



DRBD SDS

Open Source Software defined Storage for
Block IO - Appliances and Cloud
Philipp Reisner



DRBD's use cases



Cloud Ready

Storage is one of the most critical components in any cloud environment. It **has to be easy to provision, highly reliable, cost effective** and ideally running on commodity hardware.

DRBD9 is in **OpenStack**, so it integrates **seamlessly** with your Private **Cloud Environment** and **runs on commodity hardware**, other cloud OSs and virtualization platforms.



High Availability

The principal goal of a high availability solution is to **minimize or mitigate the impact of downtime**.

DRBD **seamlessly replicates data** transparent to your applications and databases, **eliminating single points of failure** within your IT infrastructure.

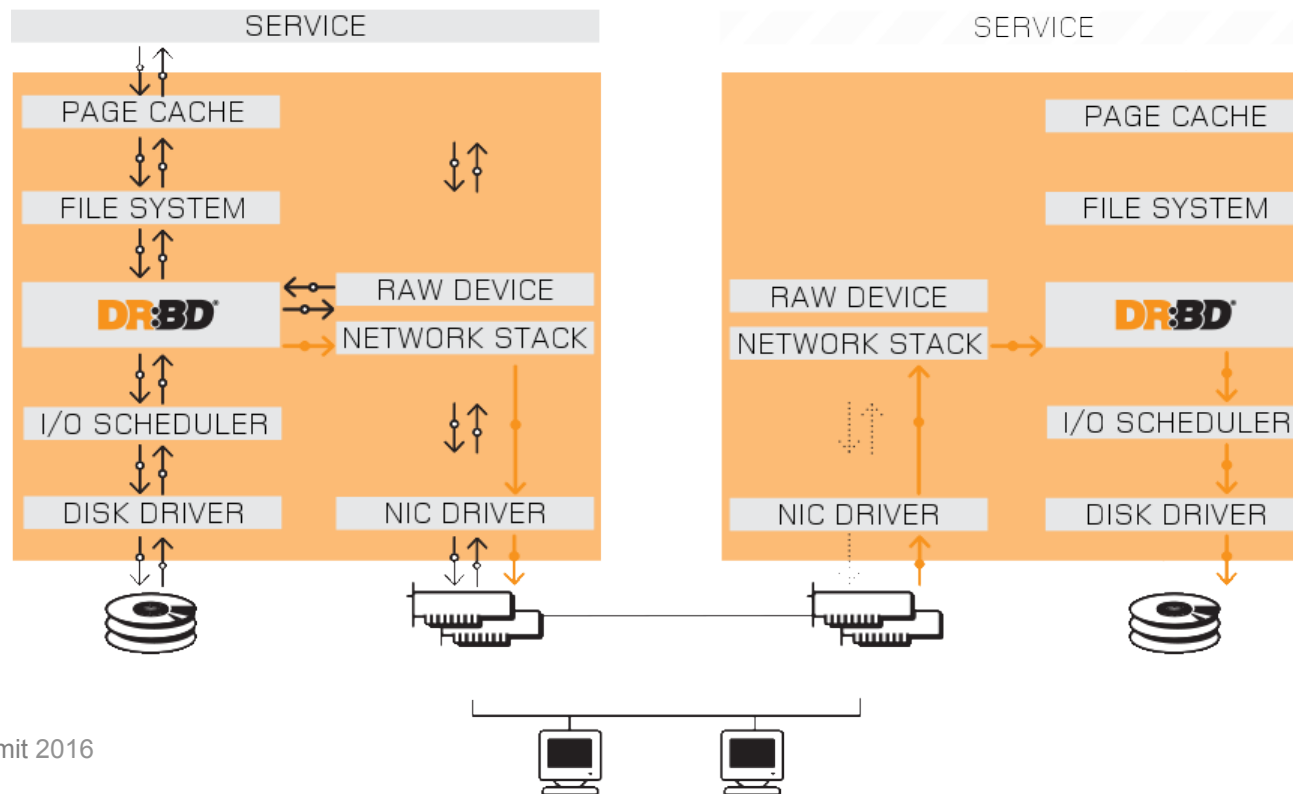


Disaster Recovery

The principal goal of disaster recovery is to **restore your systems and data to a previous acceptable state** in the event of a failure/loss of a data center.

DRBD can **mirror data asynchronously** over long distances, forming an important building block of your **disaster recovery plan**. Geo-Clustering with automatic fail-over is possible.

In Kernel Data Path





DRBD 8.x key features

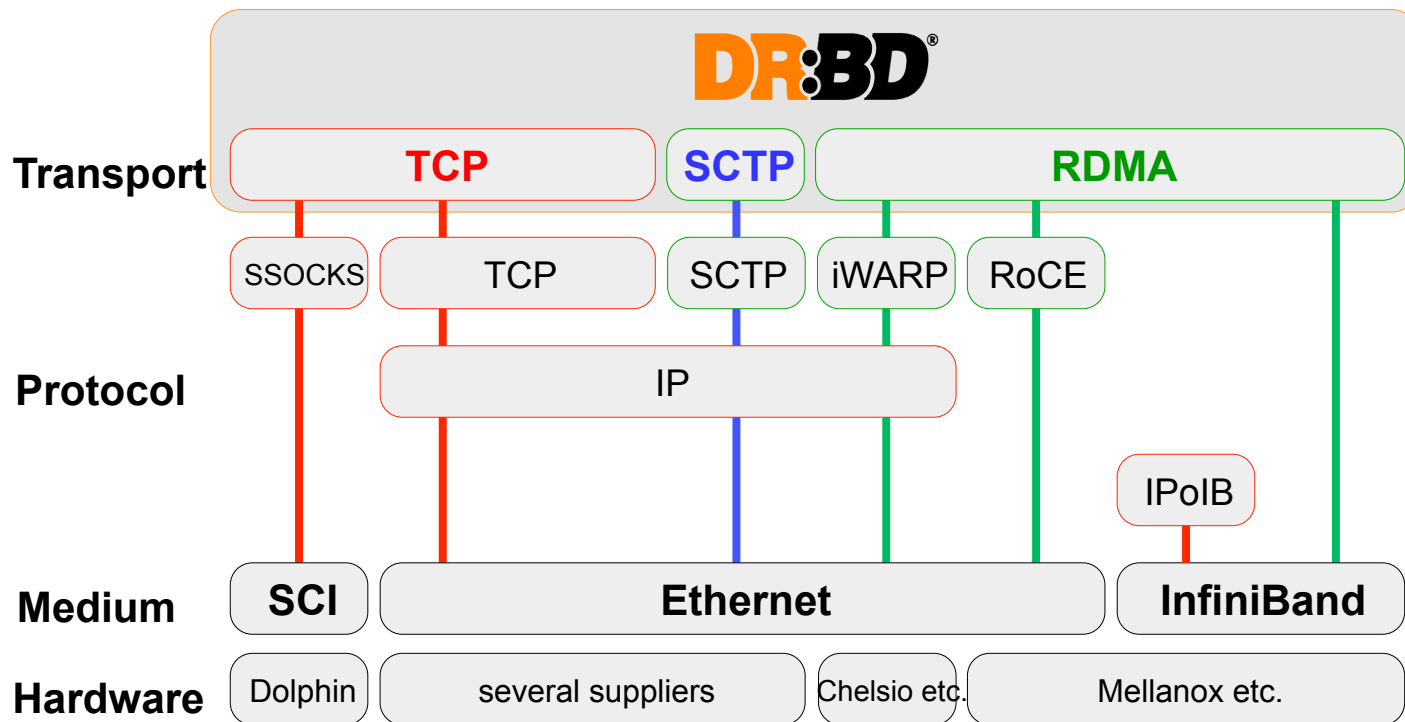
- Automatic resync after failures
- Performance Linux in kernel (160k IOPs)
- Multiple volumes per resource
- Pacemaker integration
- Synchronous and asynchronous
- In Linux upstream since 2.6.33 (2010)



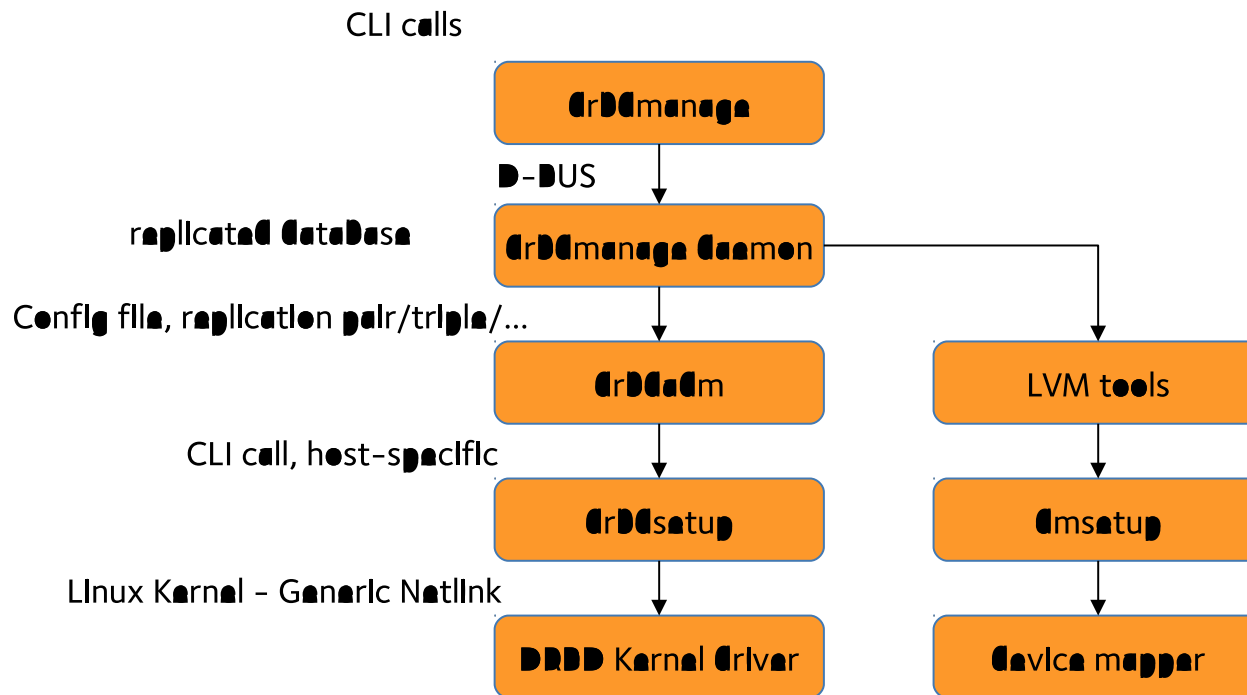
New features of DRBD9

- Up to 32 nodes per resource
- Auto promote
- Transport abstraction (TCP, RDMA, ...)
- DRBD Manage – automated control

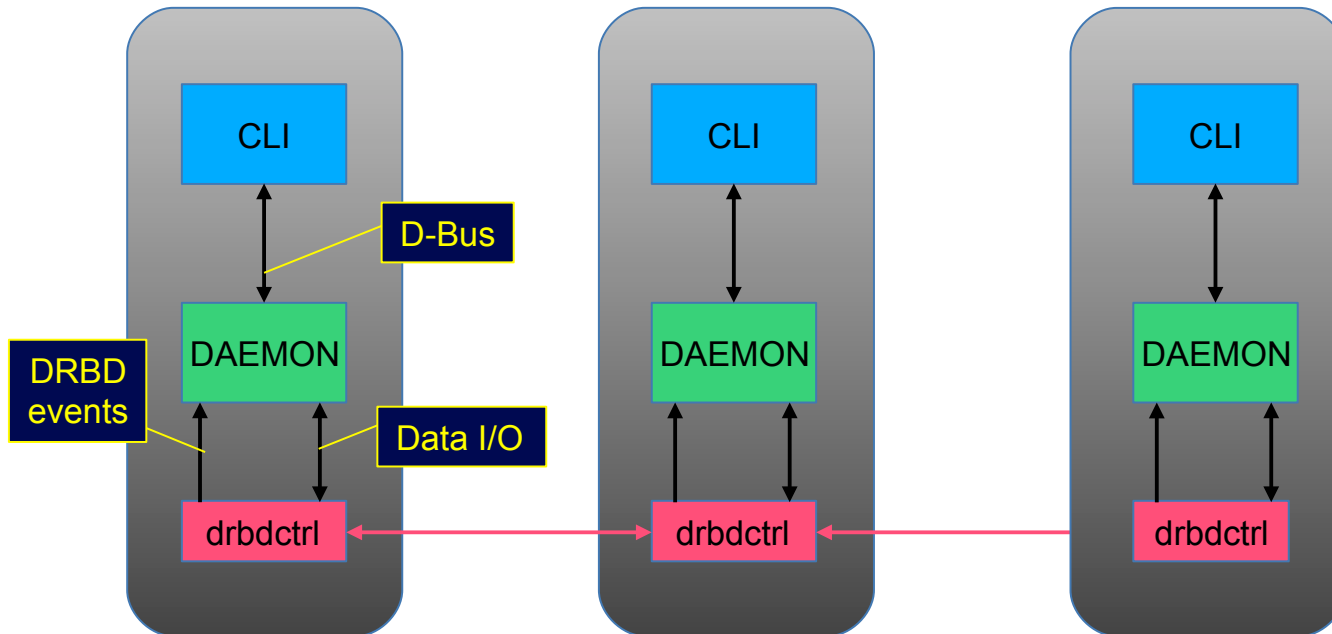
Networking Options



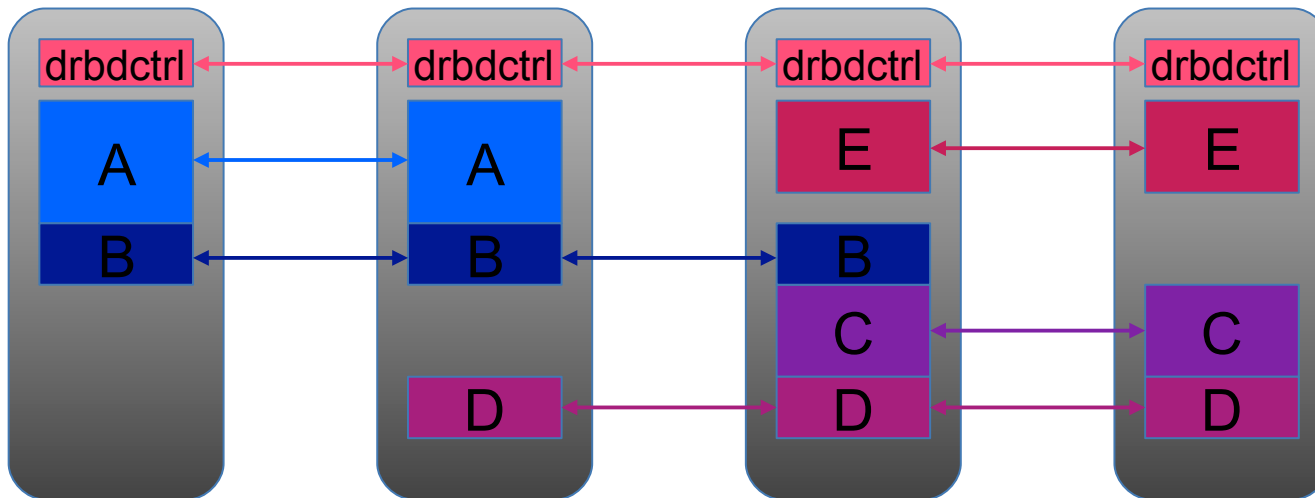
Control Path – DRBD Manage



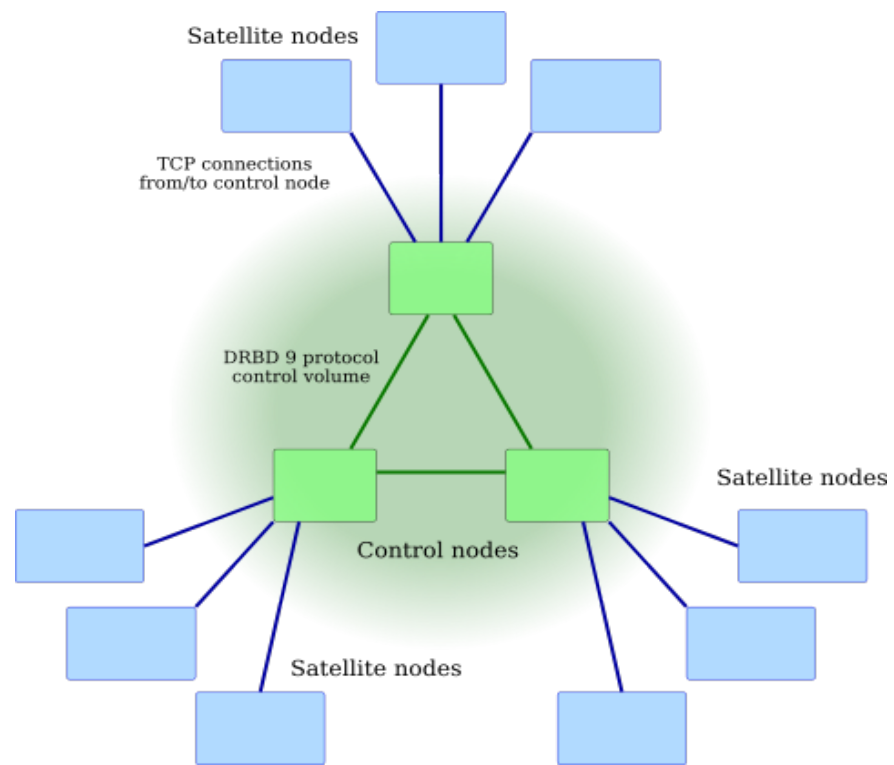
DRBD Manage Architecture



Volume Management Example

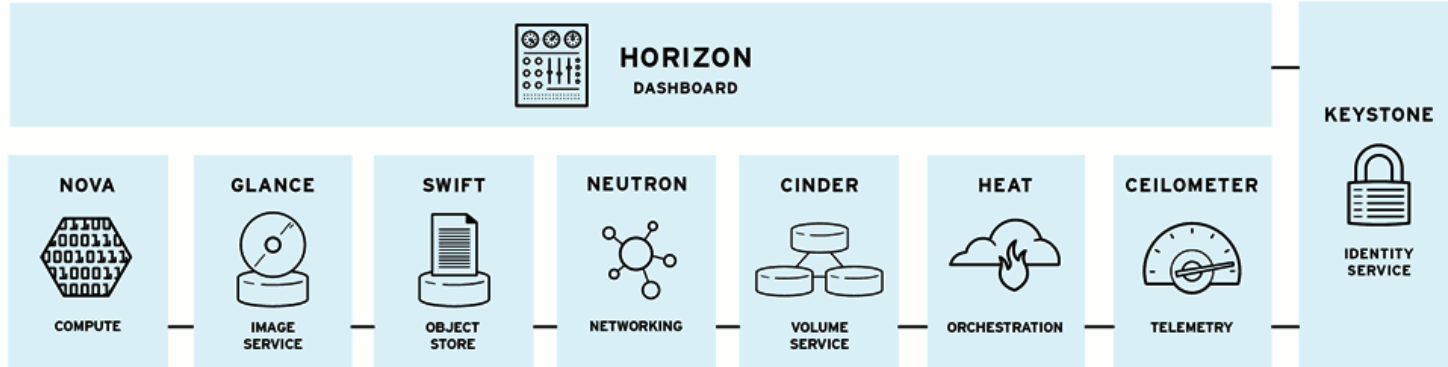


Scaling beyond 32 nodes





DRBD SDS and OpenStack



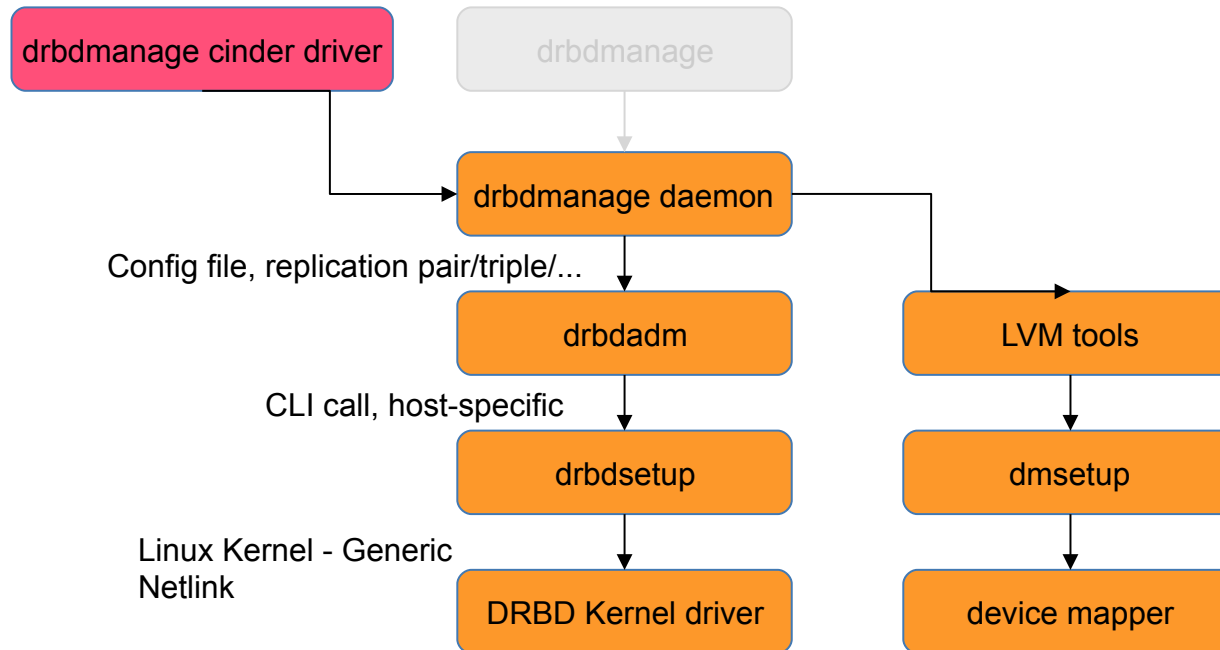
OPST0005

DRBD & drbdmanage dock on here as driver

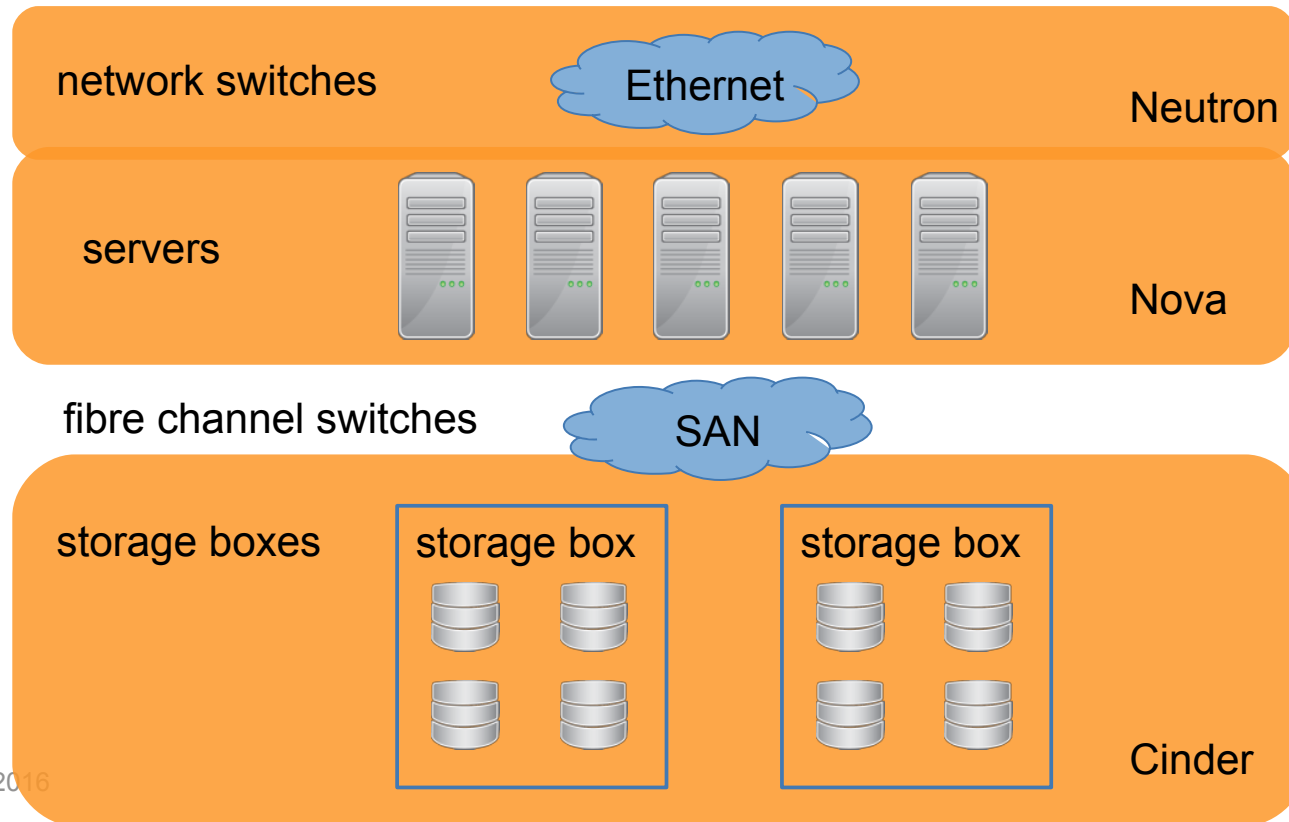


DRBD Cinder Driver

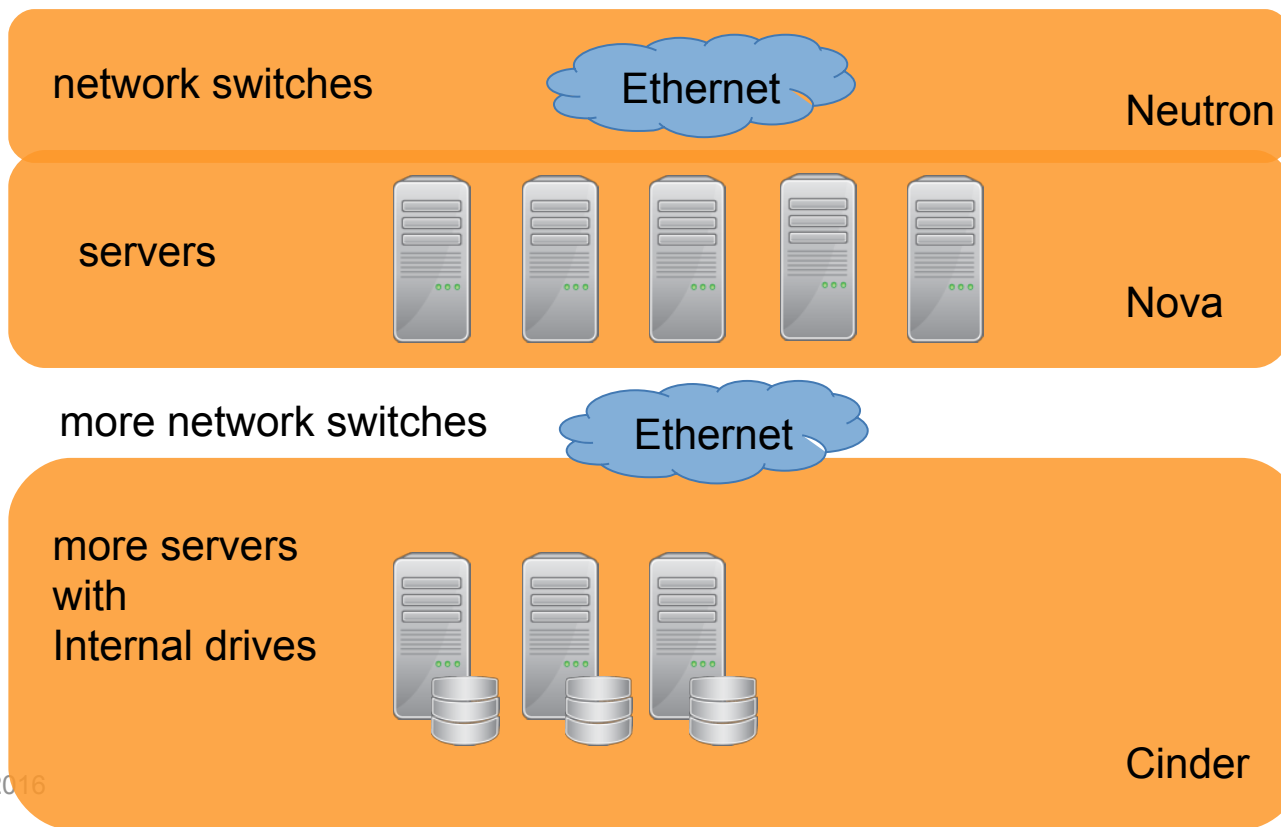
on cinder node!



OpenStack with SAN Storage

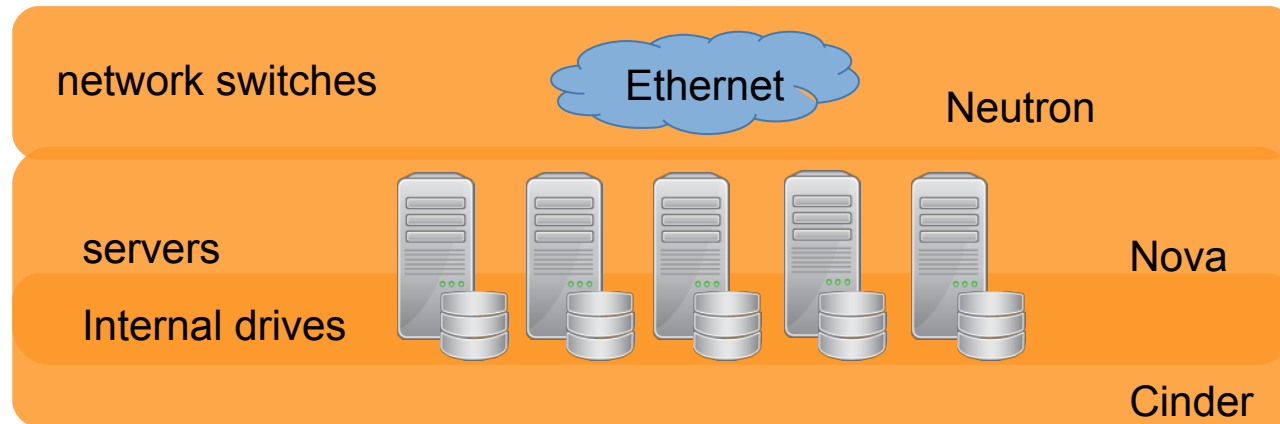


Replace SAN with DRBD SDS





Go Hyperconverged!



Low latency storage access possible by aligning nova and cinder allocations

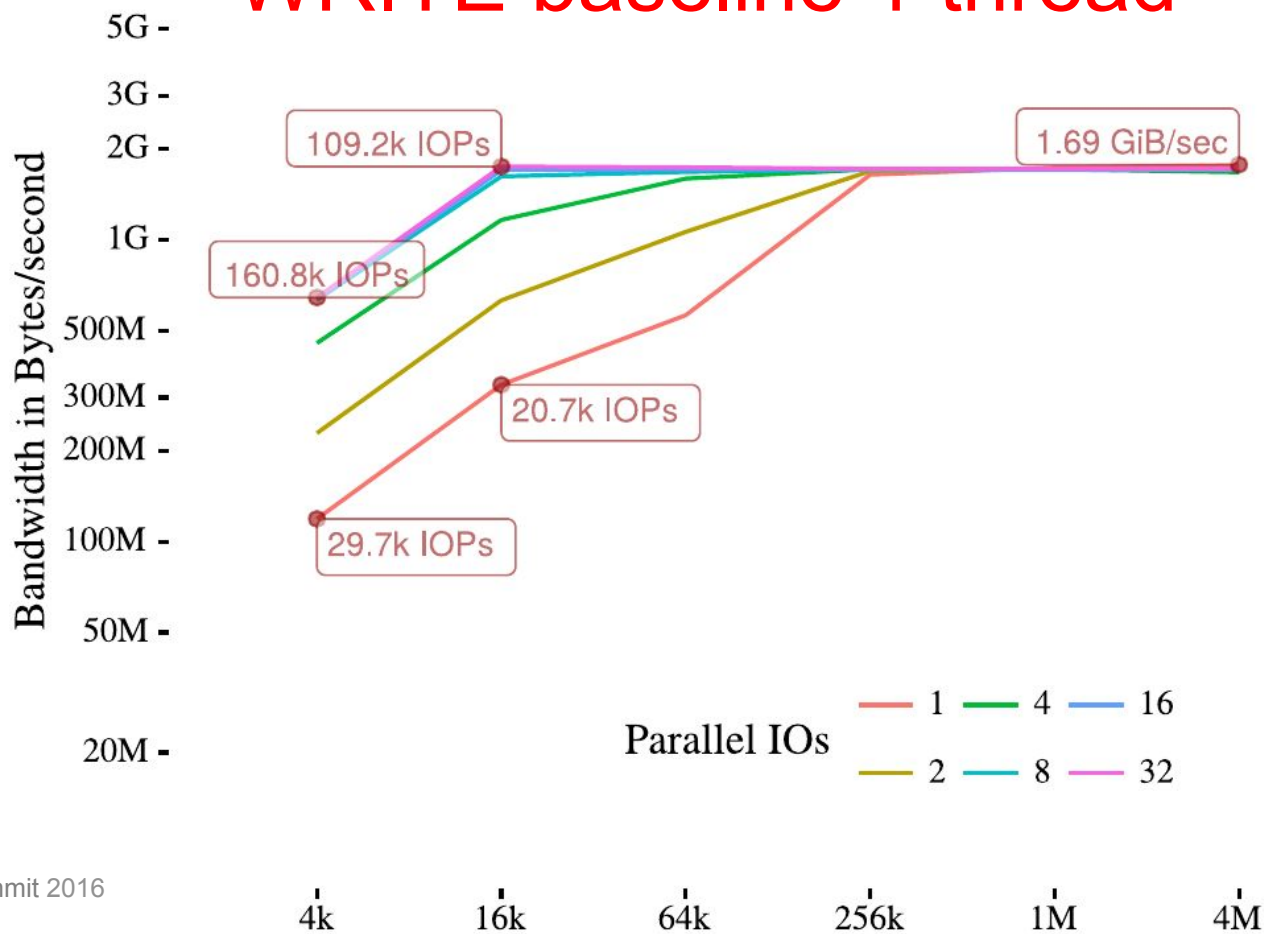


IO Performance Example

- **Hardware**
 - two IBM 8247-22L; Power8 2 sockets, 128GB RAM
 - Mellanox Connect X4 dual port; 100GBps InfiniBand
 - HGST Ultrastar SN150 NVMe SSDs, 1.6TB version
 -
- **Software**
 - Ubuntu Xenial; on bare metal
 - DRBD 9.0.1 & RDMA Transport
 - fio-2.2.10
- **Tests**
 - 100% random IO
 - Seperate READ and WRITE
 - Direct access
 - FIO

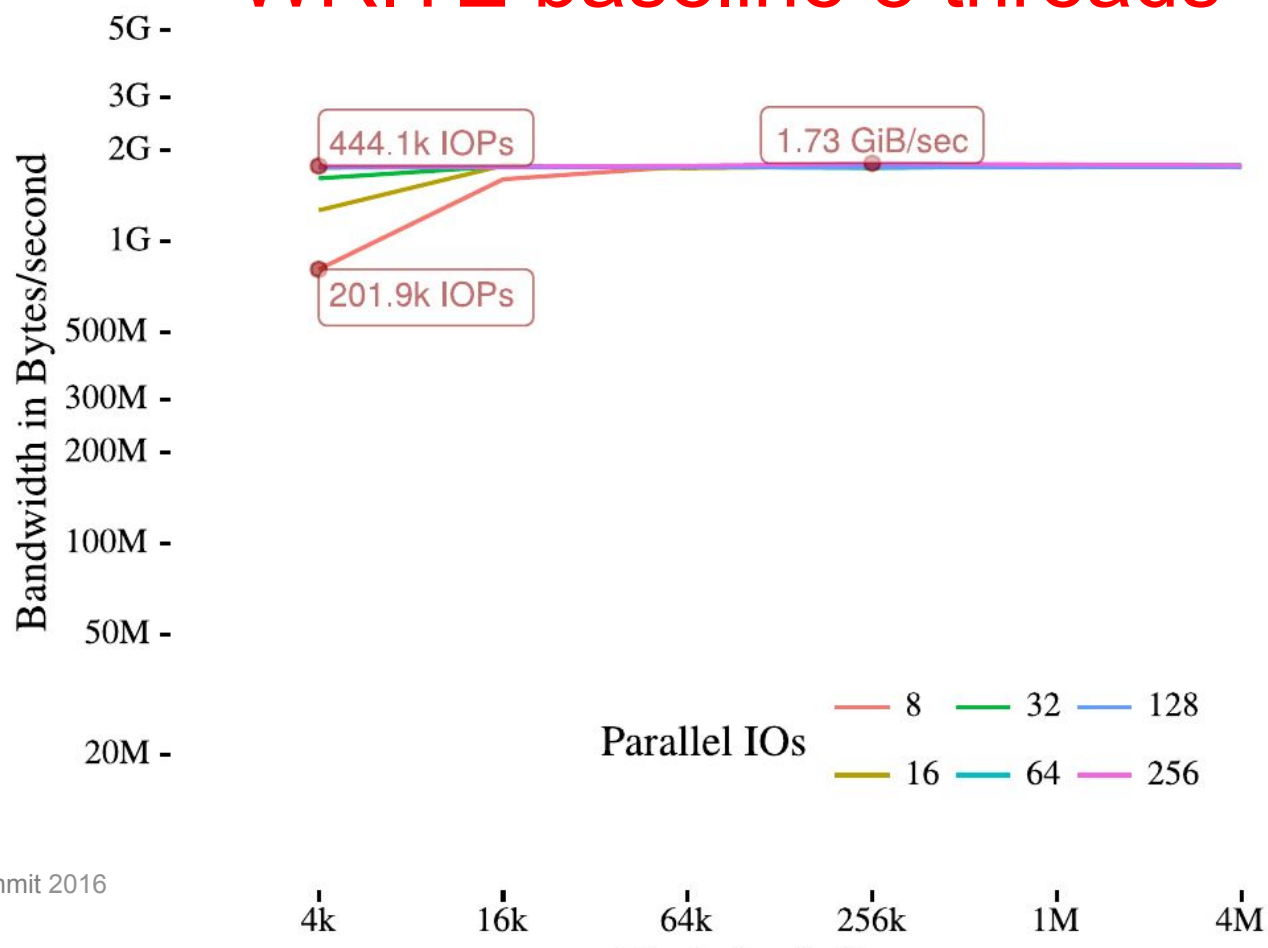


WRITE baseline 1 thread



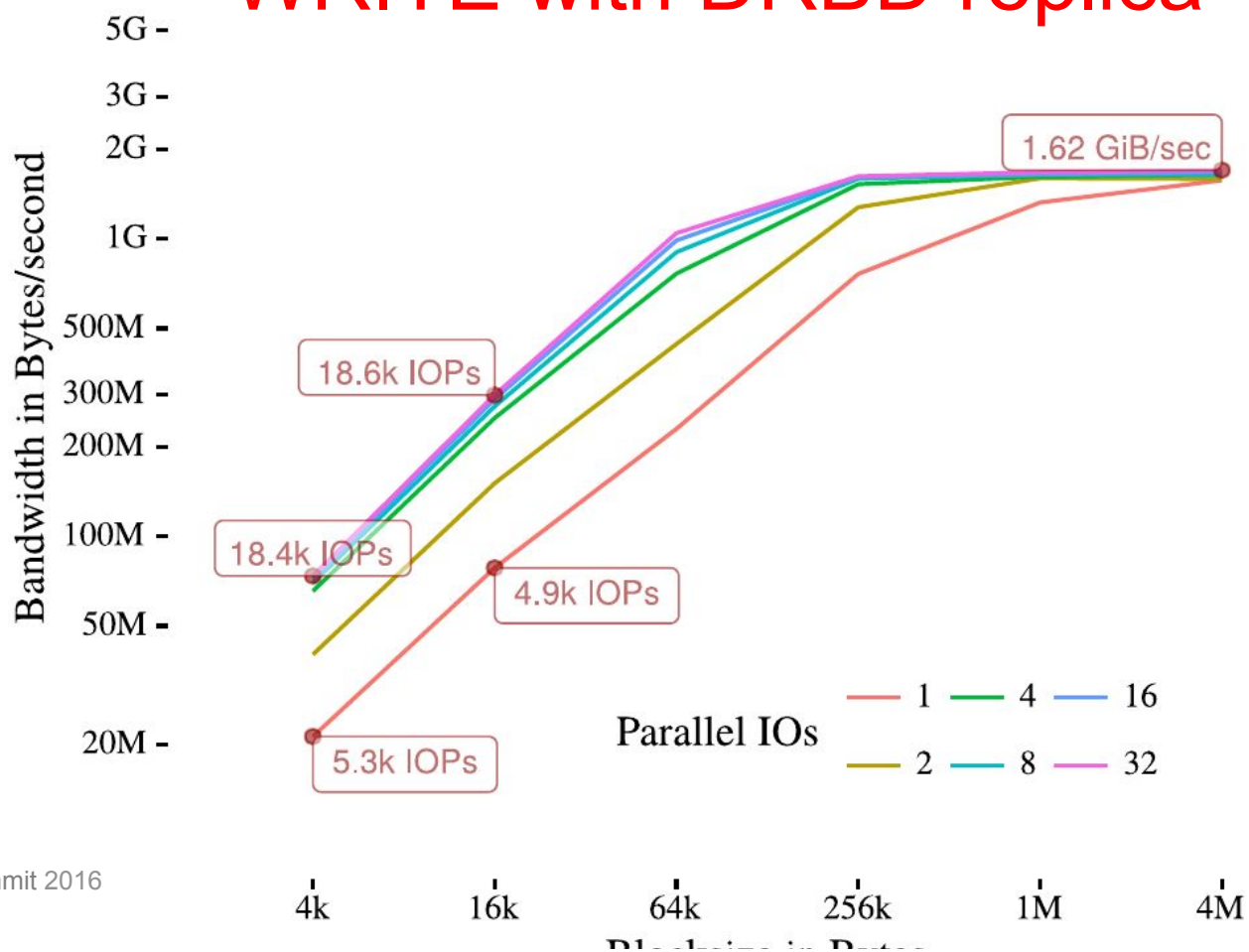


WRITE baseline 8 threads



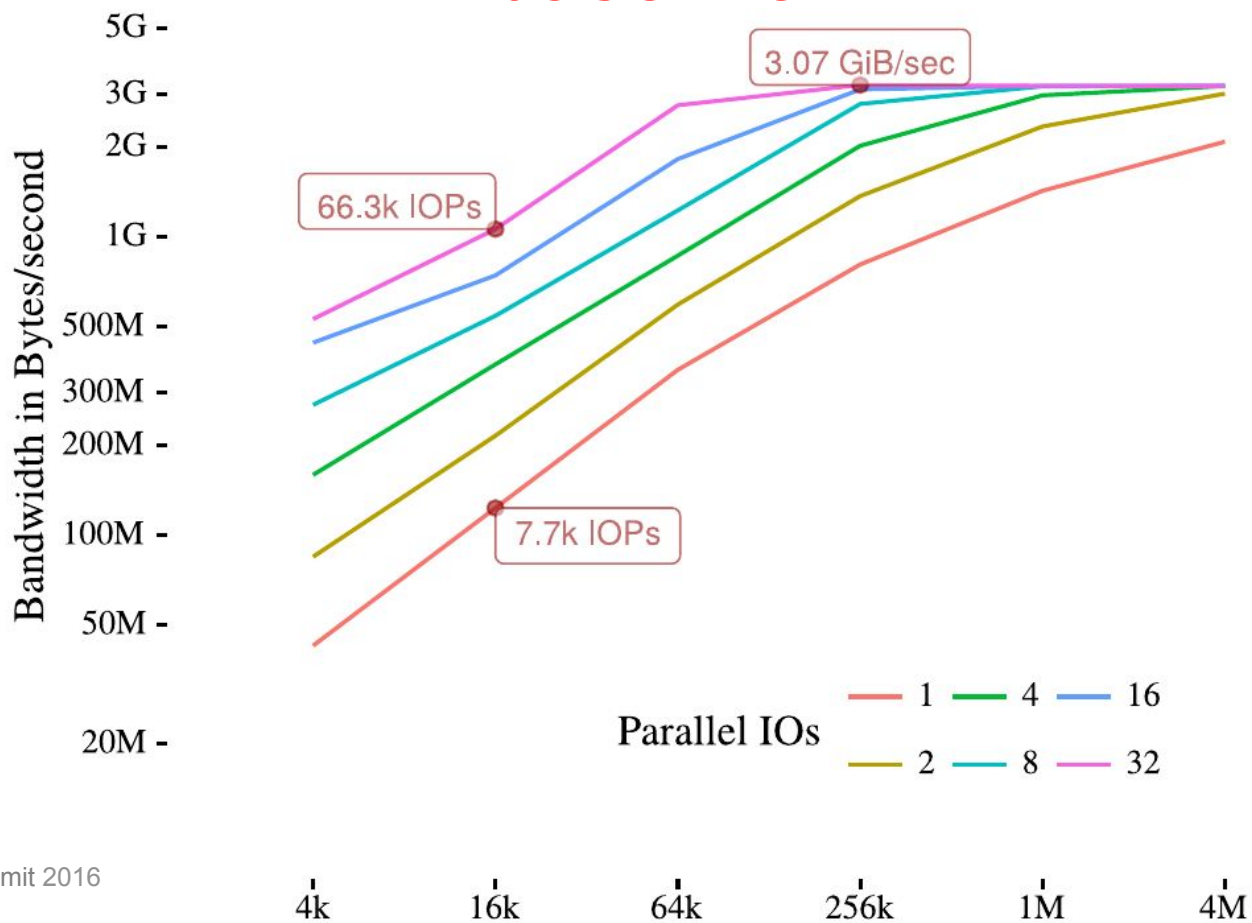


WRITE with DRBD replica



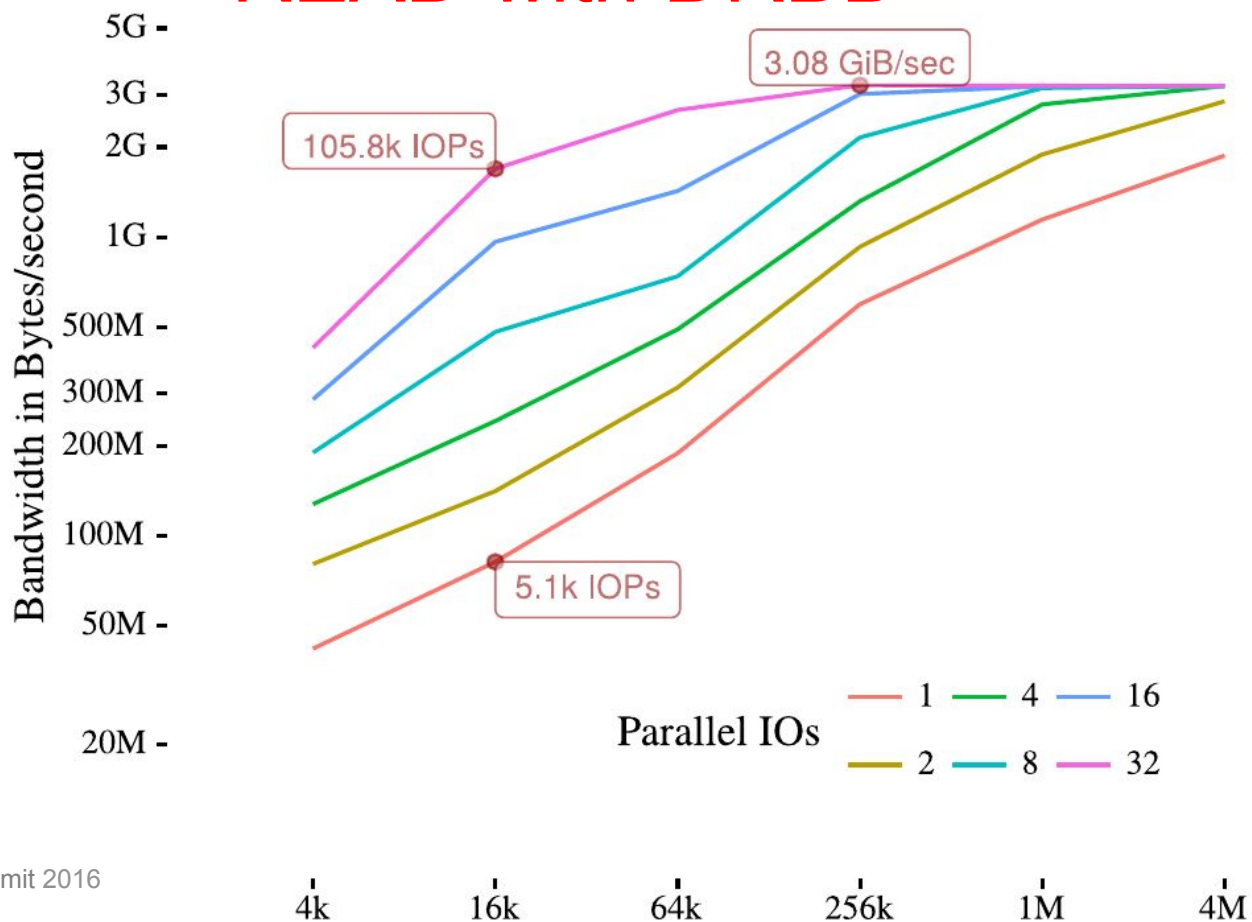


READ baseline



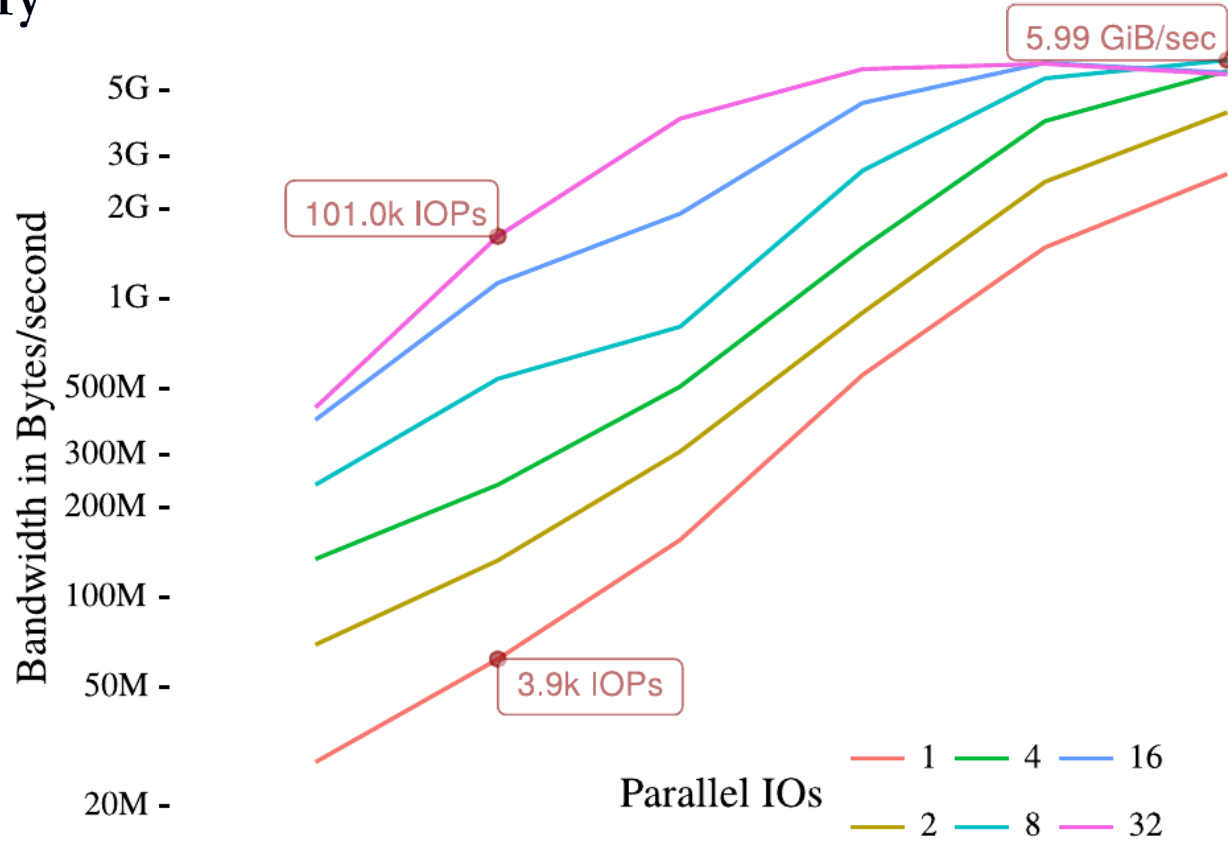


READ with DRBD





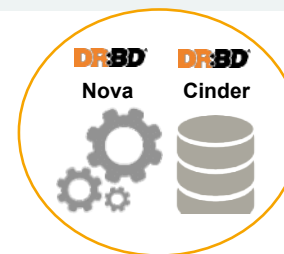
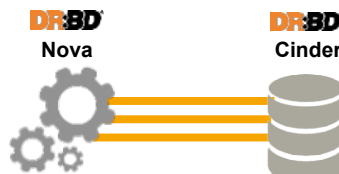
READ with DRBD's read-balancing





OpenStack ROADMAP

Q4 2015	Cinder driver	Q1 2016	Nova driver	Q2 2016	Align Cinder and Nova
<p>Benefit: Highly available storage volumes can be created via the Horizon dashboard and the cinder APIs. Snapshots supported. Replica count in cinder configuration visible as "pools"</p>		<p>Benefit: Replace iSCSI between nova node and cinder nodes with DRBD's native protocol. Improve storage write performance, enable read-balancing, faster and seamless failover.</p>		<p>Benefit: In a hyper-converged architecture, get read performance of local storage and lowest possible overhead for write accesses by aligning Cinder and Nova allocations</p>	
<p>Released</p> <p>"Liberty" October 2015</p>		<p>Released</p> <p>"Mitaka" April 2016</p>		<p>Planning stage</p> <p>Target: "Newton" October 2016</p>	





DRBD kernel driver ROADMAP

2015



OpenStack Driver

Enables OpenStack users to use their clouds on DRBD/DRBDmanage

Tokyo Open Stack Summit - October



DRBD on Power

Gives users of Power machines access to LINDIT's products

Released w/ IBM Germany - November



RDMA / InfiniBand

Multi-path support; Aggregates bandwidth of configured paths; Increases replication link availability

Released - November

2016

OpenNebula OpenNebula Driver

Enables OpenNebula users to use their clouds on DRBD/DRBDmanage

Released in May



MS Windows Support

DRBD for MS Windows

Release planned June 23



Activity log in NVRAM

Activity log in non-volatile memory. First implementation using small (1MiB) battery backed SRAM on PCI card

Release planned for June

2017

Erasura coding

Refactor DRBD to support Erasure coding
Implementation with offloading-verbs API (Mellanox)
Software only implementation



Geo redundancy for shared storage

DRBD replication of shared storage which is active on multiple nodes



Direct RDMA

Direct allows a write request to be sent from an InfiniBand HCA directly to an NVMe device



About LINBIT

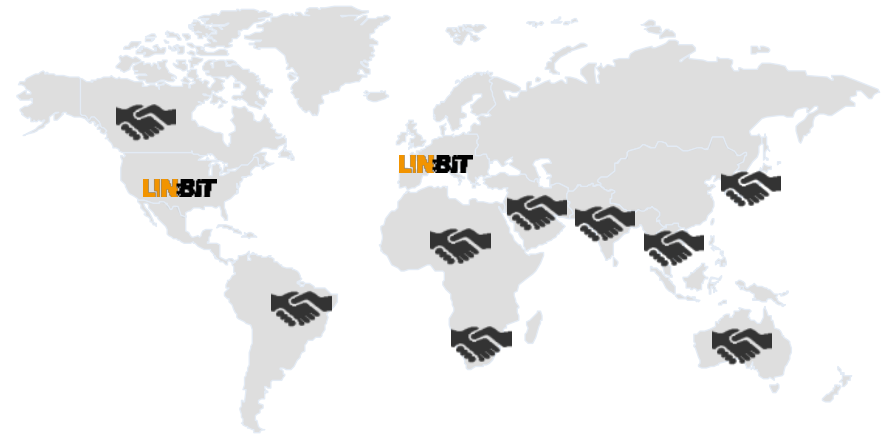


Best performing High Availability / Block device for Open Stack using common off the shelf hardware

20+ times faster than CEPH and GlusterFS*

Only replication technology exceling at both synchronized short distance and asynchronous long distance

- **Open Source DRBD** supported by proprietary add-ons
- **Hundreds of thousands** of DRBD downloads
- **OpenStack** comes with **DRBD Cinder driver**
- **100% founder owned**



References

