



WHITE PAPER



# A Cost-Effective Integrated Solution for Backup and Disaster Recovery

## Introduction

Recovery is something for which all businesses must prepare. Most acknowledge the need for comprehensive business continuity plans, but the expense and complexity of fully implementing these plans leads many enterprises to only partially implement them. Business continuity plans generally cover information technology (IT) infrastructure (applications and data), communications, people, and process recovery. This white paper discusses a cost-effective approach to providing IT infrastructure recovery solutions to meet remote (disaster recovery) and local (backup) requirements.

IT infrastructure recovery generally revolves around application and data recovery, which translates to several separate types of products: backup software that is used for local backup and restore, shipping services to transport backup tapes to remote locations, replication software that is used to replicate data to remote locations for disaster recovery (DR) purposes, and clustering products that are used to automatically recover applications when they fail (regardless of whether that is due to hardware or software failures). Operative concerns include setting up cost-effective DR plans, minimizing data loss when recoveries are required, minimizing the time associated with recovering from data and/or application failures, and making recoveries easy and reliable. Recovery point objective (RPO) is an industry metric used to target the maximum allowable amount of data loss on recovery, while recovery time objective (RTO) is another industry term used to target the maximum time to recovery, and both metrics can be applied to either applications or data. Very stringent RPO/RTO requirements can result in near zero data loss on recovery and very short recovery times. High data growth rates are compounding traditional recovery problems, in particular because dealing with ever larger amounts of data can impose performance degradation and/or disruptions and longer recovery times on application environments that are being protected.

Conventional data protection approaches have tended to rely on backup software and tape-based infrastructure. As information stores have grown, the use of tape media for local backup has led to a number of problems, including an inability to complete backups within the allotted time periods, manually intensive data protection operations, long recovery times, and poor recovery reliability. These issues have driven interest in the use of disk-based recovery technologies. Solutions built around disk-based technologies allow recovery operations to become more automated, minimize the impact of data protection on production applications, result in faster, more reliable recoveries, and can significantly improve DR performance relative to plans which rely on shipping backup tapes to remote locations.

If you are interested in moving to a more automated set of processes as the foundation for your DR plan, your current recovery processes are causing problems with production impacts, RPO/RTO, or recovery reliability, you are looking for easy, cost-effective ways to improve application availability, or you are thinking about investing in a virtual tape library (VTL), InMage Systems offers a disk-based recovery solution, targeted for use in heterogeneous environments, that combines application and data recovery into a single platform that can be used for DR, local backup and restore, automated application failover and recovery, or any combination of the three.

### Storage Realities for Mid Market Enterprises

Several trends are impacting data protection planning – both remote and local – for mid market enterprises today:

- Unprecedented data growth rates in the 50% - 60% range that are expected to continue for at least the next 5 years
- Increasingly stringent RPO/RTO requirements driven by evolving business and regulatory mandates
- IT groups that are oversubscribed and underbudgeted, leading to strong requirements for cost-effective solutions that are easy to use

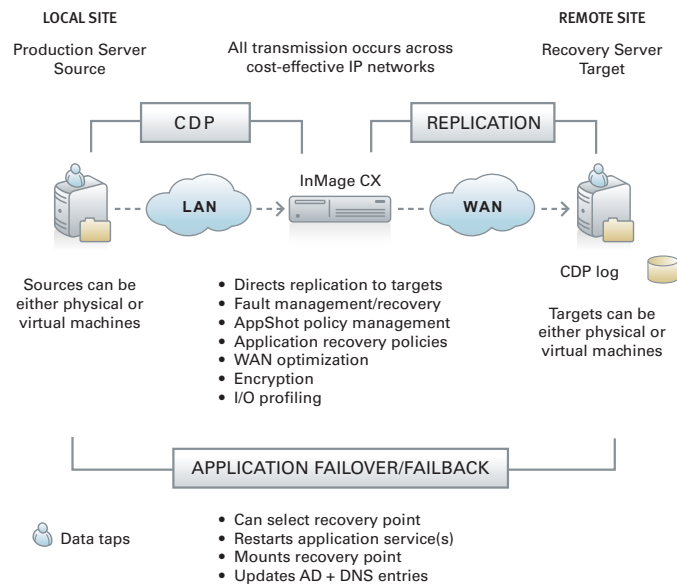
Most mid market enterprises have implemented some form of local backup which is tape-based, and it is increasingly apparent that tape-based solutions don't cope well with the above trends. Despite known issues with DR plans that depend upon regularly shipping backup tapes to remote locations by truck, scalability, overhead, and/or cost issues may have prevented evaluation of replication technologies which can clearly improve DR RPO/RTO. In an effort to improve overall application availability, clustering solutions have often been considered and even tried, but cost, complexity, and reliability issues are stalling wider deployments.

Managing heterogeneity is a fact of life for mid market enterprises, which generally have a variety of disparate server, storage, network, and application technologies. Solutions which impose vendor lock-ins can limit the ability to integrate new functionality and technologies. Options which can preserve existing investments in hardware and software while at the same time providing maximum flexibility for integrating newer hardware and software products ultimately allow IT managers to deploy the most cost-effective IT infrastructure.

Finally, technology deployments built around application solutions can provide richer functionality that is easier to deploy and use for a given application, particularly for smaller IT groups. There is a risk, however, when horizontal solutions (e.g. backup, DR) are application-specific and can only be used for a single application environment – it can lead not only to higher costs but also to additional management complexity as each application solution imposes its own learning curve on already overworked administrators. What would be most interesting are solutions which are application-aware, offering rich functionality specific to each application, and yet are deployed and managed within a common management framework that spans multiple application environments.

### InMage: A Single Platform To Comprehensively Manage DR, Backup, and Application Recovery

InMage offers InMage, a software-based recovery solution, deployed in the network, which combines remote DR, local backup, and application failover (both local and remote) into a single platform. Built upon a foundation of disk-based recovery technology, InMage definitively resolves backup window, RPO/RTO, recovery reliability, and automated application recovery issues with a solution that can stand alone or be seamlessly integrated as a "front end" to existing tape-based infrastructures. Application-aware functionality is available for a number of key enterprise applications, plugging into the InMage core platform (called the CX) and leveraging a set of centralized management practices that are common across all protected applications. InMage's architectural design flexibly accommodates Windows, Linux, and Unix servers, physical and/or virtual servers, heterogeneous storage and storage architectures, and any application environment on a platform with a low entry price point that scales cost-effectively to configurations protecting hundreds of servers across multiple locations.



**Figure 1.** To make it easy to understand InMage's theory of operation, this figure shows a configuration with one production server source and one recovery server target, although in production deployments there are generally multiple sources and multiple targets. The storage owned by the sources and targets can be of any type (DAS, SAN, NAS). InMage captures data from sources using CDP across a LAN, then replicates them to targets using asynchronous replication over IP.

## Simple, Cost-Effective DR

DR best practices recommend that DR sites be located at least 200 miles from primary sites to ensure that a widespread catastrophe like a tornado, earthquake, or hurricane cannot wipe out both sites. If tapes are being regularly shipped to a DR site and a recovery requires accessing these tapes, companies can generally expect to lose at least several days of data. Recovering applications from tape-based data stored at remote locations can also take anywhere from several days to a week.

If, on the other hand, replication is used to deploy a DR capability, data loss can be reduced to no more than a few minutes and recovery times can be on the order of minutes to hours. This is primarily because replication collects data continuously as it is created, sending it immediately to target locations like DR sites across networks instead of on the trucks that physical tapes require. Shipping physical tapes can introduce days of transit time. Asynchronous replication can meet long distance DR requirements, supporting effectively unlimited distances between primary and DR sites.

InMage leverages continuous data protection (CDP) and replication technologies to move data from source servers (the production servers) through a network-based control server called the CX to target servers. Target servers can be either local or remote, as long as there is an IP-based network connection to them from the CX. The storage attached to source servers can be any type (DAS, SAN, NAS, iSCSI, FC) and from any vendor, and the same is true for target servers (we can replicate between heterogeneous storage subsystems). This heterogeneous support provides significant deployment flexibility since any new equipment can be purchased or pre-existing equipment can be re-purposed for use in these configurations.

Applications that have large data sets can present DR challenges, particularly if there is limited network bandwidth between primary and DR sites. Because it uses a variety of capacity optimization technologies, InMage can protect very large data sets while using minimal bandwidth. Once an initial baseline is established, it collects only data changes. As it moves this data between source and target servers, InMage employs WAN optimization to minimize the amount of data that has to be sent to a given target to support recovery operations. WAN optimization algorithms include built-in TCP optimizations that minimize the amount of data that has to be sent across the network and compression, while bandwidth shaping is performed by administrator-defined quality of service (QoS) policies for co-existing with other production traffic on the network. If a disaster occurs, InMage supports failover of application services to the DR site so that business operations can be quickly restarted despite a catastrophic failure at a primary site. Production operations can be run from this DR site until the primary

site can be restored, at which point application services can be failed back while keeping all data changes incurred at the DR site. In sending this data back to the primary site, InMage uses delta differencing technologies which determine the minimum number of blocks that have to be sent across the network to re-create the DR site data states at the primary site, another technique which minimizes network bandwidth requirements. Taken together, these technologies makes very efficient use of available bandwidth, allowing us to protect large application environments while meeting stringent RPO/RTO requirements with minimal bandwidth.

InMage offers several key advantages relative to more conventional replication-based DR products. First, we offer a more comprehensive recovery capability that includes not only the data but extends all the way up to the application. At the push of a single button, InMage will fail application services over to the target location. Instead of the manual process that conventional replication products require to get an application back up and running, we have fully automated that process, making it faster and more reliable.

Second, we can enable that automated recovery to occur from any retroactively selected recovery point, given our underlying CDP foundation for local data capture. This recovery granularity for a DR solution allows us to recover from data corruption problems with minimal data loss, perform better root cause analyses of failure events, and potentially recover an application to any previous point in time that may be desired for reporting or maintenance purposes (a quarterly close, a pre patch state, etc.).

Third, InMage gracefully manages transient network outages and transparently recovers from them without any data loss, ensuring that protection is always continuous.

Fourth, InMage supports mixed physical/virtual server configurations when setting up source and target server definitions. The use of virtual servers as recovery targets can significantly lower the cost of setting up DR configurations, and the fact that we support all three major virtual server environments – VMware, Hyper-V, and Xen – provides maximum deployment flexibility.

And fifth, InMage's "hot site" recovery capabilities support much better recovery performance than "cold site" options. That translates to faster, easier recovery, which is what you want when time is of the essence.

## Eliminate Backup Impacts While Improving Recovery Capabilities

InMage moves enterprises away from the "point in time" orientation of backups that impact production operations. Since data is being continuously captured as it is created, backup windows become a thing of the past, yet enterprises have

more granular access to recovery points. InMage effectively moves backup operations off production servers so that they do not impact business operations.

This approach can significantly minimize the need for tape-based infrastructure. Since most local recoveries are performed from relatively recent backups, companies that use InMage to keep several days to several weeks of changes resident on disk will end up performing most recovery operations from disk, resulting in faster, more reliable recoveries that, given InMage's granular data capture and recovery point selection capabilities, can meet much more stringent RPO/RTO requirements. Companies that before were performing a single weekly full backup with six incrementals may decide to dump data to tape far less frequently, possibly once a week, only after data ages out of an administrator-defined data stream retention period, after data has been replicated to a remote site, or only after data reaches a certain defined age at the remote site. Remote offices that previously were dealing with tapes can now move to tape-less operations that rely on disk-based recovery that is centrally managed. Enterprises that do not keep "backup" data around for more than several months may in fact be able to move to completely tape-less backup operations.

### Application-Aware Solutions

InMage supports a number of different application-aware recovery solutions for key enterprise applications, including Microsoft Exchange, SQL, and SharePoint, as well as Oracle, MySQL, Blackberry Server, SAP, and any Windows, Linux, or Unix file systems, among others. Multiple application solutions can run on a single CX although each application solution includes application-specific integration, particularly for installation, tracking application-consistent recovery points (known as **AppShots**), and for managing application failover/failback. This integration results in a solution that produces fast, reliable application and data recovery but is managed under a set of policies (installation, deployment, marking AppShot recovery points, creating and using AppShots, failover/failback, etc.) that are common across all application environments. InMage's ability to manage multiple application recovery solutions under a common management framework provides a compelling alternative to application-specific solutions that support only one application type.

There is a big difference between application-specific and application-aware solutions, and the latter offers significant benefits:

- They make comprehensive recovery solutions for specific applications easy to buy and deploy
- They leverage application-specific domain expertise in our engineering groups to provide recovery capabilities tailored to application-specific requirements, producing better overall products

- Application-specific functionality is not a set of custom scripts put together in the field - it is fully productized, meaning that it is developed, tested, documented and supported by InMage, a fact which results in a reliable solution with predictable performance
- Multiple application solutions can be combined on a single CX but managed through a single secure, web browser-based GUI (a command line interface is available as well) to produce well integrated solutions that are easy to install, deploy, and manage

AppShots have other administrative value than just as optimum recovery points. They make it easy to cut off copies of production data sets at any time without impacting production operations. These AppShots can be mounted on non-production servers to support business intelligence tasks and other reporting functions as well as test and development operations. Since AppShots can be retroactively created to represent any desired point in time, they can be used to time shift administrative operations to optimize operator productivity.

### InMage: Unique Hybrid Recovery Technology

InMage uses an architecture that is unique in the industry. Key concerns this unique architecture was built to address include cost-effective scalability across a wide range of configurations from several to several hundred servers, non-disruptive deployment (in terms of both host impact and in failure modes), and an ability to accommodate heterogeneous servers and storage as well as a variety of different storage architectures. Granular data capture and selection of recovery points was a critical design tenet, both for minimizing the impact of data protection operations on protected servers but also in providing the maximum flexibility in selecting the optimum recovery point to meet a variety of different RPO/RTO requirements.

### Granular Data Capture and Recovery

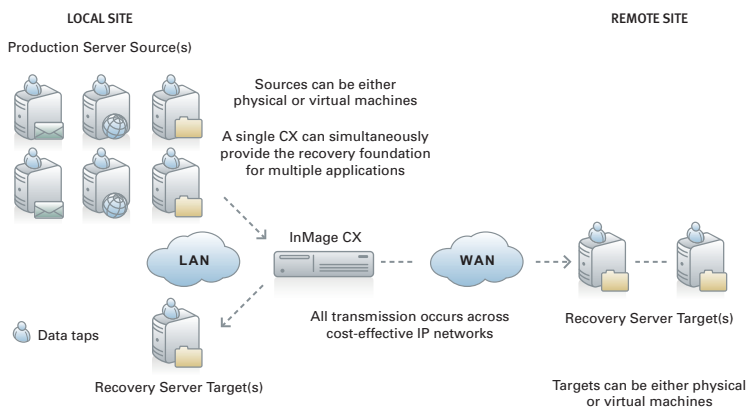
Conventional backups are performed periodically, and as data sets grow this "point in time" approach presents problems. InMage completely avoids the problems associated with trying to back up an ever increasing amount of data through static network links and within given backup windows by collecting changes continuously, as they occur. This effectively spreads the "backup" throughout the day and provides excellent protection for even large data sets across network links with very limited bandwidth. The changes appear to InMage as a data stream, and this stream is annotated with time stamps as well as any number of markers that denote application events of importance when it comes to recovery or other business operations. These markers, also referred to as bookmarks, can be inserted into the data stream either manually or through policies to mark any number of potential recovery points, such

as application-consistent points created using application APIs (e.g. VSS for Windows applications, RMAN for Oracle, etc.), pre- and post-patch points, or meaningful financial milestones such as quarterly closes. The administrator specifies a retention time frame for this data (e.g. 3 days, 2 weeks, a month, etc.), and any point in that data stream of retained data can be retroactively selected and used for recovery purposes.

Once a recovery point has been selected, the administrator can choose between creating a physical or a virtual disk-based image of that point. Think of it as turning a dial back to any desired point and generating the desired image. These images can be designated as read only or read/write, depending on what they will be used for. Image creation can be scheduled, event-driven, or generated upon demand without impacting production applications in any way, and are mounted on recovery server targets for use. AppShots (application-consistent points in time) are generally used for most application and data recoveries since they support the shortest RTOs, but other points that are generated based on time or other events can be valuable for root cause analysis, business intelligence, reporting, or other purposes.

### What Makes InMage Unique

To manage data capture, InMage deploys filter drivers, referred to as data taps, on each server for which protection is desired. The data taps pass writes directly through to storage targets owned by the production server while at the same time asynchronously passing them over to the local CX. The CX will then replicate that data to one or more recovery server targets, and those targets can be local and/or remote. The CX does not store data over time, it only acts as a "way station" for the data. The data is stored on the recovery server target(s). InMage uses both CDP and replication as foundation technologies.



**Figure 2.** InMage collects data from source servers at a very granular level (write by write) across a LAN, then replicates that data to one or more target servers which can be either local or remote.

Note that the data flow originates at the production server sources, flows through the CX, and then onto one or more recovery server targets. This is a unique architecture that is neither host-based nor appliance-based replication. It is a unique design – the only one of its kind in the storage industry – that imposes minimal server overhead, offers graceful scalability across a wide range, and can flexibly accommodate a number of different source and target devices. It does not suffer from the scalability and overhead problems of traditional host-based replication, does not require appliance pairs at each location like traditional appliance-based replication, and does not impose the vendor lock-in that is a disadvantage in array-based replication. It really is a hybrid model deploying elements of both host-based and appliance-based replication, which is why InMage refers to it as "hybrid recovery technology".

The CX includes the processing power necessary to provide significant functionality, such as driving the data stream marking policies, managing the creation and disposition of AppShots, minimizing the amount of data that has to be sent over the WAN through capacity optimization technologies such as delta differencing, compression, and bandwidth shaping, managing encryption for data that is in-flight across the WAN, handling fault management issues such as transient network outages and application failover/failback, and performing I/O profiling.

I/O profiling is an interesting area. The InMage Analyzer can be installed prior to production deployment to determine I/O rates across the network, which provides valuable data in understanding what sort of RPO/RTO requirements you can meet with your existing infrastructure. This allows you to deploy with confidence, knowing what your recovery performance will be. After deployment and during normal operation, the CX acts as a focal point through which the data streams of potentially multiple servers flow. InMage can monitor these I/O streams over time, helping administrators understand how the I/O load on their networks may vary at different times of the day and over time. This trending analysis provides valuable input for network capacity planning.

### Summary

InMage is targeted for environments that have at least several servers that need either remote disaster recovery, better local backup, and/or higher application availability managed through an automated failover/failback capability. It leverages low cost IP-based networks for all data capture and replication, supporting low entry price points. Its heterogeneous support allows the use of any storage devices using any type of architecture, and this investment preservation story contributes to its flexibility and cost-effective deployment. 50% of our customers use InMage to meet both DR and local backup requirements,

and find themselves able to remove heavyweight backup and other agents from many servers, freeing up CPU cycles (a particularly important issue in virtual server environments) through InMage's "host off-load" design. Since most data recoveries use the most recent data, customers generally keep several days to several weeks of data within the InMage repository, leveraging the advantages of InMage's very granular data capture and recovery capabilities as well as the benefits of disk-based recovery to handle most recoveries. If customers want to migrate older backup data into another application, such as backup or archive software, that data can be staged from App-Shots without imposing any production impacts whatsoever.

InMage is a great solution for mid market companies that are looking for cost-effective DR options, may be struggling with tape-based backup infrastructure, are considering a virtual tape library, and/or want higher application availability but are put off by the cost and complexity of conventional clustering solutions. It is the only block-based replication option available on the market that combines remote DR, local backup, and application failover/failback in a single, easy to deploy platform. Its' non-disruptive deployment, heterogeneous support, and granular recovery capabilities offer a compelling value proposition for enterprises concerned about DR, backup windows, RPO/RTO, recovery reliability, and/or application availability.

