



BSnapDiff for GPFS Plugin

Bacula Systems Documentation

Contents

1	Overview	2
1.1	Features Summary	2
2	Installation	3
2.1	Packages	3
3	Debian/Ubuntu	3
4	RHEL	3
4.1	Configuration	3
4.2	Snapshot Management	4
4.3	Estimating Performance	4
4.4	Limitations	5

Contents

Note

You can download this article as a [PDF](#)

- *Overview*
- *Installation*

1 Overview

1.1 Features Summary

Bacula Enterprise Edition (18.2.5) now supports BSnapshotDiff technology for the GPFS file systems on Linux backup clients. BSnapshotDiff compares two snapshots of a file system to quickly and accurately identify the changes required for backup. For GPFS file systems, GPFS programming interfaces are integrated into the bsnapshot Plugin. The File Daemon uses the results to properly back up file system changes without scanning the entire file system. If very little has changed since the last backup, the differences are reported quickly, regardless of the number of files in the actual file system.

This feature requires the BSnapshotDiff for GPFS File Daemon Plugin. If the plugin detects that it is backing up data in a GPFS Fileset, it enables BSnapshotDiff. All backups automatically create and use GPFS snapshots. Bacula supports both global snapshots and snapshots at the GPFS Fileset level.

2 Installation

To use BSnapshotDiff for GPFS, install both the bsnapshot Plugin and the BSnapshotDiff for GPFS Plugin. No additional configuration is required. The bsnapshot Plugin calls the BSnapshotDiff for GPFS Plugin if it is available.

2.1 Packages

Packages of the **BSnapDiff Plugin** are available for supported platforms. Please contact Bacula Systems Support team to get them.

Download the **BSnapDiff** Plugin package to the server on which the Bacula File Daemon is installed, then install using the package manager.

3 Debian/Ubuntu

```
dpkg -i bacula-enterprise-bsnapdiff-plugin*.deb
```

The package manager ensures that your **Bacula Enterprise** version is compatible with the **BSnapDiff** Plugin.

4 RHEL

```
rpm -ivh bacula-enterprise-bsnapdiff-plugin*.rpm
```

The package manager ensures that your **Bacula Enterprise** version is compatible with the **BSnapDiff** Plugin.

4.1 Configuration

GPFS Fileset and Bacula Fileset

Administrators of GPFS file systems should be familiar with the term **Fileset** when referring to a portion of a GPFS file system. Bacula also uses the term **Fileset** when referring to a collection of data to be backed up.

To distinguish between the two, this article uses **GPFS Fileset** and **Bacula Fileset**. Elsewhere in the Bacula documentation, **FileSet** refers to a **Bacula Fileset**.

Enable Accurate Mode in Job Resource

To handle hard links properly, set `Accurate = yes` in the Job resource of the Director for jobs that include GPFS file systems and/or GPFS Filesets.

```
Job {  
  Name = "GpfsJob"  
  Client = storage-fd  
  Fileset = GpfsFileSet  
  ...  
  Accurate = yes  
}
```

If Accurate mode is not enabled, or is explicitly disabled by setting `Accurate = no`, backups do not use the BSnapshotDiff plugin. The following message appears in the backup log:

```
bsnapdiff: accurate mode is not enabled, running regular incremental_
↳ filesystem backup
```

Enable Snapshots in the Bacula FileSet Resource

The BSnapDiff Plugin also requires snapshots to be enabled in the target file systems. Set `EnableSnapshot = yes` in the Bacula Fileset resource that uses BSnapDiff.

```
Fileset {
  Name = GpfsFileSet
  EnableSnapshot = yes
  Include {
    Options {
      Signature = MD5
    }
    File = /gpfs/parthto/fileset-mount-point/and-or-subdir
  }
}
```

If snapshots are not enabled, or are explicitly disabled by setting `EnableSnapshot = no`, backups use neither GPFS snapshots nor the BSnapDiff plugin. No message about this is logged because the BSnapDiff plugin is never called when snapshots are disabled.

File Daemon Configuration

On the File Daemon host, the **BSnapDiff** Plugin is enabled by default. It can be disabled in the `/opt/bacula/etc/bsnapshot.conf` file with:

```
bsnapdiff=no
```

When the plugin, snapshots, and Accurate mode are all enabled, the following message appears in the backup log:

```
bsnapdiff: incremental accelerator for GPFS enabled
```

4.2 Snapshot Management

GPFS snapshots are managed automatically. The snapshot history is stored in `/opt/bacula/etc/bsnapdiff.dat` or `/opt/bacula/working/bsnapdiff.dat`. Snapshots are retained until they are no longer needed, then automatically removed.

If a user manually removes snapshots that are still needed, a regular, non-accelerated, incremental backup is performed by scanning the file system. Once the snapshot history has been reestablished, the plugin continues to perform incremental backups more efficiently.

4.3 Estimating Performance

The table below shows timing in minutes for a series of backups with the `Accurate = yes` directive in the Job resource and `EnableSnapshot = yes` directive in the Bacula Fileset resource. The client GPFS Fileset in the examples below contains 12.5 million small files.

Ubuntu 22 with 32GB RAM

Total	Accurate Setup	Scan Backup Type
442	0 436 Full, regular Bacula Fileset backup	
187	11 176	Incremental, regular Bacula Fileset backup
12	11	Incremental, BSnapDiff backup

Results

The first backup in this example is a regular Full backup with the BSnapDiff Plugin installed but disabled. For the second backup, an Incremental backup, the plugin is enabled. This is a first regular Incremental backup that establishes the baseline for the BSnapDiff Plugin. It demonstrates the performance of a conventional GPFS Fileset scan with no data changes. The third backup is an Incremental backup with BSnapDiff acceleration active, demonstrating the performance improvement.

An Incremental backup with BSnapDiff active is about 17 times faster than an Incremental with BSnapDiff disabled. This improvement is primarily due to eliminating the overhead of reading unchanged directories and retrieving metadata of a large number of unchanged files to complete the backup accurately. Avoiding unnecessary metadata in file system memory caches also improves response times for other users of the file system.

These tests were performed on a small virtual machine with 32 GiB of RAM. The system is an all-in-one Bacula test environment with the Director, Storage Daemon and File Daemon all running locally. The GPFS Fileset was created on a small 800 GiB local GPFS filesystem with 18 million inodes.

These results may not accurately represent BSnapDiff performance in a finely tuned production environment. In a large production environment, several GPFS servers can host multiple large file systems containing millions or even billions of files. In this scenario, performance will depend on how many files are changing between incremental backups.

4.4 Limitations

Bacula Fileset Directive: ExcludeDirContaining

The BSnapDiff Plugin generates a list of file and directory changes in a random order. A file change may therefore appear before changes to one or more of its parent directories.

As a result, the ExcludeDirContaining directive cannot be used with the plugin. If this directive is used in a Bacula Fileset that includes one or more GPFS Filesets and the BSnapDiff Plugin is enabled, backups will fail with the log message:

```
bsnapdiff: Fileset option ExcludeDirContaining is not currently supported
```

Fileset Directive: HonorNoDumpFlag

As described above, the BSnapDiff Plugin generates a list of file and directory changes in a random order. A file change may therefore appear before changes to one or more of its parent directories.

As a result, the HonorNoDumpFlag directive cannot be used with the plugin. If the user sets this directive to yes in a Bacula Fileset that includes one or more GPFS Filesets and with the BSnapDiff Plugin is enabled, backups fail with the log message:

```
bsnapdiff: FileSet option HonorNodumpFlag=yes is not currently supported
```