



Isilon and FreeBSD

Peter Godman, Director of Software Engineering

Neal Fachan, Distinguished Engineer

ISILON CONFIDENTIAL



Isilon's product, "OneFS" and Isilon IQ.

- Hardware + software product.
- "Nodes" are 2u chassis with CPU, NVRAM, disks, IB, and ethernet.
- Nodes collectively are a NAS Appliance: NFS, CIFS, HTTP, FTP, WebDAV.
- Single volume, single namespace. Not namespace aggregation.
- Scalable to 1PB+, 100 nodes+.
- Traditional NAS constructs, like snapshots and protection level, are per-directory, not per-volume.
- IP is in software, not hardware.
- Complexity for the programmer, simplicity for the user.





Committed Isilon Modifications to FreeBSD

- Jeff Roberson's SMP VFS changes were sponsored by Isilon.
- Shared lookups & shared locking.
- Random bugfixes.



Uncommitted Isilon Changes

- Unmapped struct bufs.
- Unmapped struct mbufs.
- Performance tools: sleep and CPU profiler, disk request analysis.
- Stack collection, leak checking, ^t and power button stack dumps.
- Infiniband stack and device driver.
- WITNESS performance improvements.
- Large pages.
- Drivers for LSI controller, LCD & VFD panels, etc.
- VFS support for NTFS/NFSv4 ACLs.
- VFS support for Alternate (named) data streams.
- VFS support for encoding-aware filesystem.
- Recvfile() Implementation.
- Filehandle Affinity changes for NFSv3.
- Write_mbuf, read_mbuf VOPs.
- Support for SMART stats collection from SATA disks.
- “Fail points”, a flexible error injection system.
- NFSv3 configuration system that allows per-mount attributes.
- Panic-time condensed state saving.



Desired Changes & Upcoming Work

- Removal of VFS locking of vnodes, or refinement.
- Bigger window of optimization for disk request reordering than that which the bufdaemon currently allows.
- Integration of unmapped mbuf and buf support.
- A better solution to the partitioning between the vnode and page caches. We'd prefer that we can cache 64 million inodes **or** 32GB of data, not always 32 million inodes **and** 16GB of data.
- TOE support.
- NFSv4.
- IPv6, IPsec, zeroconf.



Observations about Working with FreeBSD

- We've been lucky that features have arrived roughly when we needed them.
- Careful development and clear path of FreeBSD minimizes churn and merge headaches for us.
- The FreeBSD community is very helpful! Thank you!