Application use cases for the HP Serviceguard Storage Management Suite

Introduction

HP Serviceguard A.11.17 tightly integrates with technology from Symantec (formerly VERITAS) to provide a Cluster File System (CFS) for HP-UX Serviceguard configurations. The HP Serviceguard Storage Management Suite offers seven bundles that include software from HP and Symantec providing customized functionality and benefits to individual applications.

The Serviceguard Storage Management Suite bundles contain software from HP, including:

- HP Serviceguard
- HP Serviceguard Manager
- HP Serviceguard Extension for RAC (SGeRAC)
- Enterprise Cluster Master Toolkit (ECMT)

The Serviceguard Storage Management Suite bundles also contain the following software from Symantec:

- Storage Foundation Standard and Enterprise (includes VERITAS Volume Manager (VxVM) and VERITAS File System (VxFS))
- Storage Foundation for Oracle® (Standard and Enterprise)
- Storage Foundation Cluster File System (includes VERITAS Cluster Volume Manager (CVM) and CFS)
- Storage Foundation for Oracle RAC

While the VERITAS software is also available from Symantec without the high availability software from HP, the software products comprised in these bundles are fully integrated with each other, thoroughly tested by HP, supported by HP, and only available through HP.

Purpose of document

The goal of this white paper is to show which applications might benefit from the Serviceguard Storage Management Suite, based on use cases. The use cases described in this document are application examples that can take advantage of the features provided by the Serviceguard Storage Management Suite. This white paper also addresses the decision criteria for selecting among the seven bundles.

You should be familiar with HP-UX and HP Serviceguard in general. For further information, see the "Related documents" section.

Definition of terms

Term	Definition
CFS	Cluster File System—A file system that is available for concurrent read and write access to multiple or all nodes of a cluster.
CVM	VERITAS Cluster Volume Manager—Enables multiple cluster nodes to access volumes concurrently.
DMP	Dynamic Multipathing—Active/active storage path failover solution from VERITAS.
DTS	Disaster Tolerant Solutions—Multi data center HP Serviceguard configurations that protect against the loss of one complete data center.
ECMT	Enterprise Cluster Master Toolkit—A collection of Serviceguard integration templates for popular third-party applications.
EOE	HP-UX 11i v2 Enterprise Operating Environment (OE)—Designed for database application servers and logic servers, this OE contains the HP-UX 11i v2 Foundation OE bundles and additional applications such as HP GlancePlus Pak to enable an enterprise-level server.
FOE	HP-UX 11i v2 Foundation OE—Designed for the demands of web servers, content servers, and front-end servers, this OE includes applications such as HP-UX Web Server Suite, Java™, and Mozilla Application Suite.
HP-UX 11i v2UD2	HP-UX 11i version 2 September 2004 update—First common operating system for HP 9000 and Integrity servers; also known as 11i v2 update 2 (output of swlist for the OE will show B.11.23.0409); minimum operating system release for HP Serviceguard Storage Management Suite.
ISV	Independent software vendor.
MCOE	HP-UX 11i v2 Mission Critical OE—Designed for the large, powerful back-end application servers and database servers that access customer files and handle transaction processing, this OE contains the Enterprise OE bundles, plus applications such as Serviceguard and HP-UX Workload Manager to enable a mission-critical server.
SG A.11.17	HP Serviceguard release A.11.17—The first release of HP Serviceguard to support the HP Storage Management Suite, including CFS with HP 9000 and HP Integrity systems.
SGeRAC	HP Serviceguard Extension for Oracle Real Application Cluster (RAC)—A software package that enables you to integrate Oracle 9i and 10g RAC into HP Serviceguard.
SGCFS	HP Serviceguard Cluster File System—T2775BA.
SGCFSO	HP Serviceguard Cluster File System for Oracle—T2776BA.
SGCFSRAC	HP Serviceguard Cluster File System for RAC—T2777BA.
SGSM	HP Serviceguard Storage Management—T2771BA.
SGSMO	HP Serviceguard Storage Management for Oracle—T2773BA.
SGSMOP	HP Serviceguard Storage Management for Oracle Premium—T2774BA.
SGSMP	HP Serviceguard Storage Management Premium—T2772BA.
SLVM	HP Shared Logical Volume Manager delivered as part of SGeRAC.
OEM	Original equipment manufacturer.
QoSS	Quality of Storage Service.
VxFS	VERITAS File System.
VxVM	VERITAS Volume Manager.

Related documents

The following documents provide valuable information on the technology discussed in this white paper. For some of the documents, multiple versions are posted. In these cases, refer to the latest version.

- http://docs.hp.com → High Availability → HP Serviceguard
  - *HP Serviceguard Version A.11.17 Release Notes*
  - *Managing Serviceguard, 12th Edition*
- http://docs.hp.com → High Availability → HP Serviceguard Storage Management Suite
  - *HP Serviceguard Storage Management Suite Version A.01.00 Release Notes*
- http://docs.hp.com → High Availability → Serviceguard Extension for Real Application Cluster (Serviceguard OPS Edition)
  - *Serviceguard Extension for RAC Version A.11.17 Release Notes*
  - *Using Serviceguard Extension for RAC*
- http://docs.hp.com → 11i v2 (under "By OS Release") → HP-UX 11i v2 Operating Environments
  - HP-UX 11i v2 Foundation Operating Environment
  - HP-UX 11i v2 Enterprise Operating Environment
  - HP-UX 11i v2 Mission Critical Operating Environment
- http://docs.hp.com → 11i v2 (under "By OS Release") → VERITAS Volume Manager and File System
  - *VERITAS 4.1 Installation Guide*
  - *VERITAS Flashsnap Point-In-Time-Copy Solutions Administrator's Guide*
- http://docs.hp.com → 11i v2 (under "By OS Release") → VxFS
  - *VERITAS File System 4.1 Administrator's Guide*
  - *VERITAS File System 4.1 Release Notes*
- http://docs.hp.com → 11i v2 (under "By OS Release") → VxVM
  - *VERITAS Storage Foundation 4.1 Oracle Administrator's Guide*
  - *VERITAS Storage Foundation 4.1 Release Notes*
  - *VERITAS Volume Manager 4.1 Administrator's Guide*
  - *VERITAS Volume Manager 4.1 Migration Guide*
  - *VERITAS Volume Manager 4.1 Troubleshooting Guide*
  - *VERITAS Volume Manager 4.1 Release Notes*

Further information about VERITAS Storage Foundation products is also available from Symantec/VERITAS at http://www.veritas.com/Products/www?c=category&refId=120.

History of the relationship between HP and VERITAS (Symantec)

HP and VERITAS look at a successful 15 plus year history during which both companies worked together especially in the file system and volume management field. The following time table gives an overview of the major steps of the relationship.

Year	Steps of partnership
1989	HP starts reselling the entire, multi-platform VERITAS product line.
1993	HP OEMs VERITAS File System as Journal File System (JFS) and Online JFS in HP-UX.
2000	HP includes the base versions of VERITAS File System (JFS) and VERITAS Volume Manager with HP-UX. The full version of VERITAS Volume Manager is made available. The full version of VERITAS File System is included in the HP packaged software bundles.
2001	HP OEMs VERITAS Cluster Volume Manager (CVM) and integrates it with Serviceguard and SGeRAC.
2002	VERITAS and HP sign an Integrated Support and Services Agreement for VERITAS products. HP and VERITAS enhanced the OEM agreement for future releases of HP-UX and VERITAS products.
2004	VERITAS and HP sign a new agreement to include VERITAS CFS and Storage Foundation products. VERITAS and HP create a migration program for Tru64/TruCluster customers.

HP Serviceguard Storage Management Suite overview

In addition to the traditional packaging of Serviceguard as a stand-alone product, part of Application Release (AR), or part of the Mission Critical Operating Environment (MCOE), seven software bundles are available through the Serviceguard Storage Management Suite offering. All of these include Serviceguard, ECMT, a version of VERITAS Storage Foundation 4.1 (Standard, Enterprise, Database, CFS), and additional products that provide enhanced storage management and database integration capabilities. HP Serviceguard Cluster File System for RAC bundle also includes SGeRAC. Each of these bundles is designed to provide a comprehensive set of software tools suited to a particular application environment.

Besides the new HP Serviceguard Storage Management Suite, Logical Volume Manager (LVM) and Shared Logical Volume Manager (SLVM) configurations continue to be supported with Serviceguard in the same cluster and on the same node. However, CFS support is only available by installing one of the Serviceguard Storage Management Suite bundles 5 through 7 or their equivalent MCOE add-on bundles.

Earlier versions of CVM (3.5) and CVM 4.1 as individual products are also still supported with Serviceguard.

Additional software in some of the bundles includes enterprise tools, such as Quality of Storage Service (QoSS), FlashSnap, Checkpoints, and additional Database Tools, including Oracle Disk Manager (ODM) for Oracle and SGeRAC, for instance. The bundle product numbers are as follows:

- Bundle 1: HP Serviceguard Storage Management (SGSM)—T2771BA
- Bundle 2: HP Serviceguard Storage Management Premium (SGSMP)—T2772BA
- Bundle 3: HP Serviceguard Storage Management for Oracle (SGSMO)—T2773BA
- Bundle 4: HP Serviceguard Storage Management for Oracle Premium (SGSMOP)—T2774BA
- Bundle 5: HP Serviceguard Cluster File System (SGCFS)—T2775BA
- Bundle 6: HP Serviceguard Cluster File System for Oracle (SGCFSO)—T2776BA
- Bundle 7: HP Serviceguard Cluster File System for RAC (SGCFSRAC)—T2777BA

The following table shows a mapping of the Serviceguard Storage Management Suite and the products included in each bundle.

Product	T2771BA	T2772BA	T2773BA	T2774BA	T2775BA	T2776BA	T2777BA
HP Serviceguard A.11.17	✓	✓	✓	✓	✓	✓	✓
HP Serviceguard Extension for RAC A.11.17							✓
Serviceguard Manager A.05.00	✓	✓	✓	✓	✓	✓	✓
Enterprise Cluster Master Toolkit B.03.00	✓	✓	✓	✓	✓	✓	✓
VERITAS Storage Foundation, Standard 4.1	✓						
VERITAS Storage Foundation, Enterprise 4.1		✓					
VERITAS Storage Foundation for Oracle, Standard 4.1			✓				
VERITAS Storage Foundation for Oracle, Enterprise 4.1				✓			
VERITAS Storage Foundation Cluster File System 4.1					✓		
VERITAS Storage Foundation Cluster File System for Oracle 4.1						✓	
VERITAS Storage Foundation for Oracle RAC 4.1							✓

The individual bundle contents are further described in the *HP Serviceguard Storage Management Suite Release Notes*. An overview of the bundle contents is in Appendix A: Product and bundle matrix. All bundles contain HP Serviceguard, Serviceguard Manager, and the Enterprise Cluster Master Toolkit (ECMT). The HP Serviceguard Cluster File System for RAC bundle also includes SGeRAC.

HP Serviceguard Storage Management Suite and HP-UX 11i Operating Environments

Relevant HP-UX 11i Operating Environment content

HP-UX 11i offers a choice of OEs all built on a common core, and each OE provides increasing levels of robust capabilities, from Foundation Operating Environment (FOE) to Enterprise Operating Environment (EOE) to Mission Critical Operating Environment (MCOE). These operating environments are completely integrated, tested, and supported by HP and are available on both HP 9000 and HP Integrity server environments.

Some of the products included in the OEs are also available through the HP Serviceguard Storage Management Suite.

For the FOE, the following products are included:

- Base VERITAS Volume Manager (VxVM)
- Base VERITAS File System (VxFS) also known as Journal File System (JFS)

For the EOE the full version of VxFS (also known as Online JFS) is added on top of the FOE contents.

The MCOE contains the following products on top of the EOE:

- Serviceguard and Serviceguard Manager
- Enterprise Cluster Master Toolkit (ECMT)

The HP Serviceguard Storage Management suite always includes the full version of the VxVM and VxFS even if a customer uses an OE that only includes the base versions.

HP-UX 11i Operating Environment add-on bundles

HP is creating specific add-on bundles for some of the OEs to eliminate the product and price duplication.

Add-on bundles with HP Serviceguard

The add-on bundles for MCOE are based on the Serviceguard Storage Management Suite bundles 4 through 7 (SGSMOP, SGCFS, SGCFSO, and SGCFSRAC):

- MCOE + HP Serviceguard Storage Management for Oracle Premium—T2794BA
- MCOE + HP Serviceguard Cluster File System—T2795BA
- MCOE + HP Serviceguard Cluster File System for Oracle—T2796BA
- MCOE + HP Serviceguard Cluster File System for RAC—T2797BA

If you are interested in the MCOE and HP Serviceguard Storage Management bundles, HP recommends ordering these four products.

Bundles without HP Serviceguard

If you are interested in the Oracle Disk Manager capability without high availability, you can take advantage of the functionality through the VERITAS Storage Foundation for Oracle products. These products will be resold through HP and branded as HP Storage Management for Oracle. The following options are available:

- With EOE:
  - EOE + Storage Management for Oracle Premium—T2792BA
- Stand-alone with FOE:
  - HP Storage Management for Oracle—T2789BA
  - HP Storage Management for Oracle Premium—T2790BA

HP Serviceguard Storage Management Suite bundles with VxVM and VxFS

The HP Serviceguard Storage Management Suite will be released in two phases. The first phase includes the full version of VxVM 4.1, VxFS 4.1, and the Storage Foundation products.

The following four bundles support clusters of up to 16 nodes but differ among each other in regards to which VERITAS Storage Foundation they include:

- VERITAS Storage Foundation, Standard
- VERITAS Storage Foundation, Enterprise
- VERITAS Storage Foundation for Oracle
- VERITAS Storage Foundation for Oracle, Enterprise

Bundle 1: HP Serviceguard Storage Management

The HP Serviceguard Storage Management (SGSM) is an integrated product bundle that includes Serviceguard high availability clustering with the full versions of Symantec's VERITAS Volume Manager (VxVM) and File System (VxFS) as part of VERITAS Storage Foundation, Standard. This entry-level bundle includes the following individual components:

- HP Serviceguard A.11.17
- HP Serviceguard Manager A.05.00
- HP Enterprise Cluster Master Toolkit B.03.00
- VERITAS File System 4.1
- VERITAS Volume Manager 4.1
- VERITAS Dynamic Multipathing (DMP)
- VERITAS Enterprise Administrator (VEA)

This bundle is suited for users with non-Oracle database environments who want to use the basic VERITAS Storage Foundation products in their Serviceguard cluster but do not have a need for CVM, CFS, or any of the premium Storage Foundation components.

Users looking for features that are not available in a Serviceguard cluster with LVM might choose this bundle. One important high availability feature—DMP—that comes with this and all other bundles is active/active path failover to storage devices that support this feature. The DMP feature of VxVM provides path failover and load balancing. It balances I/O across all available paths between a server and a storage array. The LVM alternate physical links feature, PVLink, provides only active/passive path failover, but no load balancing.

Like all the other bundles that are released in phase one, this bundle does not include CVM or CFS. For this reason, this bundle is well suited for failover applications but not for active/active multi-instance applications.

The failover applications can be customized, custom integrated, or a single-instance database. If you deploy a single-instance Oracle database, HP recommends bundles 3 or 4 because they include special features for Oracle database environments.

Users who would like to standardize volume managers on multi-vendor hardware might choose bundle 1 even if they are not looking for any of its unique features.

This bundle is suited for configurations for which the cost is a major factor.

Bundle 2: HP Serviceguard Storage Management Premium

The HP Serviceguard Storage Management Premium (SGSMP) is an integrated product bundle that includes Serviceguard high availability clustering with Symantec's VxVM and VxFS. This premium bundle also includes Quality of Storage Service and FlashSnap technologies that are part of the VERITAS Storage Foundation, Enterprise. An overview of the features and benefits are listed in Appendix B: Feature, function, and benefit matrix. This bundle includes the following individual components:

- HP Serviceguard A.11.17
- HP Serviceguard Manager A.05.00
- HP Enterprise Cluster Master Toolkit B.03.00
- VERITAS File System version 4.1
- VERITAS Volume Manager version 4.1
- VERITAS Dynamic Multipathing (DMP)
- VERITAS Enterprise Administrator
- Quality of Storage Service (QoSS)
- Instant Volume Snapshots (FlashSnap)
- Storage Checkpoints
- Disk Group Split and Join (FlashSnap)
- Fast Mirror Resync

The SGSMP bundle is suited for users with non-Oracle database environments who would like to utilize the premium features included in the SGSMP bundle. These premium features differentiate the SGSMP from the SGSM. The features are listed in Appendix B: Feature, function, and benefit matrix and described in detail in the *VERITAS FlashSnap Point-In-Time Copy Solutions Administrator's Guide* listed in the "Related documents" section.

Like SGSM, this bundle is also well suited for failover applications and not for active/active multi-instance applications.

Besides the premium features included in SGSMP, all advantages and disadvantages of the SGSM also apply to this bundle. The decision point between SGSM and SGSMP will be the premium features compared to cost.

Users who would like to use point-in-time copies of their data for backup purposes or to refresh the data for decision support systems might choose this bundle because of the functionality such as FlashSnap.

Bundle 3: HP Serviceguard Storage Management for Oracle

The HP Serviceguard Storage Management for Oracle (SGSMO) is an integrated product bundle that includes Serviceguard high availability clustering with Symantec's VxVM and VxFS, plus special database accelerators for Oracle that provide the performance of raw mode with the manageability of a file system. This bundle includes the following individual components:

- HP Serviceguard A.11.17
- HP Serviceguard Manager A.05.00
- HP Enterprise Cluster Master Toolkit B.03.00
- VERITAS File System version 4.1
- VERITAS Volume Manager version 4.1
- VERITAS Dynamic Multipathing (DMP)
- VERITAS Enterprise Administrator
- Oracle Disk Manager for Oracle 9i and 10g (ODM)
- VERITAS Storage Rollback

SGSMO is the entry-level bundle designed for Oracle single-instance database environments. It adds database accelerators and other useful functionality for Oracle databases to the contents of the SGSM bundle.

While this bundle also works for non-Oracle applications (those listed for the SGSM bundle), the specific functionality added to SGSMO can only be exploited by Oracle databases, and therefore, an Oracle single-instance database is the designated application use case for the SGSMO bundle.

The database accelerator ODM is the feature that makes this bundle most attractive for single-instance Oracle databases. The ODM interface is an application programming interface (API) that Oracle co-developed with VERITAS for incorporating underlying volume managers and file systems. The implementation of the ODM interface in a file system or a logical volume manager provides many benefits, including simplified file administration, improved file integrity, and reduced system overhead that leads to increased performance. The performance improvements are based on the following:

- File System Direct I/O

  With this feature, Oracle disk I/O bypasses the HP-UX buffer cache to provide the following benefits:
  - Improved system efficiency—Oracle uses its own buffer (System Global Area (SGA)) to cache disk I/O and has better knowledge than the HP-UX kernel about which Oracle data objects must be buffered. Hence, it is more efficient to devote memory to the SGA than to the HP-UX buffer cache for Oracle databases.
  - Improved I/O performance—Without File System Direct I/O, each Oracle I/O would be buffered twice: once in the SGA and then in the operating system buffer cache. In addition to using the memory in an inefficient way, this double buffering also consumes additional CPU resources (an additional memory to memory system call is required). Using File System Direct I/O avoids double buffering and its negative influence on I/O performance.

- Asynchronous File System I/O (/dev/async)

  Normally, direct I/O is synchronous, which can significantly downgrade write performance. For example, an application would block on a write I/O until the write completes.

  With Asynchronous File System I/O, the direct I/O is also asynchronous in the same way as it is when the database is deployed on raw volumes.

Users who are interested in the VERITAS Storage Foundation for Oracle features but have no need for high availability can choose HP Storage Management for Oracle (T2789BA) instead of SGSMO.

Bundle 4: HP Serviceguard Storage Management for Oracle Premium

The HP Serviceguard Storage Management for Oracle Premium (SGSMOP) is an integrated product bundle that includes Serviceguard high availability clustering with Symantec's VxVM and VxFS, plus special database accelerators for Oracle that provide the performance of raw mode with the manageability of a file system. This premium bundle also includes Quality of Storage Service and FlashSnap technologies. This bundle includes the following individual components:

- HP Serviceguard A.11.17
- HP Serviceguard Manager A.05.00
- HP Enterprise Cluster Master Toolkit B.03.00
- VERITAS File System version 4.1
- VERITAS Volume Manager version 4.1
- VERITAS Dynamic Multipathing (DMP)
- VERITAS Enterprise Administrator
- Quality of Storage Service (QoSS)
- Instant Volume Snapshots (FlashSnap)
- Storage Checkpoints
- Disk Group Split and Join (FlashSnap)

- Fast Mirror Resync
- Oracle Disk Manager for Oracle 9i and 10g (ODM)
- Storage Rollback
- Storage Mapping to a LUN Level
- Database FlashSnap

The content of this SGSMOP bundle is a combination of SGSMP and SGSMO, plus Storage Mapping to a LUN Level and Database FlashSnap.

SGSMOP combines the value of bundles 2 and 3, and therefore, the target applications for this bundle are Oracle single-instance databases that could benefit from database accelerators and the premium Storage Foundation features.

Users who do not want to use any of the HP Serviceguard line of products included in this bundle but are interested in the VERITAS Storage Foundation Premium features and the database accelerators can select the HP Storage Management for Oracle Premium (T2790BA). In combination with HP-UX EOE, they can order T2792BA.

NOTE
If you are interested in the MCOE, HP recommends ordering T2794BA instead of T2774BA because some of Serviceguard Storage Management for Oracle Premium products are already included in the MCOE.

For existing HP-UX users who already have Serviceguard or the MCOE and want to upgrade to a Serviceguard Storage Management Suite bundle, contact your HP representative for more details.

HP Serviceguard Storage Management Suite bundles with CVM and CFS

The second release phase of the HP Serviceguard Storage Management Suite will add three more bundles. They all include CVM and CFS, and some of them include Oracle specific tools and database accelerators. Bundle 7 also includes specific components for the Oracle RAC environment. These three bundles will initially support clusters of up to four nodes. HP plans to extend the maximum cluster size to eight nodes in the near future.

CFS-specific application integration considerations

CFS will offer new functionality to applications but provide new challenges in integrating them with CFS and exploiting the features delivered with CFS. The topics discussed in the next three sections apply for multi-instance active/active applications using a CFS:

- Comparing node-specific and cluster-wide files
- Synchronizing data updates on CFS
- Performance and scaling with CFS

The fourth and fifth sections discuss topics important for single-instance failover applications using a CFS:

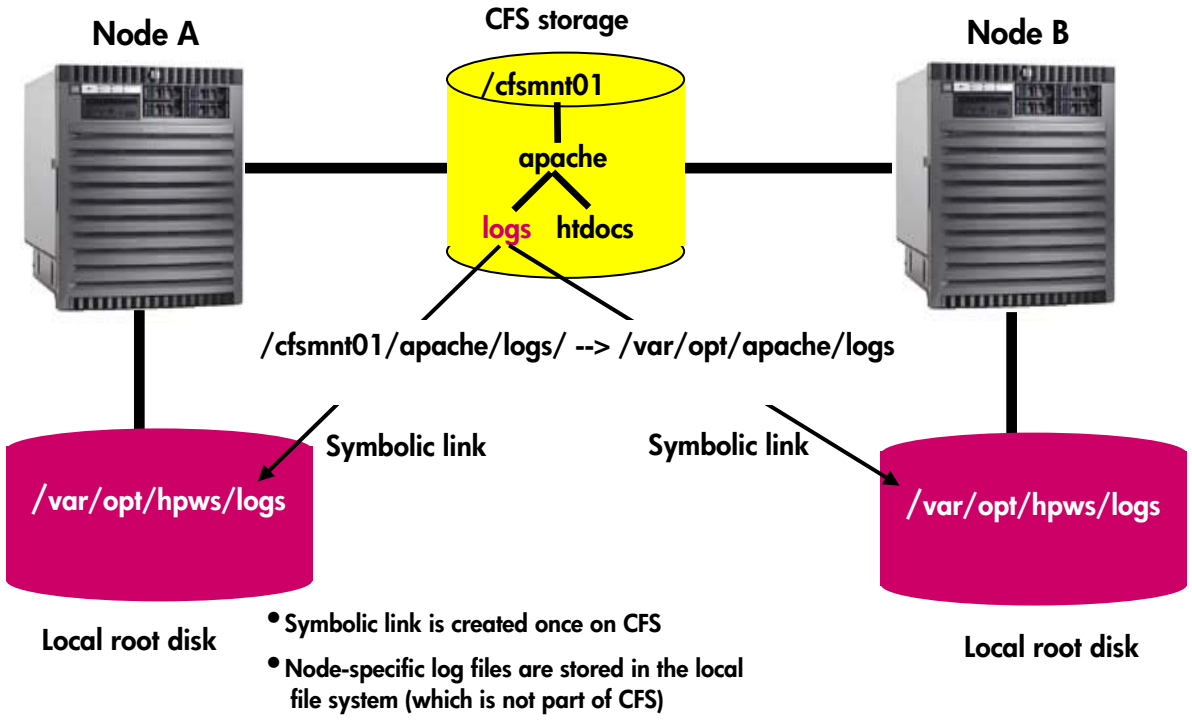
- CFS primary placement for failover applications
- Shutdown for single-instance failover applications in CFS clusters

Comparing node-specific and cluster-wide files

One of the first steps in integrating an application with CFS is to decide which files are node-specific and which are cluster-wide and shared among all nodes. This task is application-dependent. While for some files, you might have the choice between storing them locally or on CFS, for others (like the data files of a database), the location is predetermined—it is definitely on the shared file system.

An easy way to deal with node-specific files is to place them on the private file system of a local node.

Sometimes, applications expect a node-specific file at a certain path location. Apache, for instance, expects the error_log file at logs/error_log, relative to the ServerRoot directory. If the ServerRoot directory has been configured on CFS to share the configuration files, a symbolic link must be used for the logs directory to point from the shared CFS location to a local file system (in –s /var/opt/hpws/logs /cfsmnt/apache/logs). Otherwise, all nodes would write to the same error_log file, and the administrator could not identify which message came from which server.



Cluster-wide files are best placed on CFS. The application data files usually belong to these kinds of files. Other cluster-wide files (like application configuration files) could be stored locally on node-specific file systems but would need to be kept in sync manually. Because those files might change frequently, storing them on CFS is probably the better choice.

For application executables, the decision is often made in favor of local storage because they are updated less frequently than configuration files. Storing application executables locally on each node also increases the application availability slightly. Independent software vendor (ISV) applications might provide installation utilities that consider these points and provide cluster-wide installation options.

Synchronizing data updates on CFS

Multi-instance applications that access the shared data on CFS from multiple nodes concurrently must synchronize their write access to the data. All three CFS bundles provide a single file system schema that is cache-coherent and Portable Operating System Interface for UNIX® (POSIX)-compliant, which means that different processes running on different nodes of the cluster can access the CFS concurrently.

The "lockf" system call is available for applications to synchronize concurrent access of data within a file between multiple processes. With CFS, this "lockf" functionality extends from processes on a single node to processes on all cluster nodes that have a CFS mounted. Those locks are advisory only according to the POSIX standard. The HP-UX proprietary "S_ENFMT" option to enforce locking is not supported on a CFS. With "lockf", a range of a file can be locked with the granularity of a page (greater than or equal to 4 Kb) to increase parallel write access to the same file.

Performance and scaling with CFS

Another aspect to consider when implementing an application in a CFS cluster is the I/O performance. If an application uses the CFS only from one node to take advantage of possible improved failover times and ease of administration, the I/O performance is expected to be equivalent to the performance on a local VxFS.

The expectations are different when the CFS data is being accessed concurrently by multiple nodes. For each CFS, one cluster node is the primary and the other nodes are secondary nodes. The VERITAS CFS is a multi-reader/multi-writer file system, which means that all nodes can read and write user data directly to the file system without routing it through the CFS primary node. The node mounting a CFS first is elected as primary for that file system. If a primary for a CFS fails, the role moves to another cluster node automatically. You can also move the primary role manually by executing a command. File system metadata like file size, name, and time stamp can also be read by all nodes but only written by the primary node of that CFS. Keeping this in mind, the I/O performance and scaling of multi-instance applications depend on their I/O characteristics:

- Applications that do mostly read I/Os are expected to perform and scale very well. No noticeable degradation in comparison to applications doing I/O to a non-CFS is expected.
- The performance of applications that also do many writes depends on the I/O patterns of the nodes across the cluster. The current version of CFS has implemented a range locking functionality to allow multiple processes to write to different regions of the same file at the same time.
- Applications that perform many file system metadata updates are expected to not perform as well on a CFS than on a non-CFS. For example, file creation and deletion and file size expansion and truncation operations are performed exclusively by the CFS primary.

  A special mount option (`noatime`) directs the file system to ignore file access time. This option reduces file system metadata updates and should be used with applications that do not require file access time information.

  Each mounted file system has its own CFS primary, and by default, the node that mounts the file system first becomes its primary. One way to mitigate the performance impact is to distribute the CFS primaries equally across the cluster nodes. For instance, if a four-node cluster shares eight file systems, each node could be assigned as primary for only two file systems. The "`fsclustadm`" command with the "`setprimary`" option can be used to alter the primary node of a CFS.

If an application does those kinds of I/O operations only on temporary files that are not relevant to other instance of the application running on other nodes of the cluster, then those files could be configured to be stored on node-specific storage to avoid performance issues.

CFS primary placement for failover applications

For failover applications that only run on one node at the time, the location of the CFS primary is also important. The best performance is expected if the CFS primaries of all file systems an application is using are co-located with the application package on the same node and not evenly distributed among all cluster nodes.

Shutdown for single-instance failover applications in CFS clusters

Single-instance failover applications might benefit from a CFS cluster by reduced failover times because some of the tasks necessary during the failover process in an VxVM/VxFS environment are not needed with CVM/CFS (for example: importing disk groups, mounting file systems). The concurrent availability of the CFS on all cluster nodes can also introduce new challenges.

In a non-CFS cluster, the data of single-instance failover applications is moving with the application package from node to node. The volumes are exclusively activated only on the node that currently runs the package, which prevents the application data from being altered concurrently by multiple nodes. The Serviceguard package control scripts ensure that on package failure or shutdown the storage is made inaccessible on the node the package failed or was halted.

This behavior changes with CFS. The application data is accessible on multiple nodes at the same time even if an application package is not running on a specific node or not running on any node in the cluster. If an application does not shut down completely on one node before it is started on another node, data corruption could result. To address this potential issue, the start and stop functions of single-instance failover applications require special attention in a CFS cluster.

Application shutdown procedures require an analysis regarding their robustness when CFS is used to store the application data. HP provides the following recommendations for robust application shutdown within the cluster:

- Before the application shutdown procedure returns the control back to the Serviceguard control script, it should verify the success and report a failure if not all application processes died. A failure code directs Serviceguard to disable the global AUTO_RUN flag of this application package and prevents it from starting up on any other node without operator intervention.
- Serviceguard package control scripts use the "fuser" command to terminate all processes having a mount point open before it unmounts the exclusively mounted file systems and deactivates the exclusively activated volume groups. The same "fuser" command could be used in the application-specific shutdown function.
- You can temporarily make the CFS volumes unavailable on application shutdown by specifying the mount point and disk group within the same "cmhaltpkg" command as the application package. However, this process might not be suitable in all cases because the file system would require operator intervention to be made available again on the node where the package was halted.
- As a last resort, the `NODE_FAIL_FAST_ENABLED` flag can be set to `YES` to trigger Serviceguard to halt the node on a package failure. This process works when a package fails or when it fails to shut down.

Serviceguard package management provides interfaces and mechanisms to integrate robust application shutdown procedures. It is up to the application integrator to choose and utilize them to prevent unwanted concurrent access to the application data.

The following pseudo code lists the major components of a robust shutdown procedure that should be added to the `customer_defined_run_cmds/customer_defined_halt_cmds` functions of the package control script.

```
# PREPARATION: On application/package start memorize application process
#              names or process identifier (PIDs)
#              keep the list updated with monitor function if needed
#
# STEP 1: Graceful Shutdown Attempt
#         Call proprietary application shutdown command first.
#         e.g. “shutdown immediate” for Oracle database or
#              “stopsap r3” for R/3 application server
#
# STEP 2: Wait for a configurable amount of time (grace period) to give
#         native application shutdown procedure a chance to clean up.
#         e.g. while [ grace period not over ]
#                check for processes by name or PID
#                if [ none are still running ]
#                   exit 0
#                sleep 1
#              done
#              grace period expired, not all process halted proceed
#              with STEP 3
#
# STEP 3: Try harder (kill -9) and wait for another configurable amount
#         of time (kill_period) for processes to die.
#         e.g. kill processes
#              while [ kill_period not over ]
#                check for processes by name or PID
#                if [ none are still running ]
#                   exit 0
#                sleep 1
#              done
#              kill period expired, not all process halted, return
#              failure code to Serviceguard package control script
#
```

Bundle 5: HP Serviceguard Cluster File System

The HP Serviceguard Cluster File System (SGCFS) is an integrated product bundle that includes Serviceguard high availability clustering with Symantec's VERITAS Storage Foundation CFS, which adds the VERITAS CFS and VERITAS CVM for concurrent data access from multiple servers to the content of SGSMP. This bundle includes the following individual components:

- HP Serviceguard A.11.17
- HP Serviceguard Manager A.05.00
- HP Enterprise Cluster Master Toolkit B.03.00
- VERITAS File System version 4.1
- VERITAS Volume Manager version 4.1
- VERITAS Dynamic Multipathing (DMP)
- VERITAS Enterprise Administrator
- Quality of Storage Service (QoSS)
- Instant Volume Snapshots (FlashSnap)
- Storage Checkpoints
- Disk Group Split and Join (FlashSnap)
- Fast Mirror Resync
- VERITAS Cluster File System 4.1 (CFS)
- VERITAS Cluster Volume Manager 4.1 (CVM)

HP recommends the SGCFS bundle to Serviceguard users who want to take advantage of the CFS functionality in a non-Oracle database environment. All premium storage management features introduced with the SGSMP bundle are also included in the SGCFS bundle, along with the CVM and CFS.

Generally, CFS can be used for any application that requires sharing of files, such as for home directories and project files, web pages, and multi-instance applications. CFS is also applicable when you want highly available standby data in predominantly read-only environments in which you must access data or when you do not want to rely on Network File System (NFS) for file sharing.

The following applications use CFS features in different ways:

- Single-instance failover applications, including non-Oracle single instance databases, might benefit from using CVM or CFS because of the potential for improved end-to-end failover time when compared to a cluster with VxVM/VxFS volumes and file systems. Without CFS, the volumes and file systems must be made available on the failover node before the application can start. These steps are unnecessary in a CFS configuration.
- Decision support systems that use point-in-time copies of data from systems running on different cluster nodes can use FlashSnap functionality on data shared by the CFS. In this way, the decision support system can run on a different node than the primary system, and CFS and FlashSnap create the point-in-time data and make it available on the decision support system node.
- For CFS in combination with FlashSnap, using the point-in-time copy of a production system for backup purposes on a different node reduces the performance impact on the primary system during the backup process.

- The following multi-instance applications might benefit from a CFS by improving scalability through adding additional instances to additional cluster nodes:
  - Web server farms

    Web servers are particularly suitable to CFS cluster because their application is typically read-only. Moreover, with a client load balancing front end, a web server cluster's capacity can be expanded by adding a server and another copy of the site. A CFS-based cluster greatly simplifies scaling and administration for this type of application.

    Before CFS, scaling a web server beyond one node meant choosing between two sub-optimal solutions:
    - Using NFS to share the data among the nodes had the advantage that only storage for one set of data had to be provisioned and NFS kept the data in sync automatically among all nodes. The drawback was the performance.
    - Using individual copies of the data for each web server provided adequate performance, but at the price of provisioning the storage multiple times and putting manual processes in place to synchronize data updates among the nodes. HP currently uses this setup internally for the http://www.hp.com website and successfully tested CFS to reduce the storage costs by two-thirds and retain the level of performance they currently have.
  - File servers

    Two or more servers can be connected to the same CFS storage and share file system data to increase availability and scalability. Some examples are:
    - Multiple DNS servers in a CFS cluster are all accessing the same CFS data, but each of them only serves a certain subset to the clients.
    - Streaming media servers can easily scale up, adding nodes to the cluster serving the same files on a CFS.
    - Highly available NFS servers could benefit from CFS. At initial release of the HA-NFS toolkit, the complexity of configuring cross-mounts can be avoided using CFS to share the data within the cluster and NFS to the outside. In the future, HP plans that highly available NFS servers will also be able to exploit CFS for scalability.
  - Multi-instance shared-nothing databases, such as Informix XPS or DB/2, might benefit from CFS in terms of ease of administration.

The SGCFS is the target bundle for TruCluster (including Advanced File System (AdvFS) and Logical Storage Manager (LSM)) users who want to transition a non-Oracle environment from Tru64 to HP-UX.

NOTE
If you already are interested in the MCOE, HP recommends ordering T2795BA instead of T2775BA because some of Serviceguard CFS products are already included in the MCOE.

For existing HP-UX users who already have Serviceguard or the MCOE and want to upgrade to a Serviceguard Storage Management Suite bundle, contact your HP representative for more details.

Bundle 6: HP Serviceguard Cluster File System for Oracle

The HP Serviceguard Cluster File System for Oracle (SGCFSO) is an integrated product bundle that includes Serviceguard high availability clustering with Symantec's VxVM and VxFS, plus the VERITAS CFS and VERITAS CVM for concurrent data access from multiple servers for Oracle single instance environments. This bundle also includes special database accelerators for Oracle that provide the performance of raw mode with the manageability of a file system. This bundle includes the following individual components:

- HP Serviceguard A.11.17
- HP Serviceguard Manager A.05.00
- HP Enterprise Cluster Master Toolkit B.03.00
- VERITAS File System version 4.1
- VERITAS Volume Manager version 4.1
- VERITAS Dynamic Multipathing (DMP)
- VERITAS Enterprise Administrator
- Quality of Storage Service (QoSS)
- Instant Volume Snapshots (FlashSnap)
- Storage Checkpoints
- Disk Group Split and Join (FlashSnap)
- Fast Mirror Resync
- Oracle Disk Manager for Oracle 9i and 10g
- Storage Rollback
- Storage Mapping to a LUN Level
- Database FlashSnap
- VERITAS Cluster File System 4.1 (CFS)
- VERITAS Cluster Volume Manager 4.1 (CVM)

Compared to the SGCFS bundle, the SGCFSO bundle adds database accelerators and other useful tools, such as Database FlashSnap for Oracle databases. This bundle is primarily designed for Oracle single-instance database environments. However, it can also be used with all applications discussed in this paper.

This bundle also works with Oracle RAC installations, but SGeRAC must be ordered separately because it is not part of this bundle. Some of the functionality delivered with this bundle does not work with Oracle RAC configurations such as Database FlashSnap. However, the Database FlashSnap functionality, in combination with CFS, enables you to easily create a database replica that can be used on other cluster nodes for backup or decision support systems.

The SGCFSO is the target bundle for TruCluster (including AdvFS and LSM) users who want to transition an Oracle environment from Tru64 to HP-UX.

NOTE

If you are interested in the MCOE, HP recommends ordering T2796BA instead of T2776BA because some of Serviceguard Cluster File System for Oracle products are already included in the MCOE.

For existing HP-UX users who already have Serviceguard or the MCOE and want to upgrade to a Serviceguard Storage Management Suite bundle, contact your HP representative for more details.

Bundle 7: HP Serviceguard Cluster File System for RAC

The HP Serviceguard Cluster File System for RAC (SGCFSRAC) is an integrated product bundle that includes Serviceguard high availability clustering and Serviceguard Extension for RAC with Symantec's VxVM and VxFS, plus the VERITAS CFS and VERITAS CVM for concurrent data access from multiple servers for Oracle RAC environments. This bundle allows the use of a CFS in an Oracle RAC environment, together with SGeRAC. This bundle also includes special database accelerators for Oracle that provide the performance of raw mode with the manageability of a file system. This bundle includes the following individual components:

- HP Serviceguard A.11.17
- HP Serviceguard Manager A.05.00
- HP Enterprise Cluster Master Toolkit B.03.00
- VERITAS File System version 4.1
- VERITAS Volume Manager version 4.1
- VERITAS Dynamic Multipathing (DMP)
- VERITAS Enterprise Administrator
- Quality of Storage Service (QoSS)
- Instant Volume Snapshots (FlashSnap)
- Storage Checkpoints
- Disk Group Split and Join (FlashSnap)
- Fast Mirror Resync
- Oracle Disk Manager for Oracle 9i and 10g
- Storage Rollback
- Storage Mapping to a LUN Level
- VERITAS Cluster File System 4.1 (CFS)
- VERITAS Cluster Volume Manager 4.1 (CVM)
- HP Serviceguard Extension for RAC A.11.17

Compared to the SGCFSO bundle, the SGCFSRAC bundle adds HP Serviceguard Extension for RAC. Some of the functionality that does not work with RAC, such as Database FlashSnap, which is part of SGCFSO, is not included in SGCFSRAC.

This bundle is primarily designed for Oracle 9i and 10g RAC environments, but it also works with single-instance Oracle databases and other applications that would benefit from the CFS.

Additional information about deploying Oracle RAC with HP Serviceguard Cluster File System for RAC can be found in the *Using Serviceguard Extension for RAC* manual referenced in the "Related documents" section.

The SGCFSRAC is the target bundle for TruCluster (including AdvFS and LSM) users who want to transition an Oracle RAC environment from Tru64 to HP-UX.

NOTE
If you are interested in the MCOE, HP recommends ordering T2797BA instead of T2777BA because some of Serviceguard Cluster File System for RAC products are already included in the MCOE.

For existing HP-UX users who already have Serviceguard or the MCOE and want to upgrade to a Serviceguard Storage Management Suite bundle, contact your HP representative for more details.

Appendix A: Product and bundle matrix

Product	T2771BA	T2772BA	T2773BA	T2774BA	T2775BA	T2776BA	T2777BA
HP Serviceguard A.11.17	✓	✓	✓	✓	✓	✓	✓
Serviceguard Manager A.05.00	✓	✓	✓	✓	✓	✓	✓
Enterprise Cluster Master Toolkit B.03.00	✓	✓	✓	✓	✓	✓	✓
VERITAS File System 4.1	✓	✓	✓	✓	✓	✓	✓
VERITAS Volume Manager 4.1	✓	✓	✓	✓	✓	✓	✓
VERITAS Dynamic Multipathing	✓	✓	✓	✓	✓	✓	✓
VERITAS Enterprise Administrator	✓	✓	✓	✓	✓	✓	✓
Quality of Storage Service		✓		✓	✓	✓	✓
Instant Volume Snapshots (FlashSnap)		✓		✓	✓	✓	✓
Storage CheckPoints		✓		✓	✓	✓	✓
Disk Group Split and Join (FlashSnap)		✓		✓	✓	✓	✓
Fast Mirror Resync		✓		✓	✓	✓	✓
Oracle Disk Manager for Oracle 9i and 10g			✓	✓		✓	✓
Storage Rollback				✓		✓	✓
Storage Mapping to a LUN Level				✓		✓	✓
Database FlashSnap				✓		✓	
VERITAS Cluster Volume Manager 4.1					✓	✓	✓
VERITAS Cluster File System 4.1					✓	✓	✓
HP Serviceguard Extension for RAC A.11.17							✓

Appendix B: Feature, function, and benefit matrix

Feature	Function	Benefit	HP Storage Management Suite bundle
Cluster File System	Provides file system access from multiple servers	Concurrent file system access and ease of management	T2775BA, T2776BA, T2777BA
Dynamic Multi-Pathing	Balances IP across all available paths between the server and the storage array	Increased availability and performance	All bundles
Storage provisioning templates	Automatically provisions and grows storage based on user-defined templates	Creates company-wide consistency, speeds storage configuration, and eliminates storage configuration errors	All bundles
Configuration backup and Restore	System automatically makes a backup of previous volume configuration	Restores from previous configuration and minimizes downtime	All bundles
Storage Expert	Analyzes volume layout and compares against best practices	Improved performance and reliability	All bundles
Online Administration	Performs volume and file system resizing (including shrinking), domain reconfiguration, defragmentation, backup, and off-host processing while the data remains online and available	Limits the amount of time disks must be offline for maintenance	All bundles
256 TB file system, 2 TB file size	Large storage configurations	Scalability	All bundles
History log	Provides a journal of issued commands	Faster time to resolution for support	All bundles
Online Intent Log Resize	Resizes the VxFS intent log while applications are online	Minimizes application downtime	All bundles
Named data streams	Associates arbitrary data with a given file	Allows annotation of specific details about a given file	All bundles
Online LUN resize	Dynamically grows or shrinks both the file system and the associated LUNs without application interruption	Reduces the amount of planned downtime for routine tasks such as adding disk capacity to an application	All bundles
64-bit file systems	Allows files and file systems to grow beyond the 2-TB limit imposed by 32-bit systems	Allows IT organizations to have fewer file systems to manage, increase support for scientific and other vertical applications that require either many files within a single file system or a smaller number of large files	All bundles

Feature	Function	Benefit	HP Storage Management Suite bundle
Portable Data Containers 4.0	Easily and quickly converts data for use on different operating systems	Makes migrating to new operating systems easy, enables multi-platform off-host processing, and reduces migration cost and time	All bundles
Hardware Assisted E-Copy	Provides the ability to move data across the SAN without having to first read it into the server memory	Allows for off-host backup or restore processing to occur, which by definition implies virtually no server CPU cycles are dedicated to this task	All bundles
Multi-Volume File System	Provides the capability to have multiple volumes (that can be comprised of multiple LUNs) across heterogeneous storage subsystems supported under a single file system	Provides the capability to have multiple volumes (that can be comprised of multiple LUNs) across heterogeneous storage subsystems supported under a single file system	All bundles
NBU Advanced Client Support	Allows for snapshot and backup/recovery of data and leverages existing customer investment in hardware and software	Provides the fastest mechanisms known today to take a snapshot and move the data to disk or tape for rapid backup and recovery; enables the backup and recovery data without utilizing production servers, by using scsi3 xcopy enabled devices or another server to move the data	T2771BA, T2772BA, T2773BA, T2774BA
Quality of Storage Service	Identifies and moves unimportant or old files to less expensive storage without changing the way users or applications access the files	Ease of management	T2772BA, T2774BA, T2775BA, T2776BA, T2777BA
Oracle Disk Manager	Provides unified I/O and file management for Oracle servers and reduces the number of system calls that Oracle servers generate; all Oracle processes use a single I/O library	Increases I/O performance, uses system resources more efficiently, and improves file management	T2773BA, T2774BA, T2776BA, T2777BA
Storage Mapping	Uses the Oracle 9i Rel 2 STORAGE_MAPPING I/O library to map logical Oracle objects down the storage stack; identifies hot spots and tunes performance of disk arrays	Simplifies management and troubleshooting of the Oracle storage environment	T2773BA, T2774BA, T2776BA, T2777BA
Storage Rollback	Rolls back a data file, tablespace, or the entire database to an on-disk backup image	Immediate point-in-time recovery and reduced database downtime	T2774BA, T2776BA, T2777BA

Feature	Function	Benefit	HP Storage Management Suite bundle
Block Level Incremental Backup Support (Requires Net Back Up product)	Backs up only changed BLIB data blocks	Faster backups and lower CPU and network utilization	All bundles
Database (FlashSnap)	Create point-in-time volume snapshots that are readable and writable for off-host processing	Enhances application performance by allowing processing on a secondary server where the cloned DB can be located.	T2774BA, T2776BA
Storage Checkpoint (FlashSnap)	Point-in-time images at the file system level	Reduced system and network overhead during backup, reduced unplanned downtime, and faster recovery	T2772BA, T2774BA, T2775BA, T2776BA, T2777BA
Disk Group Split and Join (FlashSnap)	Splits one or more volumes and deports and then imports onto another host for offline processing and later deports and imports back	Performance	T2772BA, T2774BA, T2775BA, T2776BA, T2777BA
Instant Volume Snapshots (FlashSnap)	Point-in-time snapshot at the volume level	Ensures database consistency	T2772BA, T2774BA, T2775BA, T2776BA, T2777BA
Fast Mirror Resync (FlashSnap)	Only syncs changed blocks	Reduced unplanned downtime	T2772BA, T2774BA, T2775BA, T2776BA, T2777BA
Checkpoint Restore	Restores from the point-in-time image of the database without going to tapes by putting changed blocks back to the primary file system	Faster recovery	T2772BA, T2774BA, T2775BA, T2776BA, T2777BA

4AA0-3466ENW, Rev. 2 January 2006