RecoverGuard™
White Paper

Table of Contents

Will Your DR Solution Work When You Need It?

Enterprises invest significant resources on their disaster recovery solutions in an effort to protect critical data and ensure the continuity of their business operations. Typical DR solutions involve the replication of data to remote data center sites configured with standby hardware and software and maintained through the use of DR methodologies and policies. The fundamental challenge to IT organizations is ensuring their DR resources will reliably function in the event of an outage or disaster. According to Gartner, 70% of DR solutions will not work when needed[1]. To understand why most DR solutions will not work, let's take a closer look at the reasons for DR failures.

Reasons for DR failures

The reasons for DR failures are generally based on the scale and the constant changes introduced to the IT environment. A typical enterprise may run hundreds of different applications, running on thousands of servers, with data stored on various storage platforms, in multiple locations around the world. Changes are introduced to this already complex environment on a daily basis. Since the DR solution must be kept identical, it has to cope with all these constant changes. Any small configuration change, such as addition of a new volume, or reconfiguration of replication processes, can create a gap between the production and DR environments. Even the smallest gap may prevent the DR solution from operating when needed.

The reasons for the creation of gaps between the production and the DR environments are:

- **Large Scale Implementations –** The larger the scale the higher the likelihood of mistakes that cause gaps.
- **Heterogeneous Environment -** A single enterprise may use various technological solutions, which include multiple Operating Systems (OSs), multiple databases (DBs), multiple storage platforms, etc. Consequently, the IT environment in the DR site may not stay consistent with the primary production data center.
- **Multi-Layer Dependencies –** The DR solution involves essential dependencies between multiple layers-- operating systems, storage, databases, network, servers and applications. Due to the dependencies between these layers, small mistakes may have much larger impacts.

[1] Source – Gartner 2004, Business continuity, Gaps in best practices.

- **Multiple Stakeholders –** Configuring application DR typically involves multiple subject-matter experts, such as DBAs, DR specialists, application developers and various contractors. Any miscommunication between various IT personnel, may result in the formation of a gap jeopardizing DR.

DR gaps may create the following risks:

- **Data Protection Risks –** Application data, meta-data, and data links may be jeopardized by gaps. Gaps may occur in replication, setup, sequence of procedures, accessibility, mapping, zoning, and more. Therefore maintaining the completeness of the data, and its internal structure consistency may be a daunting task.
- **Availability Risks –** Availability gaps may occur from misconfiguration of clusters and databases, incorrect mapping of replicated storage to standby hosts, and more. Existence of such gaps may result in a standby hosts' (e.g. standby and DR servers) failure to perform its role when needed.
- **Optimization Risks –** deployment gaps could result in excessive allocation of storage resources, inefficient use of the Storage Area Network (SAN) resources, or simply by not conforming with best practices.

These layers are interrelated and interdependent. For example, in response to a routine request for additional storage space made by a DBA, a storage administrator , may correctly configure and allocate two sets of new devices; one for database file storage and the other for flat files. Typically, each set will have it's own replication policy, it's own consistency group definitions, etc. Usually, such configurations will be carried out a few weeks in advance. Finally, when the DBA has permission to perform the necessary maintenance to extend the database, a new device from the wrong set could be incorrectly used to extend the database. The result is the entire replicated data set becomes unusable.

Common Practices

Gaps will be found in even the most organized and strict organizations where change management practices are strictly followed. In an effort to remedy the situation, organizations try to audit their DR on a regular basis. Operations teams use various checklists or pre-set schedules to verify normal operation (e.g. standby environment, replication mechanisms) and to make sure that predefined SLAs are met. However, systems change rapidly and often the individuals making these changes are not tasked with updating auditing procedures. As a result, a gap occurs in auditing processes.. For this reason, organizations hire external consultants to perform full-scale periodical audits and tests every few years.

Auditing

Auditing is important for maintaining the integrity of the DR solution. Auditing generally involves performing comprehensive checks according to predefined checklists. Checklists are based on the expertise of the auditor and are designed to detect vulnerabilities. Due to the high cost, the frequency in which audits are performed is usually not sufficient for maintaining the relevancy of DR procedures. Furthermore, since audits may take weeks or months to complete, the information may be out of date by the time the audit is concluded.

Testing

The goal of a test is to simulate a real disruption scenario to see if the DR solution works. Tests are extremely important as they provide a definitive answer to the question of whether or not the DR solution will work in real time. Tests involve extensive preparation to make sure they do not jeopardize availability and that data remains protected, but at the same time simulates real disruption. The actual test is preceded by a hectic period of several weeks in which the environment is re-documented, and various teams spend significant time anticipating problems and determining how to check for the existence of possible gaps.

Tests should be as close to reality as possible, but this is, in many cases, very difficult to achieve. Easily overlooked dependencies or configuration details may lead to undetected gaps. For example, a test may involve systems which are unknowingly using resources from the production environment, instead of the corresponding DR resources, such as domain services, file servers, databases, etc. In this case, these systems will function properly during the test, but may obviously fail in a real-life situation, where the production environment is not available. On the other hand, “near-real” tests may result in downtime and data loss.

Tests frequently tend to be partial because the identification of problems during the test postpone its completion. Under the best of circumstances, critical applications are tested only once or twice a year, while other applications are tested every two years (Gartner) Furthermore, since testing is often broken into smaller portions and

performed sequentially,, some application cross dependencies are never tested. As a consequence, most current testing methods are inaccurate, risky, and at best provide merely a single snapshot of an ever changing environment.

Conclusion – Automation is Needed

Many organizations find they simply cannot afford the time required to perform comprehensive audits and tests. Searching for signs of only a single known problem on one server requires hours or even days of troubleshooting. With hundreds of servers and dozens to hundreds of known issues to check, there is simply not enough time do everything. Automating the process would make it complete, accurate and most importantly, recurrent. Automation itself, however, may not be adequate.. Home-grown check-lists for uncovering gaps in DR, may include dozens of items and still not capture all potential threats to recoverability. Only a comprehensive best practices knowledge base, compiled by hundreds of DR specialists, can be effective. The checks must be based on actual mapping of the relationships and dependencies in the IT environment. Finally, it must also be unobtrusive, otherwise it will create risks to data availability in and of itself.

RecoverGuard™

RecoverGuard™ is a comprehensive DR monitoring and analysis solution which automatically and periodically scans the IT infrastructure to detect DR risks and vulnerabilities. RecoverGuard scans the storage, servers, databases, clusters and replication infrastructure while using a constantly refreshed knowledge base containing hundreds of DR vulnerability signatures. RecoverGuard ensures that critical data is always protected and gaps are closed. As an agent-less solution, RecoverGuard is completely unobtrusive and risk-free.

RecoverGuard provides the following processes:

Document - comprehensive and regularly updated documentation of all IT resources and related dependencies in the production and DR sites. Includes sophisticated visualization and reporting tools.

Detect – comprehensive detection of DR gaps and vulnerabilities based on unique, extensive knowledge base of DR vulnerability signatures.

Optimize – optimization of IT DR resources to improve efficiencies and utilization of assets.

How does RecoverGuard answer the DR challenge?

Comprehensive yet rapid scanning – of thousands of components, including storage, replication, operating systems and databases. Automated scanning of the production and DR infrastructure provides complete, up-to-date documentation; eliminating the need for laborious, error-prone manual processes.

Understands and maps dependencies – RecoverGuard analyzes what data is being replicated and how it is accessed and referenced. It tracks the data flow to establish the relations between the production applications, their various data copies, and the hosts that can access these copies.

Uses the most comprehensive knowledgebase – to discover gaps, RecoverGuard leverages a comprehensive knowledgebase which contains over a thousand vulnerability signatures, representing the accumulated experience and expertise of numerous organizations. Vulnerability signatures are linked to the comprehensive mapping of all production and DR infrastructure elements and their corresponding data flow. Continuity Software research labs are in constant communication with our customers, industry experts and vendors, to make sure each new threat signature is identified and incorporated into the RecoverGuard knowledgebase. .

Unobtrusive – RecoverGuard is an agent-less solution, making it completely unobtrusive and risk-free.

A Closer Look at RecoverGuard™

RecoverGuard includes several components that allow it to provide unobtrusive checks across the entire IT infrastructure.

Agent-less Data Collection and Scanning

RecoverGuard uses agent-less discovery technology to build topology views and collect data about the IT infrastructure. This means that the solution can be deployed immediately "out-of-the-box", without installing agents on the various network components. The data collection combines ***read-only*** inputs from various sources through standard API's, in a completely secure manner, to create a comprehensive view of the production and the DR environments:

- **Storage Resource Management Software –** such as EMC ECC and HDS HighCommand.
- **Servers and Operating Systems –** supporting Solaris, AIX, HP/UX, Linux, and Window servers through standard protocols, such as SSH, WMI, and more.
- **Storage Devices –** connection through native API as well as standard protocols such as SMI-S.
- **Databases –** connection through standard ODBC linksto collect configuration information.

The data collection includes infrastructure at remote locations covering primary and secondary sites. The collected data includes OS and application information, data layout (SAN, DAS, NAS) and more.

The discovery and scanning process is scheduled periodically to rescan the IT infrastructure and the current configuration to quickly detect any potential configuration changes before they impact data protection and availability. This allows RecoverGuard to detect and assess changes over time for ongoing data protection and monitoring.

DR Dependency Mapping

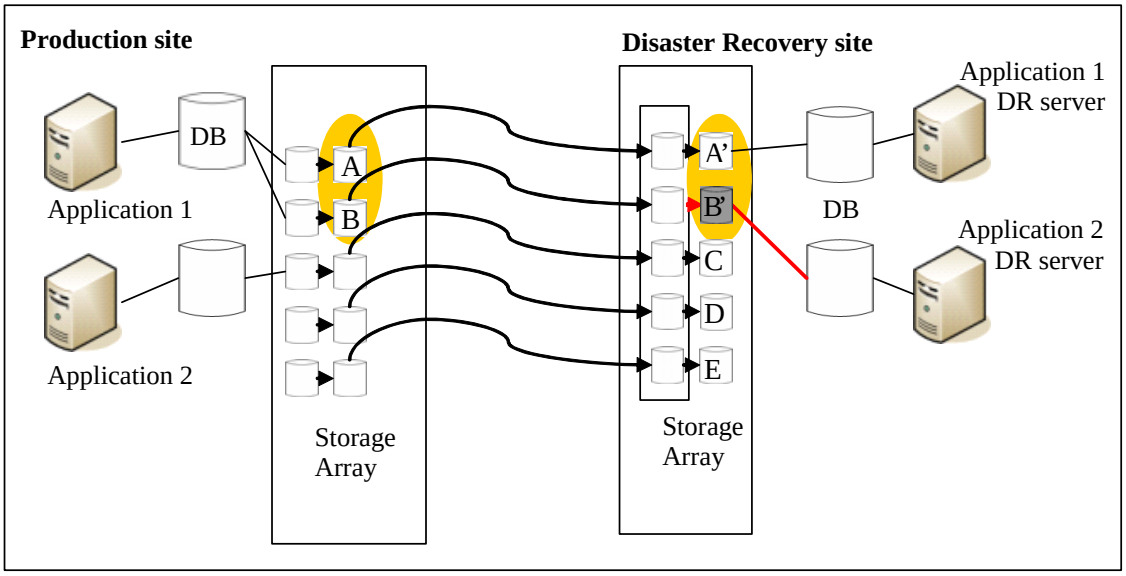
RecoverGuard maps the dependencies between the objects at the primary and secondary sites and reflects the following relationships:

- **Correlation between resources and their use –** RecoverGuard links and groups storage resources according to their usage by databases, file systems and applications. For the first time the connection between applications, databases and storage volumes is understood and documented.
- **Mapping of data set replications** – RecoverGuard follows and analyzes the replication of data sets. and creates a comprehensive layout of how replicated data maps back to database and applications within the DR infrastructure.

- **Correlation between the replicated data and DR resources**– finally, RecoverGuard correlates between replica sets and the DR resources and applications to understand the dependencies and interrelations between them.

In order to effectively monitor and manage the DR environment, it is essential to understand the various data relationships and interdependencies inherent across applications, databases and storage resources. The scanning/collection process combined with the dependency mapping process provides a clear understanding of all production and DR resources and how they are used by various applications.. The RecoverGuard data collection and dependency mapping process enables vastly improved monitoring of the environment. Armed with this information, IT personnel can manage changes proactively to ensure the continuous functioning of DR assets. For the first time, IT organization's have the ability to produce an entire dataflow document. This is a vital tool for discovering DR gaps and resolving production problems prior to DR tests or real-life disaster scenarios.

For example, consider the following situation:



In this simplified configuration, application server 1 is running a database stored on 2 SAN volumes - A and B. RecoverGuard identifies these 2 volumes as constituting a data set used by Application 1. It then follows the replication path to determine that volumes A' and B' match the replica set. As seen above, it is evident that volume B' is incorrectly mapped to the wrong server, in this case, Application 2 DR server.

If this issue remains unresolved until a disaster, Application 1 DR server will not start because it does not see the entire replica set. As a result, the database will be unable mount. After some struggles and after realizing the nature of the problem, Once the problem is identified, the system and storage administrator need to determine which of the disks are missing -- B', C, D or E . Without up-to-date documentation, it is almost impossible to find the right answer. Considering that in a real-life scenario the data set may contain dozens of disks and that a typical "B Volume" may be hiding among thousands of other drives, the magnitude of the problem becomes clear.

Gap Analysis

RecoverGuard uses a knowledgebase of hundreds of checks to identify vulnerabilities. These checks, or gap signatures, are much like antivirus signatures, only instead of identifying viruses, they identify DR gaps or vulnerabilities and suggest actions to be taken to remedy the problem. In order to find gaps, the signatures are applied on the system documentation generated by RecoverGuard. For example, a signature can look for databases, whose data files are stored in multiple SAN volumes which are replicated to remote locations; and that the *data age* of one of the replica volumes is different from that of the others. This means that the set of volumes does not represent a time-consistent copy and is therefore, most likely, corrupt.

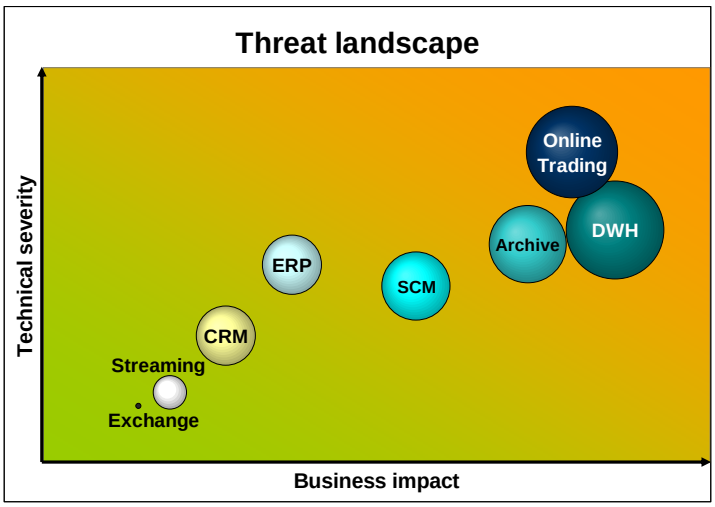
RecoverGuard's extensive knowledgebase covers the following areas:

- **Data Completeness –** checks the completeness of the data in the DR environment compared to the production environment. A gap in this category means that the data is missing from the DR environment.
- **Data Consistency -** checks that data is consistent and usable.
- **Storage Configuration** – checks for incorrectly aligned storage. For example, inappropriate association of storage volumes to device consistency groups.
- **SLA Breaches –** checks for deviations from recovery SLA policies and requirements.
- **Data Accessibility and Data Path Problems –** checks for incorrect mapping of data or applications.
- **Data tampering –** checks for inappropriate modification of data copies in the DR site.
- **Incorrect process flow** – checks for incorrect sequences of DR processes (for example, ending database backup mode before a data copy, or split, was completed)
- **Redundancy faults** – identifying areas in which redundancy has fallen below a satisfactory level (for example, too few fabric paths exist, volume groups containing a combination of protected and un-protected storage volumes, faulty or misconfigured cluster members, etc.)
- **General Faults –** checks for general faults such as deviations from best practices, vendor specific configuration issues, etc.
-

Business service alignment

RecoverGuard intelligently models customer business services, so that business impact detected risks could be easily understood. Each identified gap is assigned with a technical risk, or severity indication. By correlating the gap to the business service and the role of the involved servers, databases and applications, the business

risk level could be computed. This allows users to prioritize risks, as well as realize the exposure level of each business service.

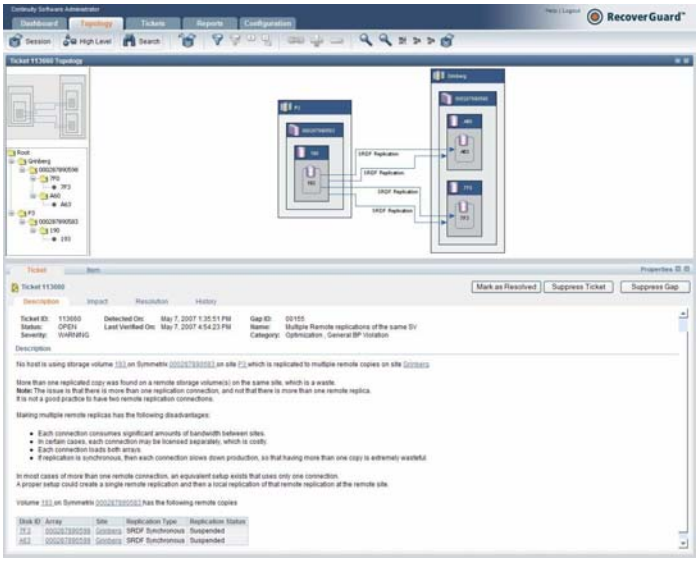


Presentation and Reporting

RecoverGuard provides various ways to intuitively visualize DR configurations and offers a multitude of comprehensive textual reports. Outputs can be used by IT specialists as well as top executives. The reports provide comprehensive insights regarding the validity of the DR environment and its efficiency status.

Below are some examples of the system outputs:

- **Gap Reports –** these reports provide in-depth information about gaps that have been identified and suggested corrective measures. Gaps can be grouped by business service, technology, impact, risk level, etc.



- **SLA and Timeline Reports** – provide powerful tools to investigate data flow and assess compliance with Recovery Point Objectives (RPO) for each business service, host and application.
- **Dashboards** – Provide qualitative and quantitative analysis of current threat levels and recent system findings. Intelligent aggregated views allow to easily identify "hot-spots" requiring immediate attention. Various views provide value to both IT professionals and general management.
- **Trend analysis**- provide views into processes, resource configuration change and DR risks
- **Optimization Reports** –provide a rich set of views revealing potential inefficiencies in storage and SAN utilization (for example, excessive SLAs, configured but unused resources, out-dated but not reclaimed replicas, and many others).

Improved Test Planning with RecoverGuard™

RecoverGuard diminishes the efforts and stress involved in testing the DR environment. Stress levels go down because with RecoverGuard, IT personnel will have much higher confidence levels that DR tests will succeed. Effort is reduced since RecoverGuard automates most of the required manual information-gathering and check-list execution. RecoverGuard also helps focus the testing on areas of interest, where most configuration changes or DR gaps were identified. In this way, instead of using rigid and outdated test schedules, DR administrators can focus their effort where it will provide most value.

The test themselves are conducted in an environment that is nearly free of gaps. This means tests will be quicker with almost no surprises and lowers risks involved with the test itself. As a result, the threat of unexpected downtime is greatly reduced and post-test corrections are eliminated. In the unlikely event of testing problems, RecoverGuard's comprehensive documentation proves invaluable in rapidly solving configuration issues in a controlled and intelligent fashion.

Optimizing the DR Environment

RecoverGuard leverages the information gathered from the IT infrastructure and the Disaster Recovery configurations to detect opportunities for improving efficiencies and lowering the total cost of DR ownership.. Such improvement opportunities can include:

- **Utilization of storage elements –** finding storage elements that are not being utilized and therefore could be repurposed for other applications. Harnessing the full power of RecoverGuard's correlation capabilities, the analysis intelligently ignores recently added storage, resources mapped to cluster standbys, etc.

- **Mis-configured replications –** finding replication jobs that are not configured according to the required SLA, utilizing too much bandwidth or are not supposed to be replicated at all. For example, synchronous replication for paging and other temporary files.
- **Excessive SLAs –** finding places where protection provided by the infrastructure exceeds the target SLA. For example, finding EMC volumes with more BCVs than required by the SLA.
- **Unused Components** – finding very old copies or ones that are no longer in use.

Summary

Organizations invest substantial resources on DR solutions to mitigate the risks of unplanned downtime. Despite best efforts at change control management, the combination of complexity, scale, and frequency of changes across production data center environments, inevitably creates gaps that prevent the DR infrastructure from properly operating in a time of need. RecoverGuard™ prevents gaps from impacting DR readiness by leveraging a comprehensive knowledgebase with thousands of vulnerability signatures to automatically and unobtrusively detect potential problems. RecoverGuard dramatically enhances data protection, while optimizing the DR environment by identifying inefficiencies..

About Continuity Software

Continuity Software is a leading provider of Disaster Recovery Management solutions. Its RecoverGuard™ software mitigates risk by monitoring your production and remote replication environments to detect data protection threats, vulnerabilities and gaps. With RecoverGuard you will be confident your data is protected and you will exceed your business continuity goals. For more information, please visit www.ContinuitySoftware.com