

Compellent Remote Instant Replay

July 2006



More than ever, businesses are facing massive challenges to both protect and move their data across multiple sites. Most critically, this challenge presents itself in the event of an infrastructure failure or an entire site outage. However, with the addition of multiple data center sites, expanding WAN investments, and increasingly heterogeneous and distributed computing infrastructures, moving production data between geographies has also become a hot-button issue. Because of both these issues of business continuity and multi-site data movement, data replication technologies have risen to the fore as a key investment consideration. Taneja Group believes that finding a robust, flexible data replication technology is now no longer optional. Unfortunately, many traditional approaches have forced customers to make difficult solution-level trade-offs in cost, complexity, data consistency, distance and performance. Customers need to develop a keen eye for evaluating which replication technologies can deliver optimal impact and ROI returns, while maximizing infrastructure flexibility for multiple usage cases.

In the following profile, we will discuss the business and IT factors driving the adoption of replication technology and examine the key challenges present in these deployments. We will then spotlight Compellent, an enterprise-class storage systems company, and their flagship replication product - Remote Instant Replay (RIR). We are very encouraged by what we see coming from the Compellent team. In particular, we believe that their RIR offering exemplifies some of the key issues of ease-of-use, integration, and flexibility that we see end-users asking for in their replication solutions.

Today's Replication Drivers

Now more than ever, applications and business data must remain online and consistent on a 24x7x365 basis, irrespective of infrastructure component failures or entire site outages. There is simply no tolerance for downtime, performance degradation, or data loss. An hour of downtime for any enterprise can mean massive revenue and productivity loss.

Over the past five years, replication technologies have enjoyed wide spread adoption. Based on our research, there are

several factors behind this trend. At the highest level, we see corporations utilizing replication technologies across a much broader swath of infrastructure than ever before. We see the four main drivers as:

1. Heightened Awareness of Outages – The tragedies of September 11th and Hurricane Katrina have driven home the need for adequate site to site disaster recovery.
2. Regulation and Compliance – State and federal governments have enacted legislations mandating businesses,

PRODUCT PROFILE

specifically in the areas of financial services and healthcare, to maintain remote, disaster recovery sites.

3. New Global Workflows – Insourcing and outsourcing have given rise to the need to synchronize data across multiple data centers worldwide.
4. Consolidation – The trend toward consolidation of IT assets exacerbates the need to safeguard those systems in the event of an outage.

Consequently, IT administrators are now faced with the daunting challenge of ensuring data protection and consistency for growing data sets. This issue is compounded by the complexity of managing heterogeneous data sources, stored across distributed storage environments, with requirements for zero downtime, no data loss and no performance impact. Not to mention, many companies now have one or more government regulatory bodies breathing down their necks on the compliance front. Needless to say, meeting all these requirements simultaneously is no IT picnic.

Today's Replication Challenges

In talking to end users, one thing becomes clear. There is a desire on the part of IT teams to protect and safeguard their applications and data more than they currently are able to do. However, disaster recovery in today's environment is a complicated equation that must weigh the risk of an outage and the importance of the application against the cost to administer and procure the technologies. From our

surveys with end users, three major challenges emerge:

1. Total Cost of Ownership
2. Management Complexity
3. Network Utilization

We will now examine each one of these challenges in greater detail.

Total Cost of Ownership: Traditionally, disaster recovery solutions, like replication, have been a costly proposition. Storage replication and mirroring technology built into storage systems from the largest storage vendors comes with a high upfront price tag, not to mention the cost of the overall infrastructure needed to support site-to-site disaster recovery. A data center must invest in duplicate storage and hardware environments in each location and high speed links connecting those two sites. The total cost of all these elements assures that only applications and their data with the highest SLAs will be protected.

We recommend that all IT purchasers conduct a thorough Return On Investment (ROI) calculation before purchasing replication solutions. Specifically, end users must understand the cost tradeoffs between using Fibre Channel or IP connectivity, including understanding the costs per network port, and the costs of individual replication and mirroring solutions.

IT purchasers should evaluate the total storage solution, not just the merits of the replication technology. Since many replication products are tied to the

P R O D U C T P R O F I L E

underlying storage system, it is important to scrutinize the investment protection of the underlying storage hardware and whether the system will be able to grow to meet unforeseen capacity and performance increases.

Management Complexity: Deploying replication technology has been fraught with pitfalls. In site-to-site recovery scenarios, two operations must occur flawlessly for a successful recovery to occur. First, the data must be copied from site A to site B. Second, the application and operating environment must be duplicated and the application restarted using the replicated/copied data set. In both of these operations, the storage administrator must deal with significant complexity. For example, synchronizing two copies of data and ensuring that the backup site always has a consistent recovery point is a non-trivial exercise.

Second, there is significant complexity in ensuring that the environment is preserved perfectly between locations. Applications on hosts must have the exact same operating environment (including OS and application images) as the primary. Moreover, on a recovery at the secondary site, the storage volumes must be quickly mounted and accessible to the proper hosts in order to hit low Recovery Time Objectives (RTOs). If this is not done properly, a lengthy recovery and consistency checking process must occur before the system can come up. In short, replication and site-to-site disaster recovery involves a complex set of manual tasks.

These obstacles are compounded by the fact that with heterogeneous storage

infrastructure and replication technologies from different vendors, a storage administrator must deal with disparate tools to configure, test, and ensure that the environment recovers properly. Often, the storage administrator must complete the recovery under the most extreme of circumstances – when an outage has occurred. Needless to say, this is when mistakes are more likely to occur.

For these reasons, we recommend that end users do a full evaluation of any replication product and establish metrics on how long it takes to accomplish specific administrative tasks, like mounting a volume on the secondary site. Although ease of use can be a soft-value criterion, it is absolutely essential that vendors are benchmarked against management complexity since replication must work flawlessly when it is needed most.

Network Utilization: Overall replication performance and Recovery Point Objective (RPO) is gated by the bandwidth and latency of the link between the two sites. Provisioning the link and determining the bandwidth requirements of any replication deployment are critical planning work items. In general, there is a direct relationship between how often data changes on the primary system and the amount of bandwidth that is consumed. The algorithm that a replication vendor uses will dramatically influence how much bandwidth is consumed and whether a lower cost link will be sufficient. Moreover, in some sophisticated replication packages, administrators can set quality of service thresholds to either throttle or speed up the replication.

P R O D U C T P R O F I L E

Meet Compellent

Compellent, headquartered in Eden Prairie, Minnesota, set out to build a next-generation storage system from the ground up – one that would be unencumbered by legacy architecture and would have a single point for managing all storage functions. Compellent believed that storage management was needlessly complex due to the myriad of disparate tools at the array level used to provision, virtualize, and manage storage. The company's flagship product, Storage Center, has been designed with the aim of eradicating complexity and simplifying storage management tasks, while providing some of the most powerful virtualization, Information Lifecycle Management, and data protection technologies available at the storage array level.

Focus on Storage Center

Before we focus on Compellent's Remote Instant Replay (RIR) technology, we must first understand the underlying platform on which it is built. Compellent's Storage Center combines a suite of intelligent, powerful storage software with a modularly scalable, redundant hardware platform, all administered from a single unified user interface. This combination at the storage system level married with easy to use replication technology truly differentiates Compellent from other storage vendors in the market today.

Technology Independence

Unlike other storage systems, Storage Center provides flexibility and choice in terms of

how hosts access data and the ability to seamlessly mix and match physical media types in the enclosure. The Storage Center supports both Fibre Channel and iSCSI access to any volume in the array. In fact, end users can have concurrent iSCSI and FC connectivity to a single volume. Other arrays cannot support Fibre Channel and iSCSI connectivity concurrently through the same controller.

Secondly, Storage Center allows customers to optimize their storage configuration to suit their needs, whether it be pure performance or cost per capacity metrics. Storage Center supports the mixing and matching of any type of disk drive, whether it be high performance Fibre Channel, mid-range Fibre Channel, Serial ATA (SATA), and eventually Serial Attached SCSI (SAS) hard disks. As a result of Compellent's design, Storage Center can quickly add new hard disk technologies without a forklift upgrade.

Redundant & Modularly Scalable

Storage Center has a clustered system architecture consisting of clustered storage controllers and fully redundant pathways to the disk. The system has no single point of failure. Adding capacity and performing upgrades (i.e., new disks or software) can be done without disruption to ongoing data access.

In addition, Storage Center has been architected to scale capacity, connectivity, and bandwidth on demand. Storage Center supports the ability to scale the capacity of a single system from a single terabyte to 100s of terabytes. Furthermore, additional ports

PRODUCT PROFILE

can be added to support higher I/O throughput rates and greater connectivity.

Feature Rich Software

Another hallmark of Compellent Storage Center is its rich software stack of storage applications. There are five main capabilities of the software worth highlighting:

- Automated Storage Tiers that migrate blocks of data within a volume from higher cost disk to lower cost disk based on usage patterns.
- Thin Provisioning that ensures capacity is consumed only when needed.
- Virtualization and Volume Management that presents virtual volumes that span different physical drive types and capacities and RAID sets.
- Continuous Snapshots for data protection.
- Remote Replication for multiple site work flows and disaster recovery.

With a brief overview of Storage Center complete, let's turn our attention to the replication capabilities of the Compellent platform – Compellent Remote Instant Replay.

Focus on Remote Instant Replay

Compellent's Remote Instant Replay (RIR) is Storage Center's method for replicating a volume from the primary system to one or more secondary systems. RIR supports a wide range of replication configurations - one system to another system (1:1), one to many (1:N), or many to one (N:1). In addition, RIR supports uni-directional and bi-directional

replication of data. RIR can scale up from the simplest of site-to-site replication deployments to the most complex data flows among multiple data centers.

Furthermore, RIR supports replication between dissimilar storage environments and configurations. Users can replicate from a primary site using FC disks to a secondary site using SATA. In short, RIR provides users more deployment flexibility and demonstrable cost savings by allowing the use of lower cost disk at the secondary site.

It is important to note that RIR replication technology is not distance limited. It can transmit data over either Fibre Channel or IP links. To that end, RIR can support replication within a data center, across a campus or Metropolitan Area Network (MAN), or even across a Wide Area Network (WAN). Typically, for replication over a WAN, RIR will transmit data over an IP network using the iSCSI protocol. Unlike other mid-range array vendors who require an FC to Ethernet converter or an external appliance for iSCSI replication, RIR provides native iSCSI replication at the controller level. While Fibre Channel connectivity might be used in a Metropolitan Area Network if performance and latency were requirements.

In addition, RIR supports two modes of replication –synchronous and asynchronous. Synchronous replication guarantees the highest level of data integrity since the two copies of the data can never be different, but it comes with a performance impact and requires very high bandwidth links.

PRODUCT PROFILE

Synchronous replication is typically used in two scenarios. First, it is used to migrate data between two different storage volumes, while ensuring that the data stays in synch across both volumes simultaneously. Secondly, synchronous replication is mandatory in disaster recovery scenarios where a Recovery Point Objective (RPO) of zero is a requirement. Taneja Group typically sees large Fortune 100 businesses, especially financial services firms, using synchronous mirroring technology to safeguard their most mission-critical systems.

The second and most common mode of replication that RIR supports is asynchronous. Asynchronous replication has essentially no impact to the overall storage system performance. From our discussion with end users, most users deploy asynchronous replication to achieve an RPO of greater than 15 minutes. Asynchronous replication constitutes the vast majority of all replication deployments.

Compellent has taken an innovative approach, marrying asynchronous replication technology with Storage Center's continuous snapshot capability. Continuous snapshots (aka "replays" in Compellent's vernacular) are an unlimited number of point-in-time copies that are taken of a particular volume. A replay or a snapshot is stored as a collection of blocks that have changed since the previous snapshot. Each of these snapshots can then be efficiently replicated across the wire to the secondary system and appear as a local snapshot. An administrator sets a policy stipulating how often the Storage Center should take a

snapshot of a given volume. The frequency of the snapshot determines the overall RPO that can be achieved.

Differentiator #1: Integration with Snapshots

By integrating snapshots and replication, Compellent eliminates the consistency problem prevalent in site-to-site disaster recovery scenarios and also reduces the complexity of bringing up a secondary site after a disaster. A snapshot represents a consistent point-in-time copy of the data. Therefore, rolling back to a snapshot ensures a consistent image from which to recover. RIR eliminates complex recovery and consistency checking procedures.

Merely replicating data does not guarantee a consistent image. If a site outage occurs during the data transfer, the secondary site can be left with an incomplete view of changing data. A lengthy file system or database recovery operation will need to be performed by the administrator before the secondary site and its applications can be up and running. Many of the major storage vendors support snapshots and replication, but do not marry the two technologies together to reduce the complexity of recovering from an outage like Compellent does.

Differentiator #2: Single Management Interface

Compellent has invested a tremendous amount of thinking and energy into building a storage system that is easy to use. Compellent's Storage Center provides a single unified control point for replication management, scheduling, and recovery

P R O D U C T P R O F I L E

verification. The Storage Center console allows an administrator to visualize both what volumes are being replicated, and their target destinations. In addition, the RIR product comes with wizard-based tools that support setup and configuration of replication policies. Even more importantly, the RIR product provides a wizard for a fast recovery of volumes on the secondary system. Unlike other storage vendors that bolt-on core product capabilities without ease-of-use considerations, Compellent has engineered the entire system around the actual administrative tasks and challenges that a user faces when deploying replication in a multi-site context.

Differentiator #3: Online Verification

RIR allows administrators to confirm recovery without downtime on a daily or even hourly basis, something that is difficult to achieve with most traditional replication solutions. RIR provides a way for administrators through a couple mouse clicks to recover and mount the replicated snapshot to any host on the secondary site. Through the user interface, the administrator can validate that the replication is working and that there are valid recovery points available in case of an outage. RIR is designed so an administrator can verify this on a minute or hourly basis if need be. This is a truly unique and immensely valuable capability of the Compellent system. Like backup, the recovery process is often overlooked and can fail at unexpected or inopportune times. By eliminating the risk of a failed recovery, Compellent has made replication easier to implement with confidence.

Differentiator #4: Bandwidth Optimization

Bandwidth optimization and throttling is a useful feature of the Compellent system that few other storage systems support. Traditionally, only high-end replication appliances provide advanced bandwidth sensing and throttling capabilities like Compellent's RIR. Since the Compellent RIR product only replicates snapshots, it is highly efficient and transports only the relevant changed blocks. Also, as a result of Compellent's thin provisioning capability, RIR only replicates used blocks of a volume. Most other storage replication systems copy the complete volume, resulting in a huge waste of system and network bandwidth. Due to the efficiency of RIR, Compellent customers find that they can use lower bandwidth links and save additional money.

In addition, the underlying Storage Center architecture applies bandwidth shaping algorithms to further improve performance across the wire and throttle bandwidth usage at a per volume level. A Compellent administrator can stipulate what level of bandwidth the RIR product can use throughout the day, week or month and set the priority for any given volume. Administrators can choose to throttle the use of bandwidth during peak usage or when a particular job may be running on the public network. As with all Compellent functions, bandwidth optimization can be setup and configured from the same, unified management interface.

P R O D U C T P R O F I L E

Taneja Group Opinion

Taneja Group believes that the data replication market is currently at a major inflection point. Driven by internal business pressures and external regulatory requirements, enterprises can no longer afford to make the sort of compromises they have made historically when implementing replication solutions. Many of the administrators with whom we have spoken have already come to this realization and are aggressively looking for viable alternatives.

Our opinion is that disaster recovery planning shouldn't have to be a harsh set of tradeoffs. Cost and complexity of the overall solution can be held in check allowing much more widespread adoption of replication technologies across a broad spectrum of applications and companies. We believe that Compellent's RIR product allows end users to navigate the complexity of DR planning with relative ease.

From the Compellent customers with whom we have spoken, they report that the Compellent Storage Center and RIR product delivers a high level of ROI and leverage across the operations of their storage infrastructure. One such user used the RIR product not only for data protection and disaster recovery needs, but also as a vehicle for streamlining production data workflows across multiple geographies. This customer was able to kill two birds with one stone – always a high ROI proposition in IT.

One of the ironies of advanced storage solutions today is that *true* ease-of-use is hard to come by. Many storage vendors have

introduced needless complexity through line extensions, disparate software products, each with their own interface. The result is a hodgepodge of utilities and products that don't truly interoperate together. We have seen that movie before and know that it can end badly for the end user.

Compellent has hit a homerun in terms of architecting the system around simplifying the provisioning, management, and protection of data. Storage Center and RIR are true examples of products that make it easier for end users to do their jobs.

We believe that Compellent's combination of robust storage software, modularly scalable hardware, and technology independence at the protocol and disk level, all wrapped with a central management interface makes it a clear winner in the midrange storage array category today and in coming years. Further, when we look to the specifics of Compellent's RIR replication functionality, the true ease-of-use story floats to the top. This product is the real deal. We have spoken to many customers of the Compellent platform and they all rave about it – a sure sign of success.

If customers have serious considerations around removing storage complexity points, delivering a strong ROI story and maximizing their flexibility across multiple usage models, there are few companies with a stronger offering than Compellent. Give them a hard look and run this platform through its paces. Our bet: It will stand up to the toughest scrutiny and pass with flying colors.



P R O D U C T P R O F I L E

NOTICE: The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document.