LA-UR-18-27425
Approved for public release; distribution is unlimited.

Title: Troubleshooting and Analyzing Network Booting Challenges Using a Raspberry Pi Testbed

Author(s): Lapid, Maria Francine Therese Ruiz
Steen, Trent David
Schild, Christoph Victor

Intended for: HPC Mini-Showcase

Issued: 2018-08-06

Troubleshooting and Analyzing Network Booting Challenges Using a Raspberry Pi Testbed

"Testing on the largest Raspberry Pi cluster in the world!"

Maria Francine Lapid
Christoph Schild
Trent Steen

Mentors:
Paul Peltz
J. Lowell Wofford

Introduction

- Background
- Testing environment
  - Hardware setup
  - Network setup
- Image distribution using Kraken
- Internet Small Computer System Interface (iSCSI)
  - What it's used for
  - What we found
- Future work

Background

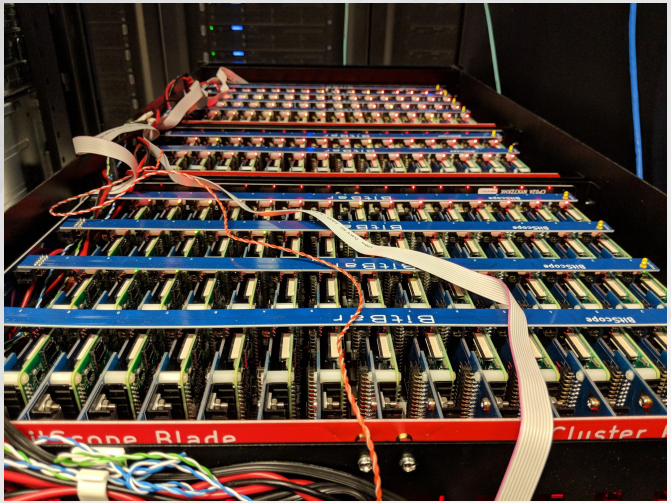
LANL operates some of the world's fastest clusters.

But…

- The lab can't do scalability tests because those computers are necessary for research
- HPC developers need an inexpensive way to do systems testing

Hence the Raspberry Pi cluster!

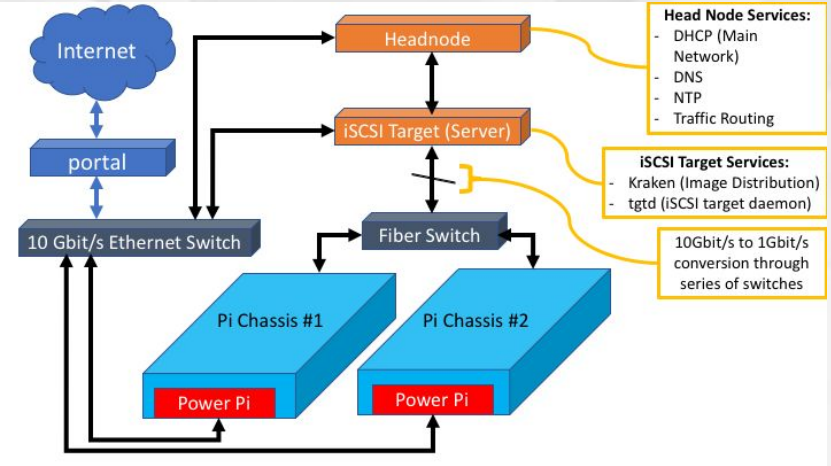
Testing Environment: Hardware Setup



- Each chassis has 150 Pis and 3 internal switches
- We have different types of links between switches and devices
  - Might be a cause of bottlenecks
- Raspberry Pis are hardwired to the Power Controller Pi, which allows remote power cycling

Testing Environment: Network Setup

- Resolving conflicts requires network isolation
- Compute Pis are on separate subnet
- Each chassis is fed by a 10 Gb link



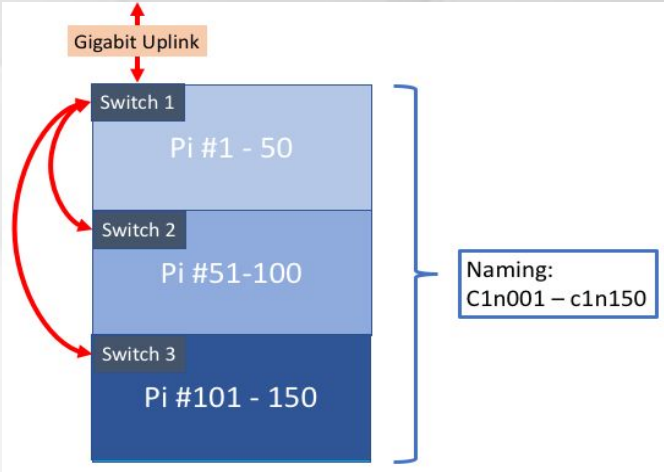


Image Distribution

- Working with Raspberry Pis poses a series of challenges
  - Not compatible with existing systems
  - Doesn’t follow standardized protocols (DHCP Handshake)
  - “Node scan” must be done manually
- Kraken solves this
  - Created in-house in Go
  - Works with oddities of Raspberry Pis (custom DHCP)
  - Includes many microservices for distribution

Our Contributions to Kraken

We contributed a few pieces of software to Kraken:

- “Nodescan” for acquiring MAC addresses
  - Listens for DHCP DISCOVER traffic
  - Builds JSON for Kraken
- “Power Cycler” for assisting in the boot process of the Pis
  - Attempts to boot as many Pis as possible through repeated, timed reboots
- Automation Scripts for data retrieval from Pis
  - PSSH, Bash Scripting, and file parsing

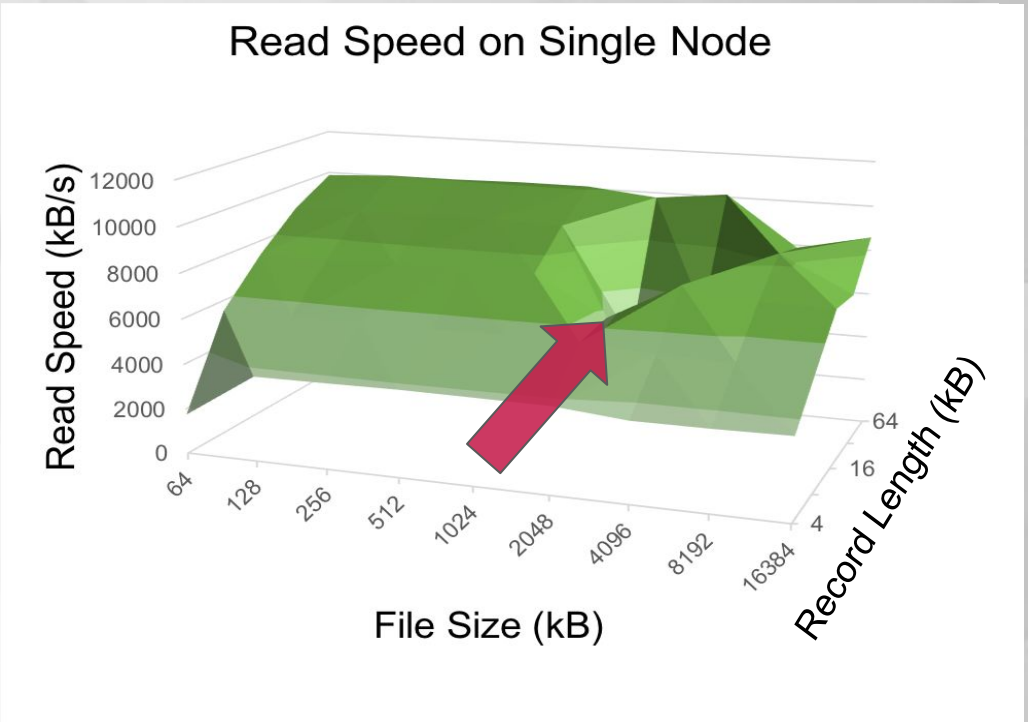
iSCSI Intro

- Internet Small Computer System Interface
- Networking storage using IP
- Enables nodes to send SCSI commands over network and see external storage as internal
- Clients = Initiators, Server = Target
- Tested scalability using IOzone filesystem benchmark
  - Interested in the “read” tests (ie, sending out an OS image)
  - Varied file size and record length

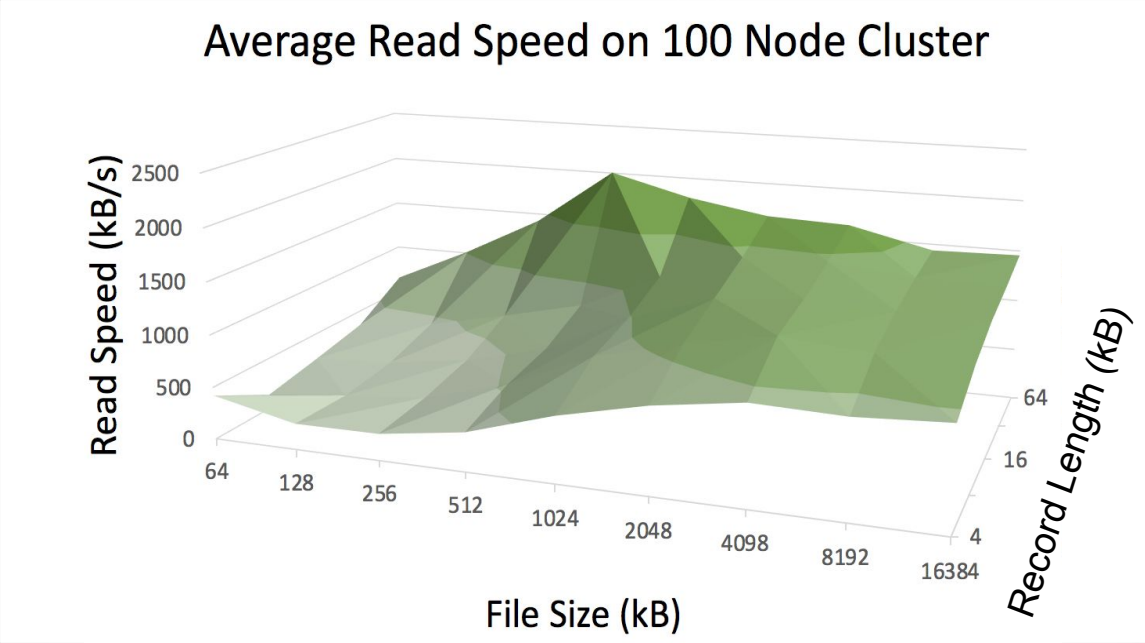
Iozone Filesystem Benchmark, www.iozone.org/.

iSCSI Test Results: Single Pi

- File size vs record length
- Mostly for comparison
- Record length has a bigger impact than file size
- Tends to slow down at file size 4 MB and record length 32 kB



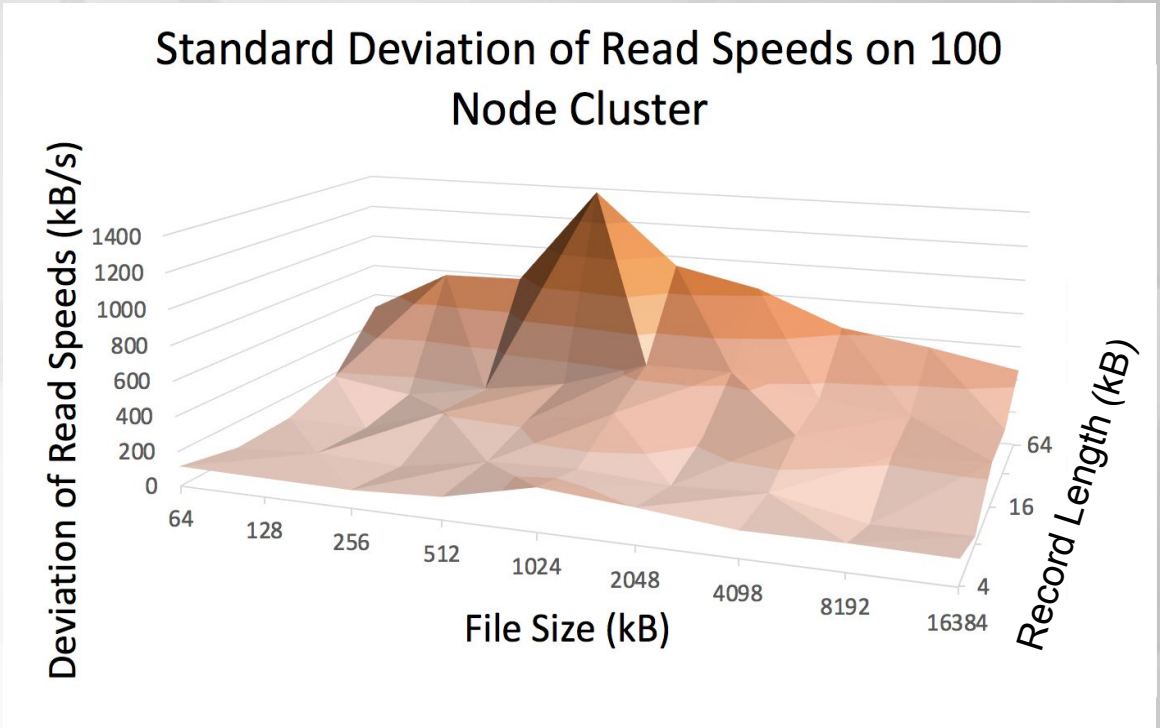
iSCSI Test Results: 100 Pis (Averages)



- Slower than Single Pi test
- iSCSI is less efficient when running many small block operations
- iSCSI performs well with larger files and larger blocks
- File size and record length seem to both have effects on read speed

iSCSI Test Results: 100 Pis (Standard Deviations)

- Variation increases with record length
  - Less consistent
- High variation shows that we stressed system



Future Work

- Test other block storage methods
- Fix scalability issues, to push storage technologies further
  - Further stress block storage systems
- Eliminate unrelated bottlenecks by exploring other configuration
  - More accurate testing

Conclusion

We believe that the Raspberry Pi cluster is a sound testbed for scalability tests.

Thanks for Listening

Francine Lapid
Texas A&M University
mail@francinelapid.com

Christoph Schild
Northern Arizona University
chris@schild.tech

Trent Steen
Dakota State University
hi@trentsteen.com