COMPAQ
Solutions Guide

April 2002
16DK-0302A-WWEN

Prepared by Enterprise Storage Group

Compaq Computer Corporation

Contents

The Compaq Disaster Tolerant SAN Solution

StorageWorks business continuity solution for Windows NT, Windows 2000 and Compaq Tru64 UNIX

Abstract

This Solution Guide describes the Compaq Disaster Tolerant SAN Solution (DT-SAN) for Windows NT / 2000 and Compaq Tru64 UNIX, a turnkey solution for the delivery of highly reliable, multisite SAN based infrastructures.

This guide is intended for use by system designers and administrators. In this guide, administrators learn about the disciplines necessary to build a DT-SAN Solution, and how the DT-SAN delivery may be achieved in their environment.

Using the DT-SAN Solution described in this guide, administrators can achieve safe, repeatable, and rapid recovery from significant losses to their SAN infrastructure, without risk to data integrity.

This guide describes the Services, Compaq StorageWorks and SANworks products, and third party products used in the solution.

The goal of this guide is to provide system designers and administrators with best practices as well as the service guidelines to implement the Compaq DT-SAN Solution.

Let us know what you think about the technical information in this document. Your feedback is valuable and will help us structure future communications. Please send your comments to: BusinessContinuityStorageSolutions@compaq.com

Notice

Disaster Tolerant SAN Service Solutions Guide prepared by Enterprise Storage Group

1 Management Summary

In the modern business world, the use of Information Technology (IT) systems to support business-critical activities is increasingly prevalent. IT is not only the process engine of a business – it can also be the supplier interface, partner interface, and customer interface. With this level of importance, the survivability of the IT environment has become increasingly significant. With the increase in globalization and telecommuting, disasters, such as complete building loss, are no longer acceptable reasons for system unavailability or loss of data. Businesses are demanding that their IT solutions provide true "Business Continuity" or "Disaster Tolerance".

The challenge of these demands is that evidence shows that 'disaster tolerant' solutions, constructed simply of two computer centers with 'mirrored' or 'vaulted' data, rarely deliver full disaster tolerance. The reason is that complexity must be introduced into the environment to achieve the basic "disaster tolerance" goals. Multisite data replication and processing can only be achieved by increased component count and heavy reliance on intersite communications, all of which require additional hardware and software.

Experience also reveals that true 'disaster tolerant' solutions cannot be implemented in isolation, but require a partnership approach between the delivery organization and the client. In particular, knowledge transfer between the implementers of the solution and the operations team can be shown to be critical, since lack of communication frequently leads to failure.

The Compaq approach to this problem is the Disaster Tolerant Storage Area Network solution (DT-SAN), producing a system that meets all of the essential criteria and can be properly managed to deliver a true disaster tolerant application platform.

The DT-SAN Solution includes carefully designed consulting services combined with state-of-the-art management tools that form a model for the effective construction and management of a true disaster tolerant solution. The underlying components are customized to reflect the unique operating characteristics and constraints of each specific environment and the specific needs of disaster tolerance. The Compaq DT-SAN Solution is constructed with minimal impact to a client's business, resulting in a highly resilient and disaster tolerant platform for applications. *The primary benefit of the DT-SAN Solution is to provide a rapidly deployable, risk-free implementation of disaster tolerance using Compaq SAN technology as the primary delivery vehicle. The final solution is manageable and supportable through simple operator tools and utilities.*

The core of the solution is built from off-the-shelf technologies. *StorageW*orks FibreChannel components, and *SANworks* software, together with properly configured off-the-shelf software for the required operating system, provide the base required to build the DT-SAN Solution. The storage area network products allow construction of a system that provides mirrored data, which is split between two data centers. This storage infrastructure is combined with a correctly configured operating system and highly configured management tools from Heroix Corporation. The result of this combination is a system, which offers leading-edge high-

availability, tremendous monitoring and manageability features – all of which are essential to a true disaster tolerant solution.

The design supports a wide range of Windows NT, Windows 2000 and Compaq Tru64 UNIX configurations. In Windows environments clustered designs (specifically the Microsoft Cluster server) and single server designs, employing servers starting from the SAN, are fully supported. However, a design that incorporates a cluster technology provides improved automatic recovery from hardware failures, while a single server design always requires operator intervention. In the Tru64 environment servers bootstrapping from the SAN are utilized.

The complete DT-SAN Solution is a modular design, both in terms of the hardware elements deployed and in the consultative steps that must be followed in order to achieve a successful result. Compaq can provide all of the services required to fulfill each of the consultancy modules, or clients can elect to use local resources for certain tasks. The aim of each project is to achieve a successful delivery, with the minimum disruption to the existing environment, and utilizing the most appropriate resources.

While the service delivery elements are modular, the software and software customization components are not. The core -SAN Solution includes the delivery and local installation of all software components, and it is not possible to omit any items.

2 Solution Overview

This guide describes the DT-SAN Solution, and also explains how the solution can be delivered into an existing SAN environment.

This guide includes the following solution specific information:

- Discussions of the issues and disciplines raised by a Disaster Tolerant design (the current section, Chapter 2);
- Discussions of the core technologies necessary for the deployment of the DT-SAN Solution(Chapter 3)
- A description of the components of work required to realize the solution (Chapter 4)
- A description of the manner in which the management tools operate (Chapter 5)
- A description of the approach used for a typical project (Chapter 6).

2.1 Disaster Tolerant SAN Approach

The principle behind a disaster tolerant solution is to produce a system where a failure or loss of a complete computer room (and consequent loss of access to all of the systems) does not result in a loss of IT services. Some customer requirements are that even a few minutes of down-time are intolerable, irrespective of the severity of the problem.

Achieving the goal of disaster tolerance proves to be much more difficult than it at first appears. **A common belief is that disaster tolerance can be provided by deploying properly redundant hardware and software**. You can duplicate hardware in a second location with mirroring or shadowing of data between locations, or you can divide systems that already contain redundant hardware. Unfortunately, experience shows that achieving proper disaster tolerance with such a simple approach is almost impossible. While such hardware duplication is an essential starting point for a disaster tolerant solution, it does not provide all of the elements needed for a successful solution.

2.1.1 Understanding "Disasters"

Generally, a "disaster" is assumed to mean a complete and permanent loss of a computer room or facility.

However, the entire permanent loss of a computer room is extremely rare. An interruption of site communications that causes loss of service, or a temporary problem with a computer room, such as power loss and air-conditioning failure, is more likely to happen. Such ***recoverable*** failures pose more challenging problems – as the likelihood is that the fault will be rectified, eventually allowing normal service to resume.

Another common misconception is that when an incident of this magnitude strikes, the effects are always instant, immediate, and catastrophic. Events, such as flood or fire, are far more likely to disable equipment in a cumulative manner, but the order of equipment failure is not likely to be predictable. Even in a situation such as a power

loss, it is highly likely that some items of equipment will continue to function for fractions of a second longer than others – and this may be long enough to cause unpredictable I/O completion, especially if asynchronous intersite connection is in use.

If the only issues of importance are those concerning complete site loss, the problem of building a disaster tolerant system would be easy. Site loss would be quickly followed by a reconfiguration of the surviving site to acknowledge the loss of the other, and applications would be restarted. No other action would be required.

However, with short-term failures, there are a range of considerations, for example:

- How are system databases reconsolidated?
- How do operators ensure that both sites do not assume "live" operational status simultaneously?
- How do operators guarantee that the optimum site is used as the "master" for any mirror copy operations after recovery?

Any of these problems ***will*** cause data loss if not handled correctly, and may even cause data corruption.

2.1.2 What Is Required of a Disaster Tolerant System?

To ensure that recoverable failures are handled as well as permanent loss failures, the system needs to satisfy the following criteria:

1) Continue offering all business critical IT services in the event of a complete failure or loss of a computer room
2) Capable of ***selectively*** continuing in only one computer room, in the event of a loss of intersite connections
3) Minimum or no loss of data in the event of failures
4) Prevent any risk of data corruption
5) Prevent issues related to the sequence of device failures
6) Provide recovery in a timely manner . For example, the time required to restore the IT services from the second computer room must be short, measurable and repeatable.
7) Reassemble the system to its full disaster tolerant specification without user disruption or damaging data.
8) Capability to backup data to tape media for offsite storage.

This list demonstrates the major issues to consider when designing a system that can provide high levels of service. The key requirements to designing a system that can reliably survive disasters follow:

- Can data integrity be guaranteed?
- Does the operator know the status of components and the system?
- Does the operator have manageability from either site?
- Can the operator recover quickly from any fault?
- Can the operator restore quickly and without disruption to users?
- Can the operator ensure user connectivity?
- Does the operator have good intersite coordination?
- Can the business avoid personnel issues?

2.1.3 Recovery

Recovery from problems requires careful consideration, especially since speed of recovery is usually a consideration.

To understand the stringent needs of the recovery processes in a disaster tolerant design, it is useful to consider the process as a three-stage event.

1. Diagnosis – do you have the relevant information to diagnose the problem correctly, accurately and quickly?
2. Recovery decision – does the diagnosis process give you enough information to allow you to decide what to do quickly and accurately?
3. Recovery implementation – is the recovery process as automated as possible to minimize the risks of recovery?

If you can diagnose the problem quickly and accurately, you can obviously speed the recovery process by presenting the information to decision makers more rapidly. Typically, the ***majority*** of the time required to recover a system is diagnosing the problem. Key issues include:

- How many places do you have to look?
- How much work is required to get ALL of the information that you need?

A high quality diagnosis benefits the decision making process by ensuring that all options are available, and allows a balanced view of the impact of each recovery option.

Finally, do not underestimate the importance of the implementation phase. A high level of automation is critical – as mistakes during system recovery makes the impact of the initial problem worse. ***In most disaster situations, such as the two site system, mistakes during system recovery cause data loss more often than the original event.***

2.2 How to Apply the Principles of the DT-SAN Solution

This section describes the key requirements that must be satisfied by a system to achieve full disaster tolerance.

2.2.1.1 Proven high availability technology

The core technology of the solution must be robust and reliable and provide the key capabilities necessary for business continuity, including the following:

- Mirroring or shadowing technology that provides complete copies of all live data, standby, and recover sites. The mirroring technology must ensure that all updates to the data are fully committed to both copies of data;
- Ability to host processing elements in both sites, such that all processing units can access the data and host the required services;
- Ability to build solutions with "no-single-point-of-failure";
- Capability to "switch" or "fail-over" operation from one site to another.

2.2.1.2 Proper planning and design

The chosen technology must be implemented correctly and with full awareness of the requirements of disaster tolerance. Single-points-of-failure in the design must be eliminated if at all possible. Infrastructure considerations must also be taken into account, such as diverse routing of intersite network connections.

The architecture should also allow for backup to tape media for offsite storage, where data can be restored from tape media when required.

2.2.1.3 Effective system management and operation

Effective monitoring and management is essential if the disaster tolerant system is to achieve its objectives. The critical features of the management tools include:

- Tools must be highly available. They must let the administrator manage the solution remotely and from multiple locations.
- Management solution must provide a fully integrated toolset for effective single point diagnosis of problems and issues. This should involve a comprehensive event collection, processing and identification capability, to maximize the speed of event detection and determination.
- Management solution must be closely matched to the hardware, firmware, and operating system releases used in the solution, to ensure the accuracy of the diagnosis capabilities.

- Management solution should present monitoring and control facilities for all aspects of the disaster tolerant environment from a single task-oriented set of tools, which includes the monitoring and control of all system, storage, network, and infrastructure components within the environment;
- Tools should be carefully tuned and implemented to ensure that the minimum of extraneous information is presented to the operations staff. When the tools are appropriate, only relevant data is presented. These tools also prevent the staff from missing critical events that may otherwise be "masked" by irrelevant data.

2.2.2 Putting The Pieces Together

The Compaq DT-SAN Solution is service engineered by the Compaq Global Services Worldwide Expertise Center for Disaster Tolerance.

The solution is based around a modular consulting package that can be customized to address a wide variety of environments, delivering a highly tuned service to the client. The modules that are included in the package address all of the needs of building disaster tolerance. The modularity of the design allows local staff or resources to be used where appropriate to ensure the best delivery solution is available for every customer situation.

The DT-SAN package also includes a comprehensive management solution that integrates several system management products on Windows NT or Windows 2000 management stations to provide a unique capability for managing a two data center solution. The DT-SAN Solution consists of a number of management stations, monitoring and controlling the application server systems. These systems may be members of a cluster or single server application engines.

The service elements of the DT-SAN Solution are described in Chapter 4, while the management solution is described in Chapter 5.

3 Considerations for Disaster Tolerance

3.1 Elements Of A Successful Solution

A good disaster tolerant solution begins with a basic computing environment that provides the required attributes. Since the specific requirements vary from solution to solution, it is difficult to provide generalizations. However, there are some features that each environment requires. The solution must have the following:

- Two locations that will be used to host identical copies of all mission critical data, and to accommodate any server capacity to provide the associated applications to users;
- A high availability inter-site network of adequate performance;
- Robust, high availability technology in each of the two locations to provide processing capability in each of the locations;
- Ability to provide real-time replication of data between two locations. The replication must satisfy the business goals for "disaster tolerance":(1) it must provide the degree of synchronization demanded by the business, and (2) it must have the ability to achieve the recovery times demanded by the business in the event of a failure;
- Alert or paging mechanism that the DT-SAN monitoring tools can use;
- Ability to backup data to tape media. Since the ability to restore data quickly and reliably is critical, the backup solution should be selected after careful capacity sizing including provision for growth. Also the time to backup and restore critical data for continued business operation needs to be considered for sizing performance requirements of the backup solution.

3.2 The Disaster Tolerant Locations

The choice of site for the Compaq DT-SAN Solution is complex, fundamentally driven by business decisions. Although most business managers want the sites "as far apart as possible", technical limitations and other compromises are considerations.

A single solution cannot satisfy all requirements. Table 1 shows a list of typical business objectives that can be achieved by various computer room geographic separations.

Data Center Locations	Disaster tolerance protection	Typical business utilization

Data Center Locations	Disaster tolerance protection	Typical business utilization
Data centers in same building	Computer room problems	Business that depends on computer systems but only if access to building is still possible.
Data centers on same campus	Computer room problems, Localized environmental problem, such as burst pipes, Power loss to one building; Limited building fires	Business that depends on computer systems only if campus is still accessible.
Data centers between 1 and 3km apart	Computer room problems, Localized flooding, Denial of access due to bomb / fire, Localized power disruption, Loss of building.	Business with moderately high DT requirement, situated in an area not susceptible to major disaster or geological disturbance.
Data centers between 3 and 50 km apart	Computer room problems, Localized flooding, Denial of access due to bomb / fire, Regional power disruption, Large disaster affecting moderate area.	Business that must maintain computing capability at almost any cost situated in an area not prone to geological problems and major flooding or meteorological events.
Data centers >50km apart	Computer room problems ,Extensive flooding, Denial of access due to bomb / fire, Regional power disruption, Major disaster affecting wide area, Geological activity; Meteorological disasters (hurricanes ,).	Business that must maintain computing capability at all costs.
Data Vault location	Data corruption, loss of data storage devices, virus & hacker attacks	Data can be restored from tape media stored at offsite data vault location. A data vault location separate from the secondary site adds an additional level of disaster tolerance and protects data from virus and hacker attacks

Table 1 – Datacenter Location Considerations

3.3 Network Considerations

A DT-SAN Solution must incorporate an intersite communications subsystem into the solution. The intersite connection is a primary system component, design considerations that fall into three categories:

- Types of links

  The types of link required are determined by the basic solution design. However, a number of generalizations can be made:

  - DRM requires two intersite connections, of equal performance;
  - A server deployment requires two extended LAN type connections, each having appropriate performance for the applications concerned. A separate, dedicated link is recommended, although the primary LANs can be used if required

  A split-site Microsoft Cluster Server based solution performs best with a private connection for the cluster communications path.
  The IP network configuration needs to be condsidered. For example, it may be necessary to configure the IP network to provide the same subnets in the server environment.

  In terms of user connections, carefully consider the provision of external user access balanced between the data centers.

- Resilience

  The design must incorporate sufficient network paths to accommodate all required availability scenarios. A single set of intersite connections is rarely sufficient. However, there are two connections, can the telecommunications provider guarantee sufficient separation of routes to make it worthwhile? A second connection running through the same trunking is unlikely to provide additional resilience over the single connection.

- Performance

  The bandwidth and latency of the intersite connections, especially in connection with the DRM connections, are critical considerations. The intersite technology selected and the distances between sites largely dictate these issues.

  The performance required by the solution must be carefully considered for each application in the environment. It is important to understand the difference between the effects of latency, which is largely influenced by the distance between sites, and bandwidth, which is largely determined by the intersite connection technology used.

  Latency directly affects application performance where single stream write I/Os are a significant issue. An application might need to perform frequent writes directly on behalf of a user. Applications, where I/O writes are buffered by the application, may be less affected by latency issues.

  Bandwidth affects any application moving large amounts of data around the system. One specific area where bandwidth has a significant impact is during resynchronization of the data sets between data centers.

When designing an infrastructure, the performance target should not be the "average state" performance. There is a serious risk that peak demand performance will not only be poor – it may cause the system to fail.

The determination of maximum distance and desired technology depends on the applications to be deployed and the manner in which they are used. The challenge is to balance the distance required by the business with the capabilities of the technology to support the applications in an acceptable manner:

- As distance increases, latency increases and the usable bandwidth of high speed links decreases.
- These characteristics affect synchronous replication significantly more than asynchronous techniques.
- As distance increases, application performance on a synchronous replication system degrades.

As distance increases, the potential for the link to be interrupted increases, because of the greater physical presence of the longer link. Therefore with longer links, the probability of having to periodically resynchronize the two sites increases.

3.4 High availability technology

3.4.1 The Data Storage Subsystem

3.4.1.1 Required characteristics

The DT-SAN Solution must be based on a storage subsystem that provides a robust, high availability storage design, with multi-path support and no single-point-of-failure characteristics.

The storage subsystem needs to support separation between sites, by the use of networking technology as described above. It must also provide replication technology to allow for simultaneous copies of data to be maintained at the two processing sites.

3.4.1.2 The Compaq solution

The Compaq solution uses the capabilities of the StorageWorks hardware product set, complemented by the SANworks software solutions. The specific technology deployed will vary between solutions, but the following elements are the key ones.

3.4.1.2.1 Fiber Channel Storage Switches

The Compaq Fiber Channel Storage Switches provide critical features in two key areas.

Firstly, they have comprehensive monitoring and management capabilities – allowing full access from an SNMP management engine. This permits full status and health-checking of the devices. The same interface also allows full control

of the switch and all of its ports. This is essential for recovery processes that may require isolation of some elements of the subsystem.

The second key feature is the ability of the switch to support long-distance interconnects, through a variety of built-in elements (for example, long range Fiber-Optic transceivers) or via support for third party interconnects, such as Fiber-Channel over IP devices.

3.4.1.2.2 HSG80 Controllers

The Compaq HSG80 controllers are industry-leading solutions for the connectivity of storage to server systems. These controllers provide a huge range of features and functions for supporting large, scalable storage subsystems.

The HSG80 provides a full bandwidth Fiber-Channel connection to the host subsystems and six Fast-Wide SCSI connections to storage devices. With sophisticated caching techniques and up to 512 MB of controller cache, the HSG80 controller can provide huge bandwidth to client systems.

The HSG80 controller has a large number of high availability features. It supports full dual-device operation –it can be paired with a second HSG80 in which case each device will mirror the others' operation. Under normal circumstances, both devices will provide processing capability for the I/O workload, but in the event of a failure of either device, the other will assume the operations of the other. Due to the fact that all operations are logged in the shared cache subsystem, even operations that were in progress at the time of failure can be completed.

The HSG80 controller is capable of supporting private battery back-up, to ensure that cache data is never lost, even in failure conditions.

The controller has comprehensive error reporting and management capability. These features can either be accessed in-band (i.e. over the Fiber-Channel fabric) or out-of-band via a standard RS423 serial port.

In the DT-SAN design both mechanisms are used for monitoring, and the RS423 port is used for control functions.

3.4.1.2.3 *SANworks* DRM

Data Replication Manager (DRM) is a storage-based data replication and workload migration solution. This state-of-the-art replication software runs on the HSG80 storage controllers and provides the functions required to maintain copies of the in two separate locations. DRM allows a host to write data to a virtual volume, hiding the fact that a copy of the data will be written to a second location. Copies can be performed fully synchronously, assuring absolute intact copies of data, and fully multi-volume locked, providing the highest guarantees of data integrity.

The transparency of DRM means that any supported operating system can utilize the replication features without modification, and any application running on the operating platform can exploit the capability, also without modification.

The DRM is an ideal solution for copying data online and in real time to remote locations via a local or an extended Storage Area Network (SAN). Using the

DRM software, data replication is performed at the storage system level and in the background to any host activity.

DRM provides the DT-SAN with the following features:

- Increased connectivity and scalability with the ability to connect more servers, storage systems and cascaded switches;
- Flexibility and ease of management through use of the cloning and snapshot capability
- Enhanced configuration flexibility with zoning and support for non-Remote Copy Sets
- Full support for data replication over longer distances via the Very Long Distance GBIC (VLDG) or other long-distance connections.

3.4.2 Servers And Operating Systems

3.4.2.1 Required Characteristics

In any disaster tolerant solution, the performance and behavior of the operating system are critical for the solution to perform as expected. Specifically, operating system behavior is key to providing transparency and manageability of recoveries.

The operating system must provide robust, predictable recovery from failures. In particular, it must be able to be able to recover its file system reliably in the event of sudden interruption of operation. The operating system must have extensive basic redundancy features and comprehensive failure reporting facilities.

Outside of these basic requirements, the Compaq DT-SAN Solution has been designed to be as undemanding as possible of the operating systems, to ensure that the widest range of solutions and operating environments can be covered.

3.4.2.2 Compaq Solutions

3.4.2.3 Windows NT and Windows 2000

The Windows NT and Windows 2000 operating systems provide the basic requirements required for the DT-SAN Solution.

They provide full checking of file systems on re-boot to ensure that incomplete file system operations can be corrected on system reboot. They provide extensive error reporting facilities that can be enhanced by the DT-SAN monitoring solution. Basic redundancy is provided by the support for multiple CPUs ECC memory support and full support for RAID storage elements.

In the DT-SAN Solution, these facilities are supplemented by the installation of the Compaq *SANworks* Securepath software that provides a complete multi-path solution to storage. This allows full no-single-point-of-failure solutions to be configured in the storage subsystem, by allowing the operating system the ability to access an individual disk drive through multiple, separate paths.

Windows 2000 and Windows NT allow a number of solutions to be configured in a DT-SAN environment, and the following sections describe them.

3.4.2.3.1 Stand-Alone Servers

The stand-alone server option implies that each Windows NT or Windows 2000

server is connected independently to the SAN, with a private bootstrap disk and a private identity. This is the simplest solution from a configuration point of view.

However, there are a number of issues around Disaster Tolerance with this approach – so while it is simple to implement, this approach makes recovery difficult.

If the storage is switched to the recovery location, how will the activated on any "recovery" systems in order to exploit the fact that the data is still available?

The problem is that identical servers are required in the second site, with the applications already loaded and awaiting the data. Even then, with many applications, there will be information in the Windows registry that must be maintained – and this may be dynamic, adding to the problem.

To help to manage this situation, the DT-SAN is implemented so that all stand-alone servers bootstrap from SAN based system disks that are themselves DRM based volumes. This approach addresses all of the problems – the primary site server will be running and maintaining the system disk, including the registry, and when fail-over is required, a system in the second site simply bootstraps from the same DRM volume, picking up all system identity and registry settings.

The only drawback of this approach is that it requires the live and recovery servers to be of identical type and configuration, otherwise problems will occur during bootstrapping the systems.

3.4.2.3.2 Local Windows Clusters

As an alternative to a stand-alone server approach, a design can be constructed with a local cluster (i.e. a cluster with both systems in the same site). This shares all of the Disaster Tolerant characteristics of the stand-alone server design, but it does offer some advantages in providing high availability.

This design will perform just like a stand-alone server in the event of site disruption or loss, however, in the event of a server failure or user network failure, the clustering technology will move the processing workload to the other server in the cluster. This means that server failure no longer has the impact that it would in a stand-alone design.

For the purposes of the DT-SAN, support is limited to the use of Microsoft Cluster Server as the clustering technology.

3.4.2.3.3 Split-site Windows Clustered Systems

Over modest distances (up to 25 km) the Windows environment offers the option of building a clustered solution that spans the two sites. In this design, the Disaster Tolerant environment is comprised of 2 systems that are members of a cluster and each member system is situated in a separate physical location.

For the DT-SAN Solution, Microsoft Cluster Server has been fully tested as the clustering component, although there are other products available. Clustering offers some advantages in ensuring that the remote site is able to pick up the application processing workload when required. Specifically, a split-site cluster offers the following advantages:

- Microsoft Cluster Server will automatically relocate and restart application processing for a range of failure conditions, such as server failure or user

network failure. This means that some hardware and software problems will be recovered automatically;

- As the Microsoft Cluster Server software is maintaining the application awareness across the cluster, it will ensure that any changes that the application makes to registry keys will be maintained across both servers that are members of the cluster. This makes cluster management far easier;
- As both servers are running and connected to the storage and network at all times, it is immediately apparent if there is a problem with the second site – so you know whether the recovery site is "ready for action".

Table 2 presents a summary of the key features of each of these approaches.

	Split-Site Cluster	**Local Cluster**	**Stand-Alone Design**
Effect on Normal Operation	Application must be supported in a Microsoft Cluster environment.	Application must be supported in a Microsoft Cluster environment.	None
Implications for Disaster Tolerance	Always know the state of all component systems	Don't know if recovery system will bootstrap when required.	Don't know if recovery system will bootstrap when required.
Manageability	Registry and environment synchronization looked after by Microsoft cluster server.	Registry maintained between cluster members – must manually ensure that remote system registry and environment is correct	Must manually ensure that remote system registry and environment is correct
Automatic recovery after server failure	Yes – Automatic within a few minutes.	Yes – Automatic within a few minutes.	No – System Reboot Required
Automatic Continuation after primary site failure	~ 20 minutes. Reboot and storage reconfiguration required	~ 20 minutes. Reboot and storage reconfiguration required	~ 20 minutes. Reboot and storage reconfiguration required
Application interruption after recovery site failure	None	None	None
Recovery after cluster/site repair	10-20 minutes. Reboot and storage reconfiguration required	10-20 minutes. Reboot and storage reconfiguration required	10-20 minutes. Reboot and storage reconfiguration required

Table 2 - Comparison of Windows Clustering Options

3.4.2.4 Tru64 UNIX

The Tru64 UNIX operating system provides the basic requirements required for the DT-SAN Solution.

Its log-based file system, AdvFS (the Advanced File System) provides full automatic recovery of the file system on re-boot with minimal checking. It provides extensive error reporting facilities that can be enhanced by the DT-SAN monitoring solution. Basic redundancy is provided by the support for multiple CPUs, ECC memory support, support for RAID storage elements, and support for multiple network paths and adapters.

Full no-single-point-of-failure solutions can be configured at the storage subsystem level, as the operating system has native support for accessing an individual disk drive through multiple, separate paths.

3.4.2.4.1 Stand-Alone Servers

The stand-alone server option implies that each Tru64 server is connected independently to the SAN, with a private bootstrap disk and a private identity. This is the simplest solution to configure.

As with the Windows systems however, there are a number of issues around Disaster Tolerance with this approach – although there are unlikely to be the same number of issues due to the fact that Tru64 does not maintain a registry of the same style as Windows.

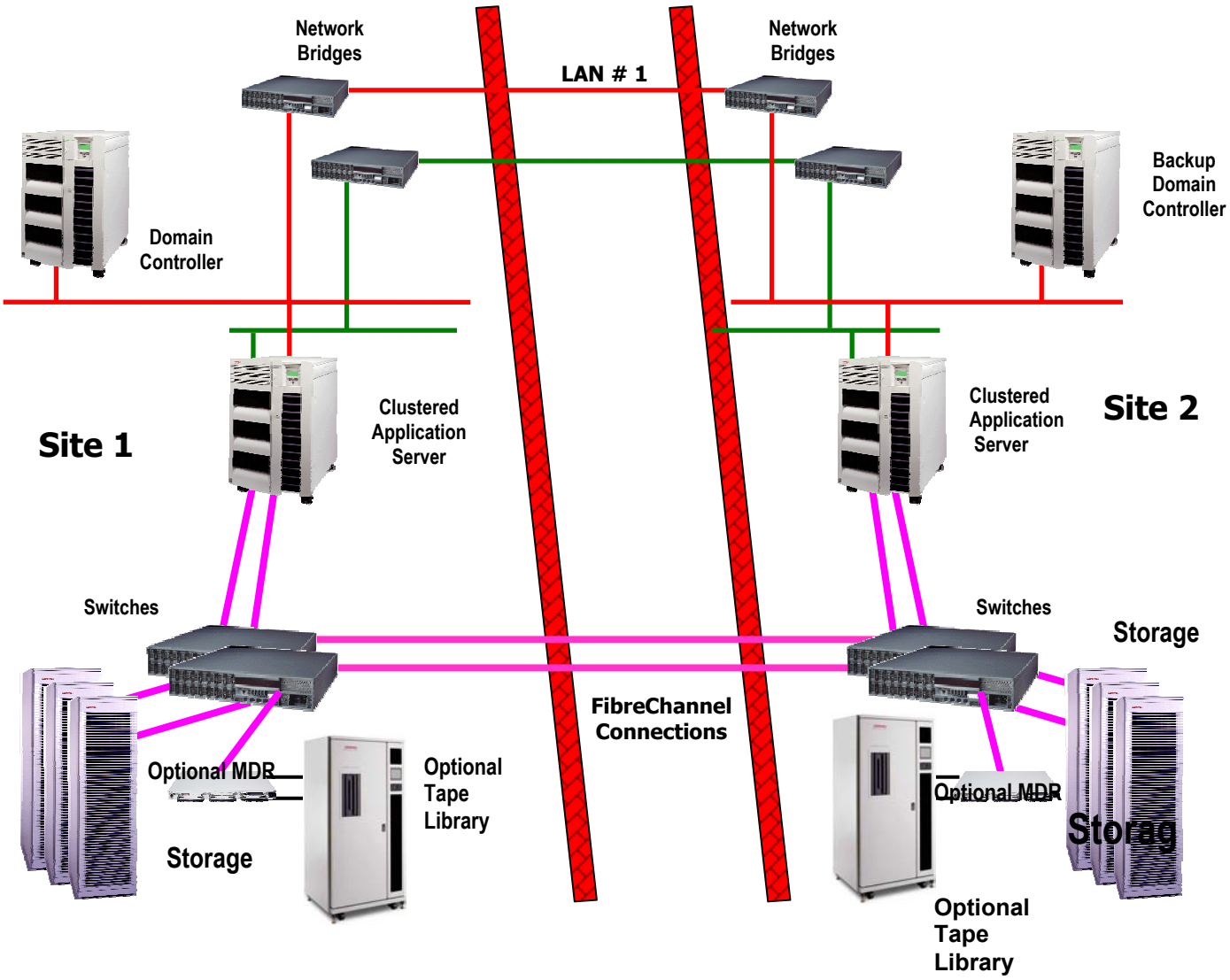
The simplest DT-SAN design is to follow the same approach as described above – that is, arrange for all stand-alone servers to bootstrap from SAN based system disks, that are themselves DRM based volumes. This approach addresses all of the problems – the primary site server will be running and maintaining the system disk, including the registry, and when fail-over is required, a system in the second site simply bootstraps from the same DRM volume. Note that unlike in the Windows case, there is some latitude here regarding server types, although they must be carefully configured to ensure that all devices present with the same device identities.

However, the DT-SAN design also supports truly independent Tru64 systems, bootstrapping from separate disks. This design requires careful installation of all application code to make sure that all required application context fails-over between sites in the event of a site loss.

3.4.2.4.2 Clustered Systems

As an alternative to a stand-alone server approach, a design can be constructed with a local cluster (i.e. a cluster with both systems in the same site). This shares all of the Disaster Tolerant characteristics of the stand-alone server design, but it does offer some advantages in providing high availability.

Figure 1 – Example of a basic SAN deployment for the DT-SAN Solution



This design will perform just like a stand-alone server in the event of site disruption or loss, however, in the event of a server failure or user network failure, the clustering technology will move the processing workload to the other server in the cluster. This means that server failure no longer has the impact that it would in a stand-alone design. Note that the recovery system can either be a separate "cluster" or a single system, depending upon the processing requirement at the recovery site.

3.5 Other Design Considerations

The core elements of the solution should be configured using the best practice principles described in other Compaq SANworks documentation (http://www.compaq.com/products/storageworks/san/entry/index.html, http://www.compaq.com/products/sanworks/drm/index.html, and http://www.compaq.com/products/storageworks/solutions/bidirectdrm/index.html). The objectives are to build a solution with dual independent SAN fabrics, meeting at the storage controllers and servers to produce a no-single-point-of-failure solution. An example of such a design is shown in **Figure 1**.

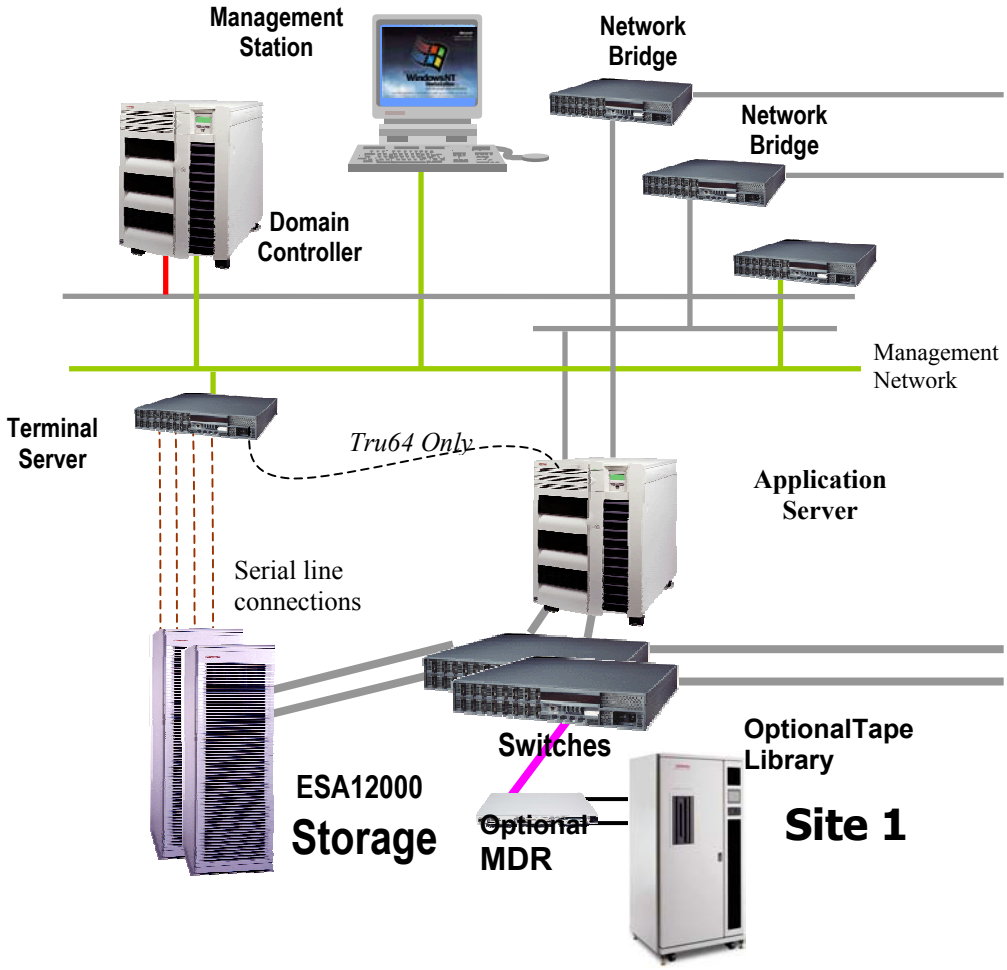


Figure 2 - Diagram showing typical serial connections for active monitoring (single site)

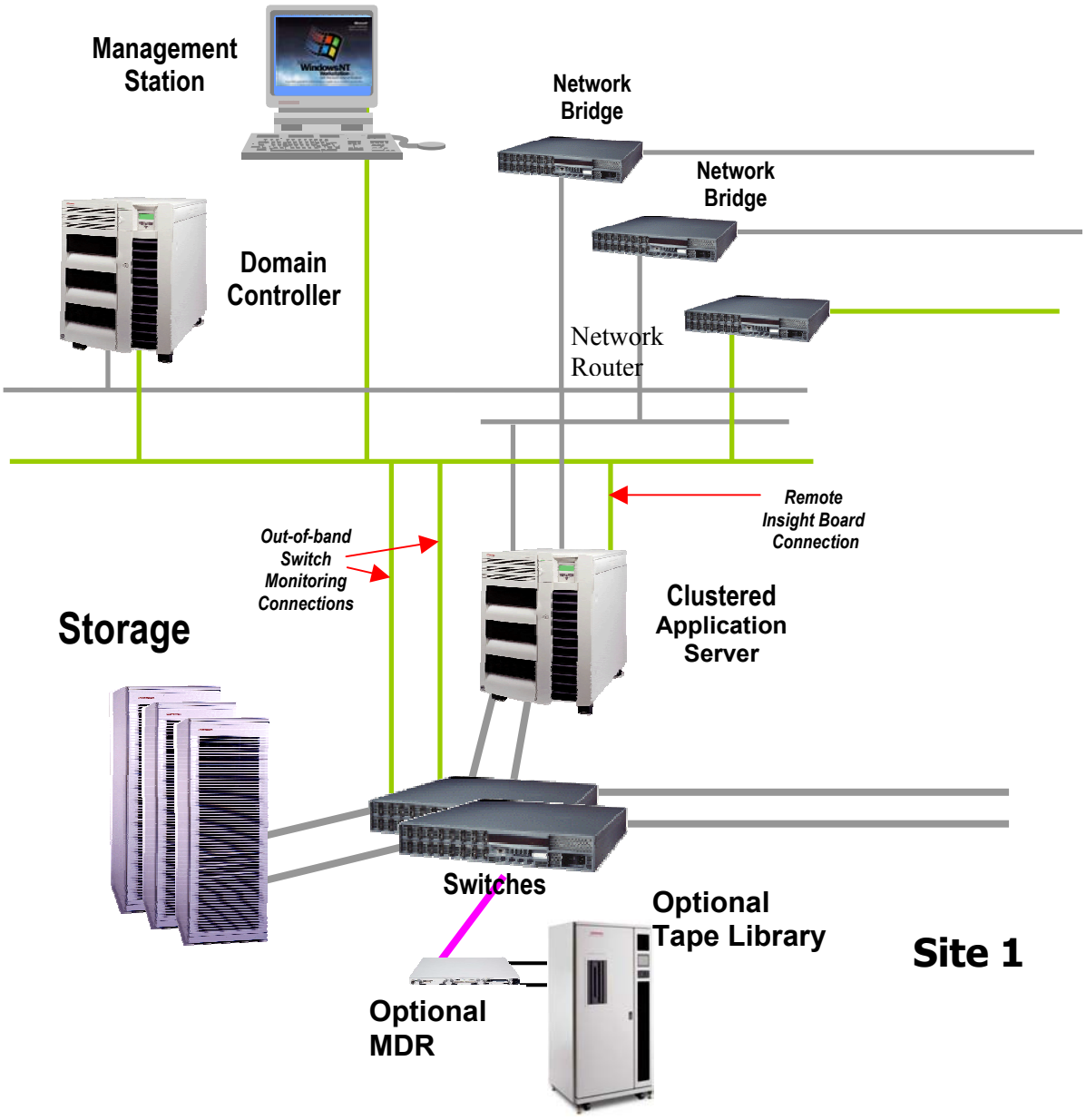


Figure 3 - Diagram showing typical ethernet connections for active monitoring (single site)

In each of the sties, the monitoring solution requires connecting to the active components. For example, the HSG80 storage controllers may be monitored either in-band over Fibre Channel, or out-of-band through the serial connection. The SAN switches may be monitored over the Fibre Channel or out-of-band through the 100 Mbit ethernet connectial . Typical connection designs are shown in Figure 2 and Figure 3.

4 Services for the Compaq DT-SAN Solution

4.1 Preparing The Environment For The DT-SAN Solution

The environment must satisfy key criteria in terms of hardware connectivity and software licensing before an effective Disaster Tolerant SAN can be constructed.

4.1.1 System and Network Software

Valid licenses for products necessary to build the basic two site solution are required for each system participating in the DT-SAN Solution. Furthermore, each of the server systems in the environment must be legally licensed to run any additional software that is a prerequisite for the application (such as SQL Server / Oracle).

In addition, each of the systems must have adequate network provisions. Each server must have a minimum of two separate network interconnects (preferably more) and must have a full range of network addresses allocated for the TCP/IP protocols to be effectively configured on the interfaces.

The provision of all licences required to install the operating system and successfully run the application is the responsibility of the client and is not part of the Disaster Tolerance SAN Solution.

In addition, each of the systems should fully conform to the standard prerequisites for the relevant operating system. The requirements for these components are described in the relevant Hardware Support List or Software Product Descriptions.

4.1.2 Management Stations

A minimum of two management stations is required for each Disaster Tolerant SAN Solution. These management stations will be used for running all of the data processing and presentation software that is part of the DT-SAN management package, as well as to run the management tools provided.

Most modern Intel systems will support the DT-SAN management package, however, the following configuration is the minimum system supported:

- PCs equipped with 900 MHz Intel Pentium III processors or better;
- Windows NT 4.0 or Windows 2000 licences (Note that required service packs will be loaded as part of the service)
- 2D graphics capability
- 512 MB main memory
- DAT backup tape
- 21” monitor
- 2 disks with a capacity of 9GB or larger

The provision of these management stations is the client’s responsibility. They are not part of the Disaster Tolerant Cluster SAN Solution.

4.1.3 Console connections

For the effective operation of the completed Disaster Tolerant SAN Solution, each storage controller, server system, and manageable network device that forms the cluster and its primary network must be monitored and managed from the SAN management stations.

In many cases, this monitoring may require the ability to collect output from the serial line console of the relevant component. For such components, the DT-SAN Solution requires a connection from the serial console line of the device to a serial line terminal server port.

Currently preferred terminal servers are the DECServer 90M or DECServer 900M models, although other TELNET capable servers can be used. Note that at least one terminal server is required in each data center, although more may be required to service the required number of ports or to provide additional management system redundancy.

The provision of these terminal servers is the responsibility of the client. They are not part of the Compaq DT-SAN Solution. The Disaster Tolerance team will advise on the number of such connections required.

4.2 The Compaq DT-SAN Solution Services

The standard DT-SAN Solution components that are typically delivered are described in this section. The service component is custom quoted to meet each customer's exact requirements and therefore the list of items may be modified to address specific requirements and demands.

4.2.1 First Consultation

The DT-SAN Solution delivery starts with a consultative session that may be one or more days, to finalize details of planning, and ensure that the environment is fully prepared for successful deployment of the DT-SAN.

The purpose of this initial consultation phase is to:

- Confirm the scope of the DT-SAN Solution and the customization required to satisfy the customer's specific disaster tolerance requirements
- Outline the DT-SAN Solution training delivery and agree on suitable dates and locations for training. Potential attendees may be identified
- Review the DT-SAN Solution configuration
- Collect any outstanding information required in order to complete the transition plan
- Visit the two main customer data centers for familiarization purposes.

After the initial consultation, you will receive a short report that summarizes the visit and includes any final recommendations regarding the implementation of the DT-SAN Solution project.

4.2.2 DT-SAN Solution Implementation Plan

A DT-SAN Solution implementation plan that includes details on all technical activities that are to take place in the delivery is one of the first major tasks in the delivery.

This plan highlights risks for each of the following activities and describes the contingency plans for each case. The client reviews the implementation plan before technical activity can begin.

Topics covered in this document include:

- A full project plan that lists the key tasks and milestones
- Names of the people responsible for critical tasks
- Technical constraints including:
    - System requirements and registry settings for all servers
    - Patch and service release level requirements
    - Hardware and firmware level requirements
    - Layered product version restrictions
    - Disk and controller reconfiguration requirements

4.2.3 DT-SAN Solution Training

It is **absolutely essential** that suitable members of staff be trained to manage and operate the DT-SAN Solution. If staff members are not trained, the investment in the solution will not be realized. The DT-SAN Solution includes training for operators and system managers. The level to which internal members of staff are trained is dependent on the individuals' responsibilities and job roles, but there should be at least two fully trained personnel.

Prerequisite knowledge for attendees of the DT-SAN Solution training includes:

- Working knowledge of management and operations for the operating system(s) (Tru64 UNIX or Windows) present in the Disaster Tolerant Environment;
- Ability to use a Windows NT 4.0 or Windows 2000 management station.

The training is either three or five days, depending upon the level of detail required. The training topics, at a minimum, include the following:

- Technical Overview
    - Description of the components included in the DT-SAN Solution and examples of how they are used.
- Operator and System Manager Overview
    - Describes the unique issues related to managing a DT-SAN Solution.
- Using the DT-SAN Solution Management Station
    - Describes the features and components of the management station and is followed by hands-on practice with the management stations.
- Recovering from failures

  - Describes the key issues relating to failure and recovery. Examples of management station error messages and data center failure and recovery are discussed.
- Business Continuity / Disaster Recovery Workshop
  - Lab exercises that require a minimum of one full day and a dedicated DT-SAN Solution for delivery. Provides an extensive hands-on session using the management stations to perform a variety of management tasks. A variety of failures are introduced into the environment to allow attendees to experience the effects and practice system recovery.

4.2.4 DT-SAN Solution Readiness Review

A readiness review is performed before the DT-SAN Solution switch to live operation, also called "go live". The review usually addresses the following issues:

- Reviewing the configuration details.
- Preparing for the management station software installation.
- Final review of the network configuration and design.
- Review of the go-live documentation.

An experienced Disaster Tolerance Consultant performs the readiness review.

4.2.5 DT-SAN Solution Management Software Installation

The DT-SAN Solution Services Management Software is installed at a convenient time during the implementation. An experienced DT-SAN Consultant installs the software and addresses the following issues:

- Installs the service packs required for Windows NT or Windows 2000, with any additional patches and device driver installation.
- Installs management station software and tools.
- Installs and configures monitoring and sampling software on the Tru64 or Windows application server systems (although this will usually be performed as part of the "go-live" support);
- Manages the DT-SAN Solution configuration database build.
- Customizes the local management stations.
- Connects RoboCentral to the management stations.
- Performs post-installation management station verification tests.

4.2.6 "Go-Live" DT-SAN Solution Support

After completion of the preparatory work, the systems can go into live operation with the new configuration changes and the new network. If you do not purchase the Build Service, this task can be deferred until the service phase of the DT-SAN, when consultants are available to assist with the switch-over.

During this phase, the introduction of the configuration changes may need to be phased in over a period of weeks or possibly months to minimize the risk of disruption to the live environment.

Here is a typical outline of the process:

- Bring a single system online in a single data center, working from a single storage array in that location.
- Carry out verification checks.
- If a cluster is used, boot the additional cluster node from the second location.
- If single server is used, configure the fall-back server to reflect the configuration of the primary server. Test the server to ensure that it can perform the role correctly;
- Start mirroring data with DRM.
- If clustering, enable applications under clustering software support for full two data center operation.

If the solution is being created or modified to operate in two sites for the first time, the Disaster Tolerant Consultant provides support for the "Go-Live" period.

One or more experienced Disaster Tolerance Consultants attends the client site during the switch to assist with any issues that may arise. Typically, the switch occurs outside of normal working hours. The consultants are also available for the day following the operation, to ensure continued smooth operation as normal workloads resume.

4.2.7 Disaster Recovery Testing

After the DT-SAN Solution and the management station platforms are successfully implemented, it is necessary to perform a complete suite of recovery tests to prove the operational effectiveness of all aspects of the DT-SAN Solution. Specific tests to be performed include:

- Restarting of the DT-SAN Solution;
- Successfully operating applications from the recovery site;
- Testing that the Management Stations fail-over correctly;
- Ensuring that a single system failure is correctly indicated and controlled recovery is possible.
- Storage Unit set recovery functions correctly.
- Complete data center failures are correctly alarmed, and recovery procedures are correct.
- Intersite link failures are correctly notified.
- Network component failures are correctly indicated.
- Terminal server failures are diagnosed (console connections failure).

It is usual for two consultants to attend for the duration of the tests. This is particularly appropriate when the data center separation is large or the configuration is particularly complex.

A variety of customer staff should be provided to ensure maximum benefit from the disaster recovery tests. A typical test involves the following staff:

- System managers and technical support.
- Network managers.
- Applications support staff.
- End users for each key application.

The tests will be scheduled for a period that is least likely to cause business disruption in the event of a problem. This is normally a weekend evening, although this is highly dependent upon the business demands on the equipment. The tests normally last between four and 24 hours. The actual duration is determined by the size of the configuration and the detail of the tests.

4.2.8 Documentation

An important part of the DT-SAN Solution is customized documentation. Documentation is required and is critical to the continued successful operation of the system.

Documentation includes the configuration document, which describes all aspects of the DT-SAN Solution configuration, and the recovery document, which contains detailed recovery process information. Both documents describe the components directly related to the DT-SAN Solution and the related network components.

Typical topics included as part of the DT-SAN Solution documentation are listed below. Note: these documents are not sufficient to document the entire DT-SAN Solution. Customer specific areas, such as additional dependencies and descriptions of feeder systems and subordinate systems, should also be addressed.

DT-SAN Solution Configuration Documentation

DT-SAN Solution documentation describes the following areas in detail, and contains detailed diagrams of the DT-SAN Solution and the storage configuration:

- A description of the system bootstrap procedures.
- Descriptions and drawings of the disk and storage configurations.
- Listings of operating system configuration details, especially any unusual features, such as specific registry or parameter settings.
- A description of the configuration of the Management Stations.
- A description of the Console Port Usage and the mapping of the terminal server ports.
- A description of the SAN fabric design, including switch port mappings and controller communication settings.
- Descriptions and drawings of the configuration of network adapters. For highly configurable network devices, a large amount of detail should be provided, such as including:
    - Port and module configuration
    - Filters and switch settings.

Disaster Recovery Documentation
This document addresses all aspects that relate to the critical applications on the DT-SAN Solution. The focus of the document is on aspects relating to system recovery after failure. The document includes the following topics:

- Data center failure detection and recovery procedures.
- Node failure detection and recovery procedures.
- Intersite link failures detection.
- Network component failure detection and recovery procedures.
- Network and storage adapter failure detection.
- Descriptions of application restart procedures.

4.3 Compaq DT-SAN Solution Lifecycle Support

4.3.1 Maintenance

The DT-SAN Solution environment is a highly complex environment. However, it is carefully designed to satisfy a specific goal. For this reason, the solution is deliberately limited in scope, to simplify maintenance and provide guaranteed behaviors and response times.

Since the DT-SAN Solution does not add invasive code to the core application delivery platform (*StorageWorks* and the associated operating systems), the solution does not introduce new failure modes. Any failure of the DT-SAN Solution monitoring removes the benefits of the monitoring, but does not affect operation of the core application servers –the environment reverts to having to be monitored and managed as though the tools were not present. This means that the DT-SAN Solution behaves in a classic ”fail-safe” manner.

However, to ensure that the solution always delivers against expectations, a rigorous approach to maintenance is critical. The following sections describe the maintenance requirements.

4.3.1.1 DT-SAN Solution Maintenance

Compaq provides a range of maintenance packages for the solution, but strongly recommends that the main systems and application components are maintained separately, as appropriate.

Compaq can provide a complete standard maintenance package for the DT-SAN Solution. The standard maintenance package is custom quoted, to match the client configuration, but the package includes all of the following:

- All update licenses and media for DT-SAN Solution components for a one year period.
- All update licenses and media for the required Heroix software for a one year period.
- Access to the DT-SAN Solution support help desk via e-mail or telephone, although these are not available on a 24 hour basis.

- Two visits per year from a DT-SAN Solution Consultant to provide upgrade assistance.

The maintenance package is renewed on an annual basis.

4.3.1.2 Other Maintenance Recommendations

It is critically important that the main system components and application code are properly maintained. The Disaster Tolerance Services team recommends the following maintenance coverage:

- Compaq extended maintenance coverage (7 by 24 hour) on all main systems, all network components, and major storage controllers.
- Firmware update services for the Compaq *StorageWorks* controllers.
- One of the following -
  - Spare storage building blocks (disks) to be used as user replaceable spares.
  - Compaq extended maintenance cover (7 by 24 hour) on the disk storage.
- Extended support, wherever possible, on all business critical software components, including –
  - Database engines
  - Transaction monitors
  - Communication products, protocols and tools
  - Applications

4.3.2 Disaster Test Services

After the DT-SAN Solution is successfully implemented, it is necessary to review the recovery procedures regularly to ensure that configuration changes are included. In addition, it is necessary to ensure that key members of staff remain familiar with the procedures, so that they can react effectively

For this reason, it is recommended that disaster tests be conducted on a regular basis, preferably twice per year.

The format of the tests should be the same as that for the test performed at the end of the DT-SAN Solution implementation.

Compaq's Disaster Tolerance Services team can provide support for these tests., Consultants can come to your site to review the procedures and to monitor, assist, and advise during the tests. However, the major effort involved in the tests must be from the client, as one of the main aims of the exercise is to ensure client staff readiness.

5 The DT-SAN Solution Management Toolset

5.1 Rationale Of The Management Solution

The successful operation of the DT-SAN Solution is dependent upon complete awareness of the state of the system at all times. In addition, effective recovery is only achievable when recovery procedures are enforced and automated to a sufficient level to remove the normal risks associated with recoveries.

In the DT-SAN Solution, these objectives are achieved by the deployment of a comprehensive management toolset.

The Compaq DT-SAN Solution employs sophisticated monitoring of all components that make up the DRM configuration – not just the storage elements, but also the servers and network components that are critical to supporting the applications and services. The DT-SAN Solution provides a fully redundant management infrastructure, including management stations residing in each computer room, both of which are fully capable of performing all tasks necessary to achieve recovery from any particular site or system outages.

This solution also provides management station ***clients*** that allow full access to all functions of the primary management stations via remote desktop systems located anywhere on the corporate LAN. This functionality includes the ability to view all events detected by the management stations and to establish sessions on the primary management stations to perform recovery activities.

The DT-SAN Solution management stations provide a significant amount of information allowing any set of failure conditions to be analyzed to determine the nature of the problem and to aide the decision making process regarding the appropriate course of recovery action. If a problem occurs that results in a decision to perform a site fail-over, the DT-SAN Solution management station includes a comprehensive fail-over manager that is invoked and initiates all tasks required to recover services.

The Recovery Manager is configured with details of all possible failure scenarios for a particular implementation, and with the customer specific recovery processes required for each of the Business Critical services hosted on the protected SAN environment. With this tool, the operator can work though each step to achieve recovery of the service. All commands required to reconfigure the storage are issued automatically along with steps to ensure the fabric is made safe from possible interference from the failed site until the operations staff is ready to combine the two sites again.

This section discusses the key features of this toolset.

5.2 The DT-SAN Solution Monitoring Tools

5.2.1 Overview of Monitoring

The DT-SAN Monitoring Toolset provides a state-of-the art solution for monitoring the availability of a DT-SAN Solution. The DT-SAN monitoring solution provides the following key features:

- Full monitoring of all availability aspects of the Storage Area Network, especially the DRM technology used for intersite data mirroring.
- Full alerting capabilities in the event of a problem or disruptive event that affects the storage infrastructure – including paging of events to storage management system or external paging system, or forwarding of events to an enterprise management solution, such as Tivoli, Unicenter, BMC Patrol or HP OpenView.
- Monitoring all networking components to ensure that all levels of the hardware infrastructure are monitored to provide a single point-of-view for all infrastructure related issues in the environment.
- Detecting all operating system events, and many hardware level events on the servers. All alerts are integrated with the storage level events to provide a comprehensive and cohesive monitoring solution that provides exactly the information required to indicate when critical problems with the environment occur.
- Monitoring critical application processes, although it is not the intention of the tools to provide full-range application monitoring. The tool is configured to monitor key availability features of the applications, such as the existence of server processes and alerting to potential looping;
- Monitoring the recovery servers for each system at all times to ensure their readiness for deployment.

Various techniques are used for monitoring different components – the rationale is to attempt to gather as much information as is possible about every element of the infrastructure. Some of the key techniques used to gather information are as follows:

- The SAN switches via SNMP polling. This monitoring includes detection of all hardware failures reported by the device, together with automatic alerting of all state changes in of the Fibre Channel ports.
- HSG controllers via the console line interface and in-band, via Fibre Channel connections. HSG monitoring and detection includes full hardware failure checking, including fans and power supplies, together with the obvious failures of any storage elements.
- Servers are monitored via a number of separate mechanisms:
    - An active agent running on the server inquires for significant server events at predetermined intervals and generates alerts on anything that might threaten the availability of the server.
    - Tru64 servers with serial line consoles are monitored over that serial line, and any events reported by this mechanism are examined and reported if relevant.

- Windows NT or Windows 2000 servers can optionally be monitored via Remote Insight Board devices.

In all cases, the DT-SAN Solution uses polling mechanisms rather than trapping type mechanisms, since these techniques provide more reliable monitoring. The monitoring system also employs a sophisticated self-checking system, including "Heartbeat" monitoring, to ensure that failed monitoring components do not compromise the effectiveness of the solution.

5.2.2 Principles of Monitoring

The DT-SAN Solution monitoring philosophy is to present information about the whole environment as clearly as possible. The tools are designed with a number of critical aims in mind:

- Information must be provided clearly, to enable easy digestion of the facts and understanding of the issues.
- Information must be provided as succinctly as possible – simple one line messages are easier to read than reams of text.
- Operators should not be presented with superfluous information that might be hiding important details – Disaster Tolerant Monitoring screens should only display information to operators that requires their attention.

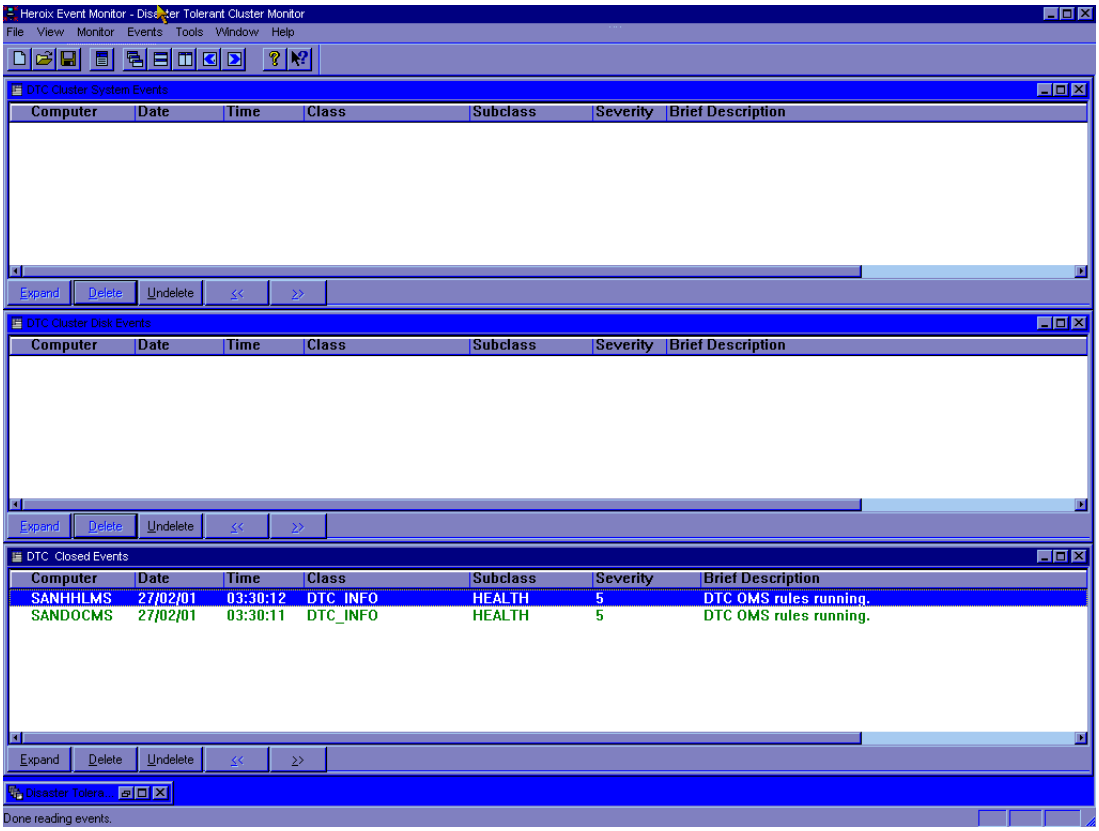


Figure 4 – The DT-SAN Solution Monitoring Window under "Normal" Conditions

To achieve these objectives, the DT-SAN Solution monitoring tools use a high degree of filtering to ensure that only important data is displayed. In addition, clear use of multiple windows enhances the ability of operators to use and interpret the tools.

5.2.2.1 The Monitoring Display

Figure 4 shows an example of the basic monitoring display, under normal conditions.

In this figure, three display windows can be seen. These windows display failure alarms (known as events) logically for interpretation by the operations staff.

The top window shows key infrastructure events, such as intersite communications, user, network, or server errors.

The middle window displays storage events, such as mirror set degradation, or any failures significant to the storage subsystem.

The lower window displays events that have occurred in the past, but which have been cleared. An example of such an event would be a server that has been taken down and then restarted. When the server was taken down, an event would have appeared in the top window (an infrastructure event). However, restarting the server clearly removes the need for operator action, so the event is "cleared" or "closed" and moved to the bottom window. This maintains a historic record of the health of the system on the screen without cluttering the key parts of the display.

Figure 5 shows the monitoring system when a major problem occurs in the environment.

In this screen, a number of events can be seen that show the environment in a serious error state. However, note that the events are color coded and prioritized. The most important events are identifiable over the incidental conditions.

Figure 6, shows an example of a detail screen. This screen is available for each event displayed on the main screens, and provides additional information including:

- Detailed description of the class of problem detected, possibly with reference to documentation or other recovery tools.
- Event history relating previous events on the same device/subsystem – allowing operations staff to correlate times and determine whether a problem has occurred suddenly, or is a result of gradual deterioration.

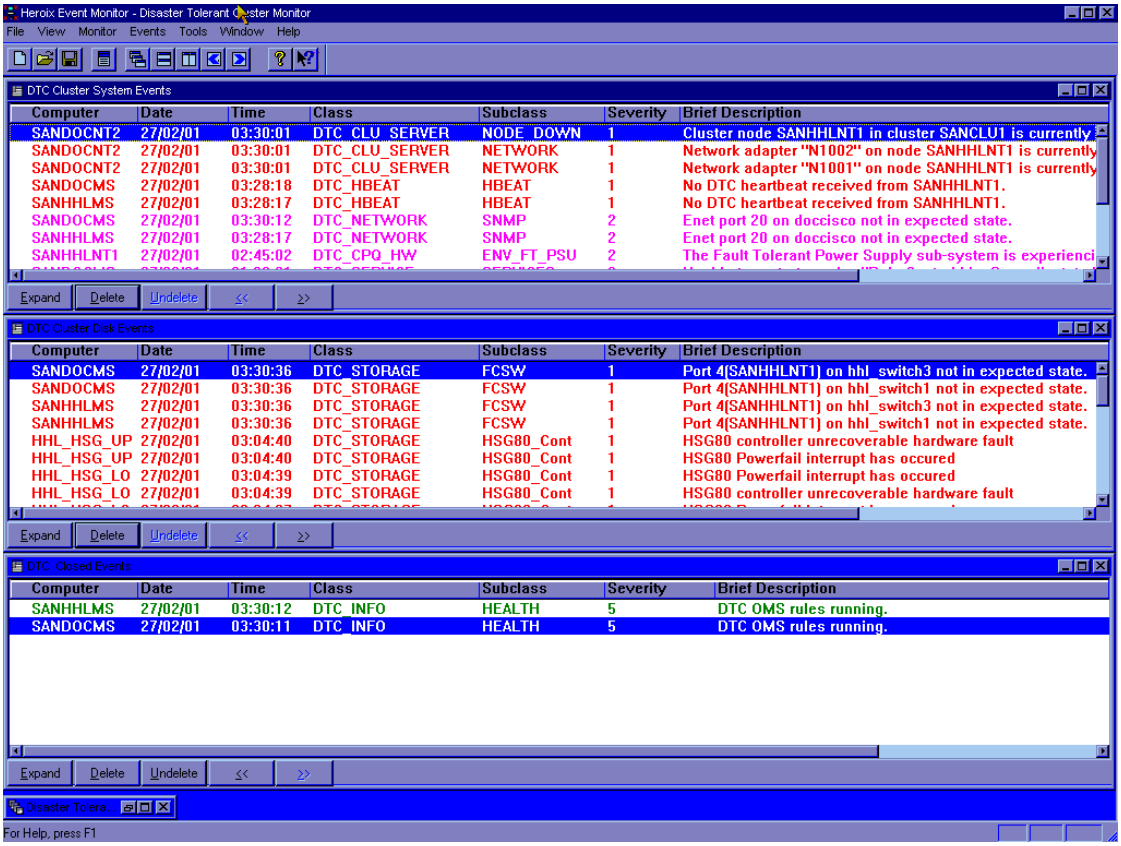


Figure 5 – The DT-SAN Solution Monitoring Tool showing example screens during a site failure

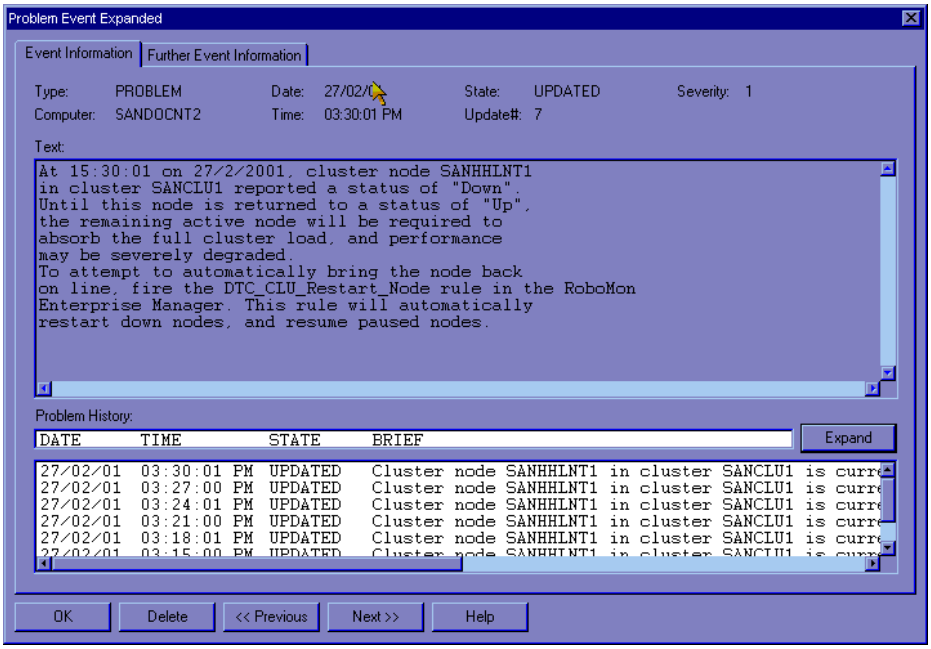


Figure 6 – The Detailed View from the DT-SAN Solution Monitoring Tool

5.2.3 Monitoring Software

The DT-SAN Solution monitoring tools use s suite of products developed and supplied by Heroix. These tools are used to provide the basic "engine" for the management station. The RoboMON tool provides monitoring and display, whilst the RoboCentral tool provides the console monitoring and console access control. RoboER provides minicomputer style console-like access to Windows 2000 and Windows NT systems, whilst RoboEDA provides communication facilities.

The following sections describe the capabilities of these tools that are utilized in the DT-SAN design. Please note that these tools have much wider capabilities that are described at the Heroix website – www.heroix.com.

5.2.3.1 The Basic Heroix Monitoring Tools

5.2.3.1.1 RoboMon

RoboMon is a system management product that monitors computers and emulates human reasoning and decision making in order to help solve system problems. Using an extensive intelligent rule based architecture, RoboMon behaves like a systems and operations manager.

The philosophy behind RoboMon is Automation - finding problems before they impact the computer systems and their users. In the DT-SAN Solution, RoboMon is configured to escalate all problems using a variety of notification methods. Note that although RoboMon has the ability to initiate automatic corrective actions, this functionality is not used in the DT-SAN design.

In today's distributed computing environments, events are generated by every system in the LAN/WAN. Suppose five events occurred per day from 100 systems. - Each event needs to be evaluated for its impact on the smooth running of your IT infrastructure. In order to allow easy problem diagnosis, this event clutter needs to be reduced. To achieve this, RoboMon uses sophisticated event filtering techniques that collect related events, report them clearly and simply, and completely remove non-availability related events. Since RoboMon continues to monitor all problems, users do not get repeated warnings about the same problem. RoboMon alerts you to the first occurrence of the event, and then updates/escalates the original problem as necessary.

Event clutter is further reduced using complex correlation facilities. Statistics from independent sources are tested together in a single condition, allowing events to be correlated before any actions are taken.

RoboMon's Event Monitor centralizes event notification in the DT-SAN, providing dynamic, real time event notification. The Event Monitor is enables system administration personnel to monitor the entire environment from a single screen. If a problem occurs anywhere in the DT-SAN, personnel are notified immediately.

The format and content of the event display is designed to display events in the most meaningful fashion. Events are grouped into different windows according to event type, such as infrastructure events and storage subsystem events. These windows provide visual and audible indications of new events.

Events of interest can be selected and expanded to view more detailed problem information. Events are typically viewed in problem mode or message mode. Problem events are automatically removed from the main displays once the condition is fixed, moving to a "closed events" window. This ensures that only outstanding problems appear in the main displays, but also guarantees that transient events are left for the operator to see. In all cases, an historical event database is available for subsequent reporting of events.

RoboMon's paging mechanism allows dial-up networking over a Hayes-compatible modem and sends a numeric code or an alphanumeric message to a pager or digital mobile phone, which accepts SMS messages.

5.2.3.1.2 RoboCentral

RoboCentral software consolidates all the consoles for every serial-line managed device in the DT-SAN. RoboCentral allows you to monitor and control all such elements from a single location, regardless of where they are located. RoboCentral offers four major features:

- Collects and logs all messages sent to the console port;
- Filters out messages of importance using advanced text searching algorithms;
- Sends important messages to the RoboMon subsystem;
- Allows full control of the device using a terminal session on the console port, also permitting scripts to be transmitted to the device via the console port.

RoboCentral runs on a Windows NT or a Windows 2000 system and does not require client software on the managed systems. The managed system needs to support ASCII data through an RS232 compliant serial port with XON/XOFF flow control and I/O buffering. The managed system needs to have this port connected to a terminal server, which supports TELNET. You can immediately monitor and manage those systems by connecting the RoboCentral system to the same terminal servers (either locally on a LAN or remotely on a WAN) and by establishing a TELNET connection between them.

In the DT-SAN design, RoboCentral is used to monitor the following components:

- Compaq Tru64 Servers;
- Bridges or routers with serial line interfaces;
- The HSG80 storage controllers.

Once a managed system is connected, all ASCII text that is sent to the serial port is collected and stored in a file specific to that system. Each managed system has its own log file containing the history of messages sent to the serial port.

RoboCentral monitors the text being gathered and looks for user defined string pattern matches. Once a match is made, an Event is generated, which is then passed to on for subsequent processing. In the DT-SAN this will be typically result in a message being passed to RoboMon.

RoboCentral also provides an interface for controlling a managed system. This is done through any ANSI terminal emulator that supports TELNET connections. The interface can either be used for ad-hoc user operations (i.e. logging into the console of a Tru64 server for maintenance) or it may be used to transmit configuration and control scripts to the managed device.

Monitoring the raw activity of a managed system can be done as follows:

- Current data can be viewed using an ANSI terminal.
- Historical data can be viewed by opening up the log file for the system directory, using the RoboCentral viewer.
- Note that in the DT-SAN design all of the event reports are forwarded to RoboMon for display purposes.

5.2.3.1.3 RoboEDA

RoboEDA (Robo Event Distribution Architecture) is a small component of software from Heroix that provides a guaranteed delivery, store-and-forward messaging subsystem. In an environment constructed from a complex arrangement of monitoring elements, delivery of messages between the components becomes a critically important and difficult task. RoboEDA is a tool that can be added into a RoboMon and RoboCentral environment, replacing the simple communication tools provided in the basic tools, to provide a range of additional functions.

In the DT-SAN implementation, RoboEDA is used for all communication between monitoring agents and elements. It adds the following features to the design:

- Guaranteed delivery of an event;
- Duplicate delivery of each event (i.e. all system events are delivered to each management workstation);
- Connection to the paging system of the client's choice;

5.2.3.1.4 RoboER

RoboER (Robo Emergency Recovery) software is a simple tool that provides an innovative addition to a Windows NT or Windows 2000 environment. RoboER provides UNIX-like console access to Windows systems.

The basis of RoboER is to offer access to systems at a level that is below the graphical user interface. This provides some tremendous benefits:

- The system can be accessed by a remote "Telnet" session in a very efficient manner. As no Graphical information is exchanged, very little network bandwidth is consumed during access.
- If the graphical interface has "frozen" for any reason (a common problem that is often perceived as a crashed system) access may still be possible to perform key operations.

RoboER provides a comprehensive range of commands that allow key emergency tasks to be completed on any Windows NT / Windows 2000 system. A suitable privileged user can perform the following tasks:

- Listing of all processes in the system;
- Various management tasks for any process, including stopping it or changing its priority;
- The shut-down or rebooting of the system.

5.3 DT-SAN Solution Recovery Manager

5.3.1 Overview

A key management component of the DT-SAN Solution is the Recovery Manager tool. This tool is used by key operations staff to recover and manage multisite based storage configurations that have ceased to operate because of an unplanned hardware failure or that have been taken out of service for maintenance or modification operations.

The primary aim of Recovery Manager is to provide rapid and safe recovery processes for DRM environments, where the actual data replication between sites is handled by the DRM firmware on the controller. The Recovery Manager provides an interface for sending instructions to the environment to achieve fail-over and recovery. For example, during an unplanned initiator site failure, Recovery Manager manages the fail-over to the target site and the subsequent recovery to the initiator site.

The scope of the Recovery Manager includes more than the storage controllers. All elements that are part of the fail-over process are manipulated, including SAN switches, servers and application processes.

Two versions of Recovery Manager, referred to as RMFULL and RMDEMO are provided as part of the DT-SAN implementation. The RMFULL version is installed on both management stations and the RMDEMO version on any number of "user" PCs.

RMFULL is used to perform "live" recoveries.

RMDEMO is used to gain experience using Recovery Manager without endangering the "live" environment and to test out configuration changes.

RMFULL and RMDEMO are native NT4/Windows 2000 (W2K) applications.

5.3.2 Design of the Recovery Manager

The configuration information used by Recovery Manager is organized into a hierarchical structure that maps directly to how users select the recoveries they are going to perform.

5.3.2.1 Configurations

A configuration is one or more HSG pairs (also known as a region) that are running DRM. For example, in Figure 7, configuration BRISTOL_DEMO consists of HSG pairs AVN and SOM.

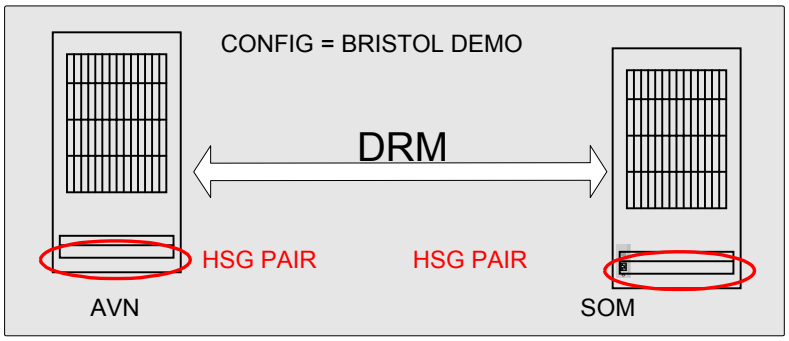


Figure 7 – Recovery Manager – Example of Simple Configuration

In Figure 8, configuration CARDIFF_ALL consists CARDIFF_LIVE, CARDIFF_DEMO and CARDIFF_TEST. CARDIFF_LIVE consists of HSG pairs, DC1LIVE and DC2LIVE. CARDIFF_DEMO consists of HSG pairs, DC1DEMO and DC2DEMO. CARDIFF_TEST consists of HSG pairs, DC1TEST and DC2TEST.

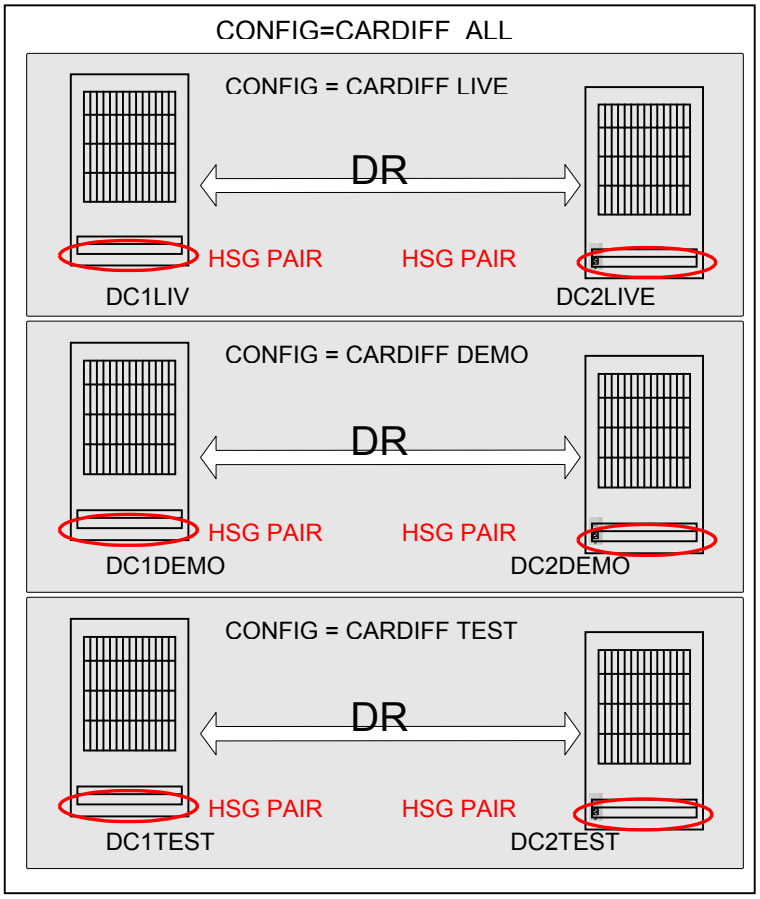


Figure 8 – Recovery Manager – Example of "All Configurations

"ALL" configurations are used to group configurations together. For example, if a site failure occurred, support staff would select configuration CARDIFF_ALL. However, if only the HSG pair DC1DEMO is failing, support staff would select configuration CARDIFF_DEMO.

5.3.2.2 Categories

The Recovery Manager is designed to provide recovery assistance in as wide a variety of failure situations as possible. To achieve this, and still have a tool that is usable, the available recovery procedures are grouped hierarchically.

Categories are the top level of grouping used to group related procedures. For example, if members of the support staff are responding to a hardware failure, they would select a procedure from the "Unplanned" category. Recovery Manager supports three categories, "Unplanned", "Planned" and "System".

- Unplanned Procedures are recovery mechanisms required when sudden events occur, such as a major component failure, site loss, or intersite link failure.
- Planned Procedures are the activities required to make the environment "safe" in order to remove components of the environment for maintenance.
- System Procedures are used to manipulate single servers or groups of servers in the environment.

5.3.2.3 Procedures

Procedures define the actions that are required to perform a recovery. Procedures are divided into stages, and stages are divided into steps. The Recovery Manager is supplied with a suite of procedures covering unplanned, planned, and system recoveries. These procedures are customized during delivery to ensure that they fully meet the client environment requirements.

Each procedure in Table 2 has a unique code, UD1, UDA3, UDF6, PD1, UOSF1, allowing easy identification for support and activity logging purposes. The first letter signifies whether it is an **U**nplanned, **P**lanned or **S**ystem procedure. The second whether the procedure relates to **D**RM or **O**THER mode.

1st Character	**Mode**	**Meaning**
U	Both	Unplanned
P	Both	Planned
S	Both	System
2nd Character	**Mode**	**Meaning**
D	N/A	Procedure relates to DRM mode
O	N/A	Procedure relates to OTHER mode
Other Chars	**Mode**	**Meaning**
A	DRM	Procedure relates to "ALL" configs
F	DRM	Procedure relates to configurations using RCS ERROR_MODE= FAILSAFE
L	DRM	Procedure uses logging disks
M	DRM	Procedure adopts minimal impact methodology
SF	Other	Procedure relates to a site flip

Table 3 – List of Recovery Manager Procedure Codes

5.3.2.4 *Users view of Configurations, Categories and Procedures*

When support staff members start the Recovery Manager, they are presented with a screen similar to Figure 9. The screen allows them to select a configuration and the recovery procedure.

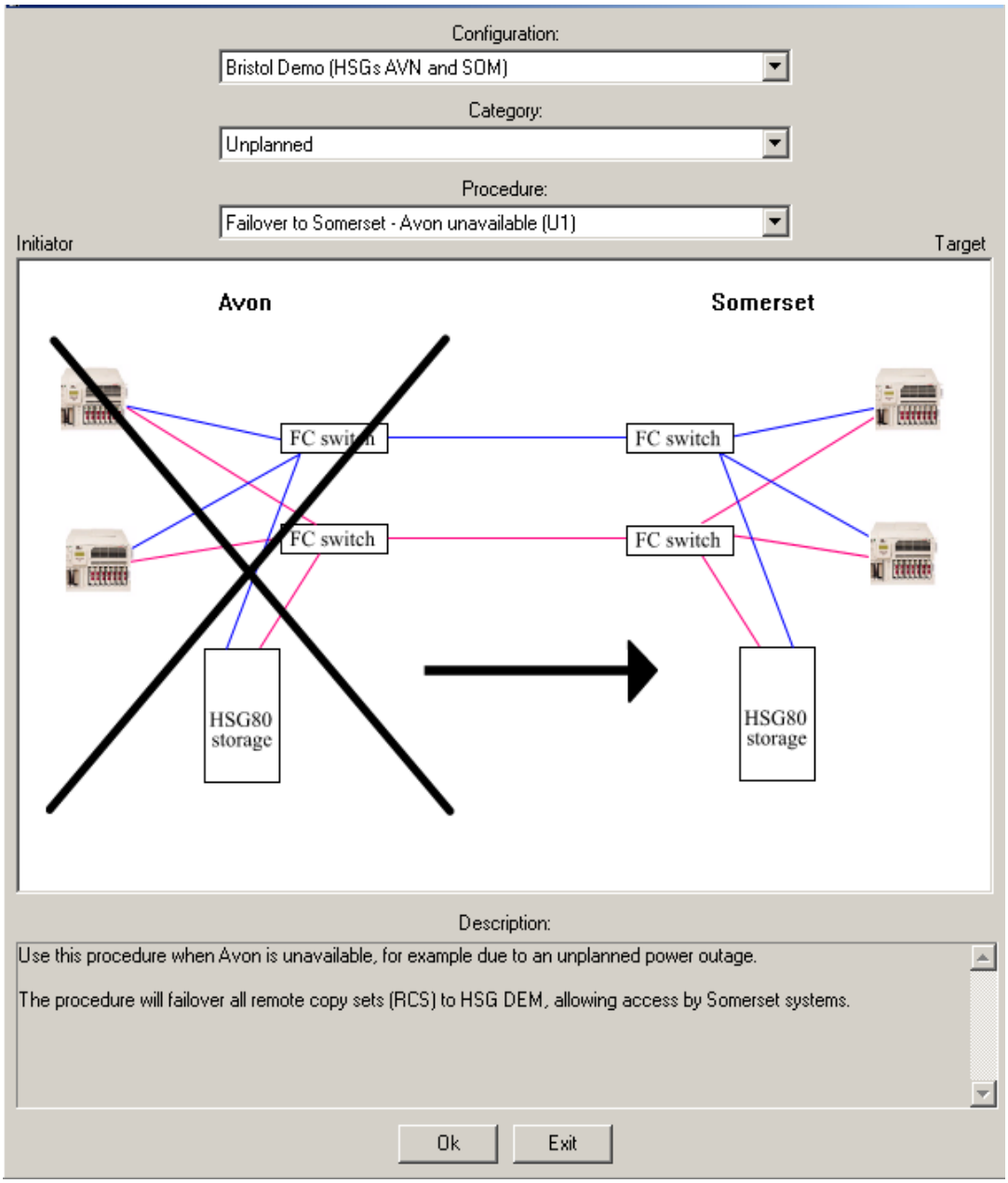


Figure 9 – Recovery Manager Initial Display

To help support staff select the correct procedure, each procedure has an associated description and simplified picture. In Figure 9, the initiator site "Avon" is unavailable and the arrows show that the procedure will fail over all remote copy sets (RCS) to the target site, "Somerset".

When a procedure is selected, a screen similar to Figure 10 is presented. The screen is divided into three sections. Section 1 (top) contains information describing the recovery, including configuration, category, and current stage. Section 2 (middle) contains the ***steps*** that have to be performed to complete the ***stage***. Section 3 (bottom) contains buttons that control Recovery Manager and optional user definable buttons that allow support staff to perform customer specific functions, such as connecting to a systems RIB board.

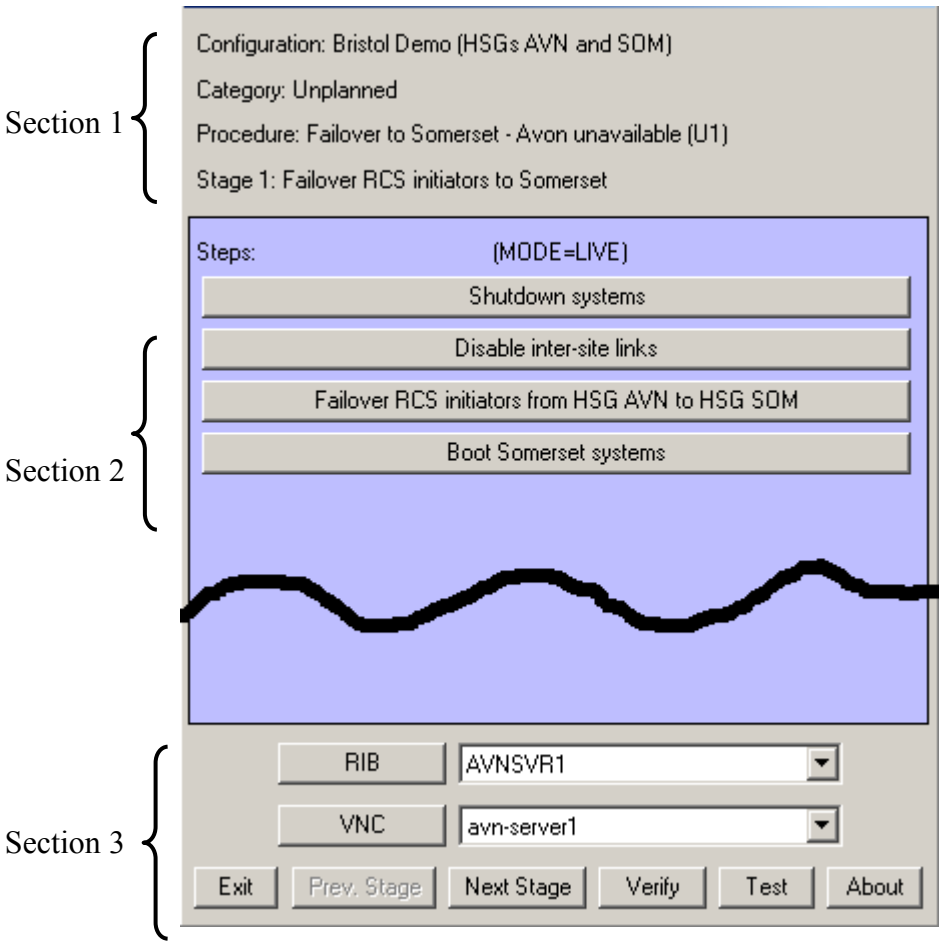


Figure 10 – Recovery Manager – Site Failure Procedure, First Screen

6 Implementing a DT-SAN Solution

6.1 A DT-SAN Project

The DT-SAN Solution is delivered as a complete implementation project led by Compaq Consultants. In this way, Compaq assumes the risk of the implementation, allowing customer staff to concentrate on ensuring that the business goals are fully satisfied by the design. It also relieves existing support staff from large time commitments during the life of the project. Compaq Consultants work closely with the customer staff in order to fully understand the customer's environment and maintain customer approval for the implementation strategy. In addition, knowledge transfer to key operational staff is ensured throughout the project.

6.2 Time Commitments

The customer support staff usually is not involved full time in the DT-SAN Solution project. The elapsed time that is required for the implementation of the DT-SAN Solution is more likely to be dependent on the following factors, rather than the availability of any staff:

- Availability of the second data center
- Availability of the intersite connections
- Deliveries of new equipment
- Constraints on the availability of the existing network and systems for the down-time required for reconfiguration tasks

If the timing for all of these points is favorable, then a typical project time scale for the migration of an existing solution is approximately three months, including the time needed to build the migration plan.

However, a completely new DT-SAN Solution can be implemented in a few weeks.

6.3 Skills Required

The precise skills required for each DT-SAN Solution varies widely. Major factors include the complexity of the existing infrastructure, detailed nature of the changes, and required implementations. In addition, as in any large project, there is some variation in the range of responsibilities that need to be adopted. Typical, medium-sized DT-SAN Solutions involve staff with the skills and responsibilities listed in the Tables 3-1 and 3-2:

Table 3-1 Typical customer staff requirements for a Compaq DT-SAN Solution

Skill / Responsibilities	Outline Tasks
Facilities	1) Data center preparation2) Cabling requirements 3) Power provision – UPS & Generator 4) Air conditioning.
Networks	Possible disaster tolerant network build or reconfiguration
Systems	Possible system build
Applications	Configuration, testing

Table 3-2 Typical Compaq staff requirements for a DT-SAN Solution

Skill	Outline Tasks
Cluster hardware engineer	Hardware installations/moves/reconfigurations
Storage Consultant or Disaster Tolerance Consultant	Determination of detailed fibre requirements Installation and configuration of storage, switches, controllers and other components
LAN, network management	Possible installation and configuration of network equipment and network management software
Operating Systems	1) Build of migration plan 2) Possible system build / migration.
Disaster Tolerance Consultant(s)	1) Installation, configuration, and final customization of DT-SAN Solution Services Software; 2) Assistance with disaster test.
Project manager	Coordinate project

Table 3-3 lists the tasks usually required in a DT-SAN Solution implementation. The "Customer Task" column indicates those tasks that a client with the relevant skills could undertake themselves. The knowledge transfer column indicates that Compaq consultants believe that they have the knowledge needed to help the customer carry out the task.

Table 3-3 DT-SAN Solution services outline

Task	Customer Task (skilled staff)	Knowledge Transfer	Part of DT-SAN Solution	Customer Resource Required
Management and Design				Various
Technical Project management	●			
Project review meetings			●	
Transition Plan Design			●	
System preparation				Systems
Verification of SAN Fabric Design	●	●	●	
Processor and system hardware preparation	●	●		
Storage Subsystem Design	●	●	●	
Storage Subsystem Configuration	●	●	●	
Build new systems	(●)	(●)	(●)	
Installation of Specific Software elements				●
(Cluster / Cluster Server Configuration)				
Application Modification				
Disaster Tolerant Connection Implementation				Networks
Priorities/filters	●	●		
Equipment installation	●	●		
Equipment configuration	●	●		
Cabling/configuration	●	●		
Test inter-site connections	●	●		
DT-SAN Software Installation and Configuration				
Installation of Operating System on Management Station			●	
			●	
Installation of DT-SAN Solution Services software on Management Station and			●	

Task	Customer Task (skilled staff)	Knowledge Transfer	Part of DT-SAN Solution	Customer Resource Required
systems Customization of DT-SAN Solution Services software				
DT-SAN Solution Go-Live DT-SAN Solution go-live		●	●	Networks Systems
Post Implementation Post implementation support Performance review	 ● ●	 ● ●	 ● ●	 Systems Systems
Documentation Disaster recovery plan System and configuration documentation	 ● ●	 ● ●		Networks & Systems

6.4 Web Links

Compaq DT-SAN Solution
http://www.compaq.com/products/storageworks/solutions/dtsan/index.html

SANworks Data Replication Manager
http://www.compaq.com/products/sanworks/drm/index.html

Compaq Tape Library Products
http://www.compaq.com/storage/tapelibrarymatrix.html

Storage Solutions
www.compaq.com/storage/solutions

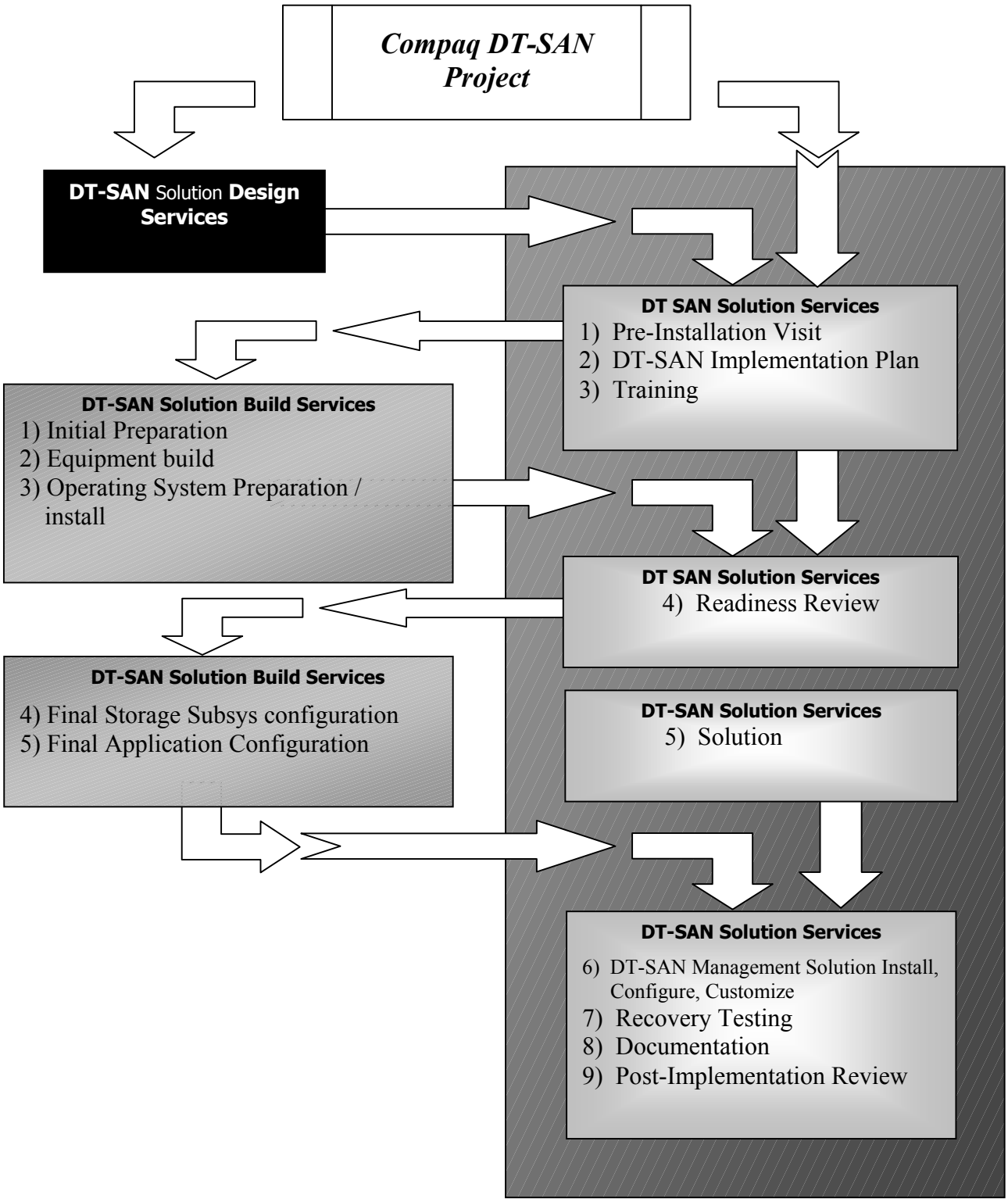
Storage Business Continuity Site
www.compaq.com/storage/continuity

Services Business Continuity Site
www.compaq.com/services/bcs

6.5 Compaq Storage and Service Offerings

Compaq offers a range of services to assist with the implementation of the complex disaster tolerant environments. If all options are chosen, the services interweave to offer a complete solution delivery, encompassing all aspects of the work.

The relationships between the component services are shown in the flowchart on the following page. If the optional services are not purchased, then the client must undertake the additional work.



Additional Components of Work (Optionally delivered as DT Service Modules)

Compaq DT-SAN Solution Services Modules