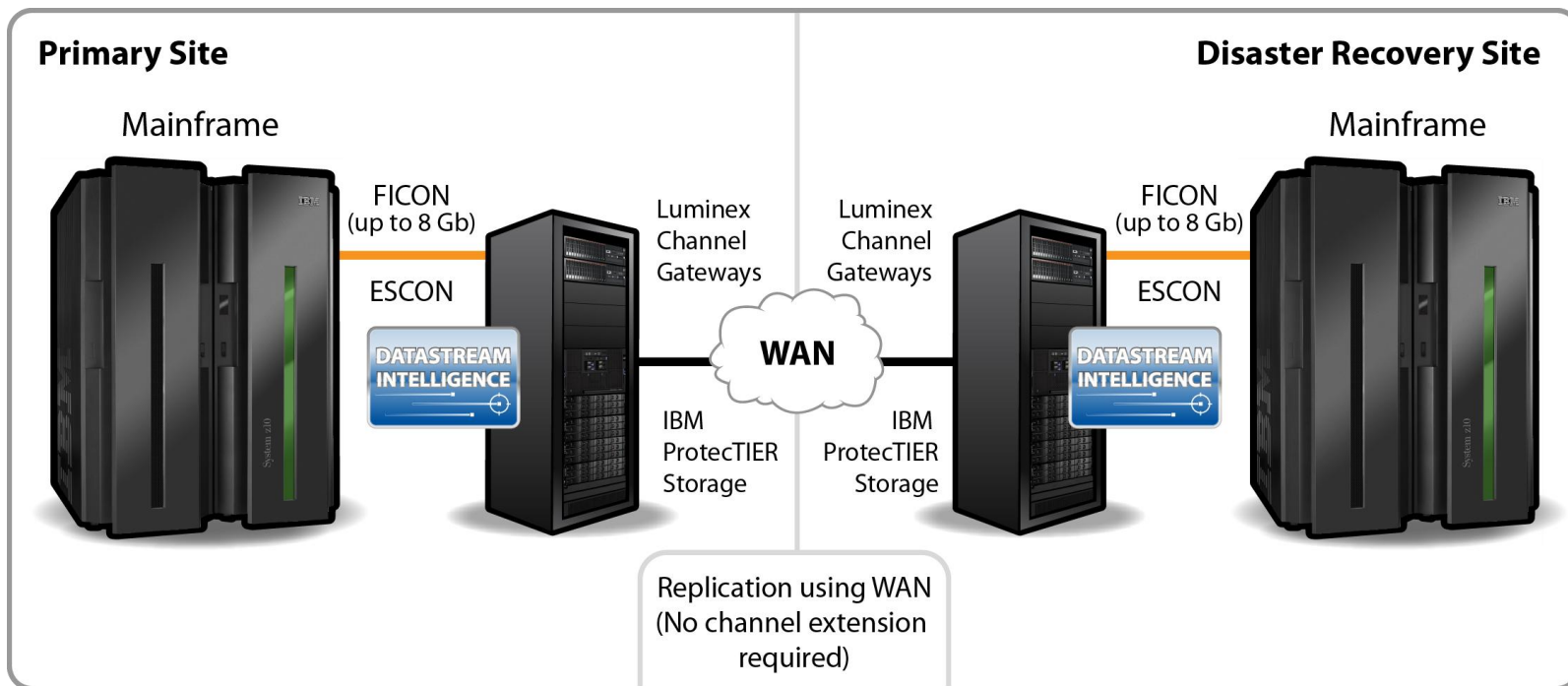
### IBM System Storage

- Network Attached Storage (Using NFS over 10 GbE)
- Mainframe tape volumes stored as standard files
- Replication for backup/DR
- RAID Data Protection
- Enterprise and modular storage systems are supported



# ProtectTIER ME

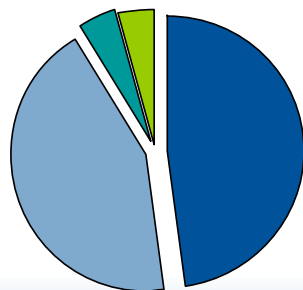
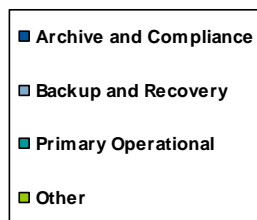
## Typical Production & DR Example



# Sizing - Storage Capacity Requirements

## TMC Analysis

Program	Description	Category	Capacity (GB)	Comp. Rate	Comp. Capacity (GB)	Dedupe Rate	Dedupe Capacity (GB)
ARCCTL	IBM DFSMSHsm Migration	Archive and Compliance	27,018.30	3.00	9,006.10	8.00	3,377.29
IDCAMS	IBM VSAM Utility (Define, Repro, Export, etc.)	Backup and Recovery	22,177.91	3.00	7,392.64	15.00	1,478.53
ADRDSSU	IBM DFSMSdss Full Volume / File Disk Backup Utility	Backup and Recovery	3,559.18	3.00	1,186.39	15.00	237.28
IFASMFDP	IBM SMF Data Offload Utility	Archive and Compliance	3,413.77	3.00	1,137.92	8.00	426.72
SAS	SAS Institute, Inc. Statistical Analysis Program	Primary Operational	1,829.80	3.00	609.93	8.00	228.73
DSNUTILB	DB2 Utilities (Load/Unload, etc.)	Backup and Recovery	1,805.20	3.00	601.73	15.00	120.35
IEBGENER	IBM Sequential File Copy Utility	Primary Operational	919.44	3.00	306.48	8.00	114.93
BAKPROD8		Other	431.96	3.00	143.99	8.00	54.00
DRCPYDAT		Other	302.47	3.00	100.82	8.00	37.81
SORT	IBM SORT Utility	Primary Operational	291.78	3.00	97.26	8.00	36.47
Other		Other	1,943.64	3.00	647.88	8.00	242.96
<b>Summary</b>			<b>63,693.45</b>	<b>3.00</b>	<b>21,231.15</b>	<b>10.02</b>	<b>6,355.05</b>



Uncompressed Capacity



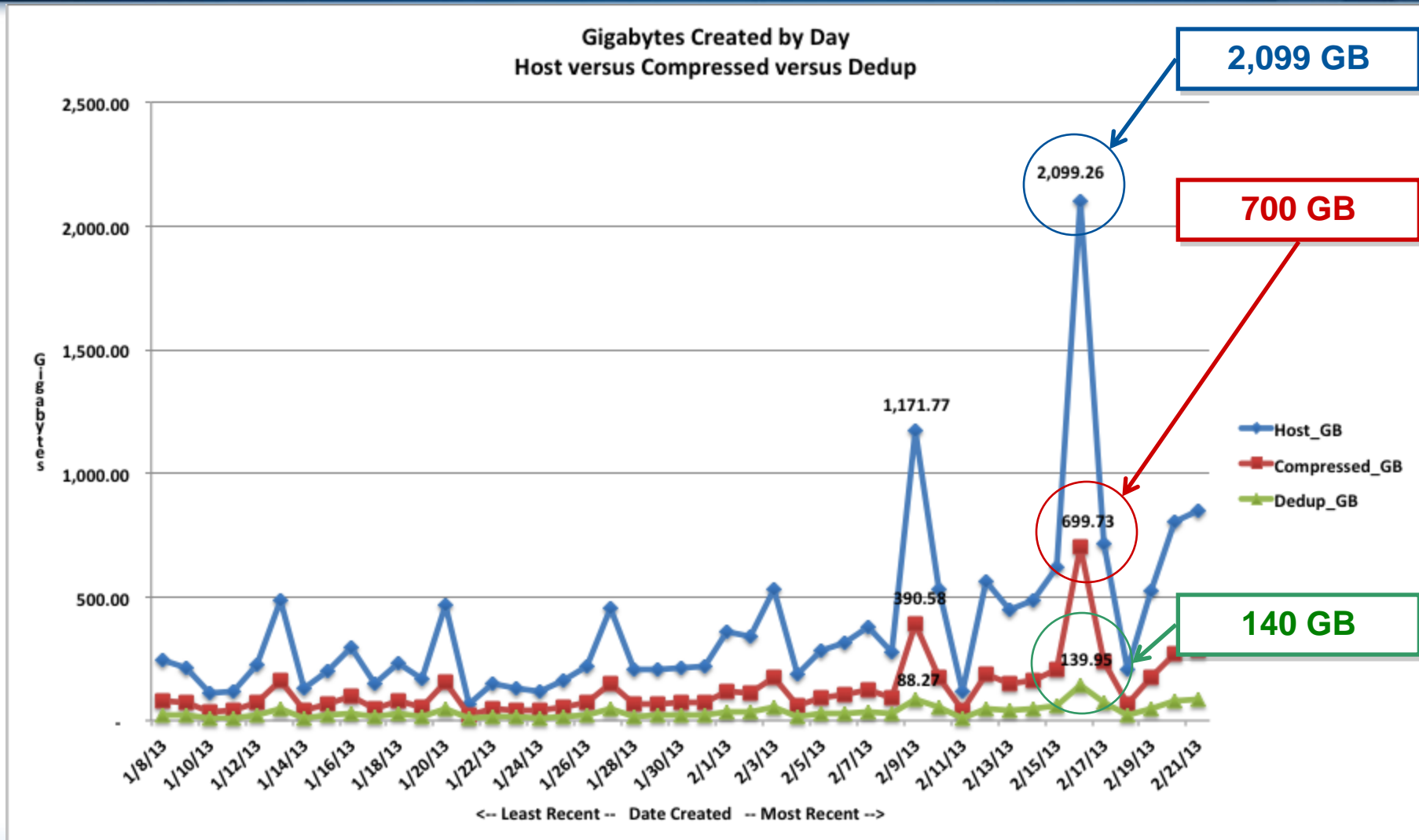
Compressed Capacity



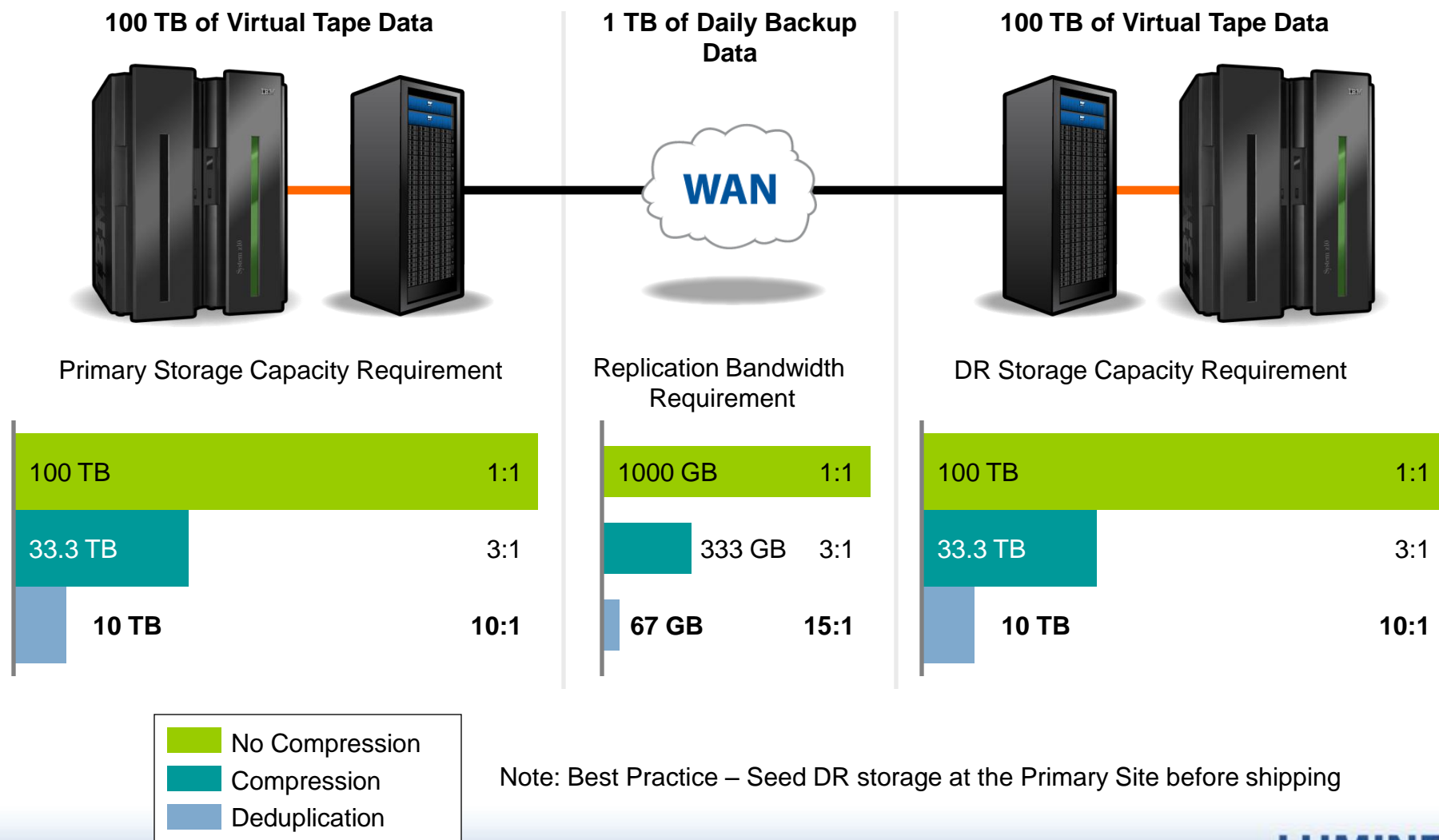
Deduplicated Capacity

# Sizing - Network Bandwidth Requirements

## TMC & SMF Analysis



# What Can Dedupe Do For You? (Example)



# How Much Bandwidth is Required?

Example: Replicate 1 TB daily

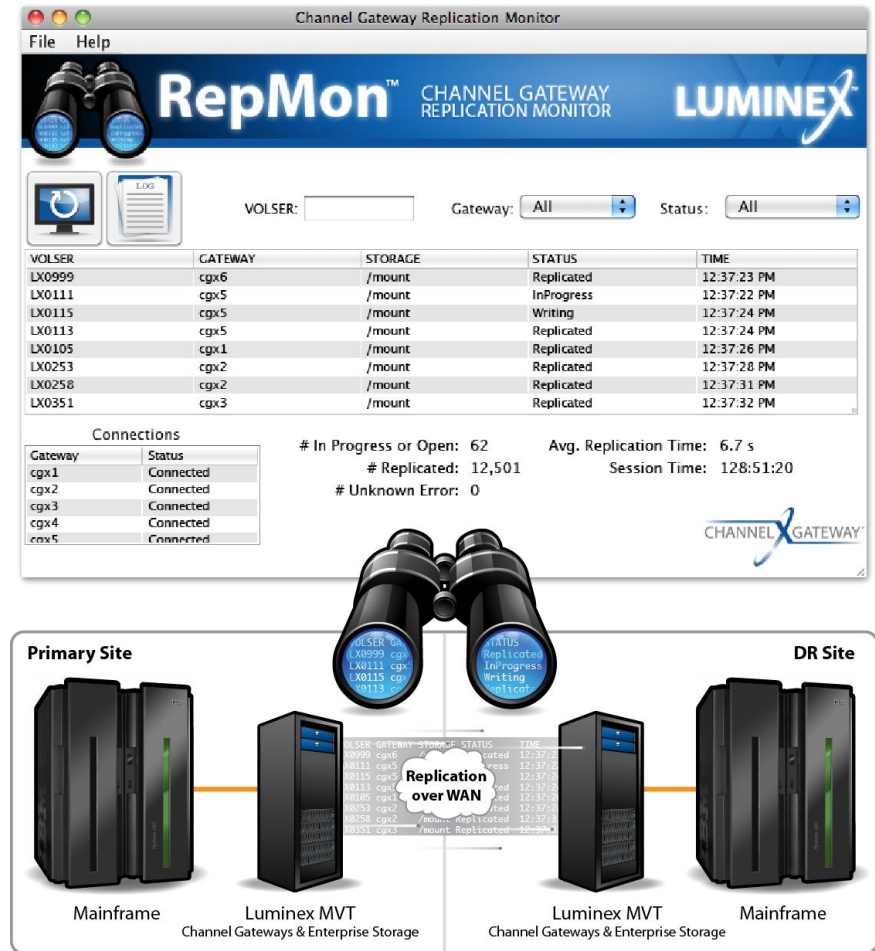
WAN Class	MB/s	GB/hr	Replication Time w/ 3:1 compression	Replication Time w/ 15:1 deduplication
DS3	5.59	19.66	16.95 hours	3.39 hours
OC-1	6.48	22.78	14.63 hours	2.93 hours
OC-3	19.44	68.34	4.88 hours	0.98 hours

If replication begins at midnight, a DS3 line will complete replication long before a courier would show up to begin transporting a physical tape.

# RepMon: Replication Monitor

Provides **real-time status monitoring and logging** of virtual tape data writes and replication to a remote disaster recovery site at the VOLSER level

- Identifies Write and Replication Status of Mainframe Tape VOLSERs
- Identifies if virtual tape data at DR is still consistent with the primary datacenter
- Provides visual and audit capabilities to confirm when backups reach DR



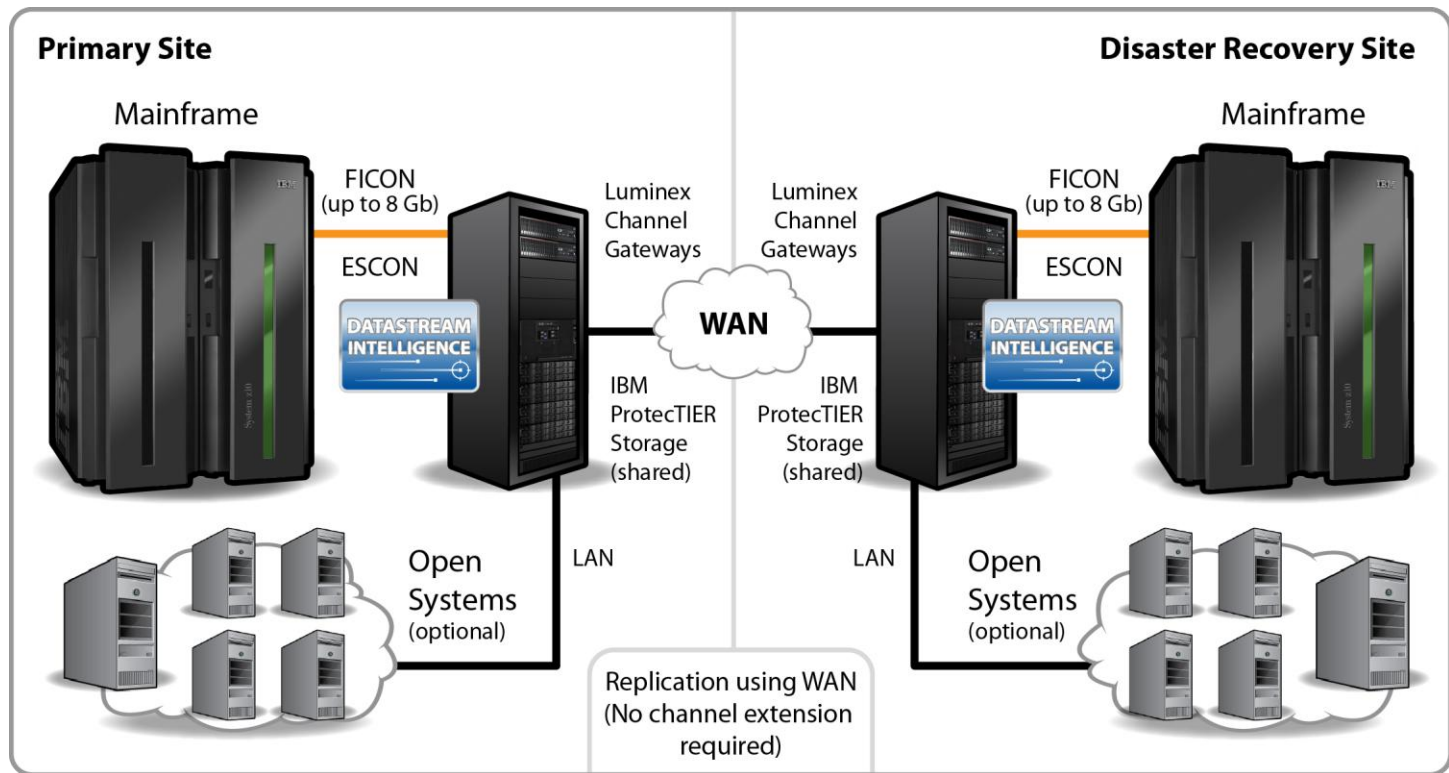


# Trends and Directions for Mainframe Data Deduplication

- Shared Storage Infrastructure for Mainframe and Distributed (Open) Systems
- True Global Data Deduplication
  - Large Central Dedupe Repository
  - Dedupe all data, not just backup
- More Capacity & Scalability

# ProtecTIER ME for Today's Trends and Directions

- Shared Infrastructure
  - Common open and mainframe backup and tape processing solution



# ProtecTIER ME for Today's Trends and Directions

## Global Deduplication

Deduplicates ALL data

**Provides superior deduplication ratios**

## Capacity

Up to 1 PB physical capacity per node

**Approximately 2x Data Domain's  
largest system**

## Open Gateway Design

Disk selection freedom

**Provides best of breed disk now and with  
future upgrades**

## Non-Hashing Integrity

Full Binary Differential check

**Completely eliminates the potential of  
data corruption due to hash collisions**

# ProtecTIER ME with IBM System Storage

## ProtecTIER TS7650G

- In-Line Data Deduplication
- Replication Options:
  - One-to-One, One-to-Many and Many-to-Many
- TS7650G can provide up to 1PB of physical storage capacity per system
  - Up to 25PB of backup capacity with deduplication

# ProtecTIER ME with IBM System Storage ProtecTIER TS7620 Appliance

- In-Line Data Deduplication
- Automated replication capabilities
- TS7620 Appliance can provide 35TB of physical storage capacity per system
  - Up to 875TB of backup capacity with deduplication

# What Can ProtecTIER ME Do For You?

- Improve all aspects of tape operations
  - Batch processing, HSM, Archiving, Backup & Recovery
- Reduce or eliminate cost of maintenance, media, shipping and off site storage for physical tape
- Improve tape data security
- Reduce data center floor space requirements
- Improve disaster recovery preparedness
  - Improved RPO, RTO and simplified their DR test process
- Reduce tape storage capacity and network bandwidth requirements
- Data deduplication can make it practical to replicate all virtual tape data to a remote DR site over a wide area network.



# Presales, Installation and Support

- Our Systems Engineers provide pre-sales consulting, tape studies for sizing throughput and capacity
- Post-sales support provides on site or remote installation and 7x24 support worldwide

**Thank You**

**Our Session # is 14337  
Enjoy the Conference!**

