InfiniDB®+GlusterFS Guide

Release: 4.6
Document Version: 4.6-1

InfiniDB+GlusterFS Guide
July 2014

Contents

1 Introduction

This guide contains a summary of steps needed to perform an install of InfiniDB with GlusterFS.

1.1 Audience

This guide is written for IT administrators who are responsible for implementing/administering the InfiniDB System.

1.2 List of Documentation

The InfiniDB Database Platform documentation consists of several guides intended for different audiences. The documentation is described in the following table:

Document	Description
InfiniDB Administrator's Guide	Provides detailed steps for maintaining InfiniDB.
InfiniDB Apache Hadoop™ Configuration Guide	Installation and Administration of an InfiniDB for Apache Hadoop system.
InfiniDB Concepts Guide	Introduction to the InfiniDB analytic database.
InfiniDB Installation Guide	Contains a summary of steps needed to perform an install of InfiniDB.
InfiniDB Minimum Recommended Technical Specifications	Lists the minimum recommended hardware and software specifications for implementing InfiniDB.
InfiniDB Multiple UM Configuration Guide	Provides information for configuring multiple User Modules.
InfiniDB SQL Syntax Guide	Provides syntax native to InfiniDB.
Performance Tuning for the InfiniDB Analytics Database	Provides help for tuning the InfiniDB analytic database for parallelization and scalability.
InfiniDB Windows Installation and Administrator's Guide	Provides information for installing and maintaining InfiniDB for Windows.

1.3 Obtaining documentation

These guides reside on our http://www.infinidb.co website. Contact support@infinidb.co for any additional assistance.

1.4 Documentation feedback

We encourage feedback, comments, and suggestions so that we can improve our documentation. Send comments to support@infinidb.co along with the document name, version, comments, and page numbers.

1.5 Additional resources

If you need help installing, tuning, or querying your data with InfiniDB, you can contact support@infinidb.co.

2 Overview

InfiniDB+GlusterFS leverages an Open Source package called GlusterFS (http://www.gluster.org/). GlusterFS is an open source, distributed file system that provides continued access to data and capable of scaling very large data. Failover is configured automatically with GlusterFS, so that if a server goes down, you don't lose access to the data. No manual steps are required for failover. When you fix the server that failed and bring it back online, you don't have to do anything to get the data back except wait. In the meantime, the most current copy of your data keeps getting served from the node that was still running.

Before attempting to configure InfiniDB+GlusterFS, you need to:

- Ensure that you are using at least InfiniDB Enterprise version 4.0 or greater.
- You must be using the same version of InfiniDB on every node in the instance.
- You must be running the exact same Linux O/S on every node in the instance.
- If you are running CentOS 6 or RedHat 6 you must have at least kernel version 2.6.32-279.
- If you are running Ubuntu 12.04 LTS you must have at least kernel version 3.2.0-31.
- You must have an installation with at least 2 Performance Modules.
- You cannot add InfiniDB+GlusterFS to an existing instance: you must start with an empty InfiniDB database. You can, however, add new Performance Modules to an existing InfiniDB+GlusterFS configuration
- You cannot change the InfiniDB+GlusterFS desired number of copies configuration after the initial setup without reloading your database and re-installing InfiniDB.
- You must configure "password-less ssh" for the root user between all InfiniDB nodes, including a loop-back to the localhost. Please reference the "Preparing for Installation" section in the InfiniDB Installation Guide for additional information.

3 Planning for InfiniDB+GlusterFS Installation

Some planning needs to occur before starting any of the Data Duplication installation. There are 3 configuration parameters that will be asked during the InfiniDB installation (postConfigure) and these need to be determined beforehand. The first two parameters deal with the configuration of InfiniDB regardless of using GlusterFS or not. They are:

- **number of Performance Modules (numpm)**
- **number of DBRoots (numdbr)**

Please reference the "Preparing for Installation" section of the InfiniDB Installation Guide for additional information on these two parameters.

The third parameter is:

- **desired data copies (ddc)** - You need to decide how many copies of the data you want to keep. In InfiniDB parlance, if you have 1 copy, then you have no duplication. Two copies would mean that all data is stored in its primary location and 1 backup location.

You need to plan where you intend to store all the copies (including the primary).

The number of DBRoots must be an integral multiple of the number of Performance Modules. Then you will need a total of:

```
numdbr * ddc
```

storage locations. These locations can be individually mountable disk partitions or simple Linux directories. The configuration script will determine the best location for this storage. Ideally, each of these storage locations will be distinct RAID devices. For a system with high IOPS requirements, this is required. Other systems may be able to function with shared locations (e.g. separate partitions on a common disk device).

It is important to understand that data duplication is not free. You will incur additional network I/O (roughly linear in the number of copies configured). While the primary data write may occur over the internal disk subsystem, the replication will use the network. You must understand how this additional network I/O will impact your installation. If you do not carefully plan for the increased network traffic, that traffic may reduce system performance to the point where it is indistinguishable from an outage. InfiniDB recommends strongly that a dedicated network be installed to handle the replication traffic. You must keep all this in mind when choosing the number of desired data copies, and choose the lowest number of copies that meets your data replication requirements. In any practical Data Warehouse application, it is virtually certain that a network infrastructure of at least 10Gb/s will be required to produce usable results.

Next, if you have chosen mountable storage, you will need to allocate

```
numdbr * ddc / numpm
```

individually-mountable disk partitions on each PM in the cluster. You also must lay down a filesystem on each of these partitions. Ext2 is sufficient, but you may use ext3, ext4, xfs, etc. Make sure the partitions are unmounted when you are done formatting them and do not add any entries for them in /etc/fstab. The configuration script will ask you, for each PM, to list the device names of these partitions so you should write them down.

If you have chosen simple Linux directories, the storage directories will be created off a common root, which defaults to `/usr/local/Calpont/gluster` for root-user and `$HOME/Calpont/gluster` for non-root.

4 Install InfiniDB+GlusterFS

First, you must install GlusterFS version 3.3.1 or later. See http://www.gluster.org/download/ for download options and installation instructions.

After installation, ensure that the gluster service is running on all nodes and that the gluster executable is visible in the InfiniDB install user path. You will be presented with a glusterfs option for storage during postConfigure.

4.1 InfiniDB+GlusterFS Installation Configuration

Install InfiniDB using the postConfigure script as instructed in the InfiniDB Installation Guide.

There are some additional prompts that will appear during postConfigure now that the InfiniDB+GlusterFS package has been installed on each PM. These prompts will display under the heading

```
===== Configuring InfiniDB Data Redundancy Functionality =====
```

Respond to the prompts according to your decisions outlined in the “

Planning for InfiniDB+GlusterFS Installation"section.

5 Adding Performance Modules

You can add Performance Modules (PM's) to an InfiniDB+GlusterFS configuration. But before adding, some initial prep work is required:

- You can only add PM's in multiples of the desired data copies value
- You must add the PM's in separate steps of *count* PM's. That is, if your desired data copies value is 2 and you want to add 4 PM's, you need to add 2 PM's and then add the final 2 PM's in two separate processes.
- You should keep in mind that the original InfiniDB+GlusterFS installation evenly spread the duplication workload among the available PM's. When you add PM's, the duplication is spread among that install group of count PM's only. There may be performance differences that need to be taken into account.

5.1 Installing InfiniDB+GlusterFS packages on new PM's

You can perform the following steps on the new PM's without stopping InfiniDB on the current cluster. Make sure that all of the prerequisites for an initial InfiniDB+GlusterFS installation are met before proceeding.

1. Ensure all the prerequisites have been met as outlined in the "Overview" section above:
   Items needing to be performed:
   - Minimum OS requirements
   - Setup password-less ssh connections to and from the new PMs being added

2. Install GlusterFS on every new PM and make sure the gluster service is running.

5.2 Running addModule Console Command

Follow the normal instructions for adding PMs (See the "Adding Modules" section of the InfiniDB Administrator's Guide), making sure to add the PMs as a group with a single addModule command. The following is an example addModule on a InfiniDB+GlusterFS with Storage configured system:

```
InfiniDB> addmodule  pm 2 srvnewpm1,srvnewpm2
addmodule    Wed Oct 24 16:08:03 2012

System is configured with InfiniDB Data Redundancy, DBRoot Storage will
will be created with the Modules during this command.
Also the InfiniDB Data Redundancy Packages should already be installed on the
Performance Modules being added and password-less ssh should be setup on those
modules.

           Do you want to proceed: (y or n) [n]: y

Number of DBRoots Per Performance Module you want to add
           Please enter: 1

Data Redundancy Storage Type is configured for 'storage'
```

```
You will need 4 total storage locations and 2 storage locations per PM. You will
now be asked to enter the device names for the storage locations. You will enter
them for each PM, on one line, separated by spaces (2 names on each line).

Storage Device Names for pm3
           Please enter: LABEL=alphdbr3 LABEL=albrk03

Storage Device Names for pm4
           Please enter: LABEL=alphdbr4 LABEL=albrk04

Filesystem type for these storage locations (ext2,ext3,xfs,etc)
           Please enter: ext2

Adding Modules pm3, pm4, please wait...
Add Module(s) successfully completed

Enabling Modules
Successful Enable of Modules

Adding DBRoots
New DBRoot IDs added = 3, 4

Assigning DBRoots

DBRoot IDs currently assigned to 'pm3' =

Changes being applied

DBRoot IDs newly assigned to 'pm3' = 3

Successfully Assigned DBRoots
DBRoot IDs currently assigned to 'pm4' =

Changes being applied

DBRoot IDs newly assigned to 'pm4' = 4

Successfully Assigned DBRoots

Run Data Redundancy Add DBRoots

Successfully Completed Data Redundancy Add DBRoots

addModule Command Successfully completed: Run startSystem command to Activate newly
added Performance Modules
```

6 Upgrading InfiniDB+GlusterFS

There is nothing unique to upgrading the InfiniDB software that is using GlusterFS. During an InfiniDB upgrade, postConfigure will recognize that GlusterFS is installed and upgrade accordingly. See the "Upgrading InfiniDB" section in the InfiniDB Installation Guide.