
OpenFest 2007

5th annual conference

Gluster Filesystem

Marian Marinov - mm@yuhu.biz
System Architect - Siteground.com

Agenda

Gluster Design

- kernel
- filesystem design

Installation

- FUSE
- glusterfs

Basic Configuration

- glusterfs-volume.spec

Gluster Filesystem Design

In the kernel

**Requires FUSE
FUSE Module**

The engine

**Server & Client
Transport Modules
Translators
Scheduler Modules**

Gluster Filesystem Design

Applications:

Server – glusterfsd

Client – glusterfs

Transport Modules:

For TCP/IP transport

transport-type tcp/server

For Infiniband transport

transport-type ib-sdp/server

For Infiniband Verbs transport

transport-type ib-verbs/server

Gluster Filesystem Design

Translators

The idea – GNU/Hurd

- **Performance translators**
- **Clustering translators**
- **Scheduling translators**
- **Other translators**

Gluster Filesystem Design

Performance translators

- **Read Ahead**
- **Write Behind**
- **Threaded I/O**
- **IO-Cache**
- **Stat Pre-fetch – still not in 1.3.7**

Gluster Filesystem Design

Clustering translators

- **Automatic File Replication (AFR)**
- **Stripe**
- **Unify**

Gluster Filesystem Design

I/O Schedulers

- **Adaptive Least Usage (ALU)**
- **Non-uniform filesystem architecture (NUFA)**
- **Random**
- **Rand-Robin**

Gluster Filesystem Design

Other translators

- **trace**
- **posix**
- **rot-13**
- **client**
- **server**

Gluster Filesystem Design

Volume Specification – Basic Configuration CLIENT

volume client

type protocol/client

option remote-port 7000

option transport-type tcp/client

option remote-host localhost

option remote-subvolume brick

end-volume

volume iot

type performance/io-threads

option thread-count 4

subvolumes client

end-volume

Gluster Filesystem Design

Volume Specification – Basic Configuration CLIENT

```
volume wb
  type performance/write-behind
  subvolumes iot
end-volume
```

```
volume trace
  type debug/trace
  subvolumes wb
end-volume
```

Gluster Filesystem Design

Volume Specification – Basic Configuration CLIENT

volume unify

type cluster/unify

option scheduler rr # check alu, random, nufa

option rr.limits.min-free-disk 5 # 5% of free disk is min

option namespace namespace-child

subvolumes child1 child2 child3 child4

end-volume

clients: client[1-n], n < 32bit number :D

Gluster Filesystem Design

Volume Specification – Basic Configuration SERVER

volume brick

type storage/posix

option directory /home/gluster-export

end-volume

volume server

type protocol/server

option transport-type tcp/server

option bind-address 192.168.1.10

option listen-port 6996 **# Default is 6996**

option client-volume-filename /etc/glusterfs/glusterfs-client.vol

subvolumes brick

option auth.ip.brick.allow *

end-volume

Gluster Filesystem Design

Volume Specification – Basic Configuration SERVER

Distributed filesystems - GFarm

? ? ? ? ? ? ? ?
? ? ? ? ? ? ? ? ?
? ? Въпроси ? ?
? ? ? ? ? ? ? ?
? ? ? ? ? ? ? ?
? ? ? ? ? ? ? ?