HEKETI FUNCTIONALITY INTO GLUSTERD2

as seen by heketi developers...

Michael Adam

hacker, Red Hat ;-)

Raghavendra Talur

developer, Red Hat

@raghavendra_t

AGENDA

1. Glusterd.Next
2. Heketi
3. Heketi features into Glusterd2 ?!
4. Demo/PoC

GLUSTERD.NEXT

1. RESTful API
   Next generation Infrastructure setup and management needs tool to tool communication
2. Abstract bricks and treat Volumes as smallest unit
   Opens possibility of many features like brick splitting, dynamic replication, automatic failover/migration
3. Easier and safer management operations
   Need to reduce involvement of admin in operations like rebalance, self heal, node & disk replacement
4. Better integration with other tools/programs
   Better coupling with Samba, NFS-Ganesha, Swift to provide more guarantees on service

GLUSTERD2

1. People started to talk about a glusterd2 project
2. There were concrete designs/ideas for the api and implementation of disk-manangement features...

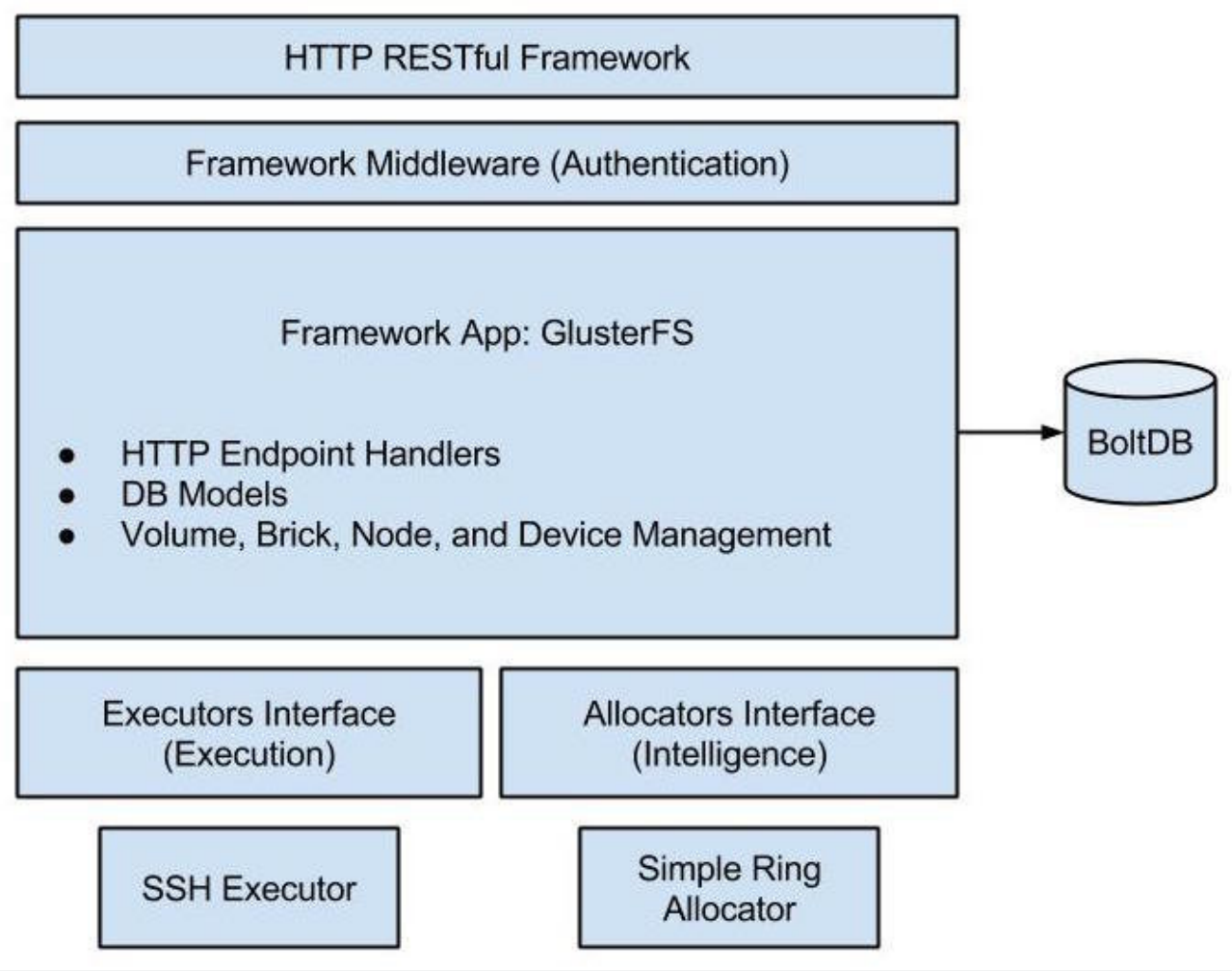
ENTER HEKETI

1. project glusterd2 was started in mid 2015
2. In early/mid 2015, Luis Pabon started an *external* project https://github.com/heketi/heketi
3. Should probably have been https://github.com/gluster/heketi ...
4. Implements several of the ideas of higher-level disk/volume-management of gd.next / gd2 (much more narrow scope)
5. heketi was put into production in mid 2016 (Aplo/CNS)
6. development can be in sync with kubernetes pace!
7. golang, ~ 28K loc, > 50% test code

HEKETI

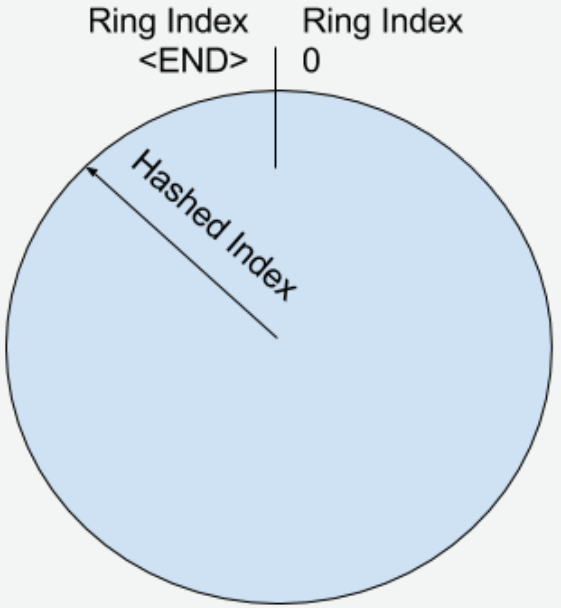
1. RESTful API
   (=> One Glusterd.Next request satisfied)
2. hide vol create complexity, disk management
   (=> Second request for Glusterd.Next satisfied)
3. implements complicated admin ops (newer)...
   (=> third request partially satisfied)
4. Manage multiple Gluster clusters
5. allow for N+1 scaling

HEKETI - DESIGN



HEKETI - DESIGN

```
[
  {Zone3, Node 192.168.13.100, Device /dev/sda},
  {Zone2, Node 192.168.12.100, Device /dev/sda},
  {Zone4, Node 192.168.14.100, Device /dev/sda},
  {Zone1, Node 192.168.11.100, Device /dev/sda},
  {Zone3, Node 192.168.13.100, Device /dev/sdb},
  {Zone2, Node 192.168.12.100, Device /dev/sdb}
]
```



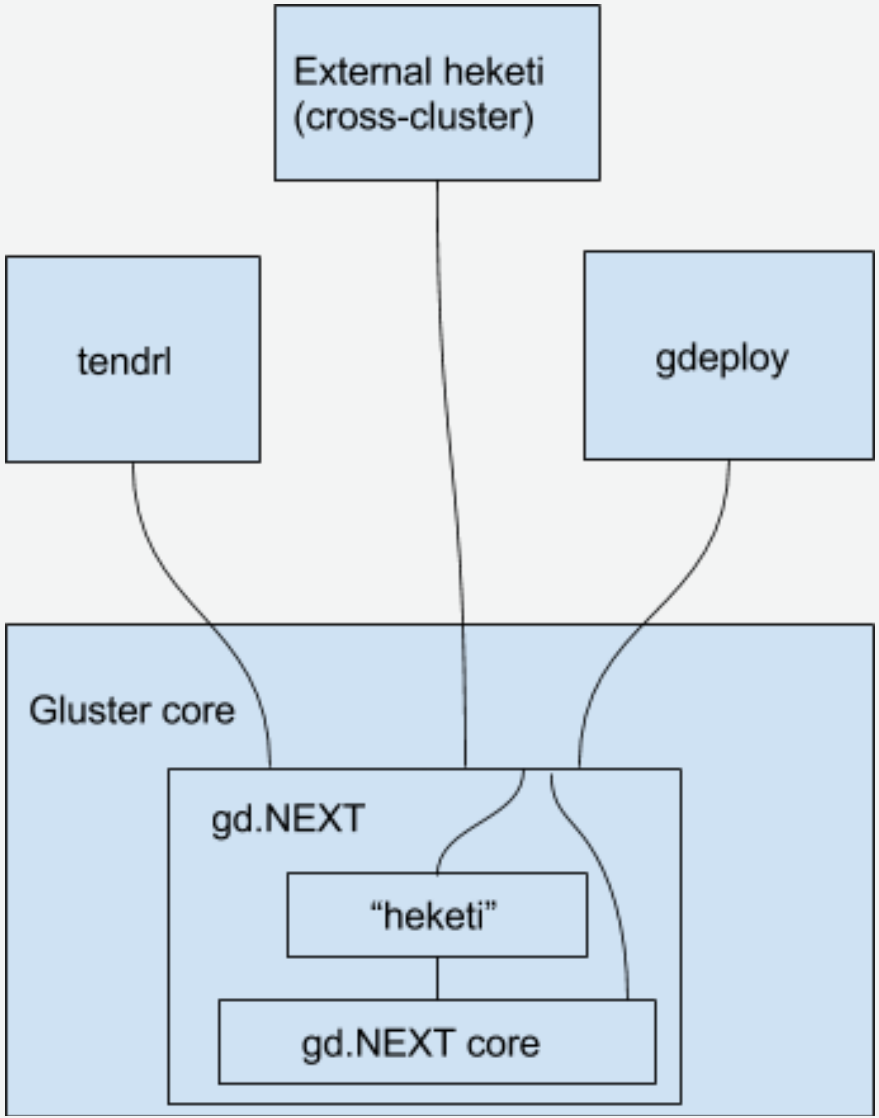
HEKETI - TRADE-OFFS

1. state additional to gluster
   => no brown-field, no cli-mixing...
2. need to mirror gluster features in heketi for support
3. Heketi is not distributed
   =>spof, but HA in kubernetes
4. Acceptance in gluster community :-)
   => almost only used by kubernetes...

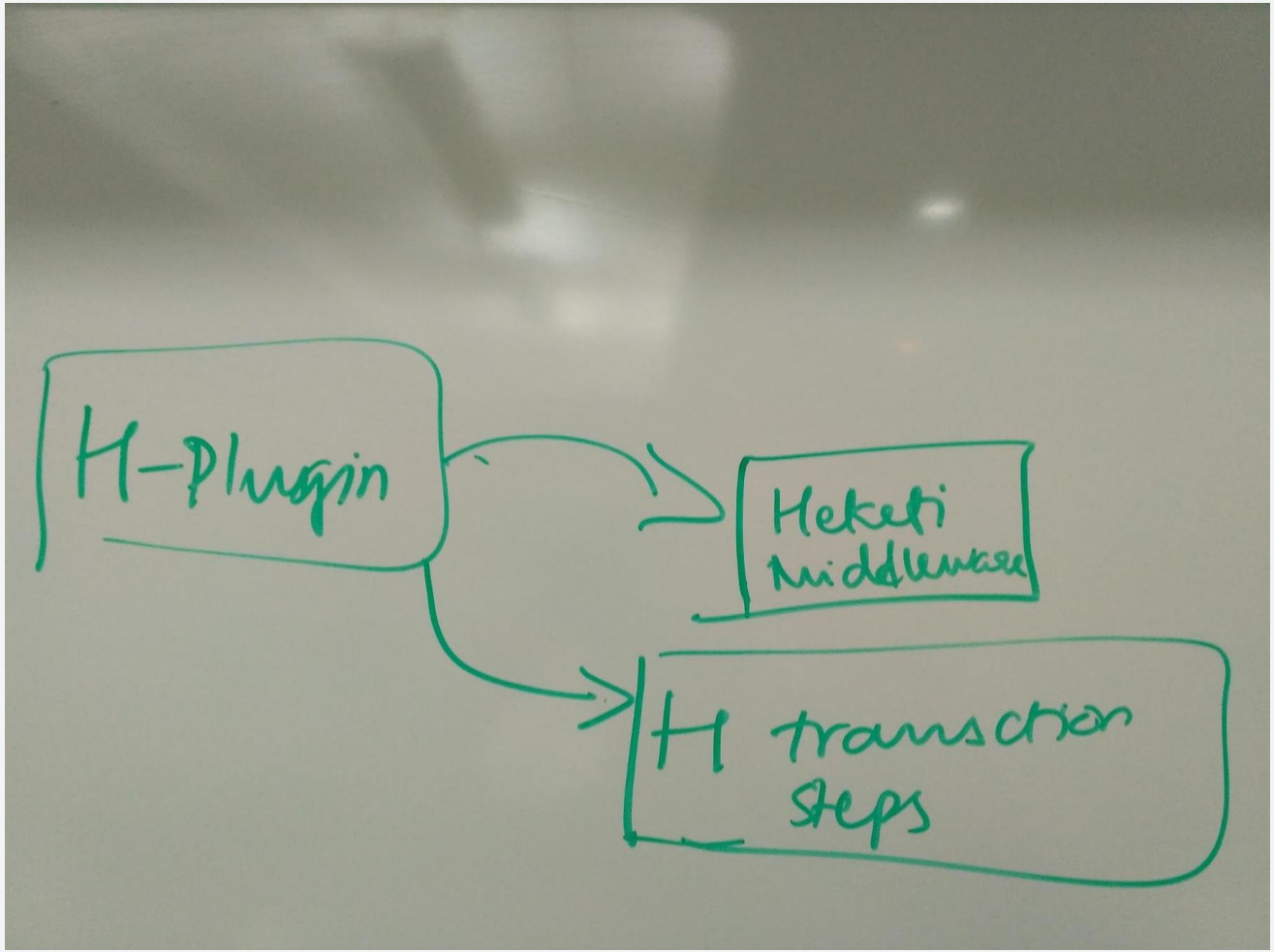
HEKETI FEATURES IN GLUSTERD2

1. Volume create should support providing bricks OR size
   maintains backward compatibility
2. Adding new volume types to Gluster becomes easier
   small changes to brick creator/selector code
3. New features related to brick selection should work
   For example, brick splitting as a feature should work with the design

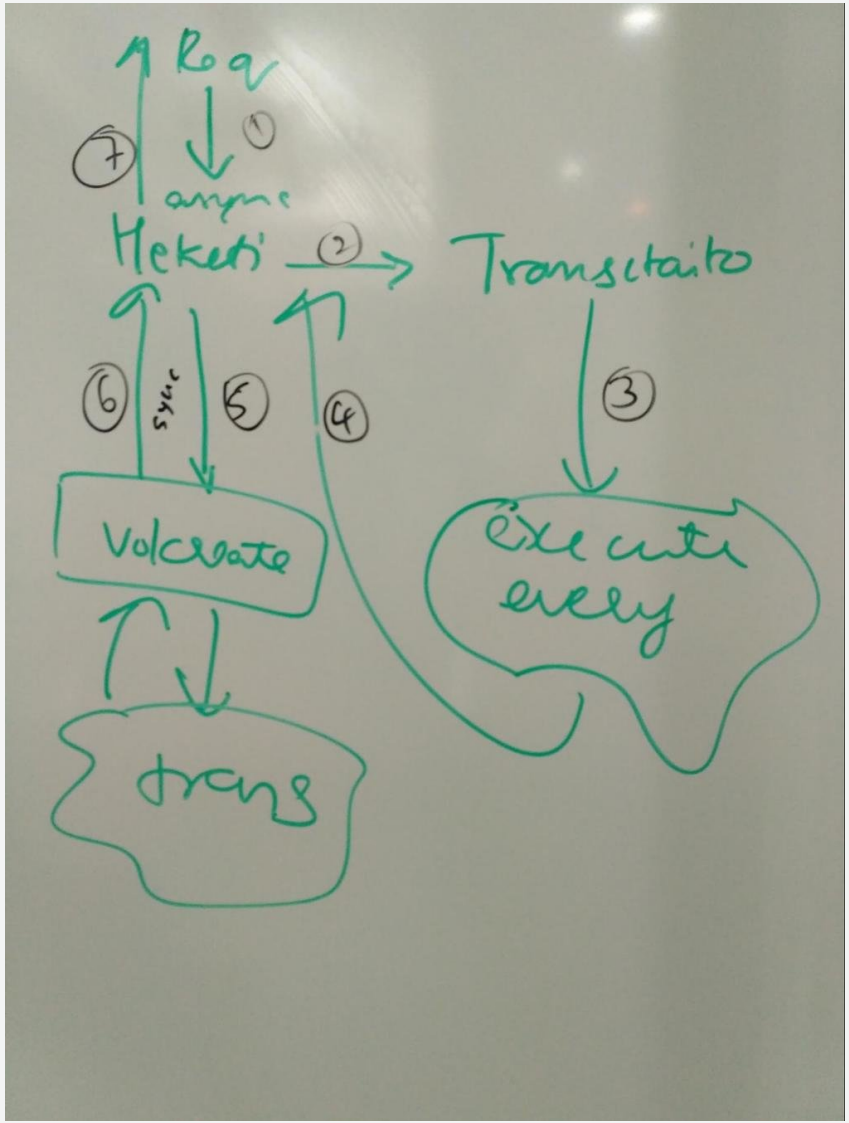
HEKETI FEATURE IN GLUSTERD2



HEKETI FEATURE IN GLUSTERD2



HEKETI FEATURE IN GLUSTERD2



HEKETI FEATURE IN GLUSTERD2

DEMO