

SONY

P r o d u c t I n f o r m a t i o n M a n u a l

PetaBack™

PetaBack™ software

Product Information Manual

C O N T E N T S

Chapter 1. PetaBack Software Concept Overview	7
1-1. Introduction	2
1-1-1. The need for high-performance data backup	2
1-1-2. How backup software functions	2
1-2. How PetaBack software functions	3
1-2-1. Features of PetaBack software	3
1-2-2. Main Characteristics	4
1-3. System components supported	6
1-4. Products	7
Chapter 2. PetaBack System Configurations	9
2-1. Stand-alone server backup system	10
2-2. Client server backup system	11
2-3. Oracle® database online backup system	12
2-4. LAN-free backup system in a SAN environment	13
2-5. Backup system in a NAS server environment (1)	14
2-6. Backup system in a NAS server environment (2)	15
2-7. Backup system in a heterogeneous and NAS server environment	16
Chapter 3. Overview of Functional Concepts	17
3-1. Basic operation	18
3-1-1. Backup Groups	20
3-1-1-1. Creation of Backup Groups	21
3-1-1-2. Reference to and modification of static attributes specified in Backup Groups	23
3-1-1-3. Volume allocation for Backup Groups	25
3-1-1-4. Customization of automatic backup schedule	26
3-1-2. Backup	27
3-1-2-1. Designated Files	27
3-1-2-2. Backup levels	27
3-1-2-3. Backup types	27
3-1-2-4. Backup retry function	27
3-1-3. Restoration	27
3-1-3-1. Files required for restoration	27
3-1-3-2. Restoration modes	27
3-1-3-3. Restoration to a specific generation	28
3-1-3-4. Directory selections for restoration	28
3-1-3-5. Restoration to another host	28
3-1-4. Emergency situations	28
3-1-4-1. When the file system for backup has crashed	28
3-1-4-2. When the install disk in the PetaBack™ server has crashed	28
3-1-4-3. When the OS disk has crashed	28
3-2. Function explanation	29
3-2-1. High-speed data transfer	29
3-2-2. Remote host backup and restoration (normal backup)	29
3-2-3. Remote host backup and restoration (LAN-free backup)	29
3-2-4. Automatic backup using a library	29
3-2-5. File size limitations	29
3-2-6. File path name length limitations	29
3-2-7. Java™ Applet-based GUI	30
3-2-8. Command-line operations	32
3-2-9. Generation administration of execution history	32
3-2-10. Administration of multiple libraries	33
3-2-11. Reporting automatic backup results by e-mail	33
3-2-12. High-speed recovery to the latest backed-up disk status	33
3-2-13. High-speed recovery of PetaBack system files after a crash	33
3-2-14. Restoring to another backup host	33
3-2-15. Volume sharing between different Backup Groups	33
3-2-16. Parallel operation for backup and restoration	33
3-2-17. Highly flexible scheduling	34
3-2-18. Immediate manual backup	34
3-2-19. A wide range of supported platforms	34
3-2-20. System construction in a heterogeneous environment	34
3-2-21. External administration of backup Volumes	34
3-2-22. Backup file information shown in a list form	34
3-2-23. Backup system using a stand-alone tape drive	34
3-2-24. Easy installation and setup	34
3-2-25. Support for Japanese file names	34
Chapter 4. Concepts and Terminology	35
4-1. Module structure	36
4-1-1. GUI	37
4-1-2. PBK client	37
4-1-3. PBK server	37
4-1-4. RMS	38
4-1-5. Backup operation example	39
4-2. Automatic backup mechanism	41
4-3. High-speed data transfer	42
4-4. NAS server backup using NDMP	46
4-4-1. A list of NDMP functions supported in PetaBack software	46
4-4-2. Backup using NDMP	46
4-5. Other	48
Chapter 5. Command Summary	49
Chapter 6. Limitations	51
Chapter 7. Glossary - Terms and Definitions	53
Chapter 8. Appendix	57

Chapter 1

PetaBack™ Software Concept Overview

PETA BACK

1-1. Introduction

1-1-1. The need for high-performance data backup

The phenomenal growth in hard-disk storage capacity and the proliferation of the Internet in today's broadband network world is generating a data explosion in computer-based storage systems. The large amounts of critical, marketable data that are now driving operations and profits in 24-hour, 7-day-a-week business enterprises require very high-performance data backup solutions. These must have the capability of automatically recovering data storage systems when disk files are accidentally lost by system or human errors, hard-disk crashes, virus infections, hacker attacks, and so forth. It is not too much to say that the continuing prosperity of today's digital world depends on the availability of high-performance backup software that has a major cost advantage over conventional backup systems.

Sony PetaBack software has been developed to fulfill this vital function.

1-1-2. How backup software functions

A simple data backup method is to copy specific data to a client's recovery media (such as CD-ROM). However, such a method is totally impractical and prohibitively expensive to administer in computer installations that are in 24x7 use. This is because it is virtually impossible to allocate systems and network time to the backup operations required to keep track of frequent changes in stored data.

The following features are often listed as requirements for efficient backup software.

- **Support for High-capacity Storage**
Sufficient space for backup against daily increases in data capacity.
- **Saving Backup Time**
Backup at a high transfer rate to save time in round-the-clock operations.
- **Multi-Platform Support**
Support for the wide range of platforms required to meet the diversity of an individual user's system architecture.
- **Provision of a Convenient Management System**
Remote administrative capabilities respond to the needs of the broadband network society.
- **High Maintainability**
Full vendor support.
- **Cost Savings**
A cost-effective back-up system as insurance against data loss.

To fulfill the above user requirements, Sony highly recommends PetaBack software as one of the most powerful quick-to-recovery backup solutions on the market.

1-2. How PetaBack™ software functions

PetaBack software is high-speed backup software that has various features and functions to meet users' demands. It reduces the time required for daily backup and performs an immediate automatic recovery in the event of a fault, enabling business enterprises to concentrate on their routines without having to be concerned with data losses. Using PetaBack software, users can save time and administrative expenses by minimizing their IT investment in backup protection.

1-2-1. Features of PetaBack software

PetaBack software provides the following important features to fulfill the needs required by the current and future data backup markets.

Supports large storage capacities

PetaBack software supports various types of tape drives and tape libraries. It can integrate a high-capacity backup system either with an AIT library or a DTF library. Using an AIT-2 library, the maximum storage capacity is up to 1.5 TB (native). When DTF-2 tape drives are integrated into a PetaSite® library system, the maximum storage capacity rises to 29.2 PB (compressed*) and 11.2 PB (native).

Saves backup time

PetaBack software maximizes the data transfer ability of tape devices, realizing the extremely high backup rates of 36 MB/s (compressed*) and 24 MB/s (native) with DTF-2 devices, and 31.0 MB/s (compressed*) and 12 MB/s (native) with AIT-3 devices.

Supports a wide range of platforms

As PetaBack software supports a wide range of platforms, including most UNIX® and Windows® operating systems, it flexibly meets a user's individual application needs. PetaBack software also enables several platforms to be combined into a single backup system, even if they are applied in a heterogeneous environment. Therefore, most users can easily add PetaBack software to their existing systems.

Provides convenient management systems

PetaBack software provides an intuitive and easy-to-operate GUI, which is based on a Java™ Applet that has remote control capability. Users can operate and manage the system from a remote host through a Web browser.

One-contact-point support and maintenance

As Sony supplies both the hardware, such as tape drives and libraries, and the backup software, a well-matched data backup system is easy to construct. It also means just one contact point for complete system support and maintenance from an investment point of view.

Cost Savings

PetaBack software supports server/client systems. Each remote client host sends files for backup over a network to the tape device connected to the server. It can integrate multiple hosts connected to a network in a single backup system. This means that tape drives and libraries are not required for each host that requires backup. Users can reduce system configuration costs and can also easily extend their existing systems in combination with PetaBack systems.

In addition to the above points, PetaBack software offers the following additional features to provide a safe operating environment for users.

Supports SAN (Storage Area Network)

LAN-free backup is available in a SAN environment when using PetaBack software, which results in a much faster, flexible backup system. Using LAN-free backup, backup can be carried out from a remote host without passing via a server or LAN network. This dramatically reduces network and server workload. Even in a heterogeneous environment, where various platforms run together, a single backup system can be constructed.

* at 2.6:1 with ALDC

1

PetaBack™ Software Concept Overview

Supports a NAS server

By supporting NDMP Ver.3 (Network Data Management Protocol), high-speed backup of the files on a NAS (Network Attached Storage) server connected to tape devices through fibre channel can be executed.

Note: As of May 2002 backup via FC bridge for AIT-2 drives had not been verified

Compatibility with HSM software

PetaBack software can share drives and libraries with Sony PetaServe® HSM (Hierarchical Storage Management) software, enabling advanced and flexible storage systems to be readily constructed.

Oracle8i™. Online Backup Function

With the Oracle8i Online Backup Function, the use of Oracle Recovery Manager enables online backup from an Oracle® database - even when the system is in operation.

1-2-2. Main Characteristics

Listed below are the main characteristics of the PetaBack software.

Characteristics	Benefits
High-speed data transfer	This can maximize the transfer capability of a tape device, e.g. at 24 MB/s (native) using Sony DTF-2 tape drives and at 12 MB/s (native) using Sony AIT-3 drives. Of course, the transfer rate can be substantially increased by adding multiple tape drives to a library.
Remote hosts' backup and restoration	Multiple hosts for backup connected to a network can be administered by a single backup system. Thus each host no longer needs its own tape drive or library, which potentially lowers the total system cost.
Differential backup and incremental backup	Backup type can be selected according to a user's needs, resulting in shorter overall backup restoration times.
Automatic backup using library	An automatic backup system without human intervention.
Unrestricted file size for backup	Operators can concentrate on their job with no need to be aware of the file size during backup.
Unrestricted length of file path name	No need to be aware of length of file path name during backup.
Java™ Applet-based GUI	Allows remote system administration through a Web browser.
Command-line operations	Script and batch files can be created to back up and restore according to user's environment, with combined use of UNIX® shell and Windows NT® command prompt.

Characteristics	Benefits
Generation administration of execution history	The period from the latest full backup to the next full backup is referred to as one generation. Users can designate a specific generation when restoring a file.
Administration of multiple libraries	In order to increase efficiency, libraries can be shared by different users for different jobs at the same time.
Reports of automatic backup results sent by e-mail	Backup results are known immediately.
High-speed recovery to the latest backed-up disk status	Quickly recovers lost files to the latest backup-up status before data loss.
High-speed recovery when PetaBack™ system file crashes	Quickly recovers to a proper operating environment in the event of a disk crash.
Restoration to a host other than the backup host	Backed-up data can be retrieved from other hosts even in the event that the backed-up host has crashed.
Multi-volume function	Multiple volumes can be spanned without the user needing to intervene.
Backup and restoration in parallel	Shortens the total time required for backup and restoration.
Highly flexible backup schedule setting	Setting of backup schedule can be customized according to user's unique requirements.
Immediate backup	In an emergency, a user can explicitly implement manual backup immediately.
Support for a wide range of platforms <i>(Please see "System Components Supported" on page 6)</i>	Users that have no need to change from their current Operating System(s) should find the PetaBack software easy to incorporate or introduce into their data backup infrastructure.
External management of backup volumes	Volumes can be picked out from a tape library for external management, and also used in other PetaBack systems.
Information on backup files shown in list form	Via either the command line, or by the GUI, files can be shown in list form.
LAN-free backup in a SAN environment	High-speed data transfer can be carried out via remote host backup and restoration, because a LAN is not used during data transfer.
System backup using a stand-alone drive	A library is not required, so a small PetaBack system can be constructed economically with a backup system that only employs a stand-alone drive.
Easy installation and setup	Dialogue-style installation and setup tools allow users to easily construct their environments by simply following the messages.
Support for Japanese file names	Files with Japanese file names can be backed up and restored, and appear on GUIs.

1-3. System components supported

Listed below are the operating systems and system components supported by PetaBack software, as of March 2002.

Operating systems

(As of March 2002)

Solaris® 2.5.1*
Solaris 2.6
Solaris 7
Solaris 8
HP-UX® 10.20* (for client function only)
HP-UX 11.00
IRIX® 6.5.x**
Red Hat® Linux® 7.1(for client function only)
Windows NT® 4.0 (for client function only)
Windows® 2000 (for client function only)

* Solaris 2.5.1 and HP-UX 10.20 will not be supported by PetaBack software versions higher than ver. 3.00.

** IRIX 6.5.8, IRIX 6.5.11, and IRIX 6.5.12 are not supported.

For the latest information, please contact your nearest Sony office/dealer.

Libraries

(As of March 2002)

PetaSite® system 8400 Series	Sony DTF Tape Library (DMS-8400B, DMS-8400D, etc)
PetaSite system 150 Series	Sony DTF Tape Library (DMS-B150L, DMS-210S)
PetaSite system 80 Series	Sony DTF Tape Library (DMS-B80L, DMS-110S)
DMS-B35	Sony DTF Tape Library
LIB-304	Sony AIT Tape Library
LIB-162	Sony AIT Tape Library
TLS-4210	Qualstar™ AIT Tape Library

Drives

(As of March 2002)

GY-8240	Sony DTF-2 Tape Drive
GY-2120	Sony DTF-1 Tape Drive
GY-10	Sony DTF-1 Tape Drive
SDX-300C/RS	Sony AIT-1 Tape Drive
SDX-400C/RS	Sony AIT-1 Tape Drive (5-inch Internal Drive)
AIT-S70	Sony AIT-1 Tape Drive (External Drive)
SDX-500C/RS	Sony AIT-2 Tape Drive (5-inch Internal Drive)
AIT-S100/D	Sony AIT-2 Tape Drive (External Drive)
SDK-700C/RS	Sony AIT-3 Tape Drive (5-inch Internal Drive)
AIT-S200	Sony AIT-3 Tape Drive (External Drive)

1-4. Products

The software and licenses required for PetaBack™ system operation are available as packages. Individual licenses are also available as options for system expansion.

All PetaBack software packaged products include a library license that covers one host, and a FZC-PBK5 client license that covers up to five hosts. So, the total host number is up to six. For the latest information, please contact your nearest Sony office/dealer.

PetaBack software packaged products

FZC-BKPE	Backup Software Standard Pack (E) For an AIT library or stand-alone AIT/DTF drive
FZC-BKPM	Backup Software Standard Pack (M) For a medium-size library system
FZC-BKPL	Backup Software Standard Pack (L) For a large library system

System expansion is available by adding the following licenses.

PetaBack software licenses

FZC-PBK5	PetaBack Client License (5 hosts)
FZC-DB1	Oracle® Database Online Backup License

Library licenses

FZC-LBE	For Sony libraries: LIB-304 LIB-162 For Sony drives: SDX-300C SDX-400C SDX-500C GY-10 GY-2120 GY-8240 For Qualstar™ library: TLS-4210
FZC-LBM	For Sony libraries: DMS-B35 DMS-B80L
FZC-LBL	For Sony libraries: DMS-B110S DMS-B150L DMS-B210S DMS-8400B and DMS-8400D
FZC-LBX1	For Sony libraries: DMS-EX150L DMS-EX210S DMS-8400D
FZC-LBX2	For Sony library console: DMS-8400C

Chapter 2

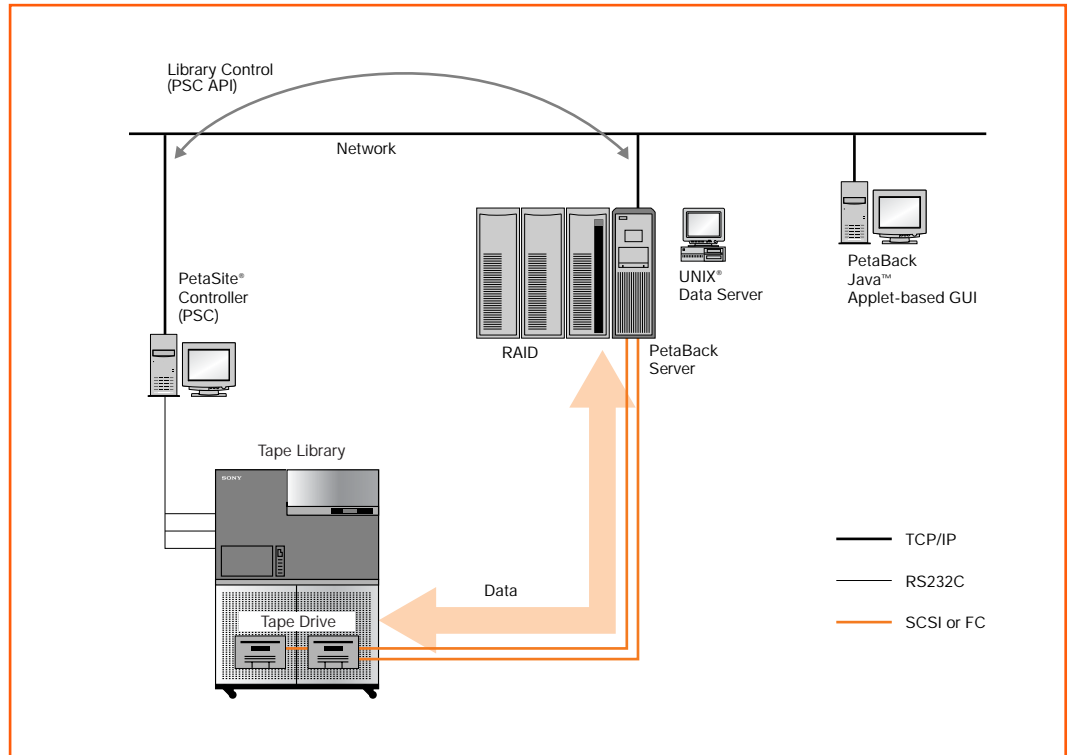
PetaBack™ System Configurations

PETA BACK

2

PetaBack™ System Configurations

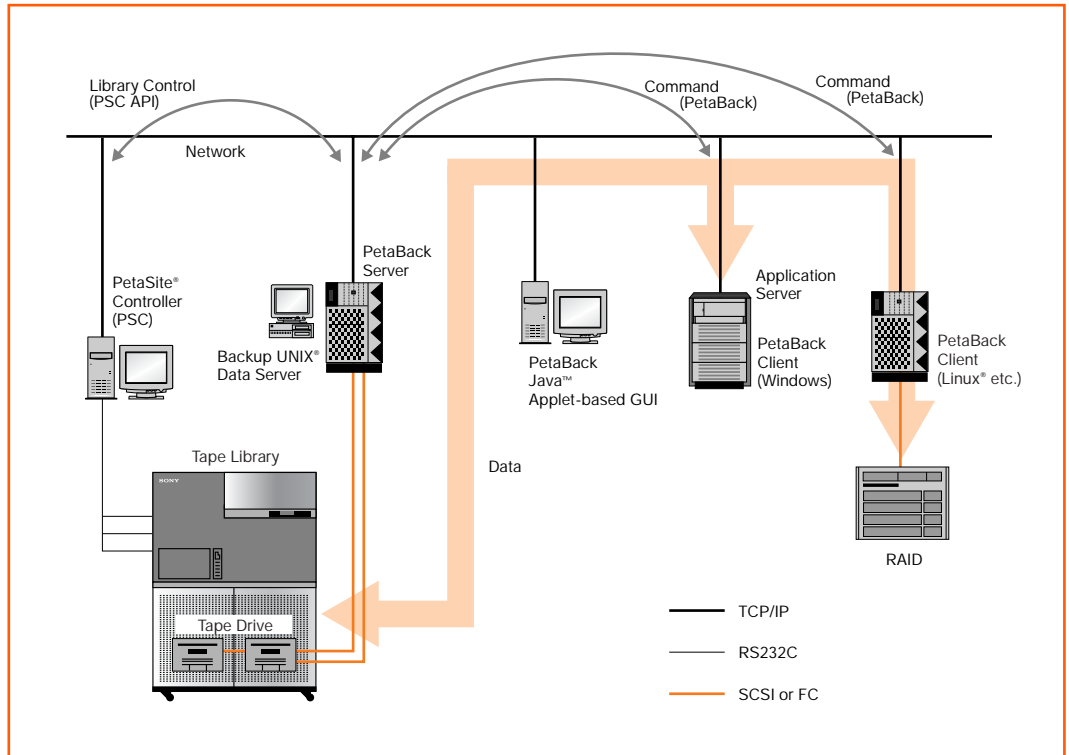
2-1. Stand-alone server backup system



Above is a typical example of a high-capacity data server for backup, directly connected to a tape drive.

In this example, a UNIX® data server on which the PetaBack software server module has been installed is called a "PetaBack Server". It controls tape library/drives and operates the backup/restoration of data in the server/RAID. The tape library is controlled via a network (PSC API) or SCSI. Using the PetaBack software Java™ Applet-based GUI on the PetaBack software server, or on a client machine such as that running a Windows® application, the System Administrator and other users can manage the set-up of backup-group creation, schedule modifications, etc., and also can operate backup or restoration functions.

2-2. Client server backup system



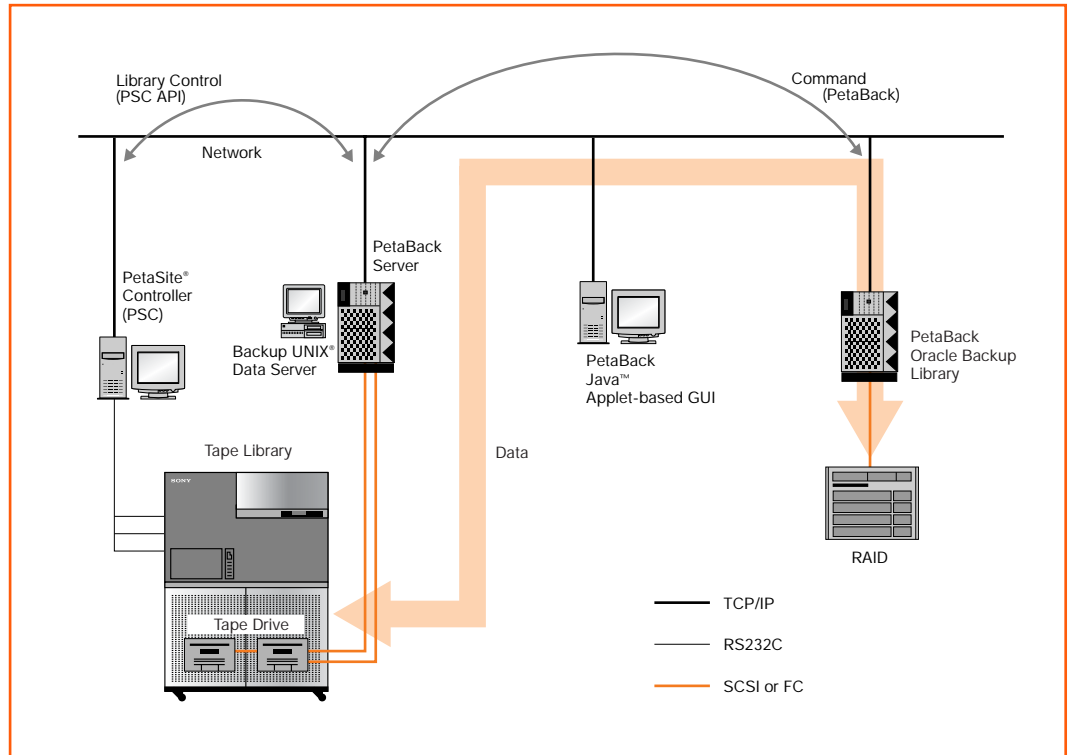
Shown above is an example of how multiple high-capacity data servers can be backed up over a network.

Here, PetaBack™ software is installed on a UNIX® server for backup, while the PetaBack software client module is installed on those data servers where PetaBack client data is to be backed up or restored. Even in a heterogeneous environment where both UNIX and Windows® OS clients exist, a single, unified backup system can be constructed.

2

PetaBack™ System Configurations

2-3. Oracle® database online backup system

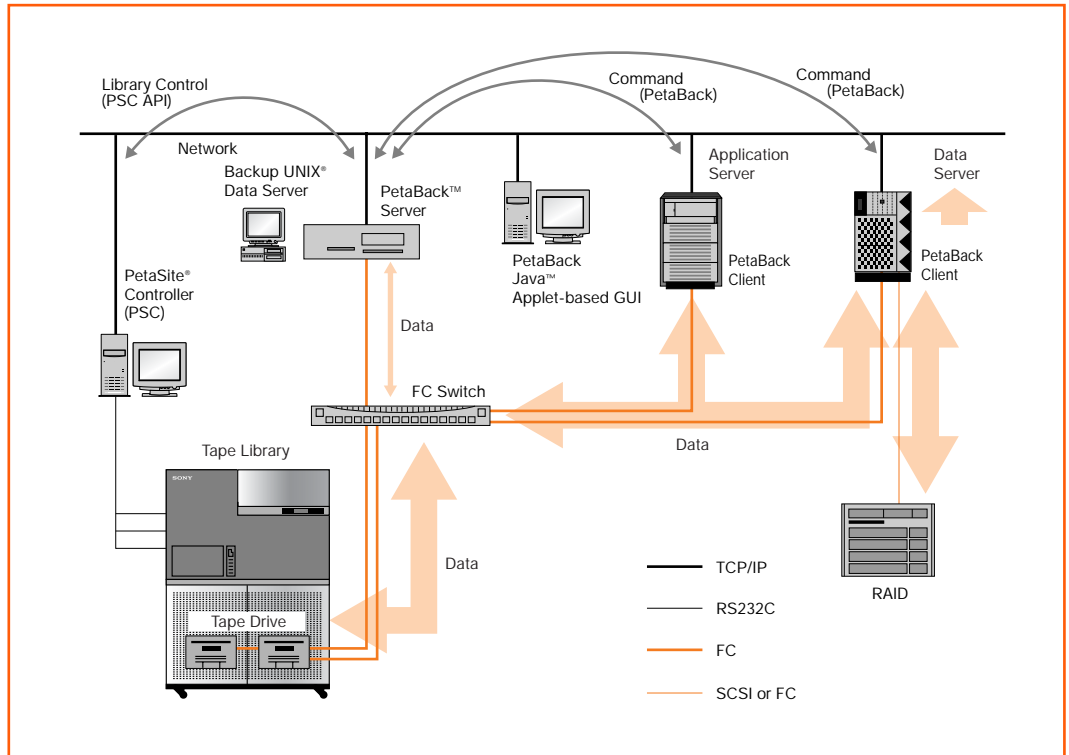


This diagram shows an example of online backup over a network using an Oracle database.

When using this configuration, prior to backing up an Oracle Database server, users need to install the optional PetaBack Oracle Online Backup software to the host running Recovery Manager. Controlled by the Oracle Recovery Manager, the Oracle database is backed up on line. The Recovery Manager also administers the backup generation and manages files.

Note: The Online Backup function does not support a SAN as of March 2002.

2-4. LAN-free backup system in a SAN environment



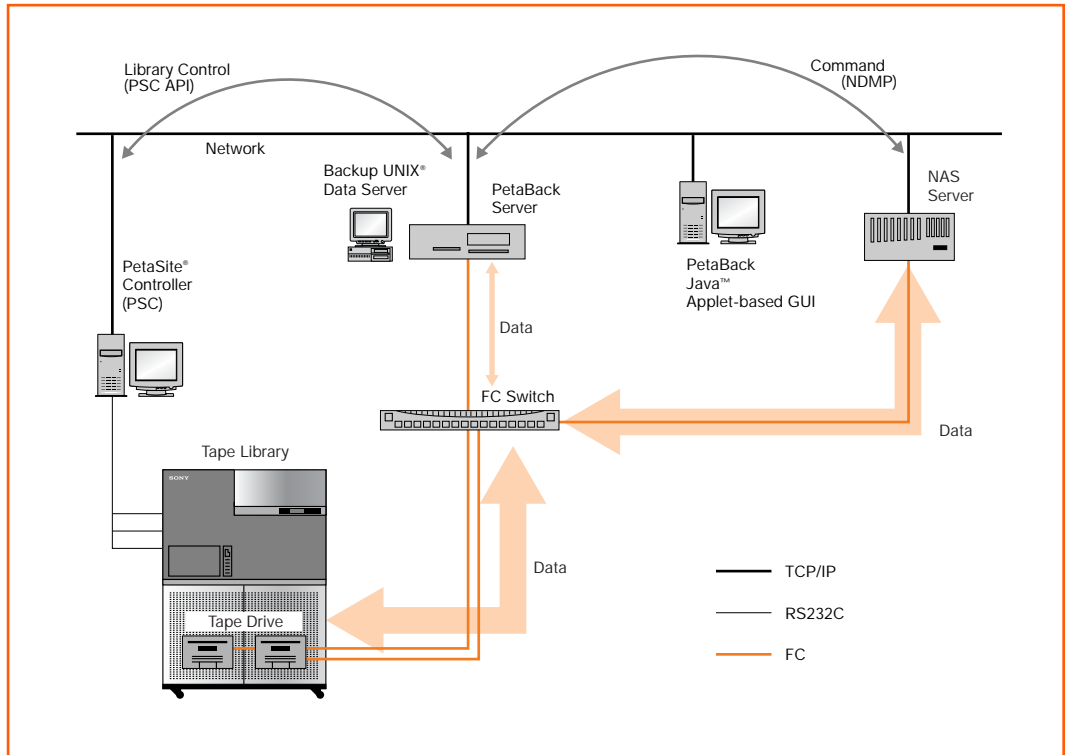
The above diagram shows that multiple high-capacity data servers can be backed up via a SAN (Storage Area Network) environment.

In this example, backup data is read and written on tape drives directly through fibre channel, without passing via a LAN (Local Area Network). This not only reduces network workload but also overhead data transfer on the backup server, maximizing the performance of the tape drives and disks installed in the server or the RAID. This application enables a single tape library to back up several high-capacity servers at very high speed, even if they are in a heterogeneous environment. The arrow shown between the PetaBack Server and the FC switch indicates data backup of the PetaBack server itself.

2

PetaBack™ System Configurations

2-5. Backup system in a NAS server environment (1)



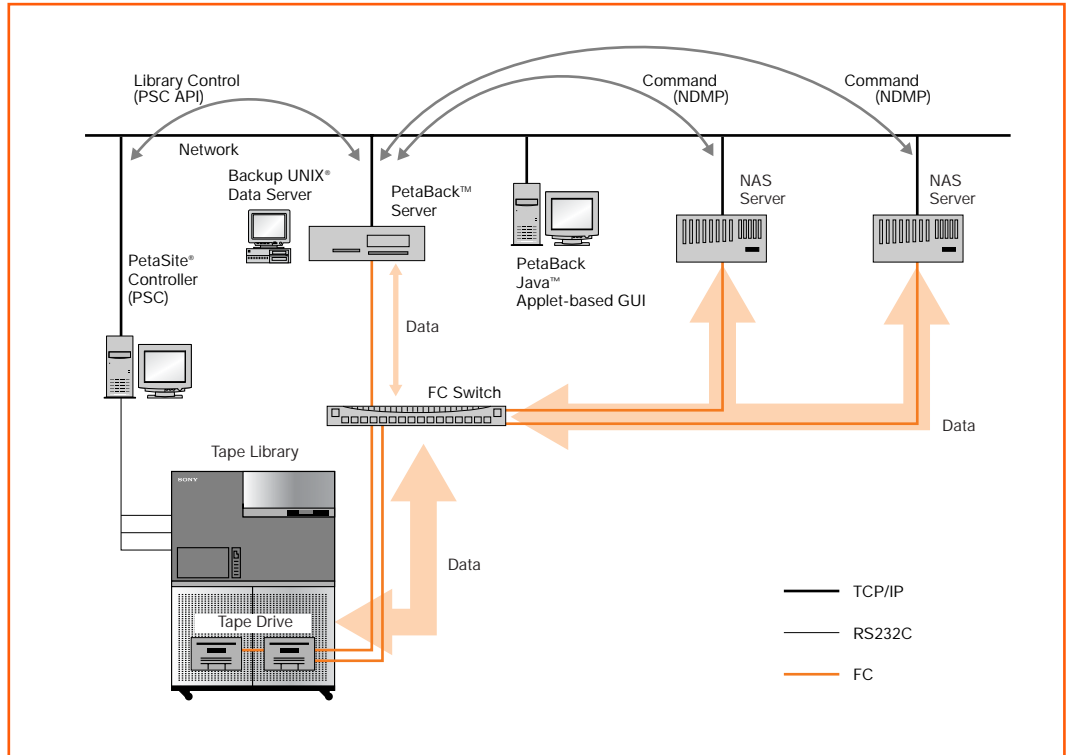
This is a typical example of how a NAS (Network Attached Storage) server can be backed up via a SAN.

Here, PetaBack software supports NDMP (Network Data Management Protocol) version 3, with which the backup and restoration of NAS servers can be implemented. Thus a PetaBack server can read and write data on a NAS server directly to and from tape drives via a SAN (FC) by sending NDMP commands to the NAS server.

The PetaBack software cannot be installed on these NAS servers, since most NAS servers use proprietary operating systems.

Note: This function is only available when a PetaBack system host uses Solaris2.6, Solaris7, or Solaris8 systems, and only supports full backup and full restoration.

2-6. Backup System in a NAS server environment (2)



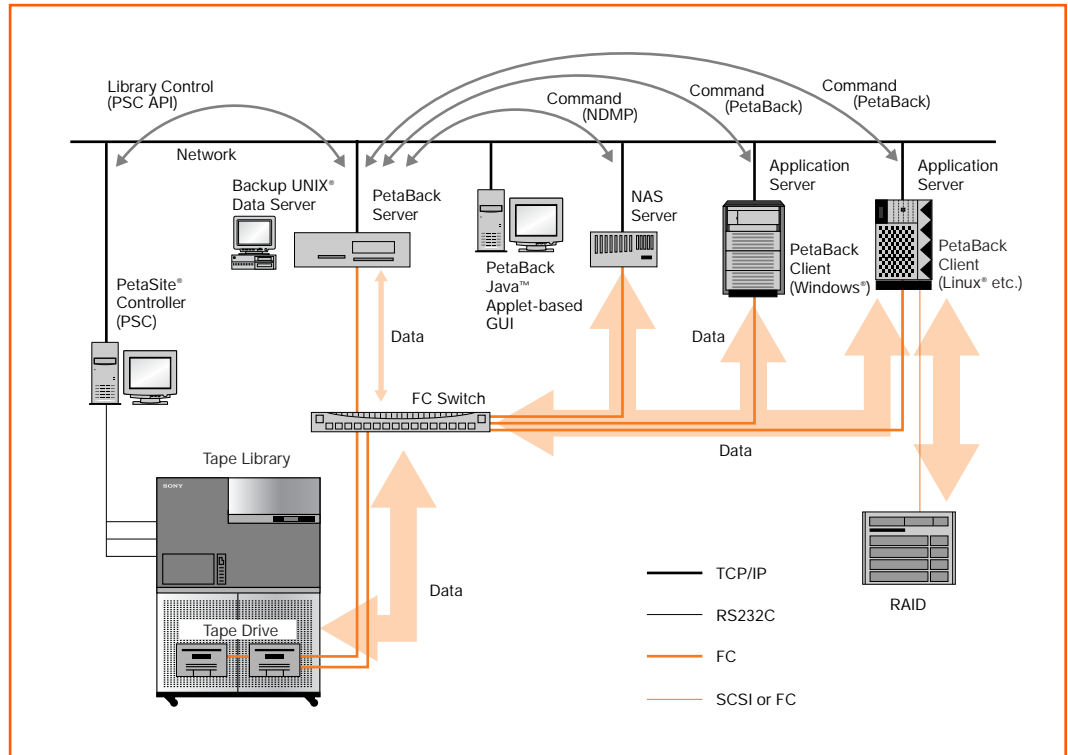
Here is a further example showing that multiple NAS servers can be backed up via a SAN (FC).

Because a NAS server is a kind of client host for the PetaBack system, a single tape library can back up multiple NAS servers.

2

PetaBack™ System Configurations

2-7. Backup system in a heterogeneous and NAS server environment



Above is an example of a typical LAN-free (SAN) backup system example in a heterogeneous and NAS server environment.

With the PetaBack system, a single tape library can back up all of the servers in a NAS server environment in addition to those in a heterogeneous environment where various operating systems exist, such as UNIX® and Windows® operating systems.

Chapter 3

Overview of Functional Concepts



3

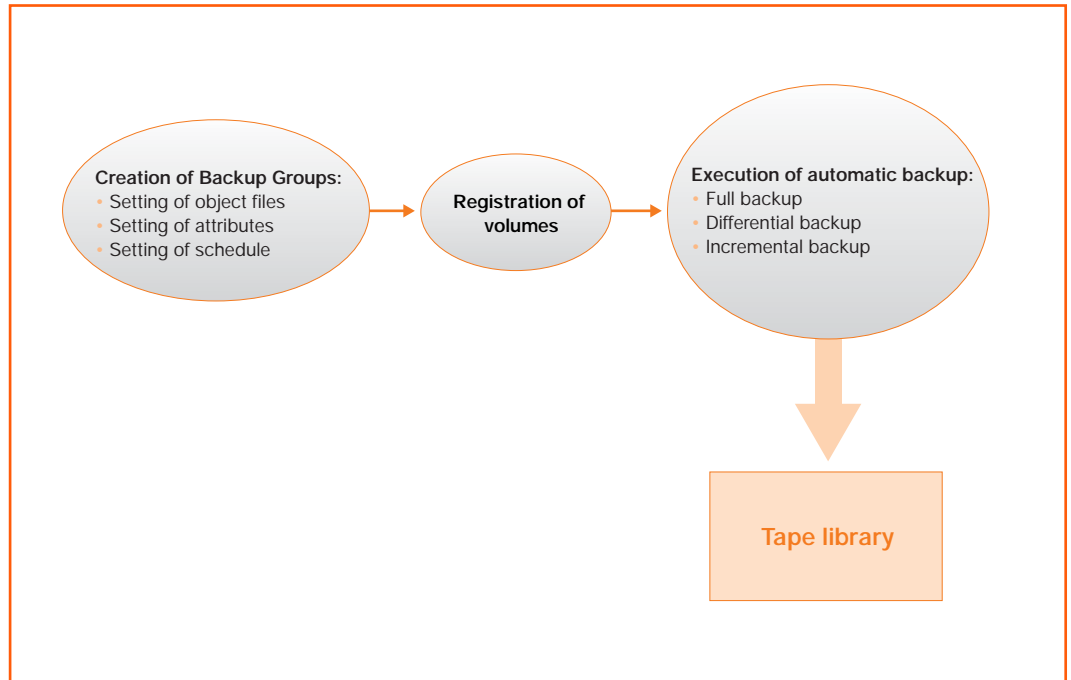
Overview of Functional Concepts

3-1. Basic operation

In PetaBack™ software, backup operation is administered by a units called "Backup Groups". To implement automatic backup for the first time, users must create Backup Groups and register volumes. Automatic backup is then carried out

according to these specified settings. The software supports three kinds of backup: Full Backup, Incremental Backup, and Differential Backup. A backed-up file can be recovered using the Restore function.

Automatic backup flow



Volume

A "volume" is a theoretical concept used to express a physical recording medium. Usually, it means memory, disk, or tape. However, in the context of PetaBack™ software, it specifically means magnetic tape.

Full backup

Unconditionally and fully backs up "Designated Files" (i.e. files specified for backup).

Incremental backup

Backs up all the Designated Files that have changed since the last backup. During restoration, all files that were backed up in the latest full backup and subsequent incremental backups are restored.

Differential backup

Backs up all the Designated Files that have changed since the last full backup. As differential backup always backs up all the Designated Files that have changed since the last full backup, it generally takes longer to implement than incremental backup. However, the restoration time required is generally shorter than incremental backup because it only has to restore those Designated Files that were backed up in the latest full backup and the latest differential backup.

The following table shows the operation examples:

	Designated files existing on hard disk	Backed up files in incremental backup	Backed up files in differential backup
Before backup	A, B		
1st backup (full backup)	A, B	A, B	A, B
2nd backup (added C)	A, B, C	C	C
3rd backup (added D)	A, B, C, D	D	C, D
4th backup (changed C to C')	A, B, C', D	C'	C', D
5th backup (added E and F)	A, B, C', D, E, F	E, F	C', D, E, F

The following is a comparison chart of backup and restoration between incremental backup and differential backup.

	Incremental backup	Differential backup
Time required for backup	Short	Long
Time required for restoration	Long	Short

3

Overview of Functional Concepts

3-1-1. Backup Groups

A "Backup Group" is a theoretical concept that defines a backup object. The backup object is defined as the pairing of a host and Designated Files. According to users' requirements,

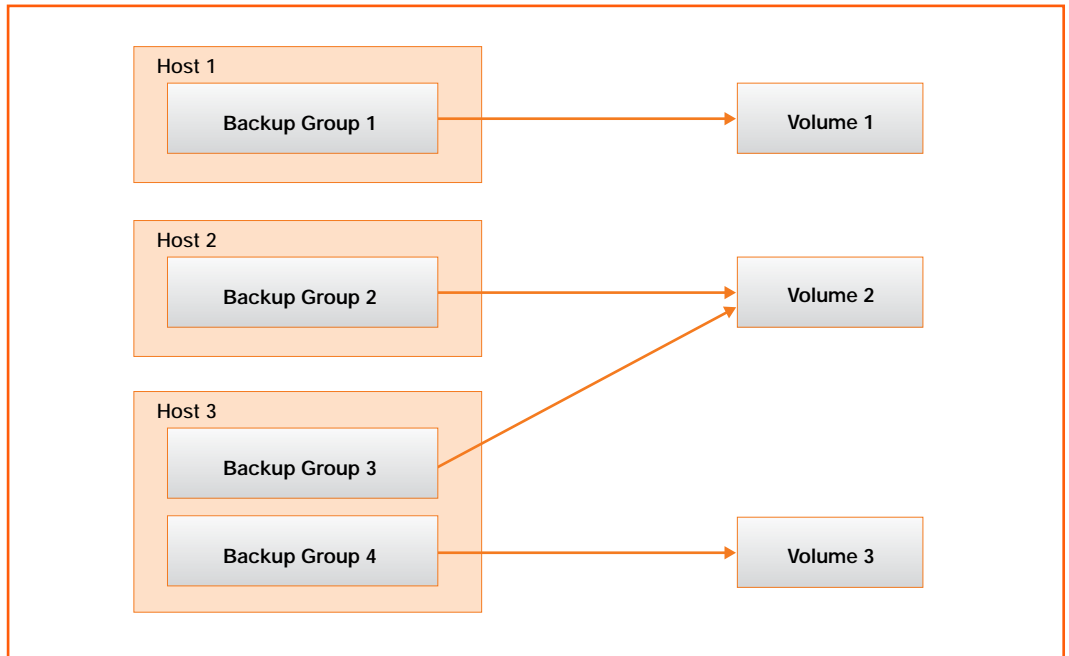
Designated Files can be categorized and a Backup Group can be defined against each of them.

A backup group can be defined as below:

- A Backup Group consists of a host and multiple Designated Files.



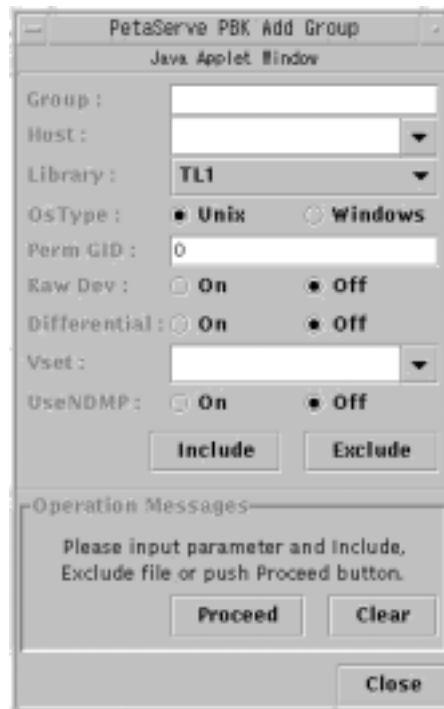
- Multiple Backup Groups can be created against a host. Also, a volume can be shared between different Backup Groups. This function is called "volume set function," with which volumes can be used more efficiently.



3-1-1-1. Creation of Backup Groups

There are two methods of creating new Backup Groups:

- Creating Backup Groups from the GUI.
- Creating Backup Groups with the *addgroup* command (a subcommand of the *pbkadm* command).



Backup Group creation menu

3

Overview of Functional Concepts

The following items can be specified during registration:

- **Group**

- **Group name**

- Up to 31 alphanumeric digits including hyphen (-) and underline (_) can be used for group names. However, PetaBack™ software will not recognize group names starting with hyphen (-) or underline (_).

- **Host**

- **Host name required for backup**

- Once the number of hosts registered to a Backup Group exceeds a total of six, the user must obtain additional licenses (see 1-4).

- **Library**

- **Library name**

- A library is an automatic device accomodating drives and volumes, and may include an auto-changer or a picker to move the volumes.

- **OS type**

- **OS (UNIX®/Windows®; default setting is UNIX):**

- Designates the type of OS for the host required for backup.

- **Perm GID**

- **A (UNIX®) *gid* attributed to the general users permitted to execute restoration.**

- Usually, restoration of backed up files cannot be executed by anyone other than by the root user. However, when UNIX's *gid* is set, the general users specified in this *gid* can also execute restoration.

- **Raw Dev**

- **Raw Device File ON/OFF (Default value is OFF)**

- A raw device file can be a Designated File , resulting in high-speed disk backup. During raw device backup, the corresponding file system should be unmounted.

- **Raw device**

- In disk devices, there are block devices for OS administration and their corresponding character devices. It allows high-speed read/write, because character device can directly access hardware without using the buffering within the kernel. This character device is called a "raw device".

- **Differential**

- **Differential backup function ON/OFF (Default value is OFF)**

- When it is ON, differential backup will be implemented. When it is OFF, incremental backup will be implemented.

- **Vset**

- **Volume set name**

- A Backup Group that is configured with a Volume set will use the volumes attributed to the volume set during backup execution. Volumes can be shared between multiple Backup Groups by designating the same volume set to multiple Backup Groups.

- **UseNDMP**

- **NDMP function ON/OFF (default value is OFF)**

- When it is ON, NDMP is used for backup.

- **Include**

- Designates the full path name of a Designated File.

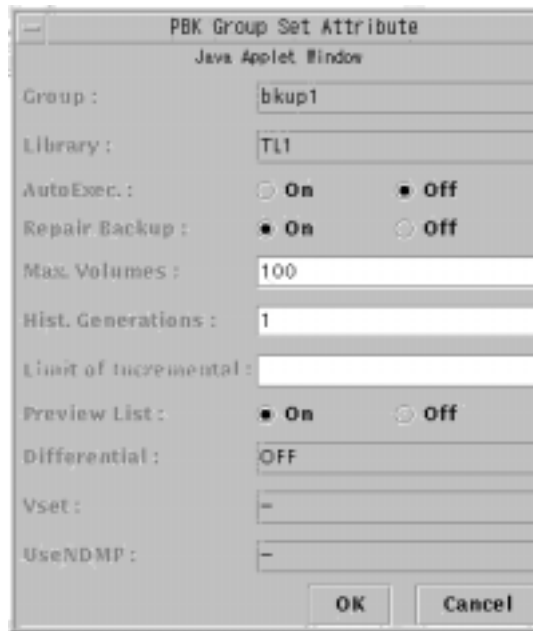
- **Exclude**

- Designates the full path name of each file or directory excluded from backup.

3-1-1-2. Reference to and modification of static attributes specified in Backup Groups

Several static attributes are defined in each Backup Group. Some of these attributes can be modified in accordance with a user's environment after creation of the Backup Group. There are two methods for modifying the static attributes specified in the Backup Groups as shown below:

- Reference and modification from GUI.
- Reference and modification using *lsgroup* and *setgroup* commands (sub commands of *pbkadm* command).



Backup Group attribution setting menu

3

Overview of Functional Concepts

Static attributes defined in Backup Groups.

The following static attributes are defined in each Backup Group:

- **Group**

Backup Group name.

- **Library**

Library name.

- **Host**

Host name required for backup.

- **Auto Exec - Automatic backup function ON/OFF. (The default value is ON)**

On each host required for backup, *osmnightly* script automatically implements backup with the *cron* function by creating *osmnightly* script at a specific time. For Backup Groups with the setting of OFF, the *osmnightly* automatically excludes them from backup.

- **Max Volumes - The maximum number of volumes that can be allocated (the default value is 100).**

The maximum value for the number of volumes that can be allocated to each Backup Group is defined and the total number allocated cannot exceed this value.

- **Hist Generations - (the default value is 1).**

The maximum number of administrative generations for the past execution history database of each Backup Group is user definable. The execution history database updates a Backup Group generation every time a full backup is performed and deletes the past execution history database. When restoring backed up files, the user can designate restoration to a specific generation within the execution history database.

- **Cycle (wday)**

The day of the week for automatic full backup (the default value is Sunday).

- **Repair Backup - Repair backup function ON/OFF (the default value is ON).**

Each time a generation is updated in the execution history database, a new volume is used for backup. If full backup is skipped for some reason during automatic backup, the execution history database will not be updated and the same volume will be kept for backup until it is completed. Once a volume is completed in this manner, another free volume will be automatically allocated until no free volumes are available. The "Repair Backup" function prevents the system from exhausting all available volumes by enabling the user to preset a maximum number of incremental backups. Once this number is reached, a full backup will be implemented and the generation of the execution history database will be forcedly updated, even on the day when an incremental backup should be carried out. When the setting is ON, the Repair Backup function is in action.

- **Limit to number of Incremental Backups**

During the creation of Backup Groups using PetaBack™ software, the default schedule specifies that full backup be implemented once every week on a specific weekday, and incremental backup on all other weekdays.

In “default schedule setting,” Repair Backup is only implemented if the last full backup was conducted more than a week ago, or if the number of incremental backups conducted since the last full backup exceeds the preset value. The incremental backup default value is 30.

- **Perm GID**

(UNIX's) *gid* designates the general users allowed to execute restoration.

- **Backup File Preview List - function ON/OFF
(The default value is OFF)**

When this function is ON, backed up file information can be immediately shown from command-line or GUI in a list form.

- **Differential**

Differential backup function ON/OFF.
(The default value is OFF).

- **Vset**

Volume set name.
(The default is blank. It is used to set up the volumes attributed to the volume set.)

- **Use NDMP**

NDMP function ON/OFF.
(The default value is OFF).

3-1-1-3. Volume allocation for Backup Groups

When a Backup Group is short of volumes, PetaBack software has two methods of getting new volumes:

- Users explicitly allocate new volumes to a Backup Group.
- PetaBack software automatically allocates new volumes to Backup Groups when necessary.

It is necessary to confirm whether or not the free volumes are sufficient for the above two cases. If they are not sufficient, the user must add the necessary number of free volumes.

Free Volumes

Free volumes are new volumes that have just been labeled and are not allocated to any Backup Groups.

There are two methods of explicitly allocating volumes:

- Through the GUI.
- Using commands (*pbkadm* command, *addvol* sub command).

3

Overview of Functional Concepts

3-1-1-4. Customization of automatic backup schedule

The following are the two methods of confirming and modifying the automatic backup schedule in PetaBack™ software:

- Through the GUI.
- Using commands (*pbkadm* command *addschr/isschr/rmschr* sub command).

The user can separately specify the following settings for full backup and incremental (or differential) backup in the automatic backup schedule:

- A specific day (e.g. the 25th of every month)
- The last day of every month
- A specific day of the week (e.g. every Sunday)
- A specific day of the week of every month (e.g. the first Saturday of every month)



Backup schedule setting Menu

3-1-2. Backup

Backup can be implemented when the creation of Backup Groups and registration of free volumes is completed. The backup function in PetaBack™ software uses a library which automatically implements the high-speed transfer of Designated Files to the tape device as a means of secondary storage. Consequently, during a certain period, the secondary storage is holding the same file data as that on the disk.

3-1-2-1. Designated Files

The two types of files can be backed up during one backup process:

- Files and directories designated by users for backup (see 1-4).
- PetaBack software's system file (system backup).

System file backup in PetaBack software is used to recover the system to the condition it was in immediately prior to a system file fault caused by a system crash.

3-1-2-2. Backup levels

Full backup, incremental backup, and differential backup are the three levels for backup.

3-1-2-3. Backup types

There are two types of backups:

- Automatic backup by *osmnightly*.
- Immediate manual backup.

Automatic backup is used in normal operations, while manual backup is usually used in an emergency.

3-1-2-4. Backup retry function

If after completion of an automatic or manual backup, the PetaBack software determines that a Designated File was not backed up for some reason (i.e. it was being updated during the backup process), this function automatically retries backup as necessary.

3-1-3. Restoration

Restoration is a function to restore backed up files to the disk of a host. In the default setting, the files are restored to the host from which they were backed up. Even if the files have been accidentally lost by a disk crash or system malfunction, the "Restore" function can restore these files.

3-1-3-1. Files required for restoration

The two types of files that can be restored are:

- Files and directories designated by the user for restoration.
- PetaBack software's system files.

The latter is performed when a PetaBack system file has been accidentally lost for some reason.

3-1-3-2. Restoration modes

There are two modes for restoration:

- A mode that restores all the files existing on the disk at the time of the latest backup (restores to the latest backup).
- A mode that restores only those files designated as required for restoration (restoration designated by user).

The former is called the "True Image Restore" function.

3

Overview of Functional Concepts

3-1-3-3. Restoration to a specific generation

When the execution history database is preserved in multiple generations, any of them can be designated to be restored.

3-1-3-4. Directory selections for restoration

Files can be restored under any directory by designating a directory for restoration.

3-1-3-5. Restoration to another host

Files can be restored to a host other than the one to which the file was last backed up. This function is called the "Cross Restore" function.

3-1-4. Emergency situations

3-1-4-1. When the file system for backup has crashed

Recovery is performed in the following steps:

- The Backup Group necessary for restoration is specified.
- The files existing on the disk at the time of the latest backup are restored.

3-1-4-2. When the install disk in the PetaBack™ server has crashed

Recovery is carried out in the following steps:

- The PetaBack software is re-installed onto the server host.
- The volume names in the *var/adm/petaback/pbklastvol* file on the server host are confirmed. The most recently backed up PetaBack system file has to be input onto the the volume.
- The volume is loaded into a tape drive and the */var/adm/petaback/psvrestore* script implemented.
- The PetaBack service is restored.

3-1-4-3. When the OS disk has crashed

Recovery is implemented via the following steps:

- The operating system is re-installed.
- The steps mentioned in section 3-1-4-2 are repeated.

3-2. Function explanation

3-2-1. High-speed data transfer

PetaBack™ software enables users to maximize the high-speed data transfer capability of DTF and AIT tape drives.

The following numbers are actual measurements:

Format	Transfer Rate (Native)	Transfer Rate (Compressed)*
Sony DTF1	12 MB/s	18.0 MB/s
Sony DTF2	24 MB/s	36.0 MB/s
AIT-1	4 MB/s	10.4 MB/s
AIT-2	6 MB/s	15.6 MB/s
AIT-3	12 MB/s	31.0 MB/s

*The above compressed transfer rates are calculated based on a compression ratio of 2.6 to 1. The transfer rates vary depending on the actual compression ratio.

3-2-2. Remote host backup and restoration (normal backup)

PetaBack software is a server client style backup software. Each remote client host can backup Designated Files to the backup server. Thus, multiple hosts connected over a network can be administered as a single backup system, and there is no need to connect a separate tape library or tape drive to each backup server.

3-2-3. Remote host backup and restoration (LAN-free backup)

PetaBack software allows LAN-free backup in a SAN environment. In LAN-free backup, massive amounts of backup data can be transferred without passing through a LAN network and a server because the remote host directly transfers data to/from the tape device via fibre channel. Thus, network and server workload can be dramatically reduced and high-speed backup through a remote host can be achieved.

3-2-4. Automatic backup using a library

PetaBack software can perform automatic backup using a tape library. In this case, data cassette loading and unloading can be automatically carried out without human intervention. Backup across several data cassettes is also possible with the PetaBack software multi-volume function.

3-2-5. File size limitations

PetaBack software has no limit to backup file size. Thus it takes full advantage of the maximum file size supported by each operating system. Satisfactory performance with file sizes from 1 byte to up to 1 terabyte has been verified.

3-2-6. File path name length limitations

As the storage space for file path name is variable in the archive format used in PetaBack software, the path name length for individual files is unlimited so that it fully utilizes the maximum file size supported by each operating system.

3

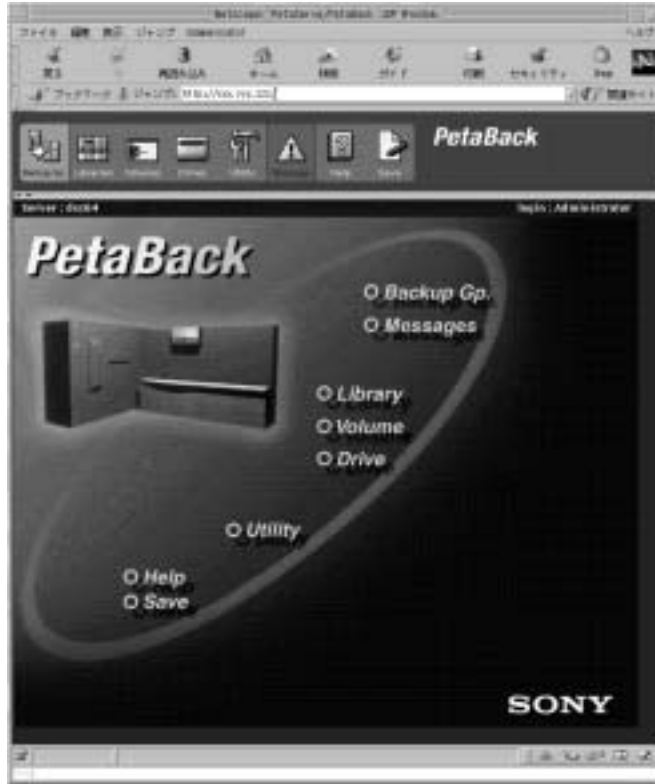
Overview of Functional Concepts

3-2-7. Java™ Applet-based GUI

PetaBack™ software provides an intuitive and easy-to-operate GUI. Users can readily use any of the functions necessary for operations such as the administration of tape volumes and libraries, the setting of the backup schedule, and the

monitoring of system status. Utilizing the Java Applet, system administration can be operated remotely through a standard Web browser.

Examples of menu screens are shown below:



Menu screen

Backup Gp

Administration and monitoring of Backup Groups.

Volume

Monitoring of all the volumes administered by the server.

Library

Monitoring of the automatic library installed within the server.

Drive

Monitoring of all the drives housed in the automatic library administered by the server.

Utility

Taking and browsing logs and changing license configurations.

Messages

Monitoring of error messages that may contain reports on serious problems.



The screenshot shows the 'PetaServe PBK Groups Manager' window. It features a menu bar with 'File', 'Edit', 'View', and 'Help'. Below the menu bar is a table with the following columns: Group, Host, OrType, Status, State, LastAction, Library, AutoExec, and Cycle. The table contains two rows of data:

Group	Host	OrType	Status	State	LastAction	Library	AutoExec	Cycle
Backup1	6004	Unit		SUCCESS	2001/10/...	TL1	OFF	Sun
test	6004	Unit		FAILED	2001/11/...	TL1	ON	Sun

At the bottom of the window, there is a status bar showing the date and time: 'Wed, 28 Nov 2001 22:41:20 (ST)' and a 'Refresh' button.

Operation menu (PetaServe® PBK groups manager)



The screenshot shows the 'PetaServe PBK Group Administration' window. It has a menu bar with 'Set Attribute', 'Assign Volume', and 'History'. The window is divided into several sections:

- Group:** Backup, **Host:** 6004
- Schedule:** A grid for selecting days of the week (Sun-Sat) and specific days (1-31). A 'Revert' button is present.
- Include Files:** A list of files to be included in the backup, such as '/log/11.txt', '/log/12.txt', '/log/13.txt', and '/log/14.txt'. A 'Revert' button is present.
- Exclude Files:** A list of files to be excluded from the backup, such as '/log/15.txt' and '/log/16.txt'. A 'Revert' button is present.
- Buttons:** 'Add...', 'Remove...', 'Refresh...', and 'Event' buttons are located at the bottom right.

Backup setting menu

3

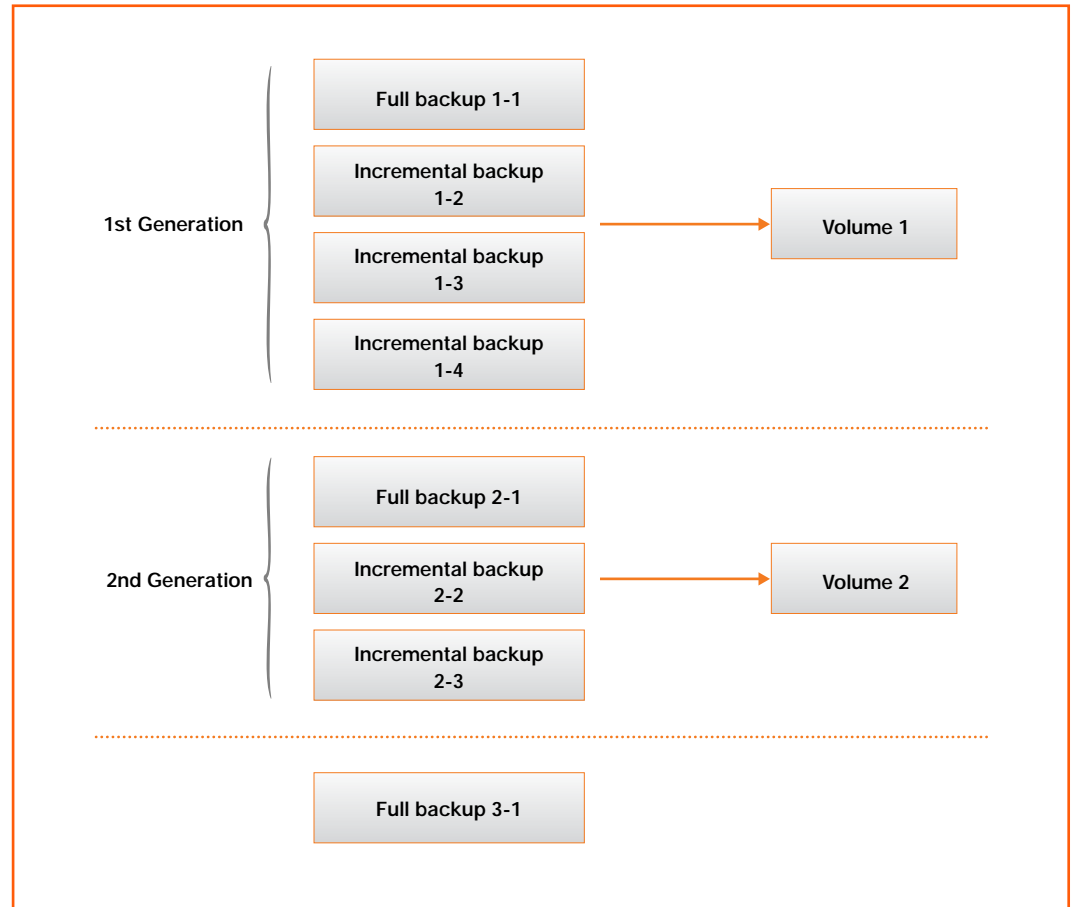
Overview of Functional Concepts

3-2-8. Command-line operations

PetaBack™ software provides a command-line user interface for backup. It can be controlled from the UNIX® shell or Windows NT® command prompt. Users can combine these to create and run a shell script or batch file. In the UNIX® OS environment, the command-line can be controlled remotely through telnet, etc.

3-2-9. Generation administration of execution history

The period from the latest full backup to the next full backup is called one generation. The history of up to 31 generations can be administrated. (The default value is two generations). Users can also designate a specific generation when restoring files.



3-2-10. Administration of multiple libraries

Multiple libraries connected to a PetaBack™ server can be assigned for use as one Backup Groups.

3-2-11. Reporting automatic backup results by e-mail

Each backup host's automatic backup completion results, including any serious faults detected by the software, can be sent to a specific e-mail address.

3-2-12. High-speed recovery to the latest backed-up disk status

If critical data files are accidentally lost, they can be easily recovered. The following is a restoration example.

- **The first backup (Full backup)**

Two files, files A and B, exist on a disk and are retained when the first of a series of backups is executed.

- **The second backup (Incremental backup)**

The user deletes file B on the disk and adds files C and D, and an incremental backup is executed.

- Files A, C, and D exist on the disk.

- **The third backup (Incremental backup)**

File A and D on the disk are changed to file A' and D', File E is added, and an incremental backup is executed.

- Files A', C, D' and E exist on the disk.

- **Restore**

When it is necessary to restore backup files, users can restore to the latest backed-up disk status. In this case, only files A', C, D' and E exist but not B.

3-2-13. High-speed recovery of PetaBack system files after a crash

Backup of the PetaBack database and configuration files must be carried out during each backup. Otherwise, the operating environment will not be immediately recovered following a crash of the hard disk on which the PetaBack software is installed. Following a system file crash, the PetaBack system file can be recovered by executing the *psvrestore* command under */var/adm/petaback*. Even if the whole system has crashed, the system can be recovered by re-installing the operating systems and PetaBack software and executing the *psvrestore* and *pbkrestore* commands.

3-2-14. Restoring to another backup host

A backed up file on a host can be restored to a designated directory on another client host. This function is only available between hosts running UNIX® operating systems.

3-2-15. Volume sharing between different Backup Groups

With PetaBack software, volumes are exclusively allocated to a particular Backup Group or generation in order to reduce any possible effect on the backup data of other Backup Groups or on generations in the event of a volume fault. However, this method lowers volume utilization efficiency. To solve this problem, users can manually implement a mode enabling volume sharing between different Backup Groups, resulting in improved efficiency.

3-2-16. Parallel operation for backup and restoration

Parallel backup and restoration can be implemented by concurrently operating all the unoccupied tape drives accommodated in the library.

3-2-17. Highly flexible scheduling

Full backup, incremental backup, and differential backup schedules can be customized in accordance with a user's environment.

3-2-18. Immediate manual backup

Usually, automatic backup is carried out according to a specific schedule. However, in an emergency users can explicitly implement immediate backup through the GUI, or a command line.

3-2-19. A wide range of supported platforms

PetaBack™ system supports a wide range of platforms including the major UNIX® and Windows® operating systems. For information regarding other PetaBack software compliant operating systems, please refer to Section 1-3.

3-2-20. System construction in a heterogeneous environment

Even in a heterogeneous environment that combines different platforms supported by PetaBack software, a single backup system can be configured. So, users can easily add a PetaBack system to their existing system at a lower cost without purchasing a dedicated host or hosts.

3-2-21. External administration of backup Volumes

Since PetaBack software administers information about each volume on a database basis, volume(s) can be removed from a tape library and administered externally. The removed volume(s) can also be used in other PetaBack systems as long as the PetaBack servers are using the same platform.

3-2-22. Backup file information shown in a list form

As backed up file information is administered on a per database basis, the information can be shown on screen in a list form and viewed at a glance through the GUI or command line. Restoration can be implemented by designating a file shown on the screen through the GUI.

3-2-23. Backup system using a stand-alone tape drive

A PetaBack system can be configured with the use of a stand-alone tape drive without the use of an automatic tape library. This type of application is called "manual library system," and allows a lower cost backup system to be constructed. The system can also serve as an entry-level system for data backup, and can then be extended to a larger-scale system by adding an automatic tape library when necessary.

3-2-24. Easy installation and setup

PetaBack software employs interactive installation and provides various set-up tools that enable users to easily build up their own operating environments - just by following the messages exchanged. Usually, it is necessary for users to explicitly show the positions of the tape drives accommodated in the library when constructing a library configuration. However, PetaBack software omits this step, because it can obtain the device information through a PetaSite® Controller to automatically create configuration files. This automatic configuration function is subject to the use of a PetaSite Controller.

3-2-25. Support for Japanese file names

When the Shift-JIS code mode is enabled, files with Japanese file names can be backed up, restored, as well as retrieved via the GUI.

Chapter 4

Concepts and Terminology

ABCDEFGHIJKLMNOPQRSTUVWXYZ

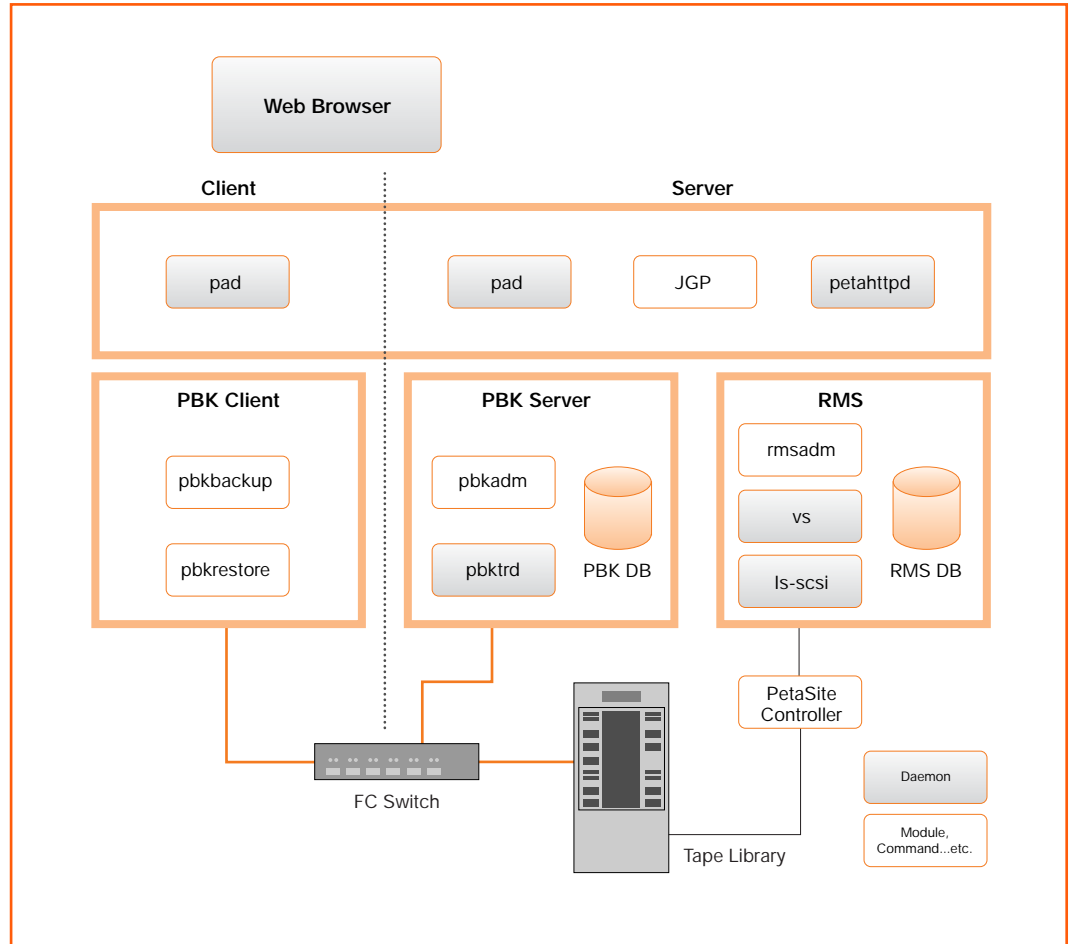
4

Concepts and Terminology

This chapter describes the concepts and technical terms relating to PetaBack™ software.

4-1. Module structure

PetaBack software modules are divided into the following categories; GUI, PBK Client, PBK Server, and RMS (Removable Media Server).



4-1-1. GUI

The GUI for PetaBack™ software consists of the following modules.

petahttpd

A dedicated JGP Web browser used solely to indicate the JGP through external Web browsers.

JGP (Java™ GUI PetaServe PetaBack)

The JGP is a Java Applet-based GUI with which the user can control PetaServe PetaBack systems operations and is shown through Web browsers (refer to 3-2-7 Java Applet-based GUI). The JGP requests *pad* for task respondent to each operation designated by users.

pad

A daemon that requests each PetaBack software module, such as PBK Client, PBK Server, and RMS, to process backup files or data upon receipt of order from the JGP.

4-1-2. PBK Client

The PBK Client module consists of *pbkbackup*/*pbkrestore* commands.

pbkbackup

A command that controls subsequent backup processes existing on the host required for backup and reads data from the disk for backup. When a backup is executed on the PetaBack server host, or, in the case of LAN-free backup, data read from the disk for backup is written directly to the tape device. However, when backup is executed from a remote host and backup data is transferred via network, the data is transferred first to the *pbktrd* daemon on the PetaBack server host.

pbkrestore

A command to control the subsequent restoration processes existing on the host required for restoration and writing data on the disk for restoration. The data is directly read from the tape device during execution on PetaBack server host or restoration via LAN-free. The data is transferred from *pbktrd* daemon on PetaBack server when executed on a remote client and transferred the data via network.

4-1-3. PBK Server

The PBK Server module is composed of *pbktrd* daemon, PetaBack databases, and the *pbkadm* command.

pbktrd

A daemon that implements actual processes upon receipt of a *pbkbackup* and *pbkrestore* request from PBK client modules. It also handles data I/O for tape devices when implementing backup and restoration from a remote client host via a network.

PetaBack database

A database directly used by a PetaBack server module is comprised of the following:

- **Backup group index**

Exists on the server to administer the indexes for all Backup Groups.

- **Backup group information file**

Exists on each Backup Group to administer the information set in backup groups.

- **Execution history database**

Exists on each Backup Group to administer the execution history for backup and restoration. When multiple generation preservation is set, it exists in each generation.

- **Volume index**

Exists for each Backup Group to administer the indexes for all volumes held by Backup Groups.

- **Volume information file**

Exists on each volume held by Backup Groups to administer volume information about backup and restoration.

- ***pbkadm***

A command on a PetaBack™ server that administers the PetaBack database so as to create or delete Backup Groups, update Backup Group attribution, and manage schedule. The complete process of PetaBack database administration from a GUI is implemented through this command.

4-1-4. RMS

RMS is the module that exists on a PetaBack server host to centrally administer and operate the automatic library and the media volumes and drives within the library. When it receives a request from the PBK server to mount or unmount media volumes, the RMS server manages the library according to the request. The RMS consists of the following modules:

- ***vs (Volume server)***

- Keeps track of the location and status of all drives and volumes and informs the user if a drive malfunctions or a volume can no longer be used.
- Sends an instruction to the library server for mounting/unmounting volumes to/from the drive.
- Operates drives on an efficient schedule using drive timers.
- Administers requests for volume mounting such as cancellation, interruption, return, position change in request queue, and priority change.
- Allocates and releases a volume on a *pbkadm* command.
- Reserves or cancels a volume and tape drive for a *pbkadm* command.
- Automatically implements initial formatting when loading a new volume into library.

- ***Is-scsi (Library server)***

- Connects to and takes an initial inventory of the library (its picker, drives, slots, and volumes), conducts any necessary hardware checks, and communicates the drive and volume location and status to its volume server.
- Inserts and ejects volumes into and from the library.
- Responds to volume server requests, mounts and unmounts volumes and marks volumes as allocated or unallocated.
- Updates the location and status of all drives and volumes under its control (a library server tracks the contents and data of its library in a database).
- Performs all shutdown tasks (updating drive and volume location and status, disconnecting from the library).
- Tracks the media life of tape volumes to prevent tape burnout.

- **RMS database**

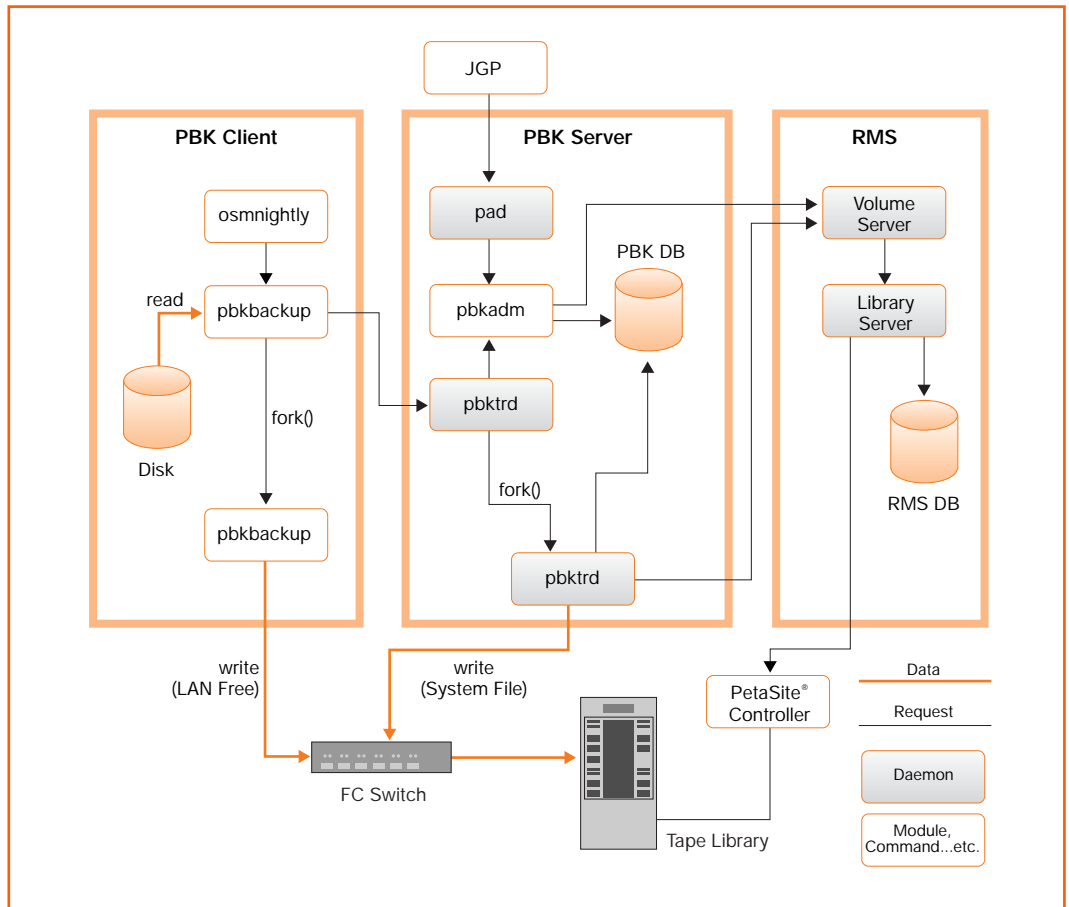
The RMS database is utilized by the volume server to administer volume information and library information.

- ***rmsadm***

Labels volumes, stops library, executes inventory, and administers the RMS database.

4-1-5. Backup operation example

The module structure shown below is an example of a LAN-free backup workflow using the automatic backup function.



• Registration of Backup Groups from JGP

- JGP requests `pad` daemon to initially register the content of the Backup Group registered by a user.
- `pad` daemon requests `pbkadm` command to register new Backup Groups.
- `pbkadm` command registers the Backup Groups on the PetaBack™ database.

- **Start of automatic backup**

- Creates *osmnightly* command using cron on a host required for backup, to start backup job at a specific time everyday.

- **Selection of volumes**

- *pbkbackup* command requests *pbktrd* daemon to select volumes for use.
- *pbktrd* daemon retrieves usable volumes from the PetaBack™ database.
- Calls *pbkadm* command when a usable volume does not exist.
- *pbkadm* command requests volume server daemon on RMS to allocate free volumes.
- *pbkadm* command registers allocated volumes on the PetaBack database.

- **Mounting of volumes**

- *pbkbackup* command requests *pbktrd* daemon to mount volumes.
- *pbktrd* daemon requests volume server to mount volumes.
- Volume server requests library server to transfer volumes.
- Library server requests PetaSite® Controller to transfer volumes.
- PetaSite® Controller sends volumes to tape drives.

- **Transfer of user-designated files and tape devices.**

- Creates sub-process for data writing by “forking” the *pbkbackup* command.
- Stores data on shared memory from disk using *pbkbackup* main process (a data reading process).
- Writes data on tape device from shared memory using the process for *pbkbackup* writing.
- Requests *pbktrd* daemon to update backup status.
- *pbktrd* daemon updates backup status on PetaBack database.

Note: For details, please refer to 4-3 (High-speed Data Transfer)

- **Backup of PetaBack system files**

- *pbkbackup* command requests *pbktrd* daemon to back up PetaBack system files.
- Creates sub-process for data writing by “forking” the *pbktrd* daemon.
- Writes data on shared memory from disk by using the process for *pbkbackup* reading.
- The process for *pbktrd* data writing writes data on tape device from the shared memory.
- *pbktrd* daemon updates backup status on PetaBack database.

Note: For details, please refer to 4-3 (High-speed Data Transfer)

4-2. Automatic backup mechanism

PetaBack™ software provides the following automatic backup methods.

- Automatically implements system maintenance every day, using maintenance script (*osmnightly*) integrated into *cron* (at function with Windows® Operating System) during installation. This function is called “Nightly Maintenance” as its default value is set at 00:16:00 during installation. Users can manually change the default setting to correspond with whatever time PetaBack software is unused.
- Nightly Maintenance confirms each day's schedule regarding the groups to be backed up on their respective hosts. If each group on the host has been scheduled for backup, Full Backup, Incremental Backup, or Differential Backup is implemented according to the setting.

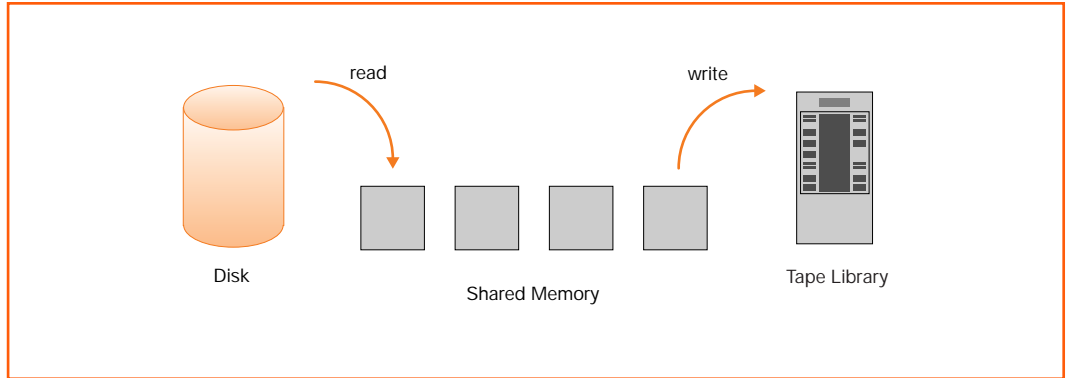
Nightly Maintenance

In Nightly Maintenance, the following operations are carried out in addition to the backup operation.

- Creates a summary file regarding the complete backup process.
- Informs the system administrator of the Nightly Maintenance results by e-mail.
- Obtains log file generations. (Refer to 4-5)
- Indicates volumes requiring recovery.
- Monitors the number of tape passes and lists the volumes that have exceeded the pass limit. The data in these volumes can be automatically copied to other volumes according to preset settings.

4-3. High-speed data transfer

PetaBack™ software is able to maximize the speed of data transfer between tape devices using the following method.



Block Size

An I/O unit's block size can be customized in accordance with the application system. The block size that is designed to apply to a tape device with a high-speed transfer rate is set at a default value of 256 kB.

Parallel Process

PetaBack software creates a dedicated read process for files on disk and a dedicated write process for files onto tape. These two processes are implemented in parallel. In this method, the shared memory works for both data delivery between the dedicated read process and the dedicated write process. The workflow is shown below:

Workflow of dedicated read process

1. Opens the file on disk
2. Reads the data on the file and writes it to the shared memory
3. Repeats step 2, depending on default block size and file size
4. Closes the file

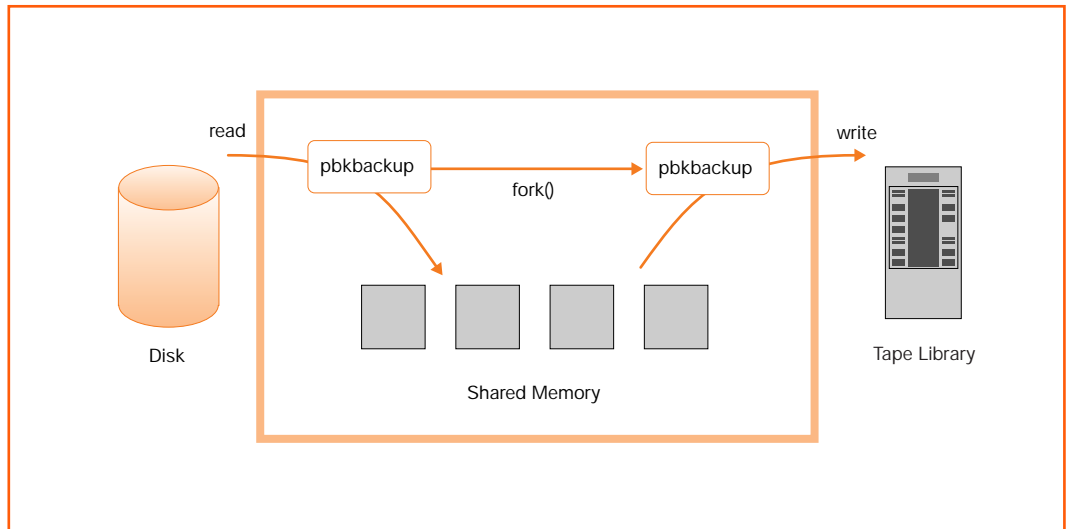
Workflow of dedicated write process

1. Opens the tape device
2. Reads the data on the shared memory and writes it on the tape device
3. Repeats step 2, depending on default block size and file size
4. Closes the tape device

Data transfer in server stand-alone systems

The *pbkbackup* command called out by *osmnightly* during backup implementation initiates a dedicated read process, reading data from disk and writing it on the shared memory. At

the same time, the *pbkbackup* command creates a sub-process through *fork()* as a dedicated write process to a tape device. This process reads the data on the shared memory and writes it to the tape device.



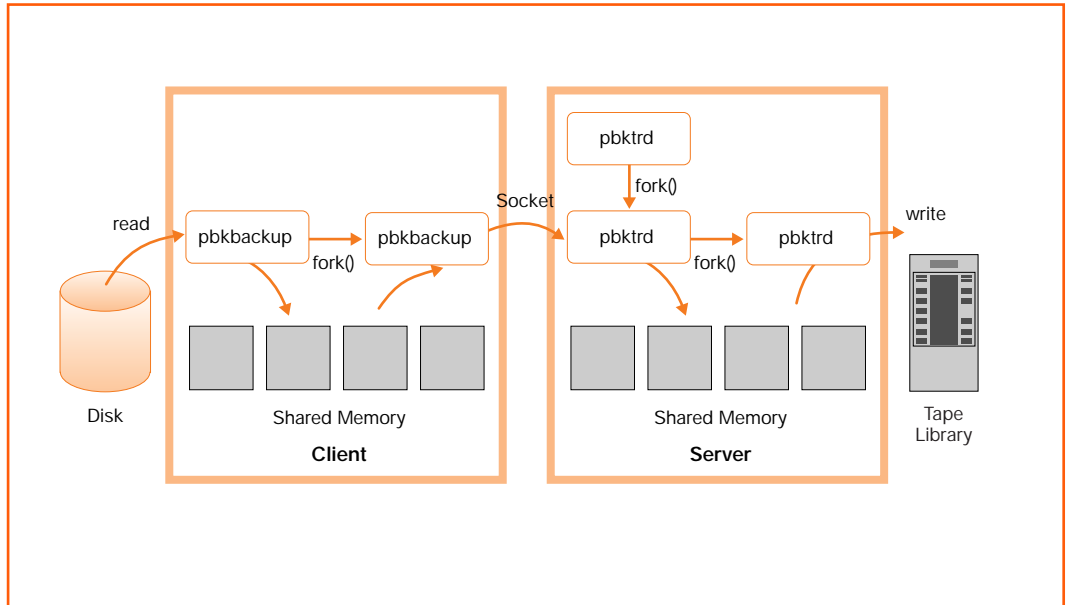
4

Concepts and Terminology

Data transfer in a server client system

The **pbkbackup** command called out by **osmnightly** during backup implementation initiates a dedicated read process, reading data from disk and writing it on the shared memory. The **pbkbackup** command creates a sub-process through **fork()**, which reads data from the shared memory and

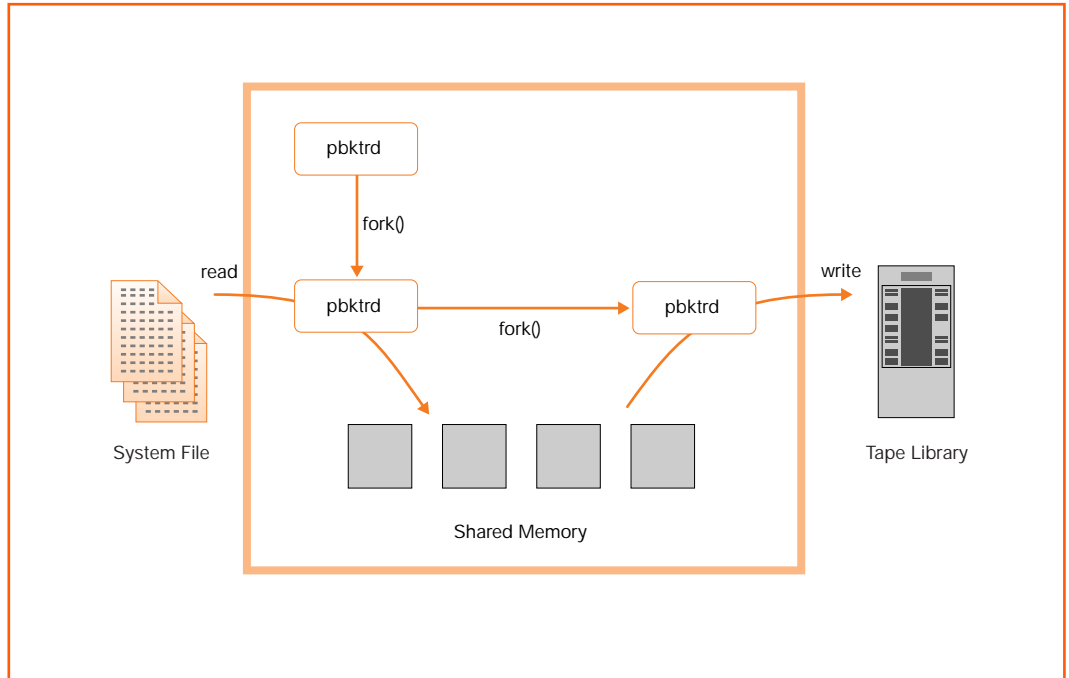
transfers the data via socket communication to the **pbktrd** sub-process created by **fork()**. The **pbktrd** sub-process then writes the data on the shared memory. At the same time the **pbktrd** daemon initiates a sub-process through **fork()**, a dedicated write process, which reads data from the shared memory and writes the data onto the tape device.



Data transfer during system file backup

During backup implementation, the *pbktrd* sub-process created from the *pbktrd* daemon through *fork()* initiates a dedicated read process, reading data from system files and

writing the data on the shared memory. The created *pbktrd* then works as a write dedicated process to read data from the shared memory and to write the data onto the tape device.



4-4. NAS server backup using NDMP

PetaBack™ supports Network Data Management Protocol (NDMP). Because of this, files on a server can be backed up at extremely high speeds when connecting to tape drives via FC.

4-4-1. NDMP functions supported in PetaBack software

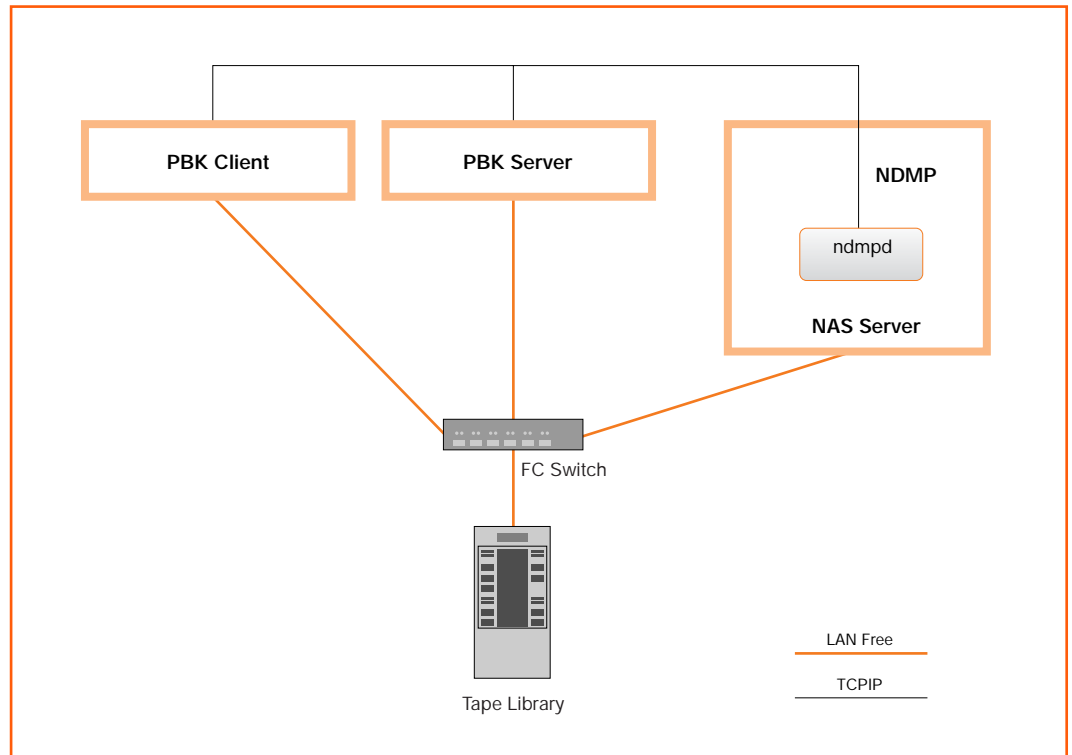
NDMP functions supported in PetaBack software are:

- Full backup function.
- Incremental backup function.
- Automatic backup function.
- Informing the automatic backup results via e-mail.
- Automatic backup scheduling function.
- Multi volumes (tape spanning) function.

- Volume sharing by multi backup groups.
- Displaying backup files through GUI/CUI in list form.
- Restoring designated files.
- Directories for restoration at the user's discretion.
- Restoring by designating backup job.
- Administering multi-generation execution history.
- Restoring by designating the generation of execution history

4-4-2. Backup using NDMP

The backup method using NDMP supported by PetaBack software is as shown below.



Only LAN-free environments supported

Backup using NDMP in PetaBack™ software is only supported in LAN-free environments.

PetaBack server as an NDMP client

When backing up files on a NAS server using PetaBack software, a PetaBack server works as an NDMP client and requests a NAS server for backup.

Backup control using the *ndmpd* daemon

The *ndmpd* daemon existing on a NAS server receives NDMP command from a PetaBack server and starts data writing.

Support for *dump* System

There are three kinds of systems - *tar*, *dump*, and *cpio* - for a NAS server backup using NDMP command. PetaBack software only supports the *dump* system.

Synchronizing the file information portion with the data portion

The conventional UFS *dump* system collates information first about files for backup and then about data pertaining to each file. The time difference between these two backup processes may result in a possible discrepancy between the file information and the file data. However, the *dump* system used in NDMP creates a snap shot when issuing a command and there is no discrepancy between the file information portion and the data portion, as the snap shot itself is backed up.

Incremental backup

The *dump* system operates over a range of 10 levels. When it designates a particular level at which to execute the process, only the files that have changed since the precise moment that the *dump* execution occurred at the level less than the designated level are archived. This level is used for NAS server backup in PetaBack software. In other words, in consecutive incremental backups, *dump* occurs at one level higher than that for the last backup. Incremental backup can be nine times consecutively executed.

4-5. Other

Archival format for PetaBack™ software

PetaBack software has adopted a unique archival format to avoid limitations traditionally imposed by the archival formats used for UNIX® standard commands such as *tar*, *cpio*, etc. As a result, users can take full advantage of the maximum file size and path name supported by each operating system.

Log mechanism

The following are the three kinds of log files used for problem analysis in PetaBack software. These secure space for the */osm/log* file system and at the same time preserves the latest logs, because the maintenance script periodically deletes old log files.

- **Alert Message**

These are error messages shown on the GUI screen in the event of trouble or a malfunction. These messages are all preserved in a *logfile*, and may be deleted manually by the user after confirmation.

- **Logfile**

The *logfile* stores logs relating to the internal operations of PetaServe® software in a file called */osm/log/logfile*. When this *logfile* size exceeds a previously specified value, nightly maintenance automatically changes the file name to */osm/log/logfile.yy.mm.dd*.

- **Tracefile**

The tracefile is stored in */osm/log/tracefile*. Maintenance is executed every 30 minutes using script integrated in *cron* during installation. The script automatically changes the file name to */osm/log/tracefile.n*, when the file size has exceeded the previously specified value (the default value is 10 MB). The last digit *n* represents the number of generations stored. The Administrator can freely set the number of generations for storage (the default value is four generations).

Methods of obtaining Logfile

PetaBack software stores all the logs regarding operations under */osm/log*. A log file can be obtained by executing the command (*fzctakelog*), or displayed on the GUI. When PetaBack software is in a decentralized client environment, a log of each compliant client can be obtained simultaneously by executing *fzctakelog* on the server or by using the GUI.

5

Command Summary

In this chapter, the main commands and *pbkadm* sub-commands used in PetaBack™ software are explained.

Main commands

<i>cron</i>	A daemon to start the process of command execution at a specific time
<i>mail</i>	A command to send mail
<i>osmnightly</i>	A command to execute nightly maintenance
<i>osmservers</i>	A command to start, monitor, and stop the PetaBack daemon
<i>pbkadm</i>	A command for PetaBack database administration, relating to the creation and structure of Backup Groups.
<i>pbkbackup</i>	A command to implement the backup process
<i>pbkedit</i>	An edit of a file defined by the variable that customizes backup and restoration of Backup Group, and also an edit of a file set for backup
<i>pbkls</i>	A command to view backup file information at a glance in list form
<i>pbkrestore</i>	A command to implement restoration
<i>psvrestore</i>	A command to restore the PetaBack system file
<i>fzctakelog</i>	A command to take logs
<i>rmsadm</i>	A command for RMS server administration relating to hardware resources such as volume, drive, and library.

pbkadm sub-commands

<i>addgroup</i>	A sub-command to create new groups
<i>addsh</i>	A sub-command to register a new schedule for a designated Backup Group
<i>addvol</i>	A sub-command to allocate a free volume to a designated Backup Group
<i>lsgroup</i>	A sub-command to show the current Backup Group at a glance in list form
<i>lssch</i>	A sub-command to show the current schedule of a designated Backup Group at a glance in list form
<i>lsvol</i>	A sub-command to show an allocated volume at a glance in list form
<i>rmsch</i>	A sub-command to delete the schedule of a designated Backup Group
<i>history</i>	A sub-command to show the execution history of a backup restoration at a glance in list form
<i>recinhi</i>	A sub-command to disable the volume from using during backup
<i>rmgroup</i>	A sub-command to delete the existing Backup Group
<i>rmvol</i>	A sub-command to release an allocated volume
<i>setgroup</i>	A sub-command to change the attribution of a designated Backup Group

Chapter 6

Limitations



6

Limitations

Maximum length of Backup Group name	31 alphanumeric digits including hyphen (-) and underline (_)
Maximum number of generation histories	31 generations
Maximum number of subsequent incremental backups before a full backup must be performed	100 times Note: 9 times when using NDMP
Maximum backup file size	Unrestricted. The maximum size supported by each OS (Operational test of files from 1 Byte to 1 TeraByte has been verified).
Maximum path name length (alphanumeric digit number) for a possible backup file	Unrestricted (The maximum value supported by each OS)
Maximum volume name length	24 letters
Maximum number of volumes	100,000 volumes
Maximum library name length	15 letters
Maximum number of libraries	20 libraries
Maximum number of clients	Managed by license
Maximum number of Backup Groups	1,000 Backup Groups
Maximum number of volumes that may be allocated to one Backup Group	100 volumes

Chapter 7

Glossary - Terms and Definitions

ABCDEFGHIJKLMNOPQRSTUVWXYZ

7

Glossary - Terms and Definitions

Auto-changer	See "Picker."
Backup Group	A theoretical concept that defines those groups that the user designates for backup. This term refers to a pairing of the host and file required for backup. According to users' circumstance, files designated for backup can be categorized against each of which Backup Groups can be defined.
Client	A host that has files designated for backup.
Cross-restore	A function that restores to a host other than the original backed up host.
Daemon	A constantly existing background process.
Differential backup	Backs up all files that have changed since the latest full backup.
Free volume	A newly-labeled volume that has not been allocated to any Backup Group.
Full backup	Unconditionally back up all files designated for backup.
Incremental backup	Backs up all the files that have changed since the latest backup.
JGP	A PetaBack™ administrative interface enables the administrator to centrally and remotely administer the data through a Web browser through a Java™ Applet that allows remote control.
LAN-free backup	A backup that directly transfers backup data to a tape device from a remote host via fibre channel without passing via a server or network.
Library	A combination of tape drives and volumes. Normally, it is an automatic appliance including an auto changer or picker to move the volumes.
Nightly maintenance	A set of maintenance jobs scheduled by cron. The script is <i>osmnightly</i> .
OS	Operating System software (Refer to "Section 1-3. System components supported" for the OS supported by the PetaBack software).
PetaServe®	A trademark for Sony HSM (Hierarchical Storage Manage) data management software that transparently migrates data from fixed disks to a mass storage repository called a "store".
Picker	A robotic device that moves volumes between slot, mail slot, and tape drives within an automatic library. See "Library."
Raw device	A character device that can directly access hardware without using the buffering within the kernel in a disk device.

Repair backup	In the setting of automatic backup, when full backup has been skipped because of some system malfunction, the generation of the execution history database is not updated. In this situation, in order to update the generation, full backup is forcedly implemented even if a time incremental backup should normally be carried out.
RMS	The removable media server that controls access to media devices and media volumes in a PetaServe® system.
Server	A host that directly connects to and administers a library.
True image restore	A function that merely restores files existing on disk at the time the latest backup occurred.
Volume	A theoretical concept meaning physical recording media, generally including memory, disk, tape, etc. In PetaBack™ software, it particularly means magnetic tape.
Volume set	A function that shares volumes between different Backup Groups, resulting in efficient use of the volumes.

Chapter 8

Appendix



Comparison of performance

The following table shows how PetaBack™ software compares to two of its competitors in the provision of important features.*

Note:

- AL: Additional license required
- C: Standard function with conditions
- S: Standard function

	NetVault™ 6.5	NetBackup™ BS 3.4	Sony PetaBack™ V3.10	Conditions affecting PetaBack V3.10 standard functions (operation)
LAN-free backup	AL	AL	S	
Support for NDMP	AL	AL	S	
Library sharing over LAN	AL	AL	C*	* Data transfer via LAN.
Drive sharing over SAN	AL	AL	S	
Oracle on-line backup	AL	S	AL	
Shared memory	AL	S	S	
Heterogeneous environment	AL	S	S	
Standard tape format	C (<i>cpio</i>)	C (<i>Gnu-tar</i>)	C* (<i>Gnu-tar</i>)	* Enhanced <i>tar</i>
Integrated scheduling	S	S	S	
Automatic media management	S	S	S	
Parallel backup/restore	Unknown	S	C*	* Parallel process is implemented on multiple tape drives.
Non-privileged user operation	S	S	C*	* Support for restoration. * Backup will be available soon.
Restore to host other than the backup host	S	S	C*	* Except that between Windows® and UNIX®.
Windows® 2000/NT support	S	S	C*	* Support for client only

* Note: The above data is based on NetVault's and Veritas' respective web site home pages as of January 2002.