GlusterFS
Current Features & Roadmap

Niels de Vos
GlusterFS co-maintainer
ndevos@redhat.com
Agenda

• Introduction in the Gluster Community
• Current stable releases
  • History of feature additions
• Plans for the upcoming 3.7 release
• Expectations of the next major release
Introduction in the Gluster Community

- Different roles
  - Users, testers, supporters, developers, editors, ...
- Different organizations
  - Products based on / containing GlusterFS
  - Service, consulting and support
  - Integration in other (Open Source) projects
Introduction in the Gluster Community

- Regular IRC meetings
- Discussions and support over mailinglists and on IRC
- Providing packages (RPMs, DEBs)
- Work with different Linux and BSD distributions to improve portability and availability
- Infrastructure hosting for Gluster related projects
  - Gerrit and Jenkins for code review and testing
  - Gluster Forge for git/wiki hosting of projects
Introduction in the Gluster Community

- Some numbers from 2014
  - Approx. 175 IRC participants
  - Two main mailinglists reach ~600 emails/month
  - 100/60 active users/devs posting to the lists
  - Around 2200 patches merged in the master branch
  - Patches of ~90 developers got included
Current stable releases

• Maintenance of three minor releases
  • 3.6, 3.5 and 3.4
• Bugfixes only, non-intrusive features on high demand
• No fixed release schedule
• Patches get backported to fix reported bugs
Features included in version 3.4

- WORM: Write Once Read Many
- Operating versions for GlusterD
- Block device translator
- Duplicate Request Cache (used with NFS)
- Server Quorum
- libgfapi for native GlusterFS support in applications
- Eager Locking
- NFSv3 ACL support
Features included in version 3.5

- File Snapshot for qcow2 files
- GFID access
- On-Wire (de)compression
- Quota Scalability
- Readdir ahead
- Zerofill
- Brick Failure Detection
- Parallel geo-replication
Quota in 3.5

• Before 3.5
  • Client side enforcement
  • Configuration in volume files would block scalability
  • GFID accesses could cause incorrect accounting
  • Only hard quota supported

• In 3.5
  • Server side enforcement
  • Better configuration management for scalability.
  • GFID to path conversion enables correct accounting.
  • Both hard and soft quotas supported
Geo-replication in 3.5

- Before 3.5
  - Merkle tree based optimal volume crawling
  - Single driver on the master
  - SPOF
- In 3.5
  - Based on changelog
  - One driver per replica set on the master
  - No SPOF
Features included in version 3.6

- Improved SSL support
- Heterogeneous bricks
- Volume wide locks for GlusterD
- Volume Snapshots
- User Serviceable Snapshots
- AFR refactor
- RDMA improvements
- Disperse translator for Erasure Coding
Plans for the upcoming 3.7 release

Feature freeze at the end of February
- Small-file performance enhancements
- Tiering, rack-aware placement and more
- Trash translator for undelete operations
- Netgroups and advanced exports configuration (NFS)
- BitRot detection
- Support for NFS Ganesha clusters
Small-file performance enhancements in 3.7

- Multithreaded epoll (transport layer)
- Caching stat and xattrs on the bricks
- Migrate .glusterfs to SSDs
- Batching of RPCs per file access
Data Classification in 3.7

- Mapping file characteristics to subvolume characteristics
- File characteristics:
  - Size, age, access rate, type (filename extension)
- Subvolume characteristics:
  - Physical location, storage type, encoding method
- User provided mapping via 'tags'
- Implemented using 'DHT over DHT' pattern
Netgroups and Exports for NFS in 3.7

- More advanced configuration for authentication based on /etc/exports like syntax
- Support for netgroups
- Patches written by Facebook developers
- Forward ported from 3.4 to 3.7
- Cleanups and posted for review
NFS Ganesha support in 3.7

- Optionally replaces Gluster/NFS
- Supports NFSv4 with Kerberos
  - pNFS support for Gluster Volumes follows later
- Modifications to Gluster internals
  - Upcall infrastructure
  - Gluster CLI to manage NFS Ganesha
  - libgfapi improvements
- High-Availability based on Pacemaker and Corosync
Plans for the next 4.0 release

- Intended for scalability and manageability improvements
- Support for multiple networks
- New Style Replication
- Improved Distributed hashing Translator
- Composite operations in the GlusterFS RPC protocol
- Coherent client-side caching
- Native ReST APIs for management and monitoring
- ... and much more
GlusterFS 4.0 – What's next?

- Code name for the release? Open to suggestions
- Submissions for feature proposals is still open!
- Implementing of key features has started
- Voting on feature proposals during design summit
  - Tentatively planned for March/April timeframe
Resources

Mailing lists:
- gluster-users@gluster.org
- gluster-devel@gluster.org

IRC:
- #gluster and #gluster-dev on Freenode

Links:
- http://gluster.org/
- http://forge.gluster.org/
- http://www.gluster.org/community/documentation/
Thank you!

Niels de Vos
ndevos@redhat.com
ndevos on IRC