Technical Report

Backup and Recovery of Perforce on NetApp Storage Solutions

NetApp and Perforce
March 2013 | TR-4142

Abstract

This technical report describes how to perform backup and recovery of Perforce® on NetApp® storage. It provides information on Perforce critical data that is important to protect and some basic procedures to back up and recover Perforce data across NetApp storage protocols. It also highlights the NetApp data integrated solutions available for high availability and disaster recovery.

TABLE OF CONTENTS

1 Introduction 3

1.1 Scope 3

1.2 Audience 3

2 Perforce Critical Data 3

3 NetApp Integrated Data Protection Solutions 4

4 Backup and Recovery of Perforce Data 5

4.1 Backup 5

4.2 Recovery 7

5 Disaster Recovery 8

6 Limitations 8

7 Conclusions 9

References 9

Version History 9

LIST OF TABLES

Table 1) Basic backup procedures for Perforce data (database, depots, and journals). 6

Table 2) Basic recovery procedures when both database and depots are corrupted. 7

1 Introduction

Perforce Software is an enterprise version configuration management application that allows engineering development teams to manage and track revisions of different types of files as a product is being developed. Thus, data protection of Perforce repositories is important to safeguard companies' digital assets. Downtime due to unplanned events such as application failures, component failures, data center failures, and regional disasters can be very costly and affect the continuity of business.

NetApp storage solutions provide customers with a wide range of data protection that provides backup, availability, compliance, and disaster recovery solutions to protect customers' critical data. The foundation of NetApp's protection software, such as SnapMirror® and SnapRestore® technologies, is NetApp Snapshot™ technology, which enables backups to be made in minutes, reducing backup windows. Also, a NetApp Snapshot copy requires minimal disk space since it only maintains and copies the set of pointers to disk blocks containing the data and not the actual copy of the data blocks.

1.1 Scope

Perforce databases, versioned files (depots), journals, and logs can be deployed on any NetApp storage protocol, including, FC, iSCSI, NFS, CIFS, and even a split configuration in which the Perforce database is on SAN and depots, journals, and logs are on NAS. This document provides methods and general procedures for protecting Perforce critical data on each of these storage protocols.

1.2 Audience

This document is intended for use by individuals responsible for the backup and recovery of Perforce data kept on NetApp storage. It assumes that readers have experience with administration of NetApp storage solutions and the Perforce application. It also assumes that the reader has read the technical reports relating to the deployment and implementation of Perforce on NetApp storage for an understanding of Perforce architecture and deployment options on NetApp storage. Please refer to the NetApp library for links.

2 Perforce Critical Data

The most critical data to protect and preserve in a Perforce environment includes:

- **Database (metadata)** – The *database files* store *metadata*, including changelists, file attributes, client workspace specifications, branch mappings, and other data concerning the history and present state of the versioned files. The database consists of several files in the top level of the server root directory. Each database file contains a binary-encoded database table.
- **Depots (versioned files) –** The depots contain the content of the files stored in Perforce. Each depot may be kept in the server root directory or on a separate file system. Typically this is a mixture of text and binary data, and each is stored differently within the depot.

Perforce provides backup and recovery mechanisms to guarantee the integrity of the database. This is achieved by reconstructing the database using the following files.

- **Checkpoint** – A text file containing all of the transactions needed to recreate the database. This is a point-in-time copy of the database. This file is much smaller than the original database and contains all the information to recreate the database.
- **Journal** – A log of updates to the database since the last checkpoint was taken. This file is truncated when a checkpoint is done and the older journal is renamed. A checkpoint with the subsequent journals is all that is needed to create an accurate copy of the database.

When a checkpoint is conducted, the database is locked and no changes are allowed. Then its contents are dumped to a file normally named *<file>.ckp.n,* where *n* is a sequence number. Before the database is

unlocked, the journal file is truncated and renamed to a file named *journal.n-1.* Afterward, versioned files stored in the depots should be backed up together with the checkpoint and journal files in order to restore the Perforce environment to a specific point in time. As long as no destructive operations are run by the Perforce administrators, a much later copy of the depot(s) can be used with an earlier checkpoint.

Perforce log files can also be backed up to maintain a full record of errors, audit events, commands run, resources used, and other information.

3 NetApp Integrated Data Protection Solutions

NetApp data protection solutions offer Perforce customers a simplified and quick approach to backing up and recovering critical Perforce data. These solutions were designed to improve the overall operational efficiency of backup and recovery.

Integrated into NetApp's storage platform are capabilities that provide high availability and resiliency in case of storage hardware failures. These include:

- Raid-DP® technology – High-performance Raid 6 (dual parity) implementation that protects against the simultaneous failure of two drives in the same RAID group.
- Active-Active – HA pair controller configuration that provides high-availability solutions during planned and unplanned downtime events.

Below are highlights of other integrated data protection solutions that NetApp provides to protect against many different types of disaster scenarios, ranging from the most common failures such as power, hardware, network, or application failures within data in the data center to the most catastrophic events, such as floods, hurricanes, or natural disasters. The solutions include the following.

- Backup and Recovery
  - Snapshot technology – Creates disk-to-disk point-in-time backup copies in native format.
  - SnapRestore – Data recovery software uses stored Data ONTAP® Snapshot copies to recover anything from a single file to multi-terabyte volumes in seconds.
- Disk-to-Disk Backup
  - SnapVault® software – Speeds up and simplifies backup and data recovery, protecting data at the block level. Also a disk-to-disk backup for NetApp FAS systems.
  - Open Systems SnapVault (OSSV) software – Leverages block-level incremental backup technology to protect Windows®, Linux®, UNIX®, SQL Server®, and VMware® systems running on mixed storage. Replication-based disk-to-disk backup for open system storage servers.
- Application-Aware Backup and Recovery Solutions for Applications
  - Snap Creator™ software – Provides a central framework that integrates NetApp Snapshot technology with applications.
  - SnapDrive® technology – Simplifies storage provisioning for UNIX or Windows platforms and automates OS-consistent backup and restore of application data.
- Tools for Backup Administrators to Simplify Processes
  - OnCommand® Unified Manager automates the management of physical and virtual storage for NetApp storage systems and clusters.
  - SnapProtect™ technology – Management software accelerates and simplifies backup and data recovery for shared IT infrastructures. Provides a single management console that allows you to create, catalog (for indexing and fast search of Snapshot copies), and manage application-aware Snapshot copies across disk-to-disk-to-tape processes.
  - Tools for Compliance

  - SnapLock® compliance software – A flexible data permanence solution for meeting strict data retention regulations or internal IT governance policies. Allows creation of nonrewritable, nonerasable volumes to prevent files from being altered or deleted until a predetermined retention date.

- High Availability and Business Continuity
  - SnapMirror – Data replication technology provides disaster recovery protection and simplifies management of data replication. Provides three modes of mirroring: sync, semi-sync, and asynchronous.
  - MetroCluster™ software – High-availability and disaster recovery software delivers continuous availability, transparent failover protection, and zero data loss.
- Archival
  - Tape – NetApp's tape backup and restore solution uses Network Data Management Protocol (NDMP) version 3 and 4, which efficiently maximize network bandwidth. NDMP-enabled commercial backup applications can be used to perform a dump backup or restore.

NetApp also provides global support and services that assist in fixing problems and/or assist in backup and recovery planning. Services that are offered include:

- AutoSupport™ and Storage Availability Audits – Monitoring and reporting technology that checks the health of NetApp storages systems on a continual basis. Provides a call-home feature that automatically contacts NetApp Support to handle or replace parts if the storage has a failure.
- Personalized support services – Availability of NetApp Support Account Manager, NetApp Support Advisor, or NetApp Resident Support engineer to provide 24/7 incident management, education, and on-site reactive and proactive support.

In addition, NetApp's data protection services leverage NetApp storage efficiency technology to reduce both storage and management costs. Core to NetApp data protection is its Snapshot technology. When a Snapshot copy is made, only new or changed blocks are transferred to disk to reduce backup windows, minimize network traffic, and reduce disk capacity.

NOTE: Some of the above features are not yet supported on NetApp clustered Data ONTAP. Refer to the Limitations section of this document for more information.

4 Backup and Recovery of Perforce Data

Regular backups of Perforce depots, databases (via checkpoints), and journals are essential in protecting valuable data and reducing any interruption to business operations. Planning for and implementing backup and recovery are important to achieve optimal availability and IT efficiency. When developing a strategy for backup and recovery there are two objectives to consider:

- Recovery Point Objective (RPO) – The amount of acceptable data loss
- Recovery Time Object (RTO) – The amount of time it takes to perform the recovery

Your criteria for the above objectives will determine the backup and recovery strategy to select. As mentioned in the previous sections, NetApp provides a wide range of data protection solutions that can assist in improving your RPO and RTO. The following sections discuss the generic steps and procedures for performing regular backup and recovery of Perforce data kept on NetApp storage protocols.

4.1 Backup

Creating a checkpoint is one of the important steps in preserving the integrity of the Perforce database. However, conducting a checkpoint can be a long process depending on the size of the database. Taking a checkpoint for a large database can be disruptive and affect the productivity of developers. With NetApp

storage solutions, *offline checkpointing* can be achieved by locking the database temporarily and taking a Snapshot copy of the database, depots, and journals, truncating the journals, and then running an offline checkpoint on the Snapshot copy from a different host. Since NetApp Snapshot copies can be taken quickly, more frequent backups are possible, enabling a more aggressive RPO and more effectively controlling the size of the journals. Additionally, the Snapshot copies can be further backed up to another NetApp storage controller using SnapVault or SnapMirror or they can be archived to tape.

The table below provides the basic steps to back up the Perforce database, depots, and journals on each of the NetApp storage protocols. The processes in between locking the database and unlocking the database described in Table 1 can be scripted and passed to *p4d –c*.

If the Perforce data is deployed entirely on SAN or in a split configuration in which the database is on SAN and depots and journals are on NAS, SnapDrive is recommended for creating the NetApp Snapshot copy. Using SnapDrive as opposed to just calling "snap create" from the appliance flushes all the data from the host memory to disk prior to taking the Snapshot copy. SnapDrive is initiated from the Perforce server host and is available for UNIX and Windows platforms. In addition, for the split configuration, Snapshot copies on both SAN and NAS using SnapDrive can be created on the same command line using the `-fs` option to maintain consistency of journals, database, and depots. For example, in UNIX, the command would be:

```
snapdrive snap create -fs  /depot  /db   /journals -snapname my_snap
```

Table 1) Basic backup procedures for Perforce data (database, depots, and journals).

SAN (FC or iSCSI)	NAS (NFS or CIFS)	Split (DB on SAN, Depots and Journals on NAS)
1) Truncate the journal (*p4d –jj*). 2) Use SnapDrive to create a snapshot of the volume(s) containing the database, depots, and journals from within a *p4d –r <p4root> -c <SnapDrive command(s)>.* The database will be locked for the duration of the command run within the –c "" and unlocked when it exits. 3) Make a LUN clone of the database snapshots. 4) Map the clone to another host not running the Perforce service and mount the LUN as "read-only" on that host. 5) Run the checkpoint on the cloned database (*p4 –jd*). 6) Back up the checkpoint on NAS volume. 7) Unmount, unmap, and destroy clone.	1) Truncate the journal (*p4d –jj*). 2) Take a snapshot of the volume(s) containing the database, depots, and journals using `snap create`. This will need to be performed within a *p4d –r <p4root> –c <snapshot creation command(s)>.* The database will be locked for the duration of the command run within the –c "" and unlocked when it exits. 3) For NAS volumes, ".snapshot" directory can be visible under the mounted volume. From a host different from where the Perforce service is running, mount the volume containing the database as "read-only" and run a checkpoint *(p4 –jd)* from the ".snapshot" directory of the database.	1) Truncate the journal (*p4d –jj*). 2) Use SnapDrive to create a snapshot of the volume(s) containing the database, depots, and journals from within a *p4d –r <p4root> -c <SnapDrive command(s)>.* The database will be locked for the duration of the command run within the –c "" and unlocked when it exits. 3) Make a LUN clone of the database snapshot. 4) Map the clone to another host and mount the LUN as "read-only" on that host. 5) Run the checkpoint on the cloned database (*p4d –jd*). 6) Back up the checkpoint on the NAS volume. 7) Unmount, unmap, and destroy the clone. 8) Back up the checkpoint on the NAS volume.

	4) Back up the checkpoint on the NAS volume. 5) Unmount the volume.	9) Unmount the volume if needed.

4.2 Recovery

Recovering a corrupted database is done with the help of the latest checkpoint and the subsequent journal file(s), including the active journal. The associated Snapshot copy of the depots can be restored accordingly. When both the database and depots have been lost or corrupted, the journal file is not required for reconstruction of the database since the versioned files restored will only reflect the depot as it existed since the last checkpoint. This section describes basic recovery procedures for corrupted Perforce data; for more details on recovery, refer to the Perforce System Administration Guide.

For the restore procedures described in Table 2, it is assumed that the checkpoint was backed up on a NAS volume. If the Perforce data is backed up on a NAS volume, use "snap restore" to perform a restore of the desired depots. Then, a single file restore can be used to restore the checkpoint file and journal file as well. If the Perforce journal and depots reside on a LUN, use SnapDrive to restore the LUNs containing the journals and depots. Then perform a single file restore of the checkpoint file.

To reconstruct only the database, repeat steps 1 and 2, below, and modify step 2d to specify the journal file (that is, *p4d –r $P4ROOT –jr checkpoint_file journal file*).

Table 2) Basic recovery procedures when both database and depots are corrupted.

SAN (iSCSI or FC)	NAS (NFS or CIFS)	Split (DB on SAN, depots and journals on NAS)
1) Do a single file restore of the last checkpoint from the NAS volume. 2) Reconstruct the database using checkpoint. a. Stop the current instance of p4d. b. Move the corrupted database to the tmp directory (best practice). c. Compare the MD5 checksum of the most recent checkpoint with the checksum generated at the time of its creation. d. Invoke p4d with the journal restore flag, specifying the most recent checkpoint.	1) Do a single file restore of the last checkpoint file from the NAS volume. 2) Reconstruct the database using the last checkpoint. a. Stop the current instance of p4d. b. Move the corrupted database to the tmp directory (best practice). c. Compare the MD5 checksum of the most recent checkpoint with the checksum generated at the time of its creation. d. Invoke p4d with the journal restore flag, specifying the most recent checkpoint.	1) Do a single file restore of the last checkpoint file from the NAS volume. 2) Reconstruct the database using checkpoint. a. Stop the current instance of p4d. b. Move the corrupted database to the tmp directory (best practice). c. Compare the MD5 checksum of the most recent checkpoint with the checksum generated at the time of its creation. d. Invoke p4d with the journal restore flag, specifying the most recent checkpoint.

e. Verify system integrity after restoration (“p4 counter lastCheckpointAction” and “p4 verify”). 3) Use SnapDrive to restore the LUN containing the last Snapshot copy depot.	e. Verify system integrity after restoration (“p4 counter lastCheckpointAction” and “p4 verify”). 3) Run “snap restore” of the depots associated with the checkpoint.	e. Verify system integrity after restoration (“p4 counter lastCheckpointAction” and “p4 verify”). 3) Use “snap restore” to restore the depots associated with the checkpoint.

5 Disaster Recovery

Although catastrophic disasters such as flood, hurricanes, and other natural disasters are less likely to occur, they can have the greatest impact on businesses. Being prepared and having a disaster recovery strategy is key to business continuity. NetApp provides SnapMirror and MetroCluster to deal with these types of disaster scenarios.

The NetApp SnapMirror solution creates an identical second set of data capable of replacing the primary set of Perforce data should something happen to the primary. Since NetApp SnapMirror is based on Snapshot technology, only changes and updates are transferred to the remote site. Other SnapMirror capabilities include deduplication or compression, which can further speed up the transfer and reduce bandwidth utilization by transferring only deduplicated or compressed data. Perforce can also take advantage of NetApp SnapMirror technology to provide a replicated copy (read-only) of Perforce file depots for access by clients in a distributed development environment. Perforce’s current software release provides on-demand replication that allows the file depots not to be natively replicated to a remote site and to leverage NetApp SnapMirror to handle the replication. This provides the added benefit of using the storage resources to do the replication instead of the host (the Perforce server), allowing the Perforce server to handle other workloads. Refer to the following technical reports for more information on SnapMirror.

- TR-3446: SnapMirror Async Overview and Best Practices Guide
- TR-3326: SnapMirror Sync and SnapMirror Semi-Sync Overview and Design Considerations
- TR-3703: Using SnapMirror for Disaster Protection with Block Access Protocols
- TR-3999: SnapMirror Startup Guide for Data ONTAP 8.1 Operating in Cluster-Mode

The NetApp MetroCluster solution is also available to help in disaster recovery within a metropolitan area. MetroCluster offers synchronous data mirroring that enables zero data loss and, with automatic failover, enables nearly 100% uptime. For more information on MetroCluster, refer to:

- TR-3548: Best Practices for MetroCluster Design and Implementation

Both NetApp and Perforce provide consulting services that assist our joint customers in disaster recovery planning. Our consultants can assist in designing the best solution for your environment. Refer to the Perforce Web site for the company’s High Availability and Disaster Recovery Planning services.

6 Limitations

As of the writing of this technical report, certain data protection features for NetApp clustered Data ONTAP version 8.1 do not yet support the following:

- Qtree SnapMirror
- Synchronous SnapMirror
- SnapVault technology

- MetroCluster

Future releases of NetApp clustered Data ONTAP will support them.

7 Conclusions

Developers require Perforce services and data to be available 24/7 because downtime adversely affects the development process, leading to potential revenue loss. Having a robust and resilient storage system such as NetApp's offers high-performance and high-availability options for development environments. NetApp Integrated Data Protection solutions assist in protecting Perforce critical data against all types of disaster scenarios, from the most common failures, such as power loss, hardware, network, or application failures within the data center, to the most catastrophic events, such as floods, hurricanes, or natural disasters. NetApp solutions simplify Perforce backup and recovery processes. They also provide an easy--to-implement and robust mirroring solution. Implementing NetApp's data protection solutions will significantly reduce an organization's risks and protect Perforce data assets in the event of a failure or disaster.

References

- "Data Protection Online Backup and Recovery Guide" on the NetApp Support site for the appropriate Data ONTAP version
  http://support.netapp.com/documentation/productlibrary/index.html?productID=30092
- "Perforce System Administrator's Guide"
  http://www.perforce.com/resources/documentation/perforce-administration-manuals

Version History

Version	Date	Document Version History
Version 1.0	March 2013	Agnes Jacob and Scott Sanford, NetApp, and Randy DeFauw and Tim Brazil, Perforce

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

NetApp provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

Go further, faster®

TR-4142-0313

NetApp

www.netapp.com