# A path to OpenAFS 2.0: rxgk, IPv6, and more

Benjamin Kaduk kaduk@mit.edu bkaduk@akamai.com

20 August, 2015



#### Roadmap History

What's in 1.87

Getting to 2.0 rxgk

IPve

. . .

# Things we want

- rxgk
- ▶ IPv6 support
- read-write replication
- per-file ACLs

rxgk?

- rxgk?IPv6 support

- rxgk?
- ► IPv6 support
- read-write replication

- rxgk?
- ► IPv6 support
- read-write replication
- per-file ACLs

- rxgk?
- ▶ IPv6 support
- read-write replication
- per-file ACLs

That's where we want to be (and where we think we'll actually be able to get to).

In order to get there, we have to know where to start — where are we coming from?

### History

People have wanted rxgk for longer than I've been using AFS!

Back in 2006, the Elders published

https://www.openafs.org/pages/no-more-des.html.

After the 1.6 release, the build system will be modified to build OpenAFS without kaserver unless it is specifically requested.

The OpenAFS Elders endorse the development of the rxk5 and rxgk security classes in order to enable the use of Kerberos 5 ciphers other than single DES for both authentication and data security between AFS clients and servers.

When OpenAFS is capable of supporting Kerberos 5 with non-DES ciphers the major version number will be changed to "2".



- ► OpenAFS is getting left behind
- World is running out of v4 addresses
- ▶ IPv6 is more and more common
  - cell networks
  - home networks, too!
  - Some ISPs charging more for v4 service
- v6 support is more than just replacing 32 bits with 128 bits

- OpenAFS is getting left behind
- World is running out of v4 addresses
- ▶ IPv6 is more and more common
  - cell networks
  - home networks, too!
  - Some ISPs charging more for v4 service
- v6 support is more than just replacing 32 bits with 128 bits
- and more than just keeping both and a way to choose
- (but we do still need to bump the vldb format to do anything)

#### IPv6

I'm not an expert on IPv6 for AFS; talk to Jeff or Mike or Simon or some number of other people who know more. But...

- privacy addresses callbacks might not be deliverable!
- larger minimum packet sizes; rx must adapt its ideas of MTU handling
- Much larger physical packets possible, but rx jumbograms suck
- fragmentation is not automatic must be prepared to resend in smaller chunks
- Will the OS let you get the ICMPv6 fragmentation needed notices?
- everybody becomes multihomed



#### IPv6

- Many, many RPCs to update
- ▶ We're already "in the middle of" and RPC refresh . . . for the past five¹ years
- rxdebug
- happy eyeballs



<sup>&</sup>lt;sup>1</sup>maybe more?

#### Roadmap

History

#### What's in 1.8?

Getting to 2.0

rxgk

IPV

. . .

#### Why 1.8?

- ▶ openafs-stable-1\_6\_x was branched on August 10, 2010
- ▶ 4658 commits on master since then, but only 2033 commits on 1.6 since then. (Most of the commits on 1.6 are cherry-picks from master, but not all; I didn't write the script to check.)
- ▶ Those other 2.5k commits have some useful features in them!
- master is actually pretty usable right now; rxgk and whatnot will probably break things temporarily. A new release gives a stable baseline to start from.
- We're also overdue for some breaking/intrusive changes that should only be done at a major release boundary.



#### What's in 1.8?

#### Already on master:

- Code cleanup/robustness (e.g., from static analyzers, compiler warnings fixes, and general refactoring) — thanks, Simon!
- libtool in the build system for more reliable shared libraries, and other build system cleanup
- pthreaded dbservers
- Greatly reduced krb5 dependencies
- More reliable autoconf logic
- Windows stuff is quite different
- Some parts of the documentation are not embarassing
- Improvements changes to the rx stack (again, thanks, Simon!)
- libroken and libhcrypto from Heimdal
- Cache bypass on Linux
- OpenAFS Portable Runtime (opr) thanks, Simon!
  - 4 D > 4 A > 4 B > 4 B > B = 40 Q C

Some afsd.fuse improvements

#### What's in 1.8?

#### Already on master:

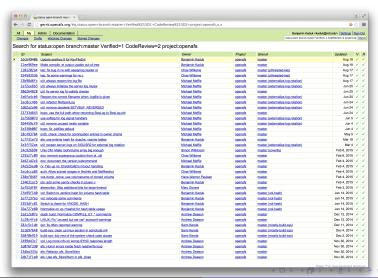
- cmd refactoring
- Something resembling a test suite framework
- New default for ihandle sync-ing behavior
- Some bosserver refactoring
- Ability to control ugen/ubik client flags more closely
- memcache improvements
- rxgen per-opcode stats
- openafs-client.conf and openafs-server.conf
- opr\_Assert() vs. opr\_Verify()
- More dbserver threads by default
- buserver uses libutil's logging (format changes)
- other changes to server logging, rotation
- CacheTruncateDaemon improvements

#### What's in 1.8?

#### Already on master:

- ptserver wire format arrays' length is capped
- bos, pts emit error messages on stderr (not stdout)
- vos remaddrs
- opr\_softsig

# In gerrit ready to be merged





# In gerrit ready to be merged

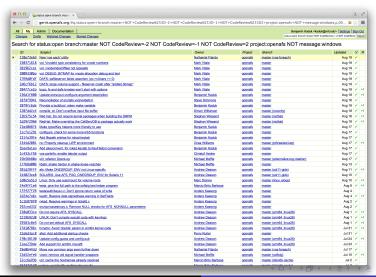
38 changes with good code review. Need to actually get merged.



# In gerrit ready to be merged

- ▶ Jenkins hash in volume, vnode, vcache, dcache hash tables
- Subnet ranges in NetInfo/NetRestrict

## In gerrit needing review

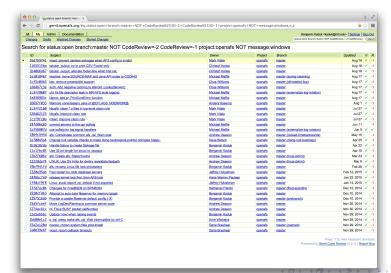




# In gerrit needing review

54 changes (just barely on two pages)

- update bos restricted mode
- encrypt inter-volser traffic
- Revert "Lockless path through afs\_linux\_dentry\_revalidate"
- afs: Don't retry timed-out RW operations forever



#### 33 changes with -1 review

- viced: prevent useless salvages when AFS config is invalid
- auth: Add negative caching to afsconf\_LookupServer()
- convert servers to the opr softsig
- Fast restart for Ubik database servers

#### 33 changes with -1 review

- viced: prevent useless salvages when AFS config is invalid
- auth: Add negative caching to afsconf\_LookupServer()
- convert servers to the opr softsig
- Fast restart for Ubik database servers
- Attempt to auto-tune fileserver for memory target
- Provide a usable fileserver default config (-X)

#### 33 changes with -1 review

- viced: prevent useless salvages when AFS config is invalid
- auth: Add negative caching to afsconf\_LookupServer()
- convert servers to the opr softsig
- Fast restart for Ubik database servers
- Attempt to auto-tune fileserver for memory target
- Provide a usable fileserver default config (-X)
- p\_osi\_sleep make afs\_osi\_Wait interruptible by ctrl-C



#### What's not in 1.8?

- kaserver (by default)
- pam modules (by default)
- ► Linux 2.4 support
- Documentation for AIX-, HP-UX, and IRIX-specific installation in the QuickStartGuide
- NFS translator
- A comprehensive test suite
- ▶ afsd -settime
- libjuafs.a just use libuafs.a
- LWP fileserver
- Cruft from MR-AFS in the bos utility
- Netscape plugin (what is it???)
- a -k argument to the fileserver
- ZFS\_BOZONLOCK\_ENV



```
Roadmap
History
```

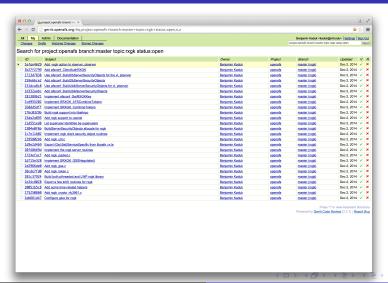
What's in 1.87

```
Getting to 2.0 rxgk IPv6
```

# Specification

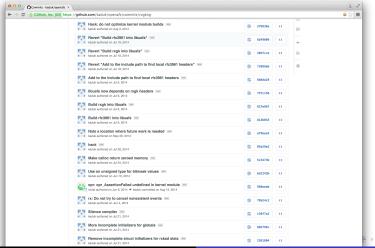
```
⊕ ⊖ ⊖ / # draft-wilkinson-afs3-mg/ ×
← → C A https://tools.ietf.org/html/draft-wilkinson-afs3-rxqk-afs-08
                                                                                           The VL RegisterAddrsAndKey RPC is described by the following RPC-L:
              struct RXGK ServerKeyDataRequest {
                  afs int32 enctypes<>;
                  opaque nonce1[20];
              }:
              struct RXGK ServerKevDataResponse {
                  afs int32 enctype;
                  afs uint32 kvno;
                  opaque nonce2[20];
              1:
              const RXGK MAXKEYDATAREQUEST = 16384;
              const RXGK MAXKEYDATARESPONSE = 16384;
              typedef opaque keyDataRequest<RXGK MAXKEYDATAREQUEST>;
              typedef opaque keyDataResponse<RXGK MAXKEYDATARESPONSE>;
              VL RegisterAddrsAndKey(
                  IN afsUUID *uuidp,
                  IN afs int32 spare1,
                  IN bulkaddrs *ipaddr,
                  IN afs int32 secIndex,
                  IN kevDataRequest *request,
                  OUT keyDataResponse *response) = XXX;
         uuidp: The fileserver's UUID.
         sparel: Unused. (Clients SHOULD pass zero.)
         ipaddr: The list of addresses to register as belonging to this
                fileserver.
```

### Implementation



### **Implementation**

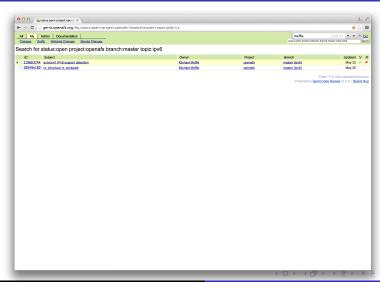
#### https://github.com/kaduk/openafs/commits/rxgkng



# Specification

'()

### Implementation



# Implementation



#### Your Feature Here

- read-write replication
- per-file ACLs

#### Thanks!