NFS Ganesha : Weather Report

Jiffin Tony Thottan
Software Engineer

Agenda

- Recap
  - nfs-ganesha
  - architecture
  - configuration file
  - helper scripts
  - last stable release 2.7.3 - April 2019

- Ongoing developments in 2.8
  - delegations
  - sticky grace period
  - labelled NFS
  - async I/O

- High Availability
  - storhaug
  - without failovers

Thanks for the Awesome Work

- ➢ Jeff Layton <jlayton@redhat.com> jlayton

- ➢ Soumya Koduri <skoduri@redhat.com> skoduri

- ➢ Kaleb S. Keithley <kkeithle@redhat.com> kkeithley

- ➢ Daniel Gryniewicz <dang@redhat.com> dang

- ➢ Frank Filz <ffilz@redhat.com> ffilz

RECAP

NFS-Ganesha

> A user-space, protocol-compliant NFS server

> Supports NFS v3, 4.0, 4.1, pNFS, 4.2 and 9P from the Plan9 operating system.

> Provides a File System Abstraction Layer(FSAL) to plug in to any storage mechanism

> Can provide simultaneous access to multiple file systems.

> Small but active and growing community ; CEA, Red Hat, IBM, SUSE are active participants

NFS-Ganesha architecture
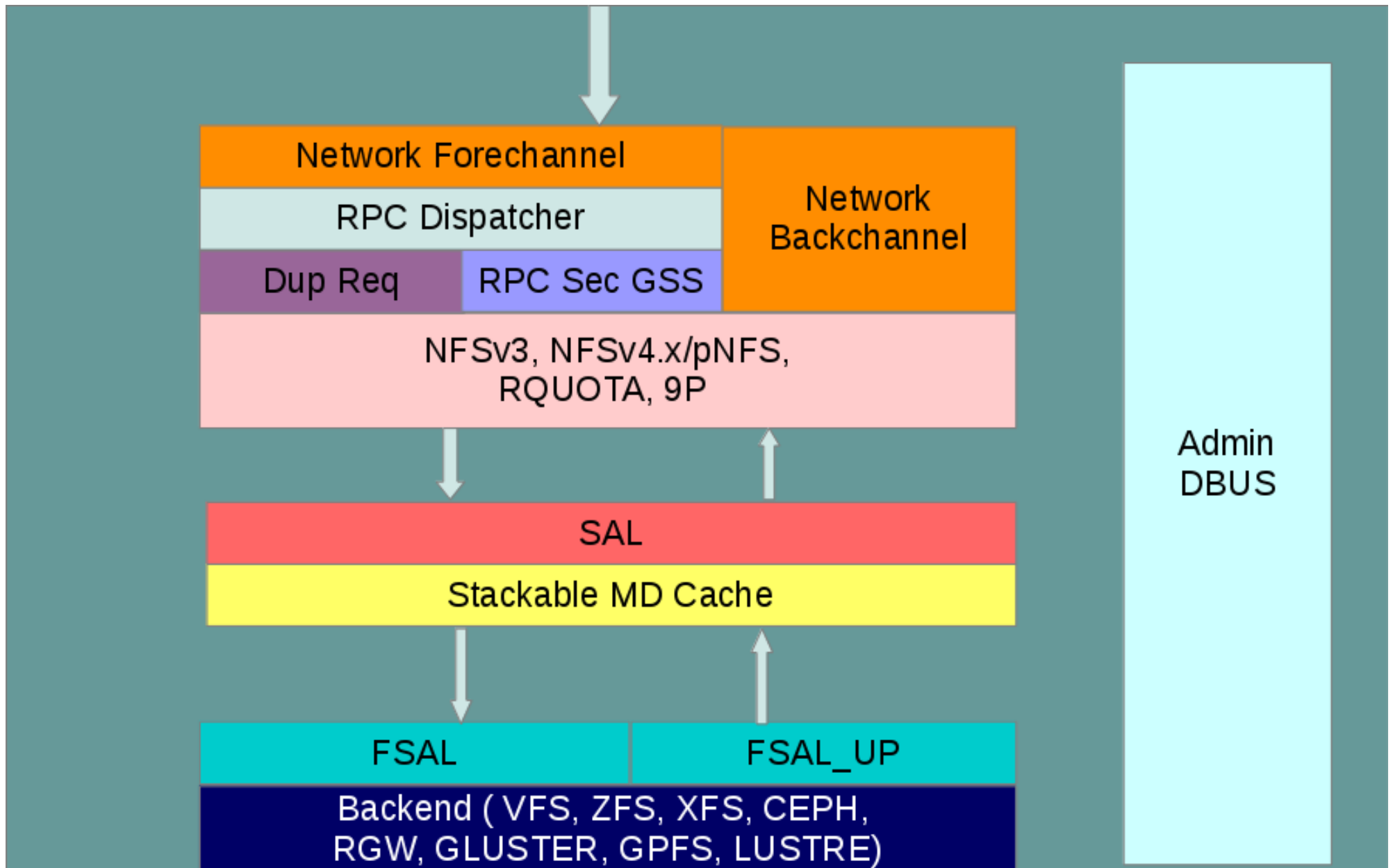
Benefits of NFS-Ganesha

➢ Can manage huge meta-data caches

➢ Dynamically export/unexport entries using D-Bus mechanism.

➢ Easy access to the services operating in the user-space (like Kerberos, NIS, LDAP)

➢ Provides better security and authentication mechanism for enterprise use

➢ Portable to any Unix-like file-systems

Configuration file aka ganesha.conf

- ➤ NFS_CORE_PARAM {}
- ➤ NFSV4 {}
- ➤ NFS_KRB5 {}
- ➤ CACHEINODE {}
- ➤ EXPORT { FSAL {} CLIENT {}}
- ➤ LOG {}

Helper Scripts : src/scripts/

- ➢ ganesha-conf.py : used to modify blocks in ganesha.conf
- ➢ ganesha_mgr.py : commandline tool for admin of nfs-ganesha
- ➢ knfs2ganesha-exports.py : converts knfs exports to ganesha
- ➢ ganeshactl.py : PyQt4 GUI tool for admin of nfs-ganesha

Last Stable release 2.7.3 - April 2019

- ➢ Pull up libntirpc 1.7.3

- ➢ Support for async I/O operations in the FSALs

- ➢ Many bug fixes (many backported to V2.6-stable)

- ➢ Strip out legacy dirent cache

- ➢ FSAL_LUSTRE reintroduced, built on FSAL_VFS

- ➢ NLM Fixups to allow Windows NFS client to work

- ➢ Grace period enhancements

Continued ..

➢ FSAL_GLUSTER implementation of V4.2 SEEK

➢ Improved V4.2 fallocate operation with CephFS

➢ Make expire_time_attr settable in EXPORT_DEFAULTS

➢ Compilation fixes for FreeBSD

➢ Implement NFS v4.1 TEST_STATEID

Ongoing Developments

Delegations

- Part of nfsv4

- Server transfer authority of file to client

- support was added from 2.2 onwards

- fully fledged write delegations support

- getattr_cbk

- currently supported FSAL: GPFS, CEPHFS, GLUSTERFS

Sticky grace period

➢ Grace period : Reclaiming window for the client

➢ For clustered file system, inflight operations can be affected

➢ Better handling of operations during grace period

➢ kinda of bit flexible time interval.

➢ with help of ref counting, flags and mutex variable

Labelled NFS

➢ part of nfsv4.2 protocol

➢ provide mechanism to use selinux context

➢ configurable option "Security_Label" in Export Block

➢ supported FSAL : CEPHFS

Async I/O

- ➤ Ongoing

- ➤ Supported if only FSAL has async api's

- ➤ Made read and write calls to vectored

- ➤ Added async ability in COMPOUND OPS

- ➤ Converted async call to synchronous with help of special call back

High Availability

Storhaug

- High Availability setup utility for NFS-Ganesha

- with help of CTDB

- Change "CTDB_MANAGES_NFS=yes" in ctdbd.conf

- Plus add few tunables

- Supported operations: setup/teardown of cluster, add/remove exports

- Currently it is written for GlusterFS volumes

HA without failover

- Multiple independent ganesha-servers serving same

- no clustering software

- if server crashes, just restarts

- node goes down, manual intervention is needed

- ganesha can be run on vm, baremetal or even in containers containers

- with help of rados objects

- written for CephFS

References

- Links (Home Page):
  - https://github.com/nfs-ganesha/nfs-ganesha/wiki
  - https://lists.nfs-ganesha.org/archives/
- References:
  - https://github.com/gluster/storhaug/wiki
  - https://jtlayton.wordpress.com/2018/12/10/deploying-an-active-active-nfs-cluster-over-cephfs/
  - https://www.ietf.org/rfc/rfc3530.txt
  - https://selinuxproject.org/page/Labeled_NFS

Contact

- Mailing lists:
    - devel@lists.nfs-ganesha.org
    - support@lists.nfs-ganesha.org

- IRC:
    - #ganesha on freenode

Q & A

Thank You