#SHAREorg

SHARE
Technology · Connections · Results

IBM Real-time Compression Extraordinary Storage Efficiency For Both Block & File

Steve Kenniston – BLE Data Protection / Copy Data

The Storage Alchemist @skenniston

The Storage Alchemist

IBM

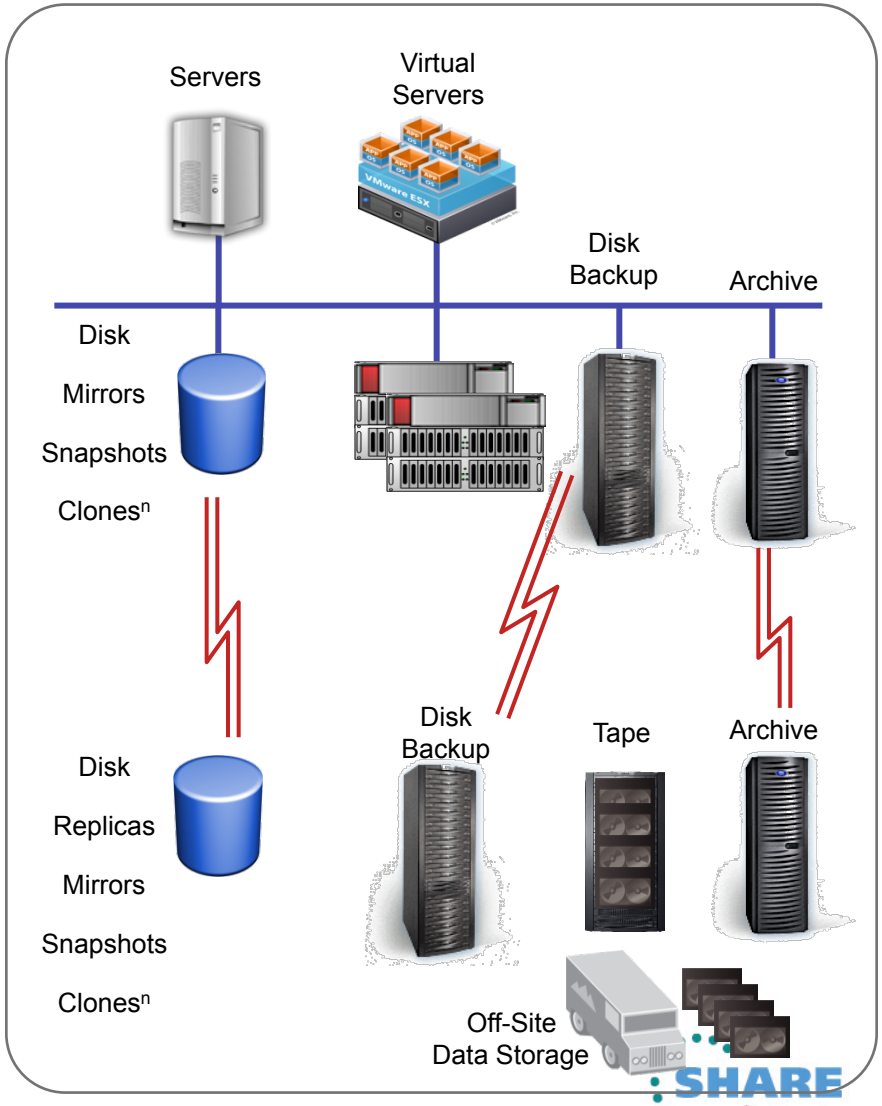
August 13, 2013

Session: 14099

SHARE in Boston

The Primary Storage Growth Challenge

- Data continues to grow 35% to 65%* for most companies
  - Growth is a factor of:
    - Mirrors
    - Snapshots
    - Clones
    - Replicas
- All of this data is being backed up
  - Of the backup data, a good deal of it is being replicated for DR purposes
- Archives of primary data are also created for compliance purposes
  - Archive data is also replicated for DR
- In the remote facilities backup tapes are cut from the disk copies for long term preservation
  - This data is hauled off on tapes and the tape 'farm' continues to grow



* Source: Wikibon

IBM Real-time Compression – Block

Hundreds of RtC clients worldwide are reaping the benefits of:

- Innovation – 35+ Patents
- Real-time Compression – up to 80%
- No impact to application/storage performance
- Adaptive to any workload
- Industries only heterogeneous compression solution

IBM Real-time Compression Helps Clients Grow into the Future

Data Growth / Usage

- **Cloud** – As companies look to leverage the private and public cloud, data movement and capacity savings are critical. Optimizing data through transparent, Real-time Compression can ensure you get the full benefit of all your cloud services
- **Analytics** – Analyze up to 80% more data in the same foot print to learn more about your business and become more competitive
- **Big Data** – Big Data environments will grow at an astounding rate. Help your 'Big Data' environment be more efficient by leveraging Real-time Compression

Real-time Compression is the storage industry's only heterogeneous compression solution available!

Evolution of Storage Efficiency

Archive	WAN Connectivity	Backup	Primary & Secondary
Tape Compression	WAN Optimization	Backup Deduplication	Real-time Compression
IBM, LTO, Quantum DLT, Ultrium LTO	riverbed, CNT, Actona, Cisco, Juniper Networks, BlueCoat	datadomain, Avamar, NetApp, Diligent, symantec., Sepaton, Exagrid	IBM, Real-time Compression

Market Impact

Technology Evolution

10 Years Ago | 6 Years Ago | 3 Years Ago | Today

Compression Without Compromise

Enhancing Storwize V7000 & SVC to Deliver Extraordinary Efficiency

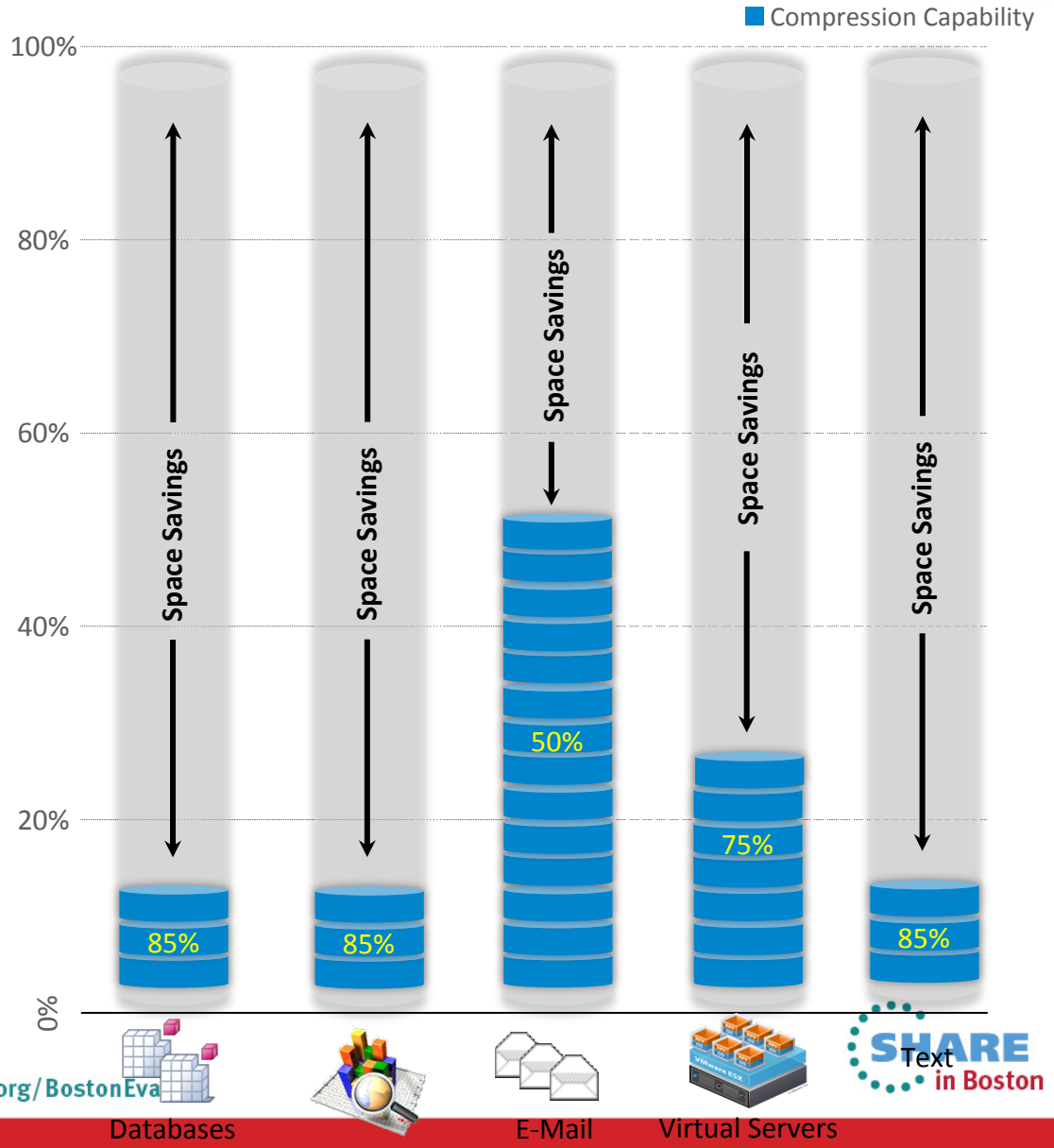
- Introducing IBM Real-time Compression for V7000 & SVC
- Innovative, easy-to-use compression fully integrated into Storwize V7000 & SVC
- High performance implementation supports active primary workloads

- Real-time Compression builds on storage efficiency advantages of Storwize V7000 & SVC
- Real-time Compression delivers up to 80% compression
- Compression helps reduce
  - Storage purchase costs
  - Rack space
  - Cooling
  - Software costs for additional functions

- Compression can help freeze storage growth or delay need for additional storage capacity

Real-time Compression – Maximum Compression

- **Real-time Compression**
  - Provides the highest level of compression for your active workloads
  - Allows you to reclaim up to 80% of your existing disk space
  - Use ***Compresstimator*** to find out how much capacity your client will save
  - Reduces your overall $/GB
  - Immediate results!



Fixed vs. Variable Input Compression

- IBM's Real-time Compression changes the game when it comes to defining how compression is performed

- Traditional Compression
  - Traditional compression starts with a fixed file that is divided up into segments
  - Each segment is then compressed and put into the compressed file
  - Due to this, compression ratios are not optimized
  - This is know as "Fixed Input" to "Variable Output" Compression
  - This also causes fragmentation inside the compressed file
  - The compressed file can actually grow over time

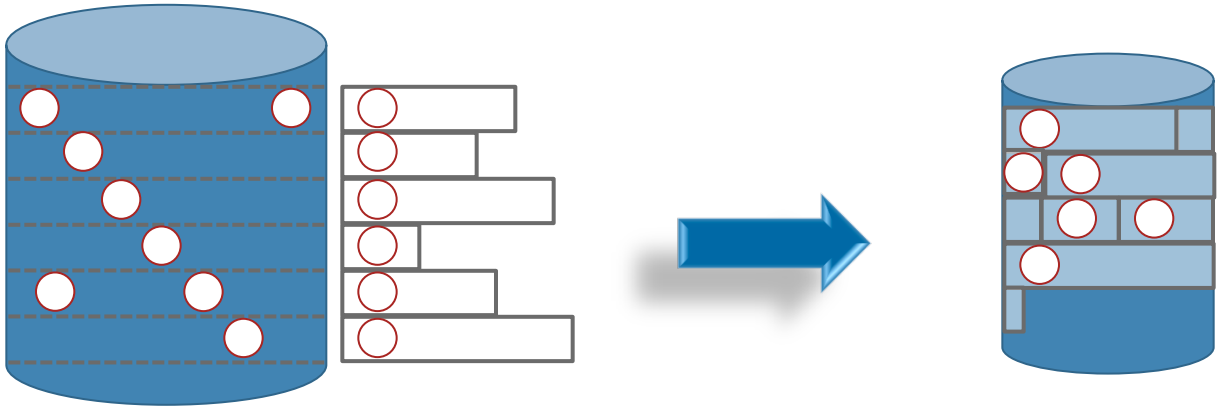
- Real-time Compression
  - Starts with "Variable Input" because we are in the data stream we see all the data
  - All "like" data is compressed together as it is written together
  - Compressed data is put into known fixed segments that represent 1 I/O
    - This is known as "Fixed Output"
  - This maximizes compression ratios
  - It is like having a journaled file system inside the compressed file
  - A map maintains where the compressed segments are
  - "Wholes" can be filled w/ new data as the map knows where the "wholes" are

Compression Technology

Fixed Input

Variable Output

Traditional



Original Data

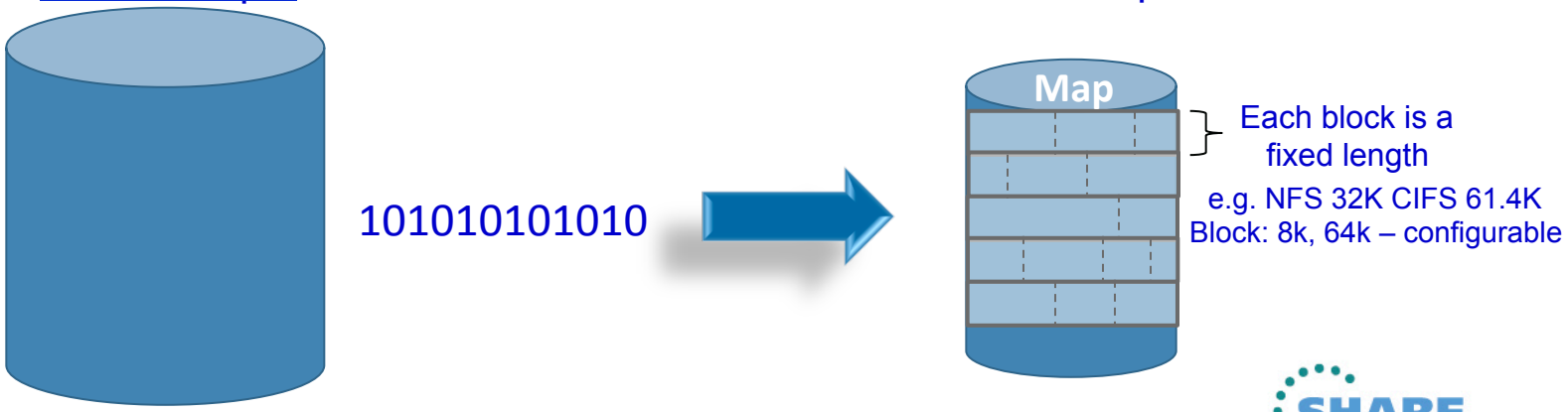
Compressed Data

Variable Input

Fixed Output

RtC

101010101010

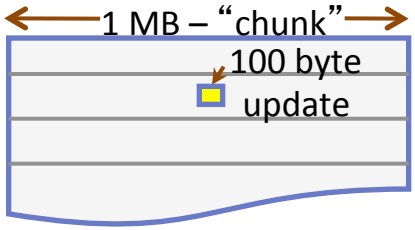


Original I/O

Compressed Data

Real-time Compression - Performance Differentiation

Traditional Approaches



Must decompress entire file to edit

Much more I/O

Much more CPU 'work'
- Decompress / recompress more data

Network impact
- Data moved
- Data backed up
- Data replicated

No way to make real-time

Traditional Compression	IBM Real-time Compression
1 MB Read	0 MB Read
1 MB Decompress	0 MB Decompress
100 Byte Update	0 Byte Update
1 MB Compress	100 Byte Compress
1 MB Write	< 100 Byte Write
Total I/O: 2 MB I/O	< 100 Byte I/O

Real-time Compression


RtC Appliance


IBM SVC


Storwize V7000/U

No need to decompress data that needs to be modified

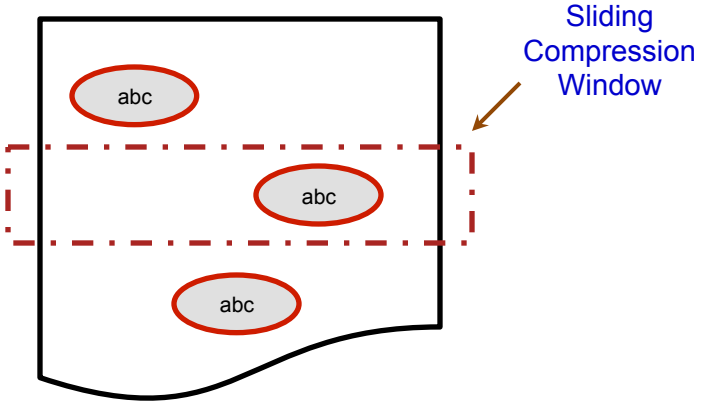
Increases storage cache (SVC)
- Compressed data in array cache
- Increase size of storage cache

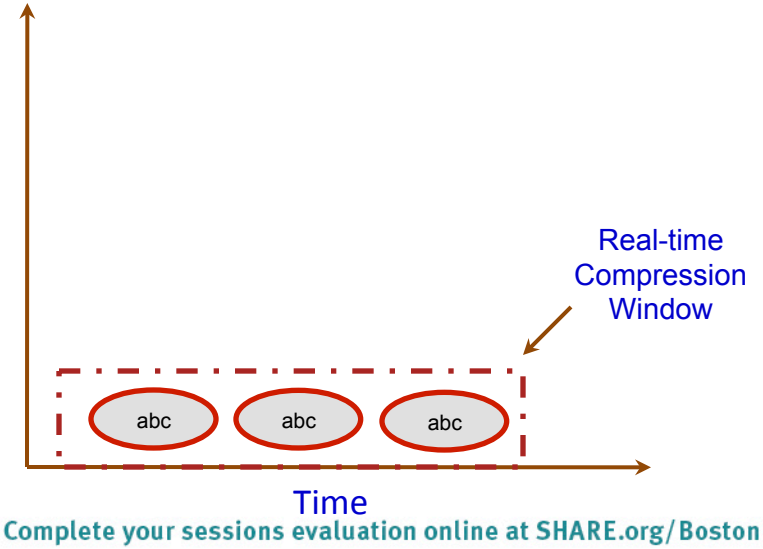
Variable data stream to fixed output
- Compressed volume has consistent, logical layout
- Variable size input yields better compression ratios
- Fewer I/O's means better performance

Real-time Compression
Less I/O Impact
Less CPU, Disk, Network Impact

Location based vs. Time based Compression

Location Based Compression





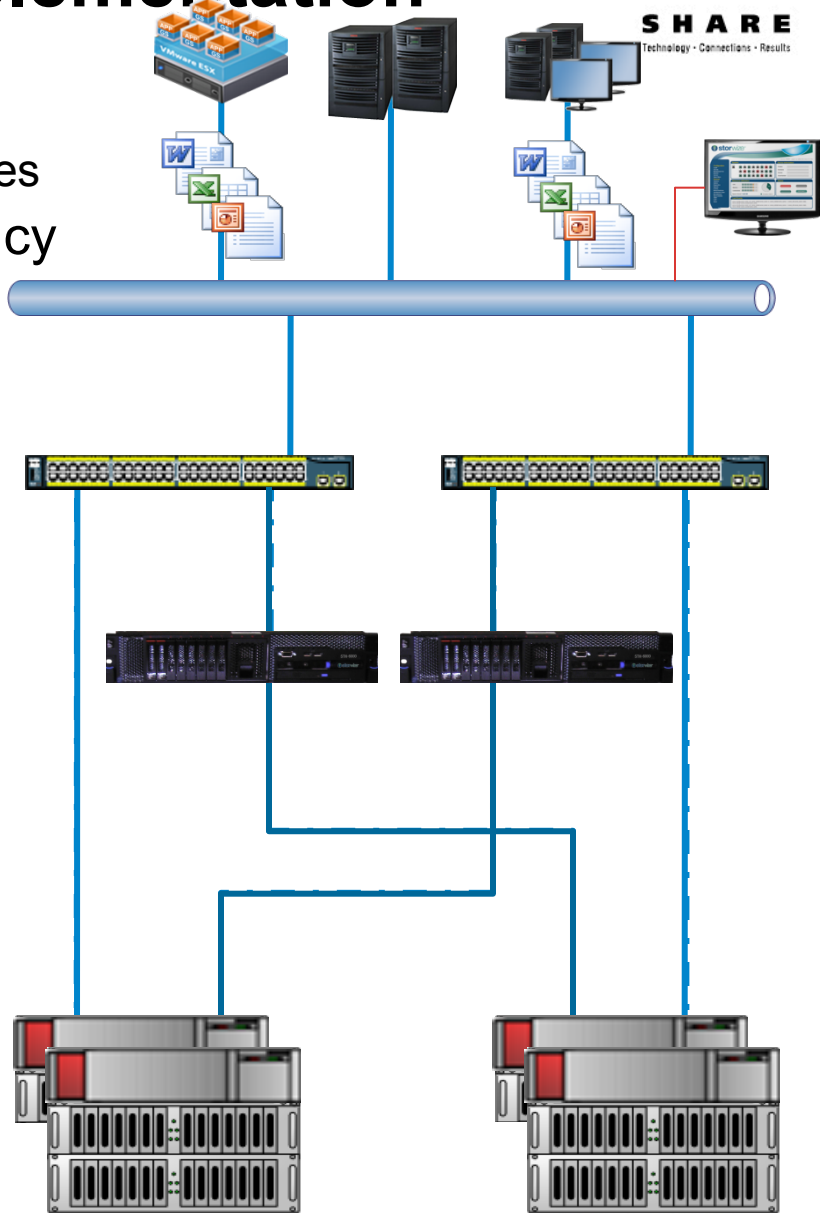
- Location based Optimization
  - Only looks at the end state of the file
  - Only deals with the locality of reference between data
  - Miss potential compression with 'sliding window' technique

- Time based Optimization
  - Relationship between data is time based
  - RtC understands how the data was created
  - Optimization is enhanced when you compress 'like' data
  - Only optimization technique that ***can*** work with deduplication

Real-time Compression Implementation

Simple Installation / Management

- Configure the Real-time Compression appliances
  - Network – Set up network transparency
  - NAS – Add connections to Filers
- Install Real-time Compression appliances in the data path
- Configure and test implementation
  - Compression Filters
    - Choose data to compress
- Begin compressing data
- *Run Compression Accelerator
  - Reduce already stored data



To see how easy it is to install watch:
IBM Real-time Compression Out of the Box Experience
On YouTube – IBM Real-time Compression Channel

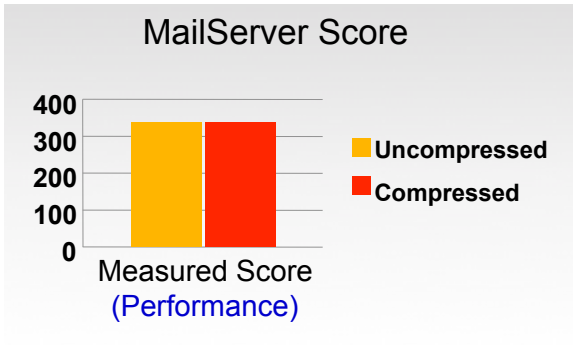
Compression in SVC/V7000 – In a Nutshell

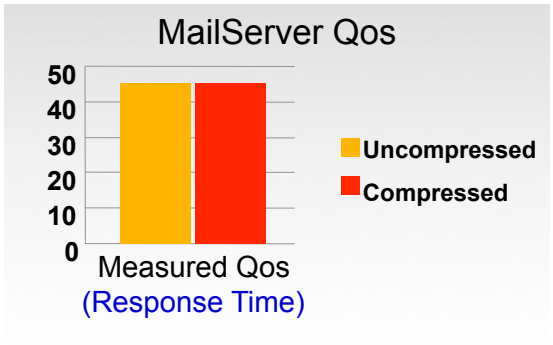
- Supported on V7000 and SVC (CF8 and CG8) clusters
- Uses the same code used in Real-time Compression Appliances
- Compression operates at a volume level
- Up to 200 compressed volumes supported per I/O group in this release
- A compressed volume is a third type of volume
  - Generic (fully allocated)
  - Thin Provisioned ('Space Efficient' term is deprecated)
  - Compressed
- A compressed volume is another kind of thinly provisioned volume
  - Only uses physical storage to store the compressed version of the volume
  - Compressed volumes are always thinly provisioned
- Specified at creation of the volume
  - No conversion in place
  - Volumes can be non-disruptively migrated via Volume Mirroring
- Compressed volumes can be source and target of remote or flash copy maps

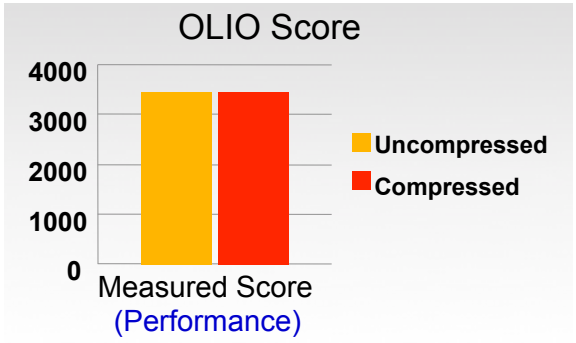
VMWare - Performance Results

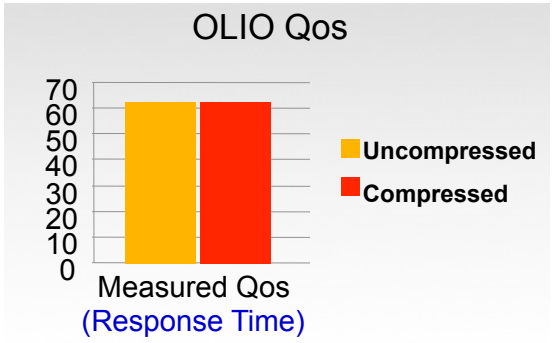
- Storwize V7000 with Real-time Compression delivers up to 4x compression while maintaining VMware and Application performance

VMware Vmark Performance Benchmarks



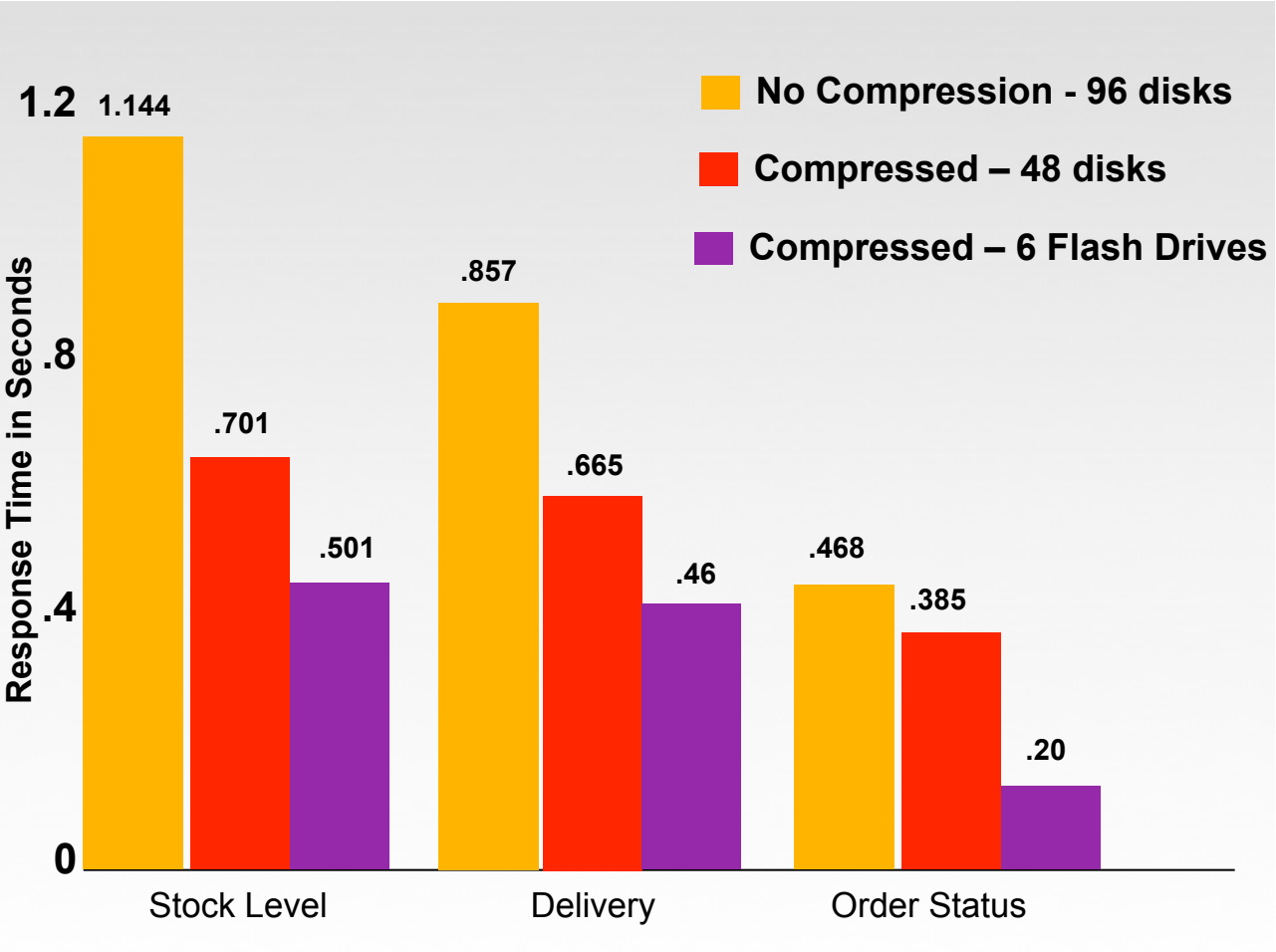


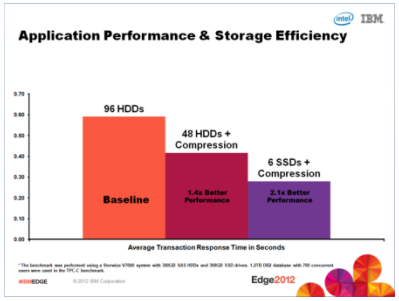




Compression without Compromise

- Storwize V7000 with Real-time Compression delivers up to 5x compression while maintaining or improving database transaction response time and overall business throughput





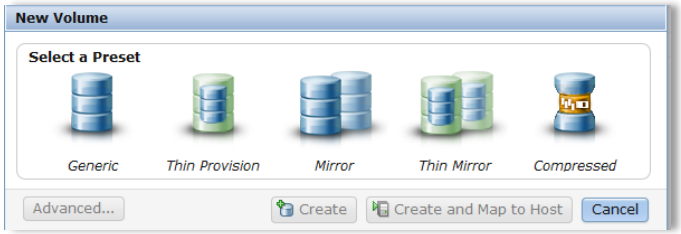
The benchmark was performed using a Storwize V7000 system with 48 x 300GB SAS disks. 1.2TB database with 700 concurrent clients were used in the benchmark. The same test was performed with compressed volumes and non-compressed volumes.

Real-time Compression – Transparent to Your Applications

- Support & Usability behaviour is the same as if using Thin Provisioning
  - If it works with thin provisioning it works for compressed volumes (e.g. auto-expand)
  - Applications are unaware they are talking to a compressed volume

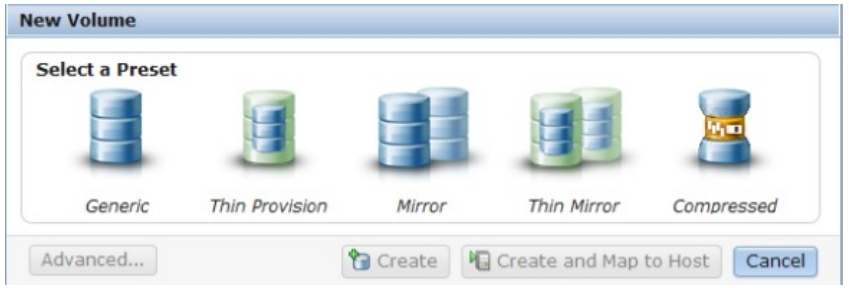
- All advanced SVC features are supported for compressed volumes including:
  - EasyTier
  - FlashCopy
  - Remote Copy
  - Mirroring
  - ISV integrated functions (e.g. VMware VAAI)

- GUI Support for creating Compressed Volumes
  - GUI has a preset for compressed volumes

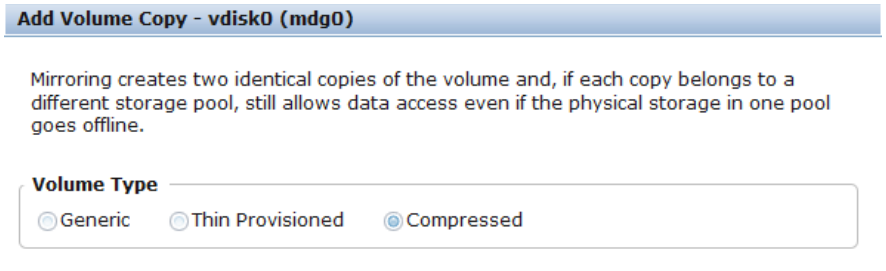


Compression Without Compromise
Easy to Use

- Real-time Compression integrates easily into Storwize V7000 GUI
  - Simply select new volume preset

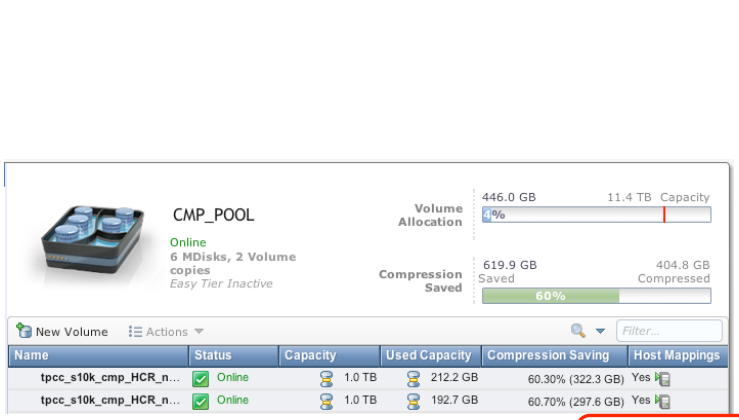


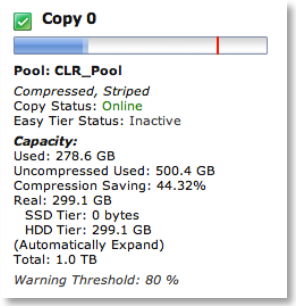
  - Up to 200 compressed volumes per Storwize V7000 control enclosure
- Straightforward conversion of existing volumes using volume mirroring
  - Convert to compressed thin provisioned and can eliminate unused space during conversion
- Nothing else to do!

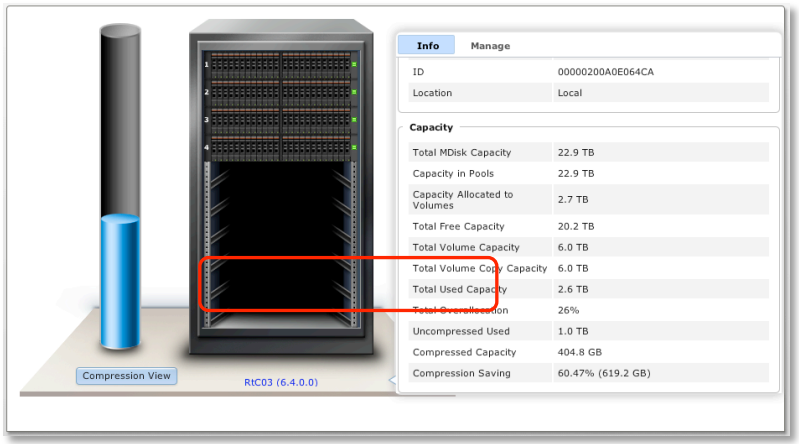


Compression Savings in GUI and CLI

- User Interface provides 3 different measurements:
  - **Virtual size** - the volume size observed by the host
  - **Compressed size** - amount of capacity actually provisioned for a volume. This includes savings from both Real-time Compression and Thin-Provisioning
  - **Uncompressed size** - the effective uncompressed size of capacity provisioned for a volume. This is the projected size of the volume without compression. It includes savings from thin-provisioning only







Comprestimator

Comprestimator is a host based utility for a fast estimation of a block device compression ratio

Objectives:

- Run over a block device
- Estimate:
  - Portion of non-zero blocks in the volume
  - Compression rate of non-zero blocks

Performance:

- Runs FAST! < 60 seconds, no matter what the volume size is
- Provides a guarantee on the estimation: ~5 % max error
  - Can improve guarantee with more samples (longer running time)

Method:

- Random sampling and compression throughout the volume
- Collect enough non-zero samples to gain desired confidence
  - More zero blocks → slower (takes more time to find non-zero blocks)
- Mathematical analysis gives confidence guarantees
- Note: the tool is estimating compression during migration of a volume into RtC (data at rest)

Availability

- Download from Fix Central
  - Comprestimator
- Supported clients:
  - RedHat Enterprise Linux Version 5 64-bit
  - VMWare ESXi 4.0, 5.0
  - AIX V6.1, V7.1
  - Windows 2003/2008 Server (32-bit/64-bit)
  - Solaris 10
  - HP-UX 11i v3

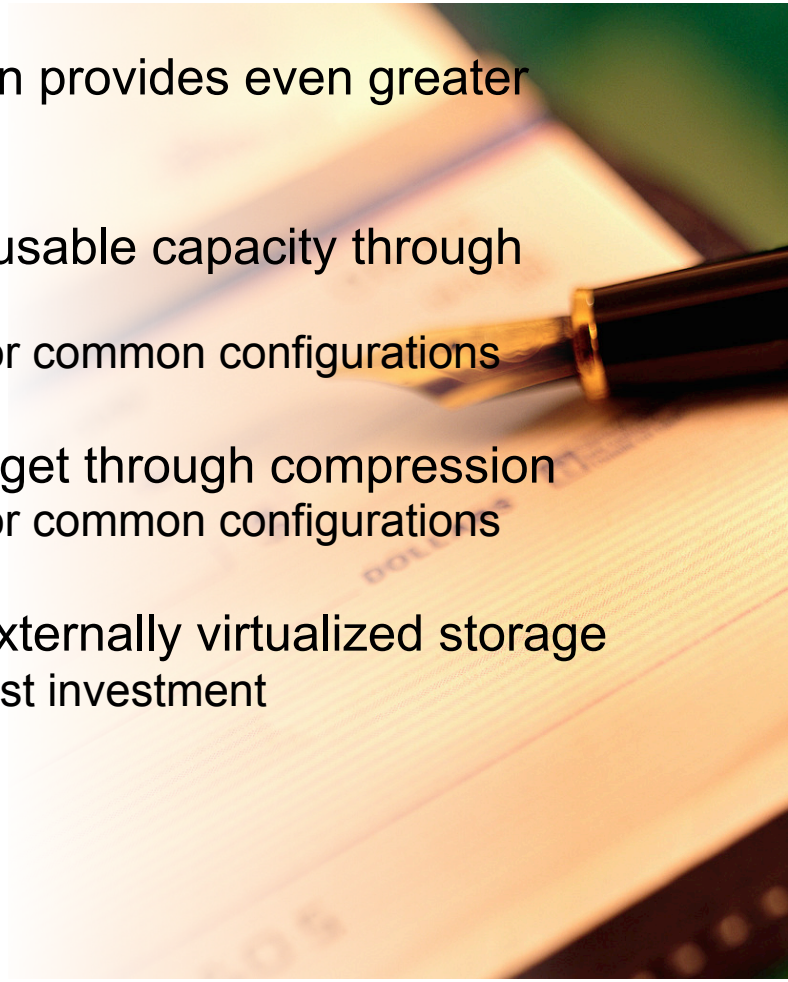
Compression Without Compromise

Advantages Compared with EMC and NetApp

- IBM Real-time Compression can be used with active primary data
  - High performance compression supports workloads off-limits to other alternatives
  - Significantly expands candidate data for compression
  - Greater compression benefits through use on more types of data
- IBM Real-time Compression operates immediately and is easy to manage
  - No need to schedule periods to run post-process compression
  - Eliminates need to reserve space for uncompressed data waiting post-processing
- IBM Real-time Compression supports all Storwize V7000 & SVC storage
  - Internal or externally virtualized storage
  - Can significantly enhance value of existing storage assets

Compression Without Compromise

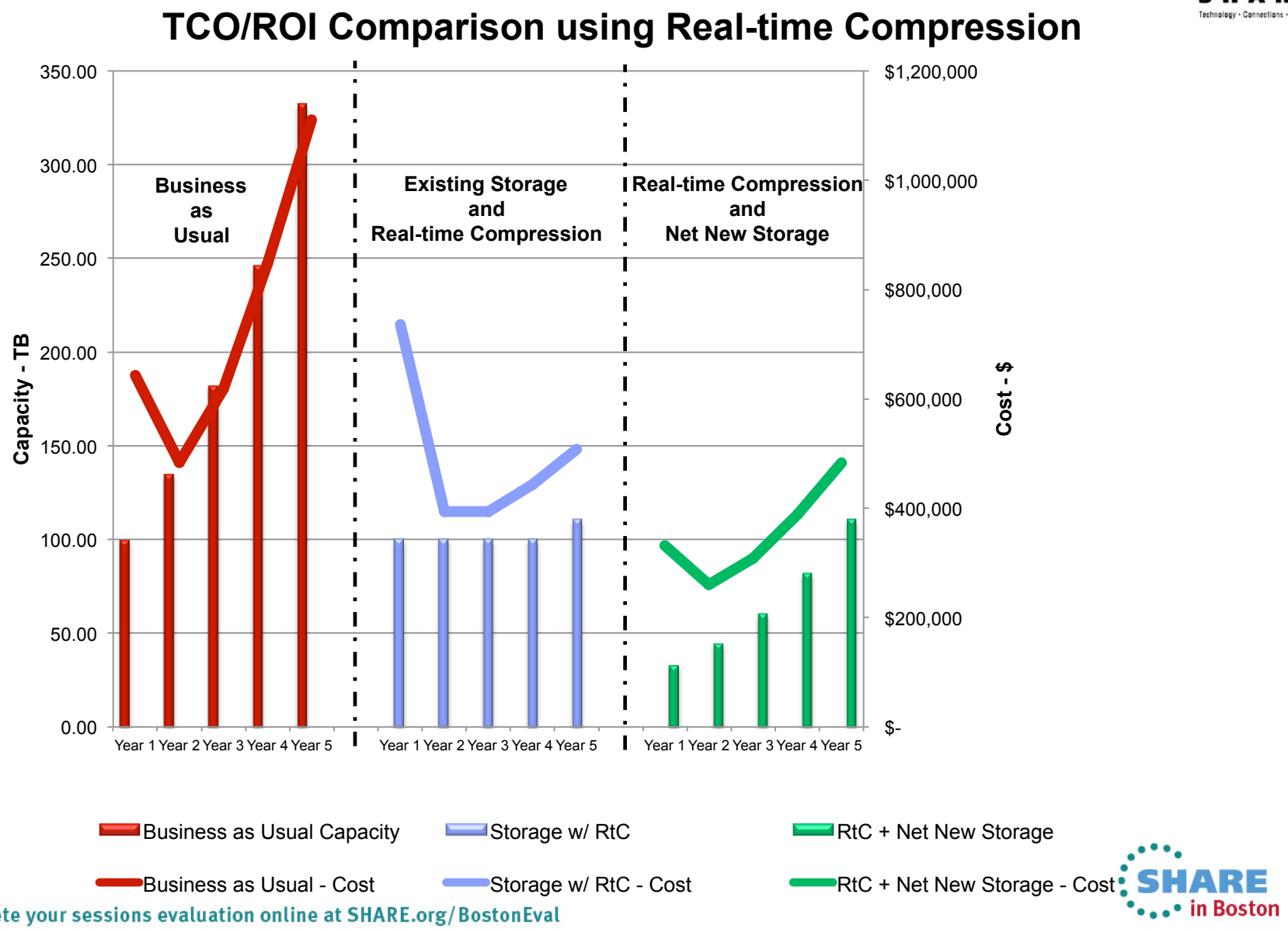
Financial Benefits



- IBM Storwize V7000 Real-time Compression provides even greater flexibility
- Maintain storage budget yet obtain greater usable capacity through compression
  - Expect approx 30%-40% reduction in $/GB for common configurations
- Maintain usable capacity but with lower budget through compression
  - Expect approx 25%-35% reduction in $/GB for common configurations
- Apply compression to existing capacity or externally virtualized storage
  - Approx double usable capacity for very modest investment
- Additional lifetime savings
  - Lower power and cooling
  - Lower software subscription and support

Based on IBM analysis. Assumes clients achieve 50% overall compression savings, 2-4 enclosure configurations

Compelling TCO Advantages



Take Control of Your Storage Budge

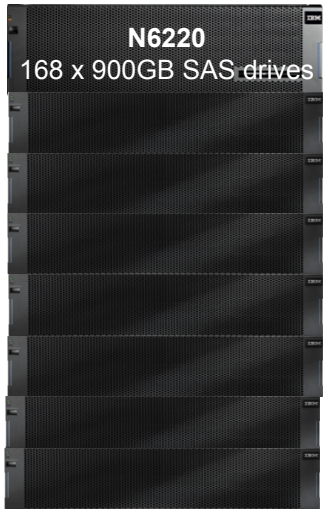
Traditional

N6210

168 x 900GB SAS drives

100TB Data capacity

List Price - **$757,880**



40% SAVING Difference Of $297,000

Smarter Storage

N6210 + RtCA

67 x 900GB SAS drives

100TB Data Storage*

2 x RtCA

SW License & 3 yr Maint

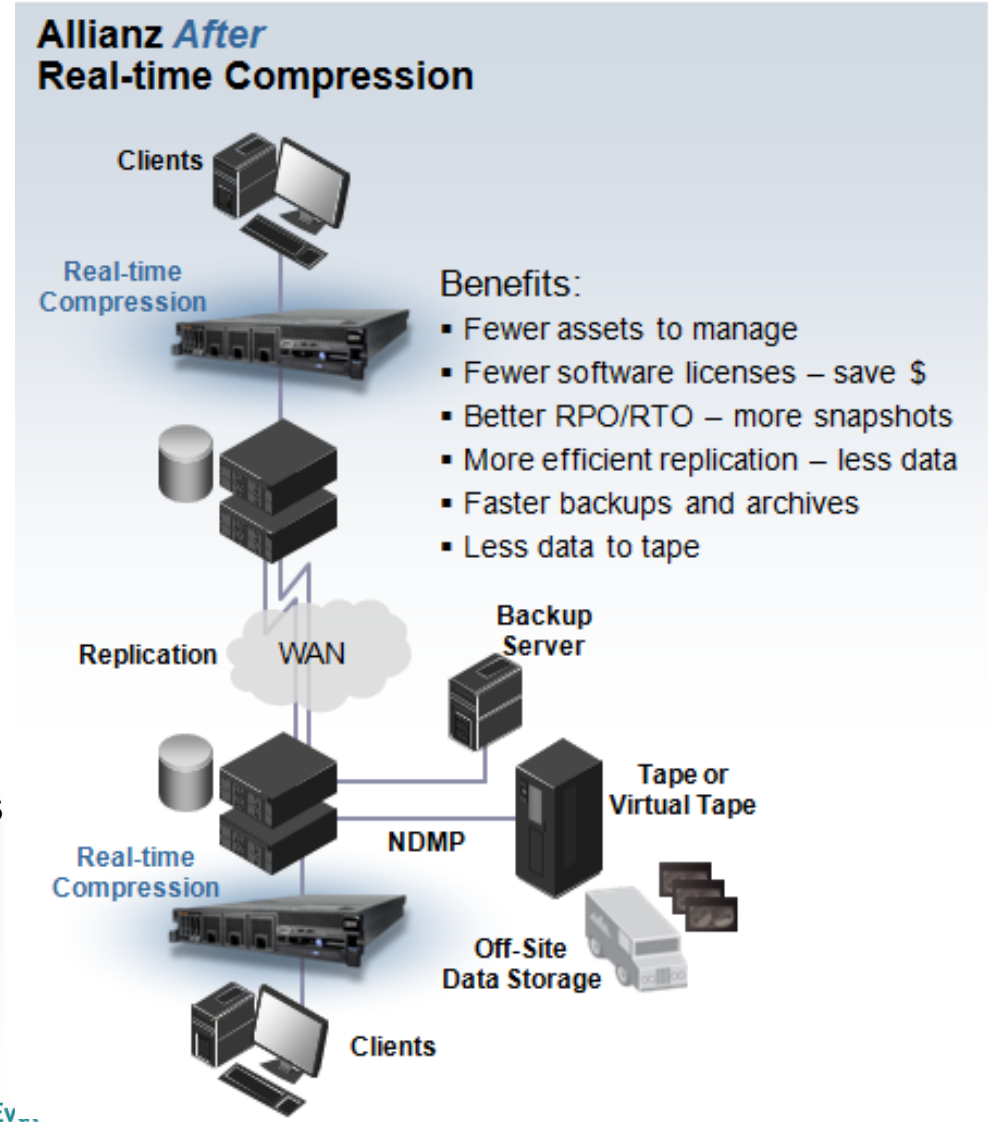
List Price - ***$460,851***

Customer Use Case

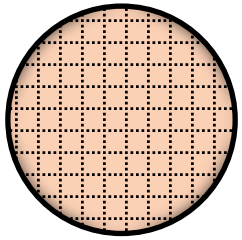
"IBM Real-time Compression has enabled us to be more efficient."

Daniel Gill, Infrastructure Analyst

- Utilizes NetBackup via NDMP for its backups to disk
- Allianz saw:
  - 55% reduction in primary disk storage
  - 65% reduction in backup disk storage
  - 35% reduction in backup times
  - More granular RPOs



IBM Real-time Compression Delivers – On All Fronts



Ground-Breaking Architecture –

Only “real time” platform designed to ensure high performance and total transparency to all applications

Maximum Compression –

Up to 80% throughout the life of the file delivering tremendous TCO/ROI. *Compresstimator* helps to speed the time to ROI by compressing existing data in place without migration.

Total Transparency & Heterogeneity –

Compression works with the V7000 as well as SVC can compression data on any storage array and works with NAS as well. Simple to implement, flexible for your requirements and requires no change to your applications.

No Performance Degradation –

Real-time Compression will not slow down your applications like other competitive technologies. Clients will not notice their data is being compressed and IT will gain significant space savings.

Where to find more information

- Real-time Compression:
  - Redpaper – IBM Real-time Compression in Storwize V7000/SVC
  - Real-time Compression Evaluation User Guide
  - Storwize V7000 InfoCenter
  - IBM Real-time Compression Best Practices:
  - Comprestimator
  - Analyst Papers
  - Product Videos

ibm.com/systems/storage/rtc