Extreme Homelabbing

Modernizing Infrastructure one Outage at a Time

Agenda

- About Us
- Challenges Migrating
- Picking Software
- Optimizing Software
- Questions

About Us

Joe

- Security Researcher and Educator
- DevSecOps / AI Security Engineer by Day
- Operations Lead for WRCCDC/CIRCUS for the last 8 years

dex

- I'm (just) Dex :3
  - they/them/any
- Professional software engineer during the day
- Hardware ~~hacker~~ "security researcher" by night
- Specializing in:

  High Performance Enterprise-grade *Jank*

The Problem

Meet WCC Comps:
The world's most overloaded VPS provider

What is this "WCC" thing anyway?

- Approximately stands for "Western Coalition of Competitions"
- Started as WRCCDC in 2008

  (Western Regional of the National Collegiate Cyber Defense Competition)
- Began with a single regional event each year
- Grew from 16–18 teams to an average of 25–30 teams
- Introduced a qualifier to accommodate increased demand
- Expanded to invitational "pickup" events with 40+ teams from across the country per event
- In 2024 began collaborating other events including PRCCDC and CIRCUS providing events for over 2000 competitors in the 2025-20226 season.

- Bridge gap between industry and academia
- Provide complex scaling challenges seen seldom elsewhere
- Ability to rapidly test new technologies at scale without risking too much
- Entirely volunteer run: paid for in chuckles and pizza
- Built for students

Why is this relevant to SCALE?

The Volunteer Factor

- On average 50% turnover on volunteers due to time commitments and real life
- “Leads” are essential and pick up tons of tasks but have very little time to do “all the things”
- Low learning curve is required to keep things functional: Trust + Training mean few can do administrative tasks
- Deploying and maintenance is difficult and time consuming which takes away from volunteers limited time – can’t simply work during work hours!

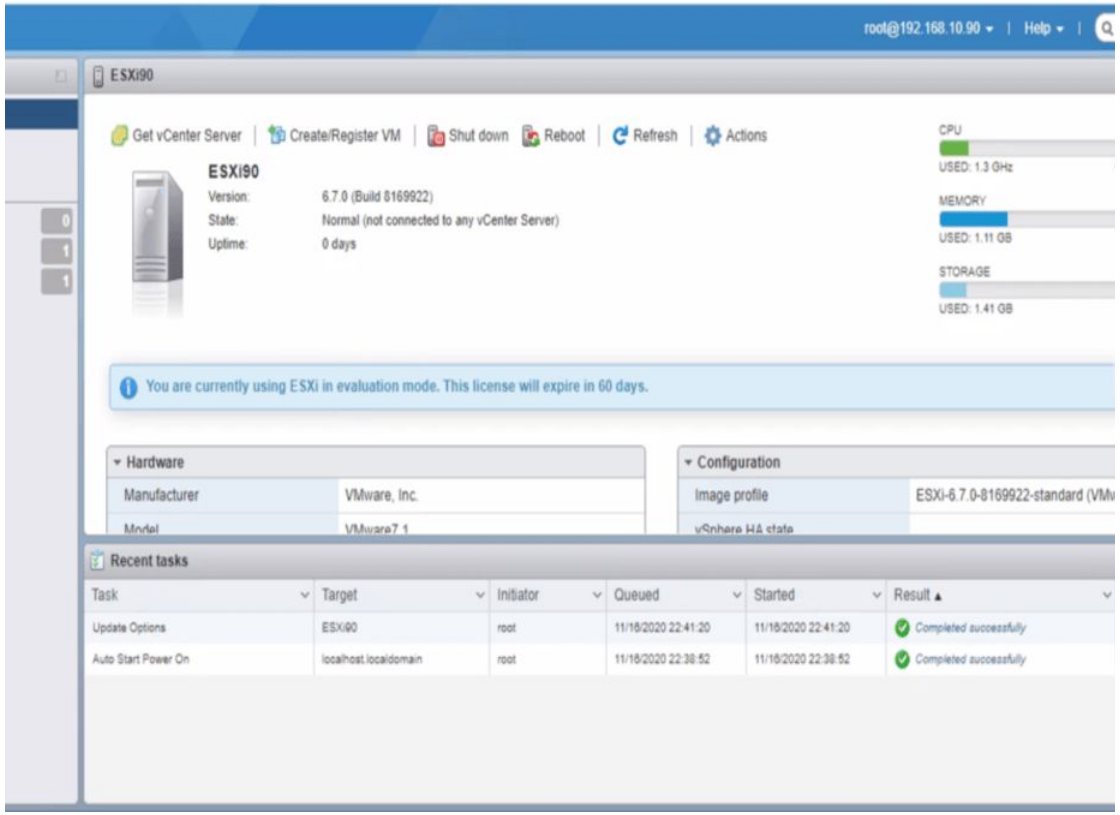
It takes a Village

The Problem

What happens when you provide non-profit services and can no longer afford licensing costs?

The Problem

What happens when you provide non-profit services and can no longer afford licensing costs?



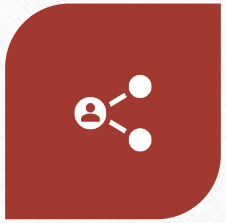
VMWare Changed the Game… Literally

- Licensing model used to include Academic Uses
  - Suddenly it didn't
  - Stuck on vCenter 7.0
- Entire infrastructure was built around VMWare including SAN designed to be used by VMWare
- SAN Requires Support Contract and License to Configure
- Microsoft / Windows licensing becoming a bigger challenge: Very Windows Centric and SSO Difficult
- Lots of proprietary resources which locked us into a specific "way" of doing things.
- Very complicated and manual networking
- Everything was A VM

Challenges of “Moving Away”

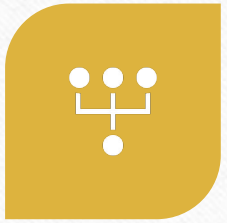
NATIVE HIGH AVAILABILITY (VMOTION)

STORAGE REDUNDANCY TIERS



EASE OF UNDERSTANDING FOR NEW/RETURNING VOLUNTEERS

MIGRATING MANY INDIVIDUAL SERVICES AND VMS INTO SOMETHING MORE MODERN



REBUILDING REQUIRES ON-SITE TIME AND GETTING A GROUP TOGETHER TO REINSTALL EVERYTHING IN ONE SHOT

Staying on Existing Setup Was Riskier

OLDER OPERATING SYSTEMS UNSUPPORTED

AUTOMATION AND TOOLING WAS ALL DESIGNED FOR ACTIVE DIRECTORY AND VMWARE

MORE AND MORE SECURITY BUGS (IN A SECURITY COMPETITION)

PALO ALTO NETWORKS (PAN) APPLIANCES WORKED BUT WERE HITTING RESOURCE LICENSING LIMITATIONS AND DIFFICULT FOR VOLUNTEERS TO MAINTAIN

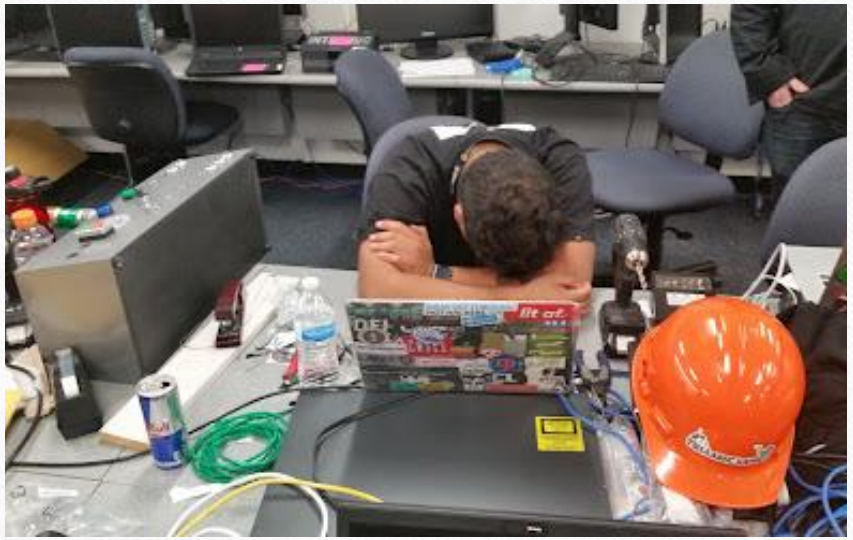
DELL SWITCHES SUPPORTED 100G BUT NOT EASILY MANAGED

OUR CISCO GEAR WAS ALSO EOL

For the Students

Research Resources we Provide

- Archive Server
- PCAPs
- In Person and Virtual Competitions



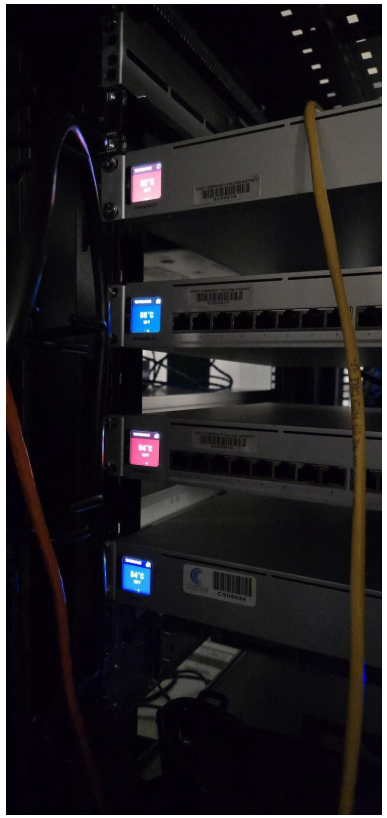
The Home Lab Mindset

“Now” is the always Perfect Time

- With changes in ability to get grants we had to keep our hardware working and secure as possible
- Home Lab is about making do with what you can and making it as good as you can
- Another important goal: Make our setup match students so they could more easily replicate and practice
- Open Source is a community, and we felt we needed to also give back
- Open Source allows us to iterate and try new things more easily then with contracts or vendors

Goodnight SAN

- SAN only supported iSCSI targets
- Proxmox has very limited iSCSI support (and ZFS over iSCSI is brittle)
- What do we do with 20TB of storage and no way to replace it? Thanks 2025
- What target do we pick? Proxmox? OpenStack? Xen? QEMU with Custom Frontend (OpenNebula)
- Limited time on site meant we had to pick and move quickly



Small Steps

- Replacing Dell gear and other devices with Unifi to simplify management and allow any volunteer to work on networking
- Originally tried to do many parallel tasks but this only slowed us down
- Compute Scales Differently when you need it for storage AND hypervisors
  - Had to order many parts from eBay and Amazon
- Volunteers donated a ton of gear

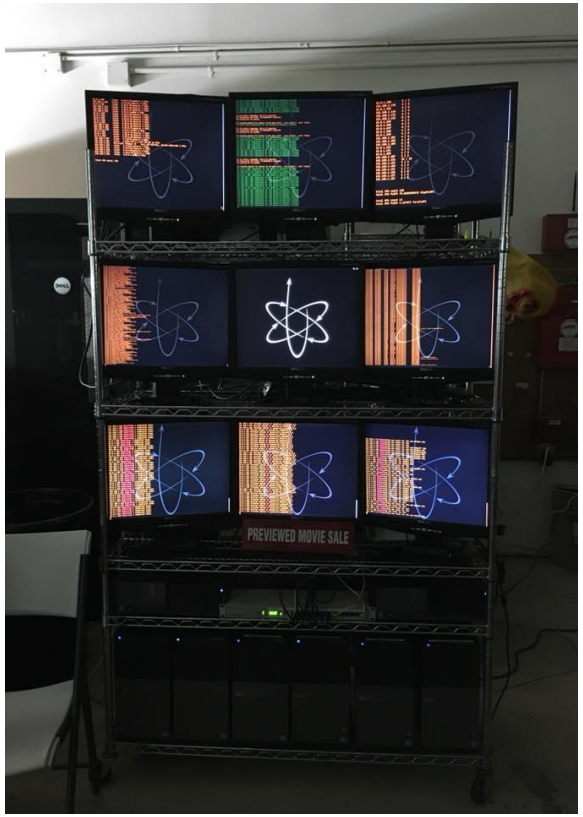
Hello NFS

- In the end we picked Proxmox
  - Students use it
  - Its easy to cluster
  - Mostly “works” like VMware*

We ended up having one catastrophic issue after another

- In total 4 on-site visits
- Lots of disk swaps for performance reasons - Finding any drives that could work
- A storage switch that wasn’t switching

One major problem: No Storage*



NUCC Inc. To the Rescue

- Provided 10x Unifi 1gig Switches
- Provided three new compute R640s
- More RAM
- Two beautiful and [old] JBOD arrays with 40TB of replacement NAS storage
- Replaced our HP Hypervisors to act as redundant storage

- Got extra gear or surplus from your work?
- Please consider donating to NUCC!

Picking Software

Core CCDC Services:

- Wiki
- VPN
- Web [Hosting] Services
- Archive Server
- Authentication
- Resource Access/VDI
- Automation Services

Support Services

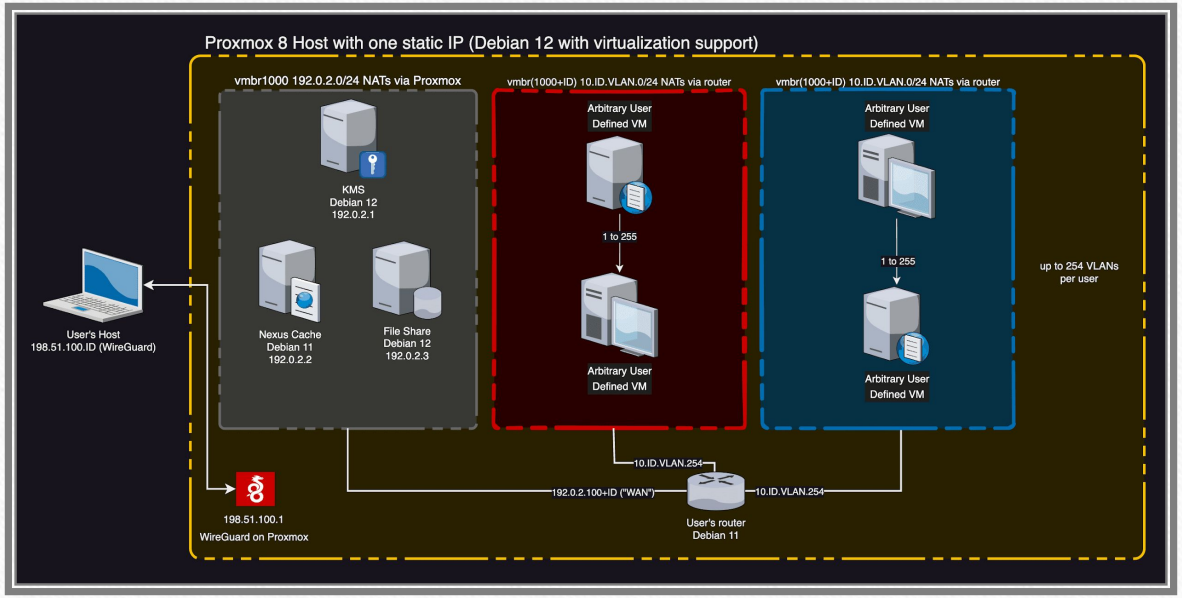
- Uptime / Service Monitoring
- Ticketing
- API / Password Vaults

Quirks Mode

- All software has trade offs
  - Tailscale/Headscale vs Netbird
  - Proxmox vs Xen
  - Dockge vs Portainer
  - Dokuwiki vs WikiJS , etc
  - Sometimes its time to build our own software (Ticket System)

Why not Ludus?

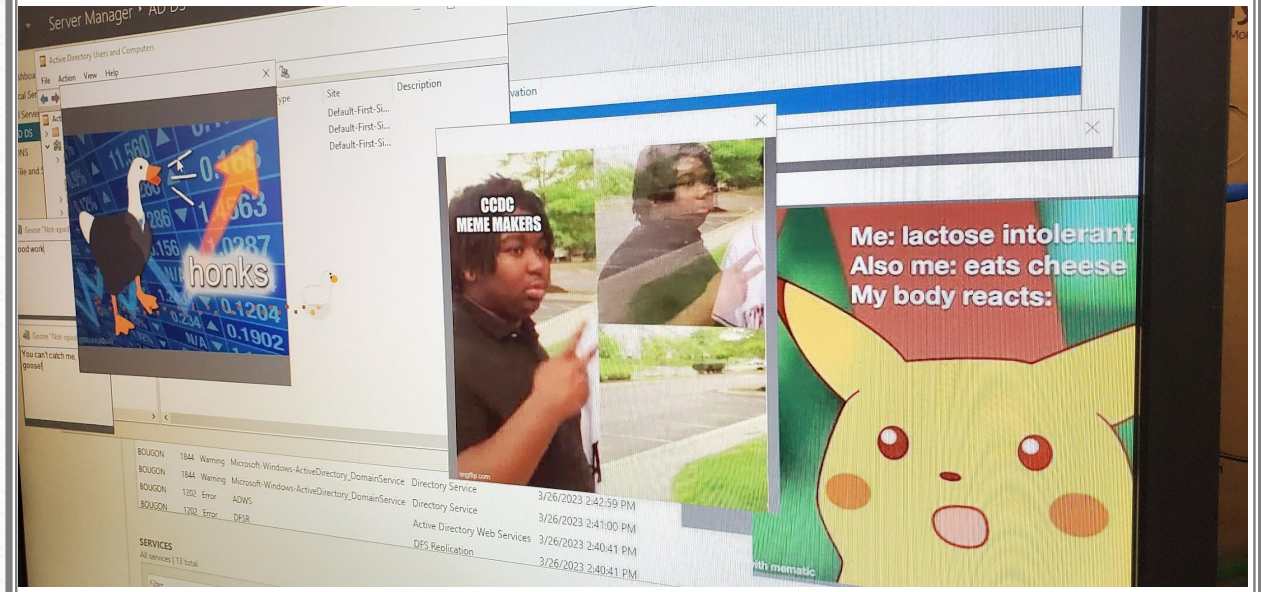
- Ludus is a lab automation “addon” for Proxmox
- Works great for labs, but scaling and creating proper entry/exit points for teams is difficult (VPN is primary method)
- Teams are not meant to access shared resources
- Proxmox API is NOT robust (Moving from PowerCLI to Proxmox API was a poor experience)
  - People online generally recommend using “qm” directly
- Retries everywhere in Semaphore



First Run: Everything Crashed

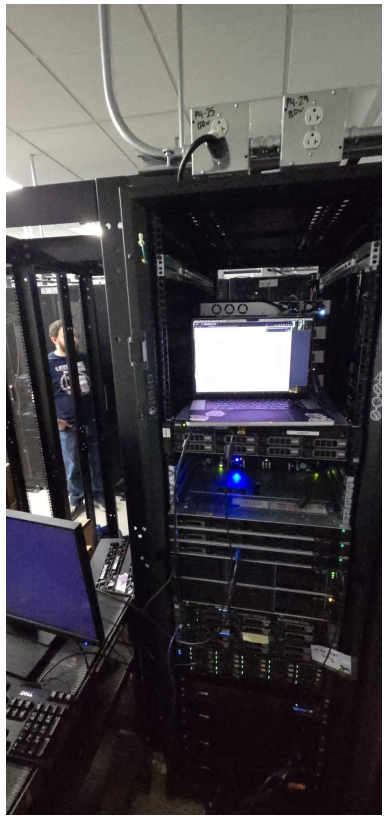
- Authentik ate more and more RAM
  - Authentik Enterprise “Ready”
  - We had to go from 8GB of RAM to 24 GB of RAM
- DB was stuck doing slow queries
- VPN was unable to stay connected
  - Panic time?

- Just throw more * at the problem



Second Run: All Problems Fixed?

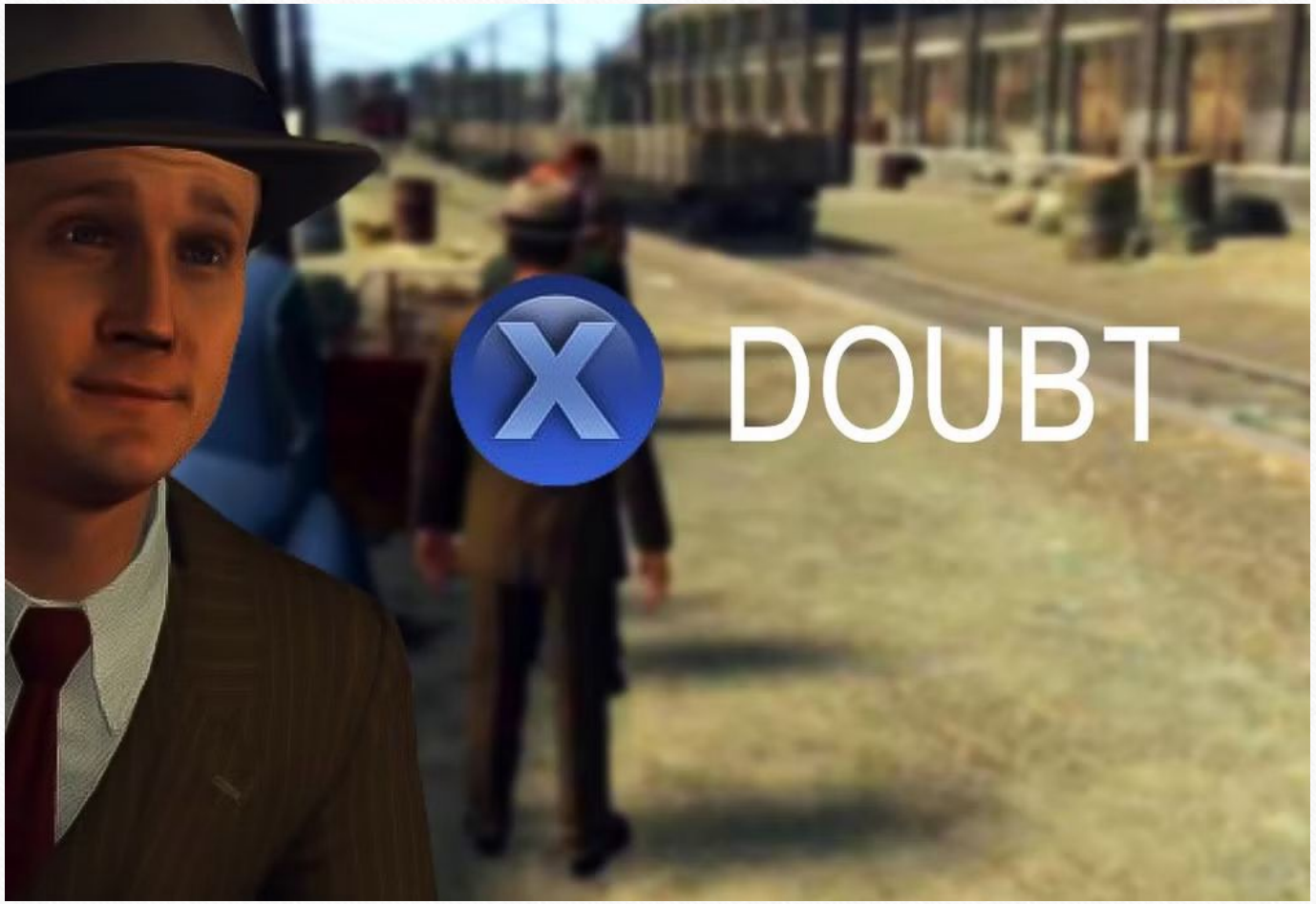
- VPN was connecting users to different accounts
- Accidentally reset scores
- At least Authentik was working….



Third Run:

- Crashed core networking
- Interfaces and Core Router Configuration Reset/Wrong
- Failed to give teams access as expected

I’m sure it’s fine now…

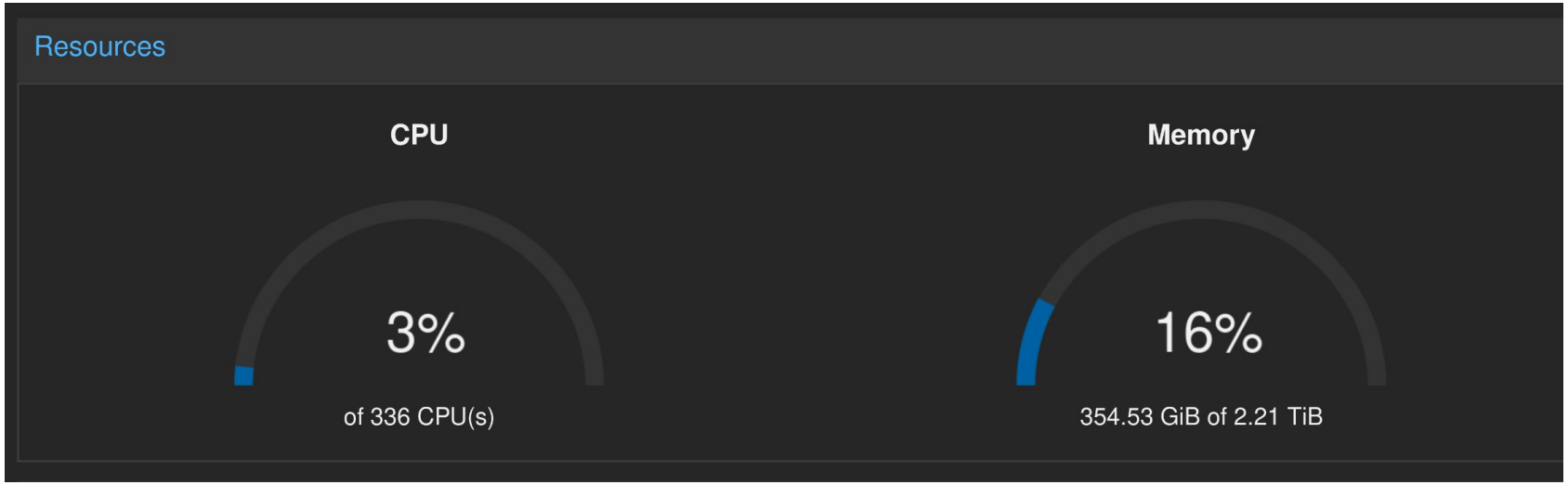


Or is it… The Resource Problem

50 Teams * 10 VMs/team = 500 VMs

2 cores/VM = 1000 CPU Cores

4 GB/VM = 2TB Memory



The Resource Problem

Limited budget. High expectations.

- Budgets are tight
- Hardware is not plentiful
- Lots of parallel compute/network activity
- Competitors/students don’t care
  - **SLOW IS SLOW**

Teacher hypes up class party all year

The Party:

2:38 PM · Aug 23, 2020

What to do?

- Give up.

"We don't have the budget."

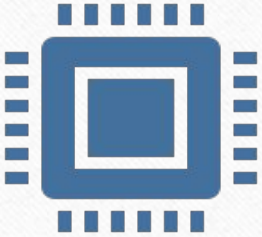
- wasabi

What to do?

Giving up is not an option.

Educational experiences are incredibly valuable.

Students depend on us.



Budget controls <u>hardware purchasing</u>

Hardware is **not** all of performance

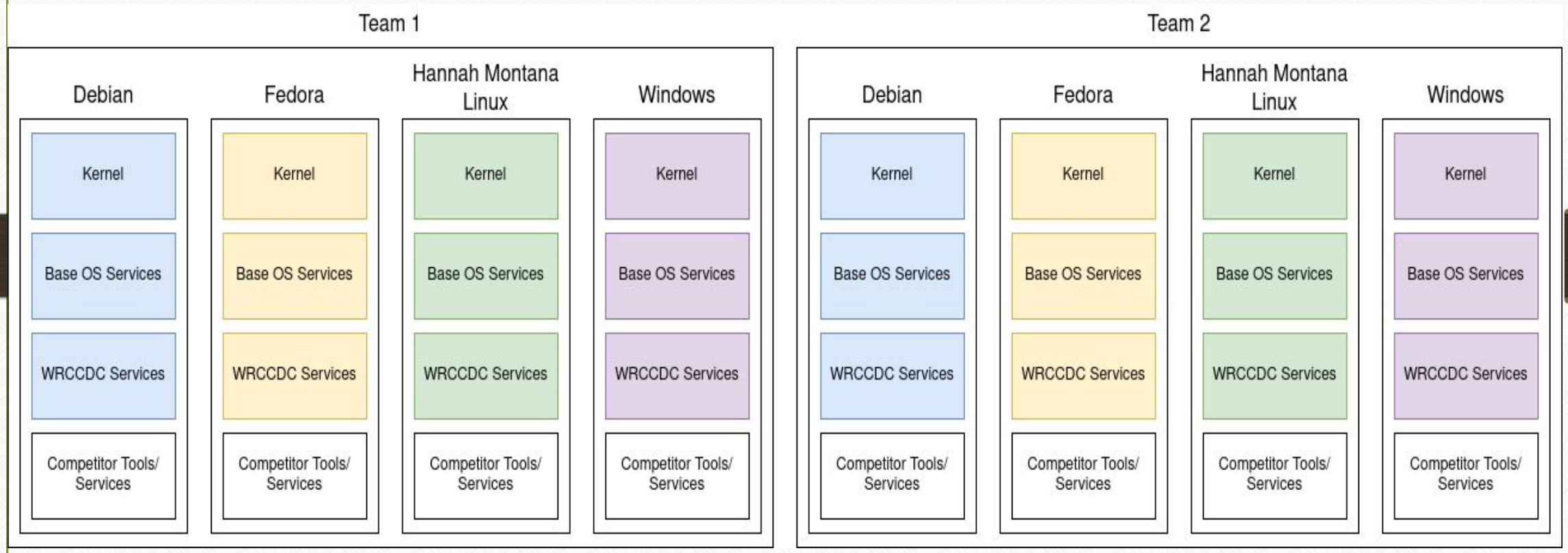
Hardware => 60-80% of performance

Software Optimizations => **Last 20-40%**

Addressing the Elephant

- Memory is expensive
  - Ask IT/ewaste orgs
- Lots of competitors = lots of VMs = lots of memory
- BUT a lot of memory pages are the same...

Kernel Samepage Merging (KSM)



Kernel Samepage Merging (KSM)

- Merges identical memory pages together
- Useful in mass environment deployments
- Enabled **by default** in Proxmox!

The Compute Problem

- What do we need?
  - Quality of Service
- When did we need it?

The Compute Problem

- What do we need?
  - Quality of Service
- When did we need it?
  - Yesterday.

Compute Quality of Service

- QoS is not magic
- QoS limts the maximum amount of resources one team can use
  - High usage of compute resources = worse experience
  - Allows for all competitors to have an equal experience
  - No 1 team hogs all resources

Compute Quality of Service

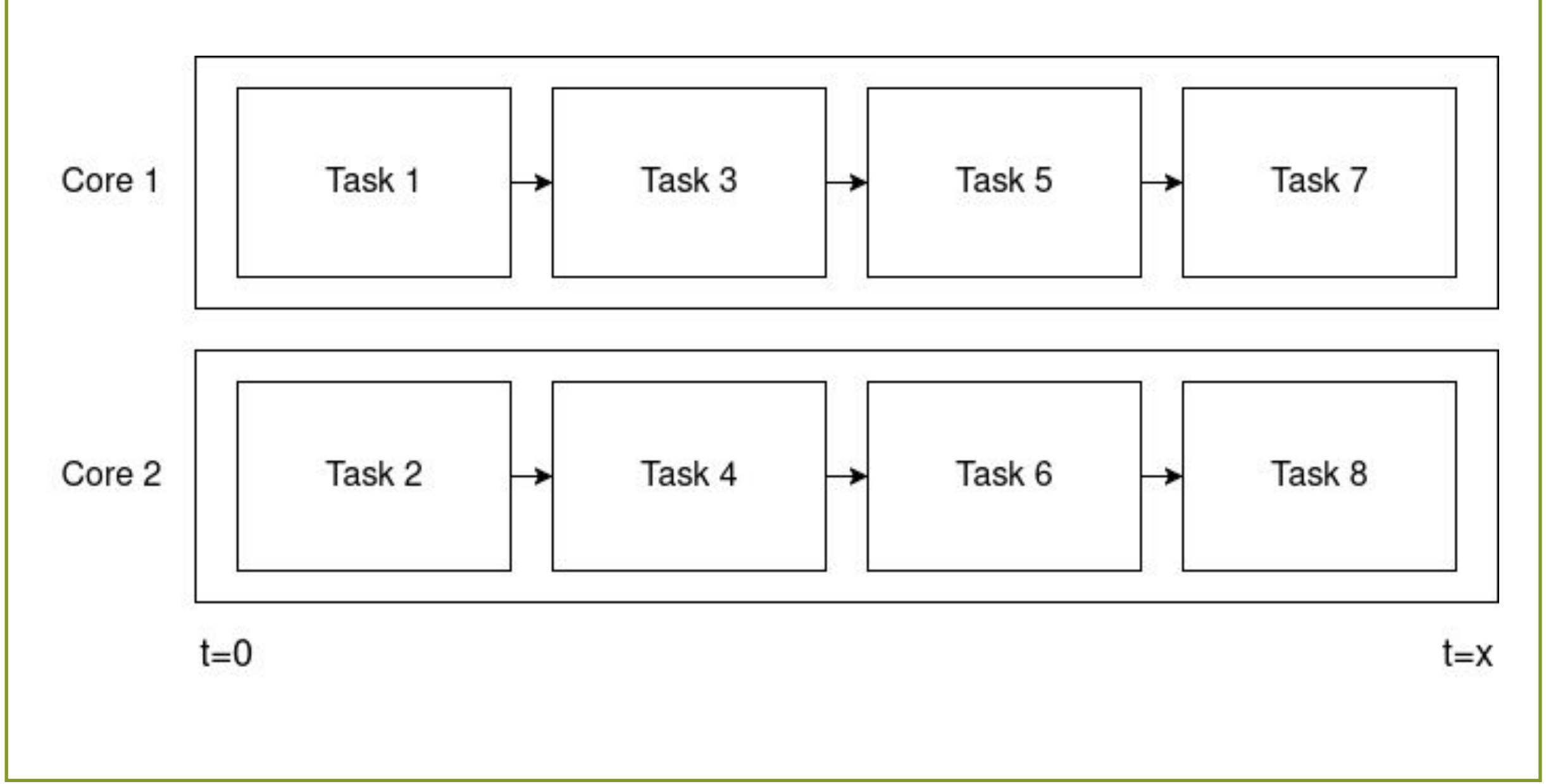
CPUQuota=

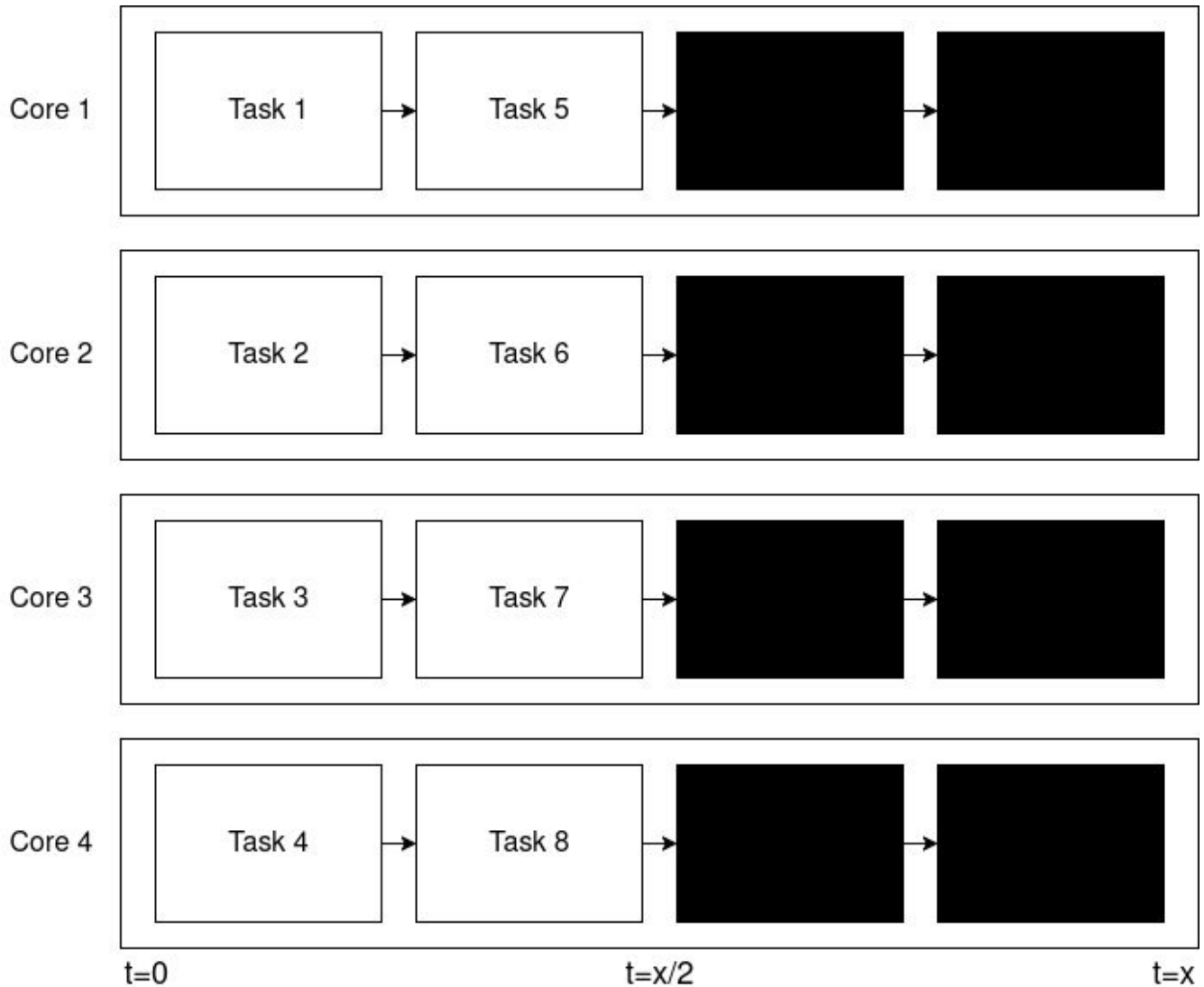
Assign the specified CPU time quota to the processes executed. Takes a percentage value, suffixed with "%". The percentage specifies how much CPU time the unit shall get at maximum, relative to the total CPU time available on one CPU. Use values > 100% for allotting CPU time on more than one CPU. This controls the "cpu.max" attribute on the unified control group hierarchy and "cpu.cfs_quota_us" on legacy. For details about these control group attributes, see **Control Groups v2**[2] and **sched-bwc.txt**[5].

Example: *CPUQuota=20%* ensures that the executed processes will never get more than 20% CPU time on one CPU.

Compute Quality of Service

- cpulimit in Proxmox = CPUQuota in systemd
  - One sets the other!
- `cpulimit = 2.0`
  - Whats the difference between this and just setting the CPU cores to 2 for the VM?


Core 1
Task 1
Task 3
Task 5
Task 7
Core 2
Task 2
Task 4
Task 6
Task 8
t=0
t=x

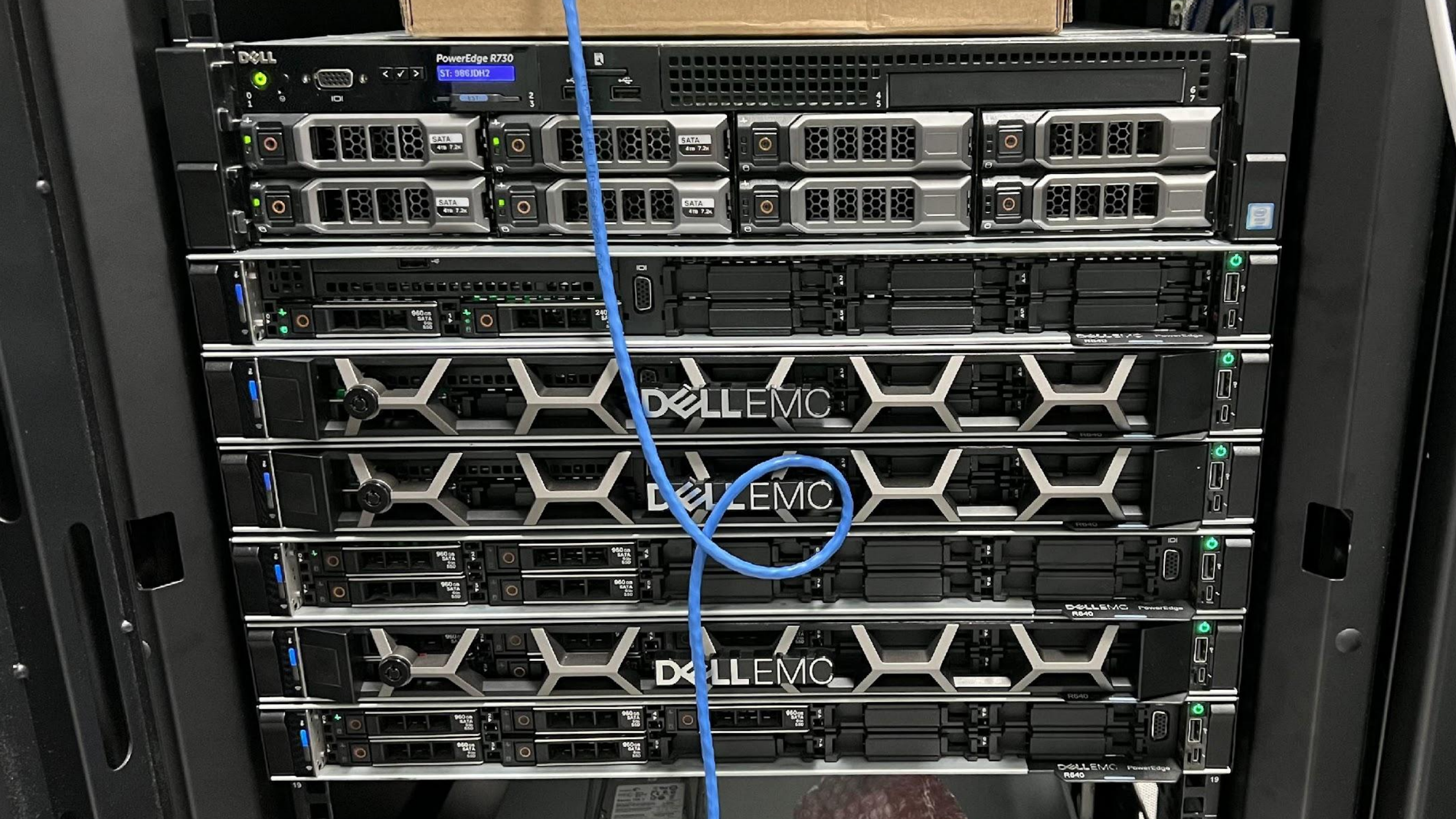

Core 1
Task 1
Task 5
Core 2
Task 2
Task 6
Core 3
Task 3
Task 7
Core 4
Task 4
Task 8
t=0
t=x/2
t=x

Compute Quality of Service

- The Linux Scheduler can now help us
  - Earliest Eligible Virtual Deadline First (EEVDS) Scheduler
  - Ensures fair scheduling across processes/threads of same priority

The Storage Problem

- Storage is expensive (esp now)
- Centralized network storage = savings
  - Faster live migrations
  - A lot easier to get wrong
    - Optimizations can get 200-400% performance uplift


PowerEdge R730
DELL EMC
PowerEdge R640

TrueNAS as a Storage Backend

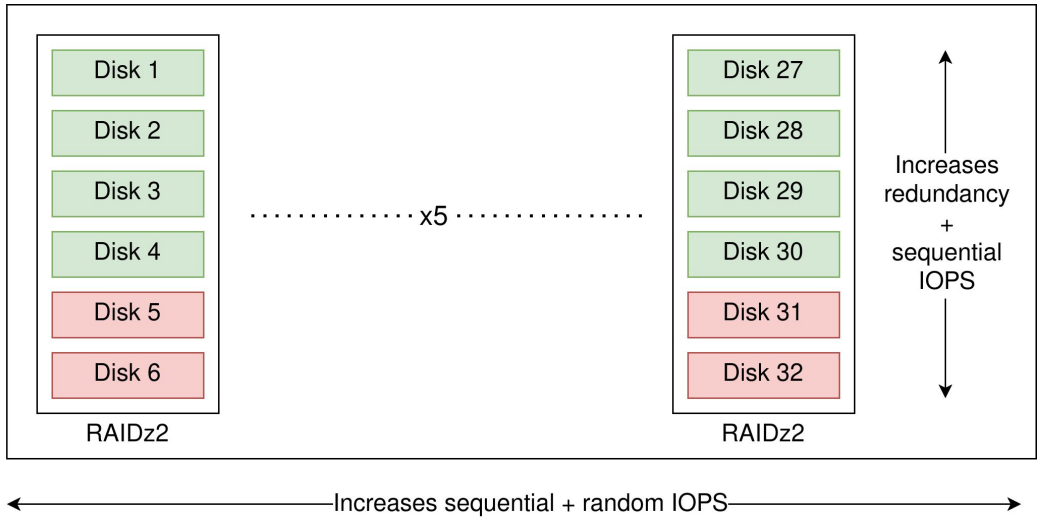
- Easy GUI-based administration of storage infrastructure
- ZFS can be performant **IF** you tune it properly, *but…*
  - Extremely hardware dependent
- Built in network storage configuration
  - NFS isn’t as bad as everyone says it is

ZFS Hardware Selection

- Buy what you can afford and tune your system for it
- CPU Cores > CPU Speeds
- VMs depend on **A LOT** of random I/O
  - Optimize for this as much as possible
  - Must compromise random IO vs storage volume on HDD
  - Not as much of a problem on SSDs
- Caching is just as important

What we run

- 2x Netapp 2.5” JBODs with 20x 1TB HDDs each = 40 disks
  - Arranged in a 7x 6 disk RAIDz2 = 8 hotspares

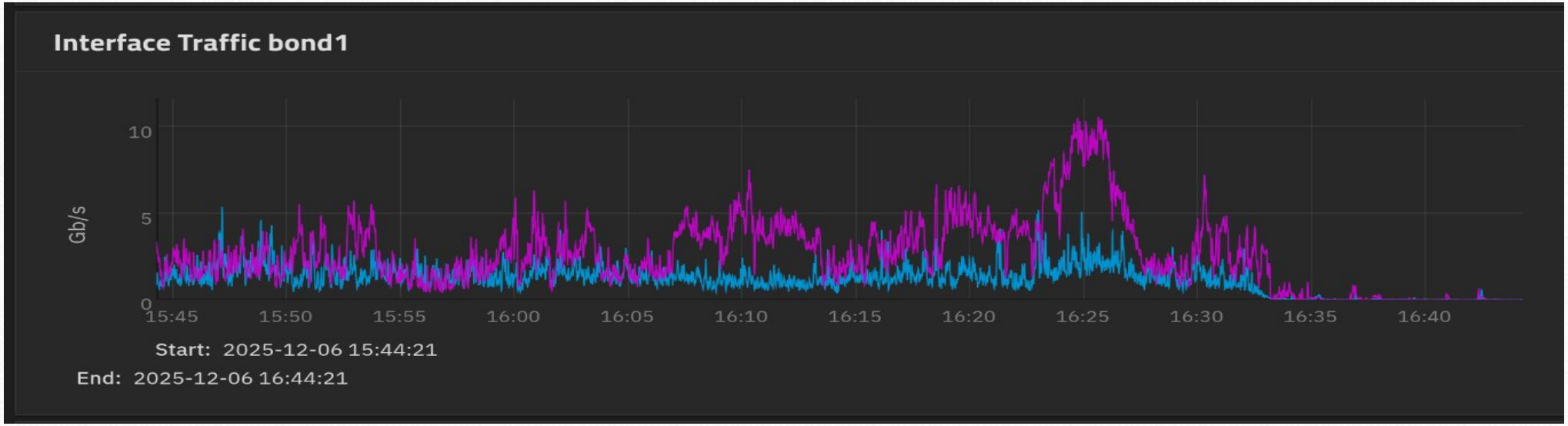


What we run

7x 6 disk RAIDz2

What we run

- 300GB of Memory as L1 cache
- Enough to fit entire competition environment
  - 2TB Intel P4510 NVME as L2 cache
- These are still cost effective!
  - 375GB Intel Optane P4800X as SLOG
- RIP Optane
- **DO NOT TURN OFF SYNC WRITES**


Interface Traffic bond1
Gb/s
10
5
0
15:45
15:50
15:55
16:00
16:05
16:10
16:15
16:20
16:25
16:30
16:35
16:40
Start: 2025-12-06 15:44:21
End: 2025-12-06 16:44:21

Intel SSD DC P4510 SERIES 1.0TB SSDPE2KX010T8 2.5" NVMe/PCIe FREE SHIPPING 🚚

Potential-Electronics (8347)
100% positive · Seller's other items · Contact seller

US $160.00 or Best Offer

as low as $27.77/mo with Klarna. Learn more

Condition: **Used**

";) Before purchasing, please SEE DESCRIPTION AND ALL PICTURES to see what is/is not included in your"... Read more

Quantity: 1 More than 10 available

Intel Optane P4800x 375GB SSD U.2 NVME PCIE HP SSDPE1K375GAP1 Solid State Drives

szdm888 (14)
88.9% positive · Seller's other items · Contact seller

US $142.28

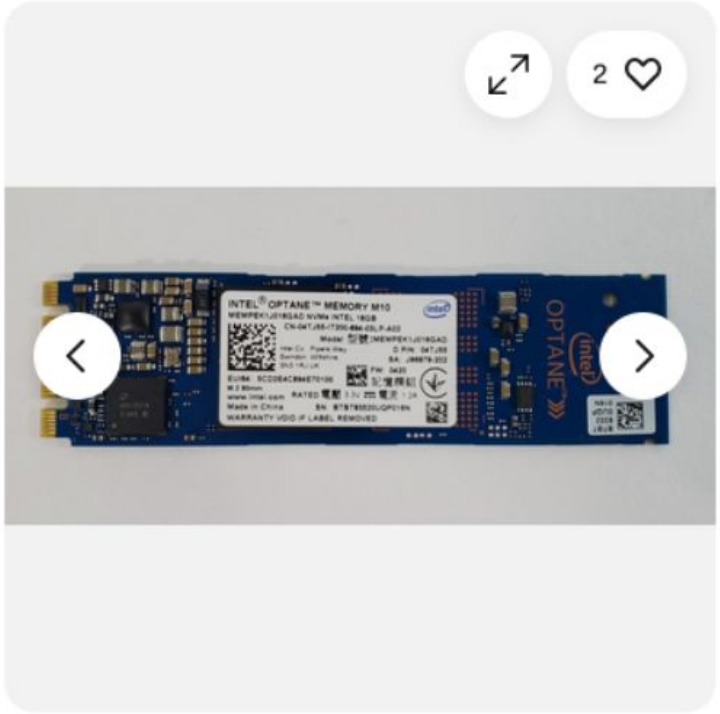
as low as $48.54/mo with Klarna. Learn more

Condition: **Open box**

"Work normally, tested before shipping"

Buy It Now

Add to cart



Intel Optane Memory M10 16GB MEMPEK1J0116GAD NVMe M.2 SSD

Finding Treasures For You Shop (1027)
100% positive · Seller's other items · Contact seller

US $9.95 or Best Offer

Condition: **Used**

Buy It Now

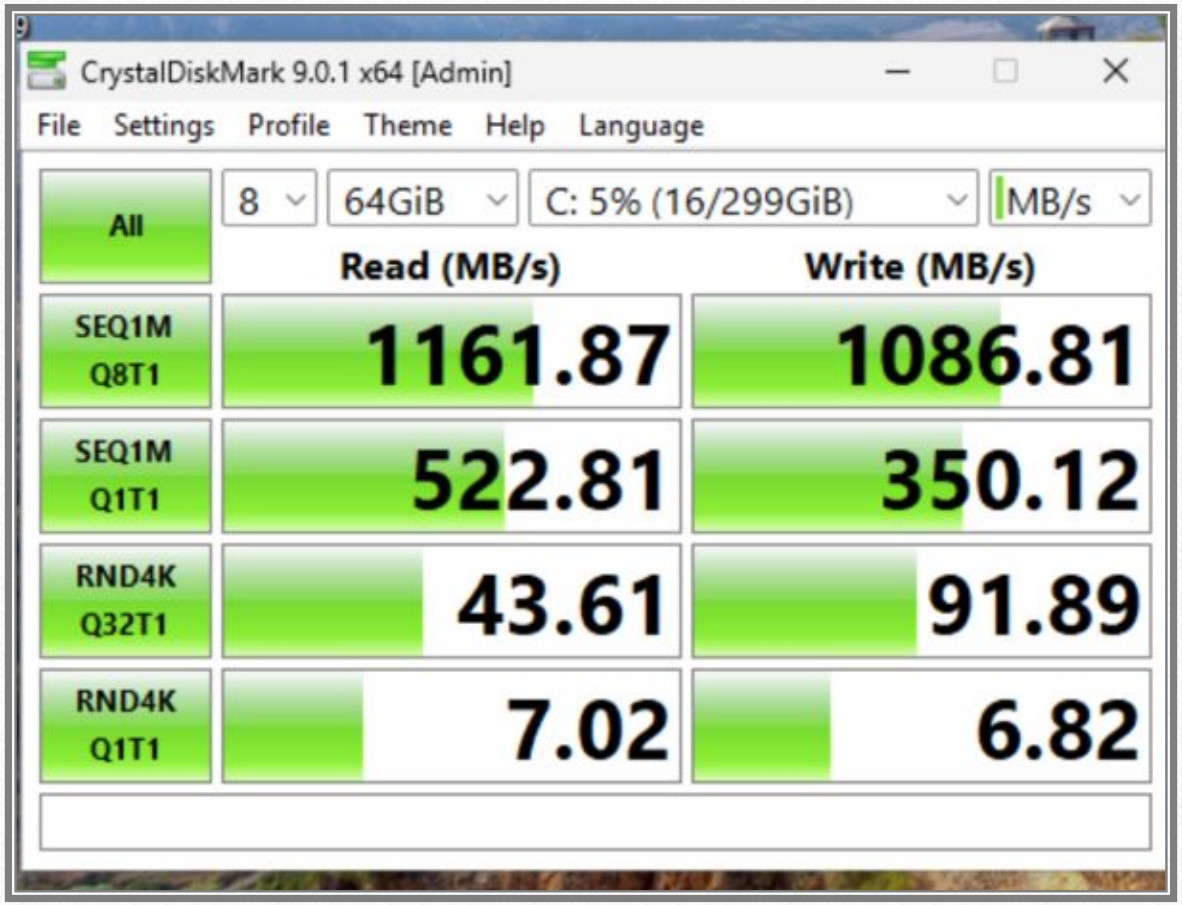
Add to cart

Make offer

Add to Watchlist

PSA: DO NOT BUY THESE

Eh.


CrystalDiskMark 9.0.1 x64 [Admin]
File
Settings
Profile
Theme
Help
Language
All
8
64GiB
C: 5% (16/299GiB)
MB/s
Read (MB/s)
Write (MB/s)
SEQ1M Q8T1
1161.87
1086.81
SEQ1M Q1T1
522.81
350.12
RND4K Q32T1
43.61
91.89
RND4K Q1T1
7.02
6.82

Misconception: NFS for Virtualization is Slow

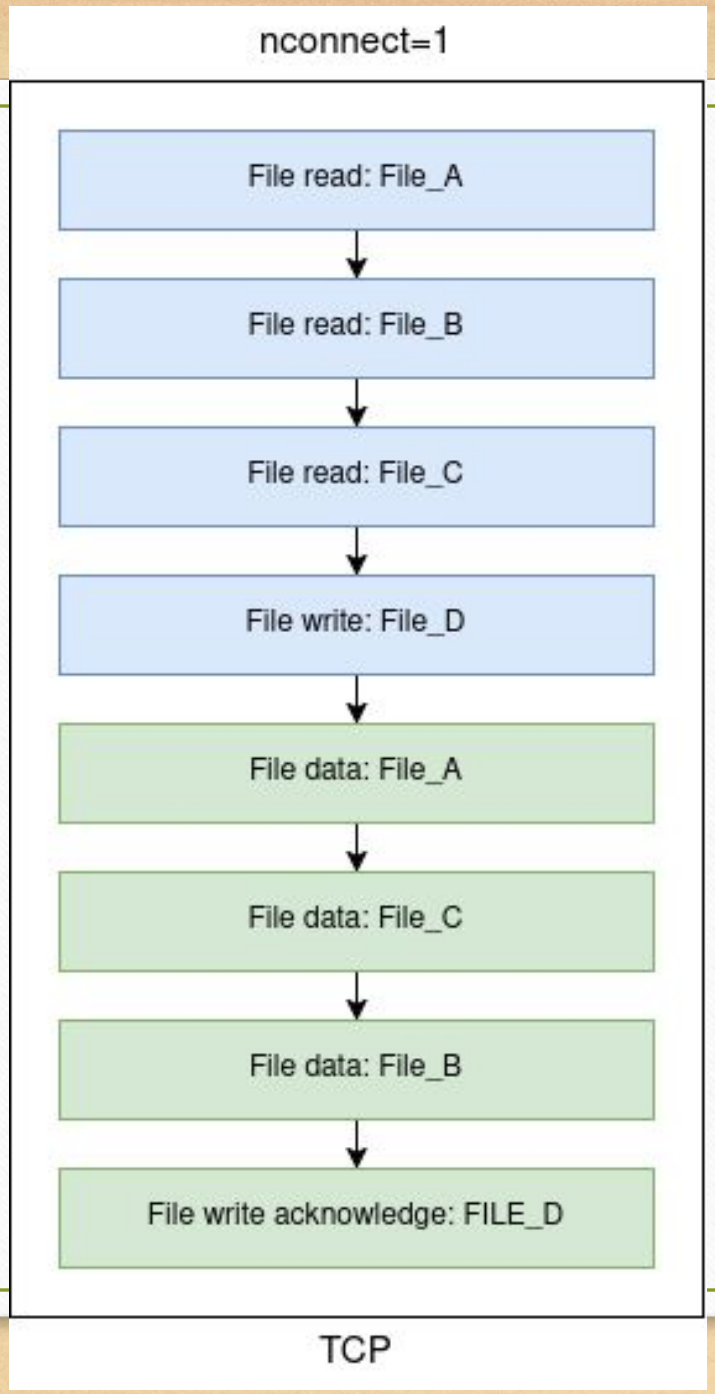
Misconception: NFS for Virtualization is Slow

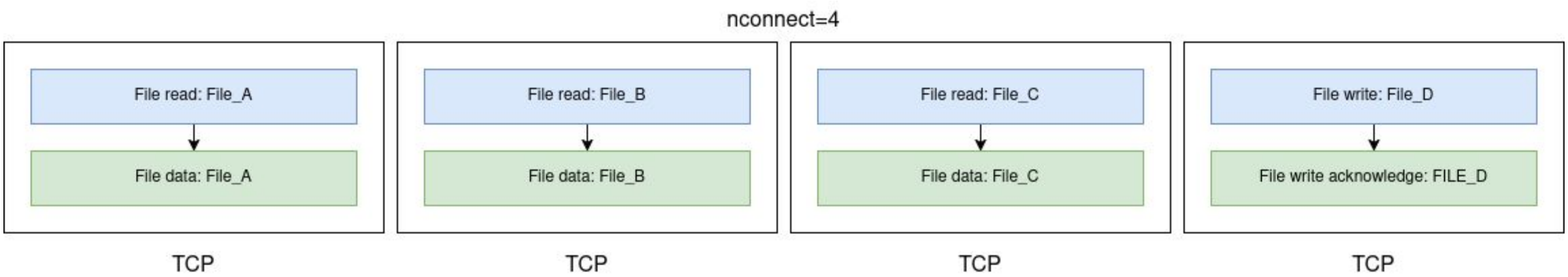
It’s not slow. The default settings suck.

nconnect=n

When using a connection oriented protocol such as TCP, it may sometimes be advantageous to set up multiple connections between the client and server. For instance, if your clients and/or servers are equipped with multiple network interface cards (NICs), using multiple connections to spread the load may improve overall performance. In such cases, the **nconnect** option allows the user to specify the number of connections that should be established between the client and server up to a limit of 16.

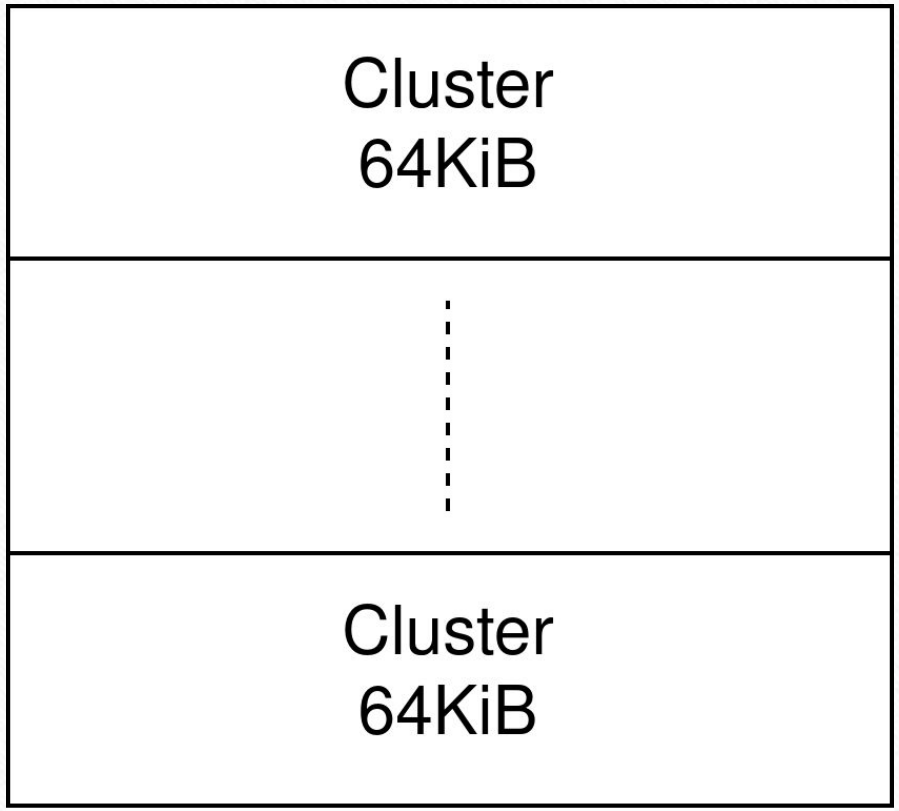
Note that the **nconnect** option may also be used by some pNFS drivers to decide how many connections to set up to the data servers.


nconnect=1
File read: File_A
File read: File_B
File read: File_C
File write: File_D
File data: File_A
File data: File_C
File data: File_B
File write acknowledge: FILE_D
TCP

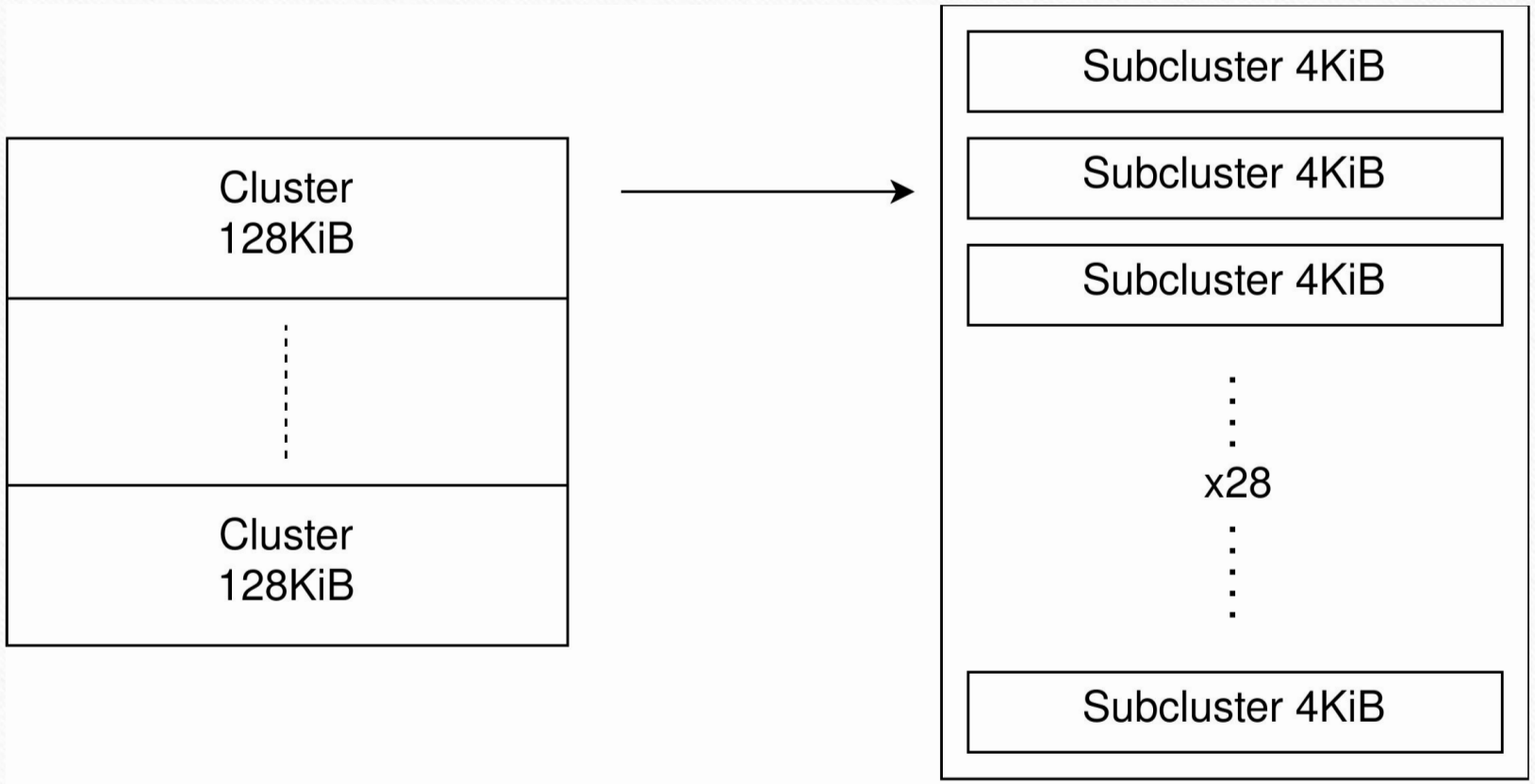

nconnect=4
File read: File_A
File data: File_A
File read: File_B
File data: File_B
File read: File_C
File data: File_C
File write: File_D
File write acknowledge: FILE_D
TCP
TCP
TCP
TCP

Lets Optimize QCow2

- Used for file-based VM storage
  - Supports snapshots in Proxmox
- Very easy to get wrong (don’t trust the defaults)
  - QCow2: cluster_size = 128kib
  - QCow2: extended_l2 = on
  - ZFS: record_size = 128kib



Default QCow2



Optimized QCow2

Intel Xeon Platinum 8173M 2.0Ghz 38.5MB 28-Core 165W LGA3647 SR37Q

MKTLLC (13184)
100% positive · Seller's other items · Contact seller

US $33.00/ea or Best Offer

Condition: **Used**
"Clean Tested Pulls. 30 Day Money Back Guarantee and 90 Day Warranty."

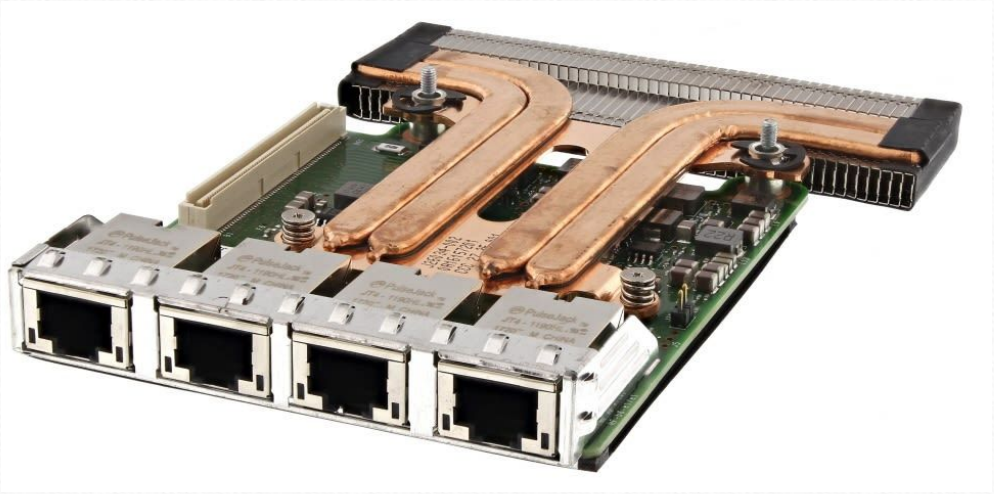
Quantity: 1 More than 10 available · 136 sold

Bulk savings:

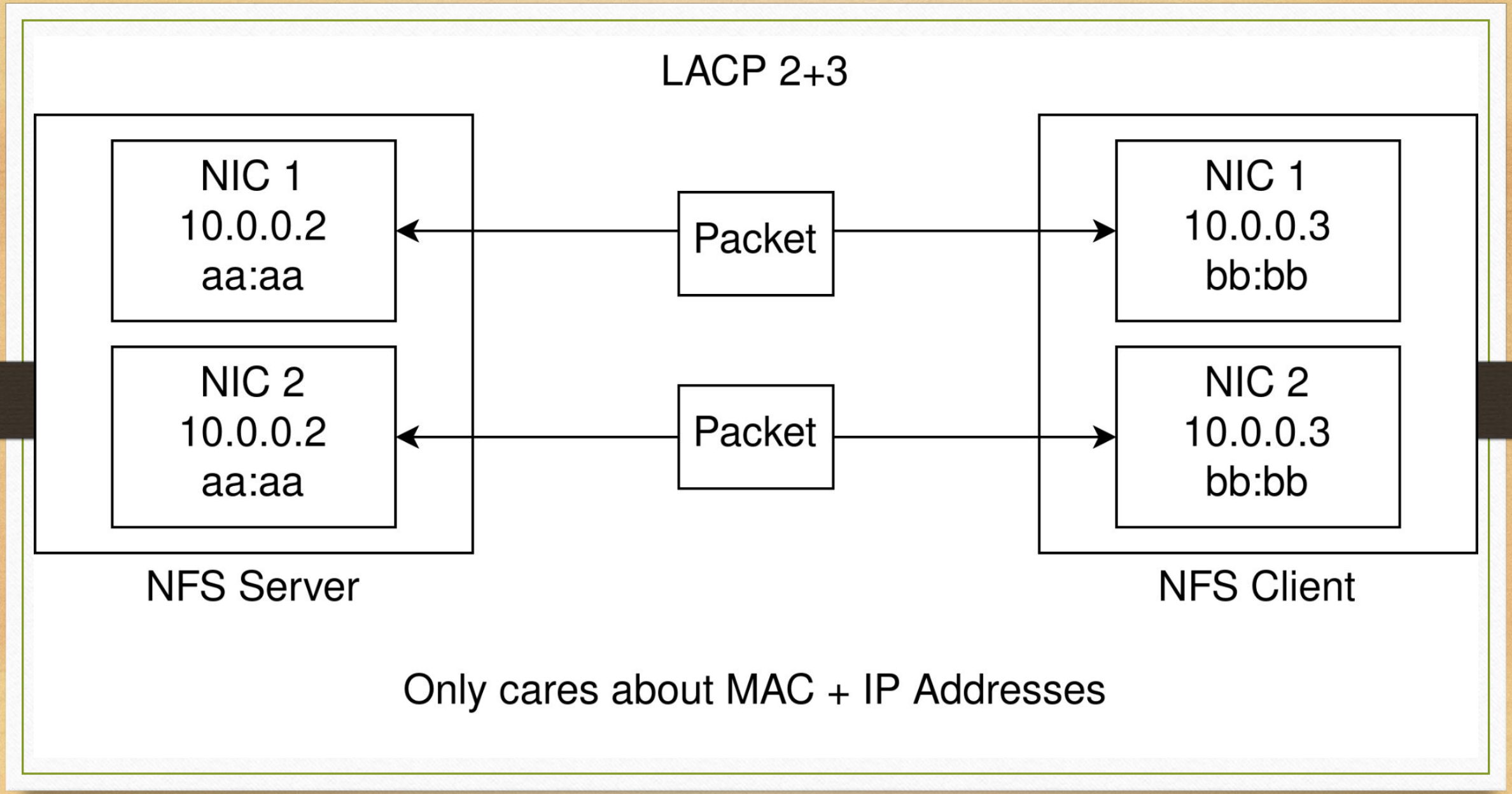
Buy 1	Buy 2
$33.00/ea	**$31.35/ea**

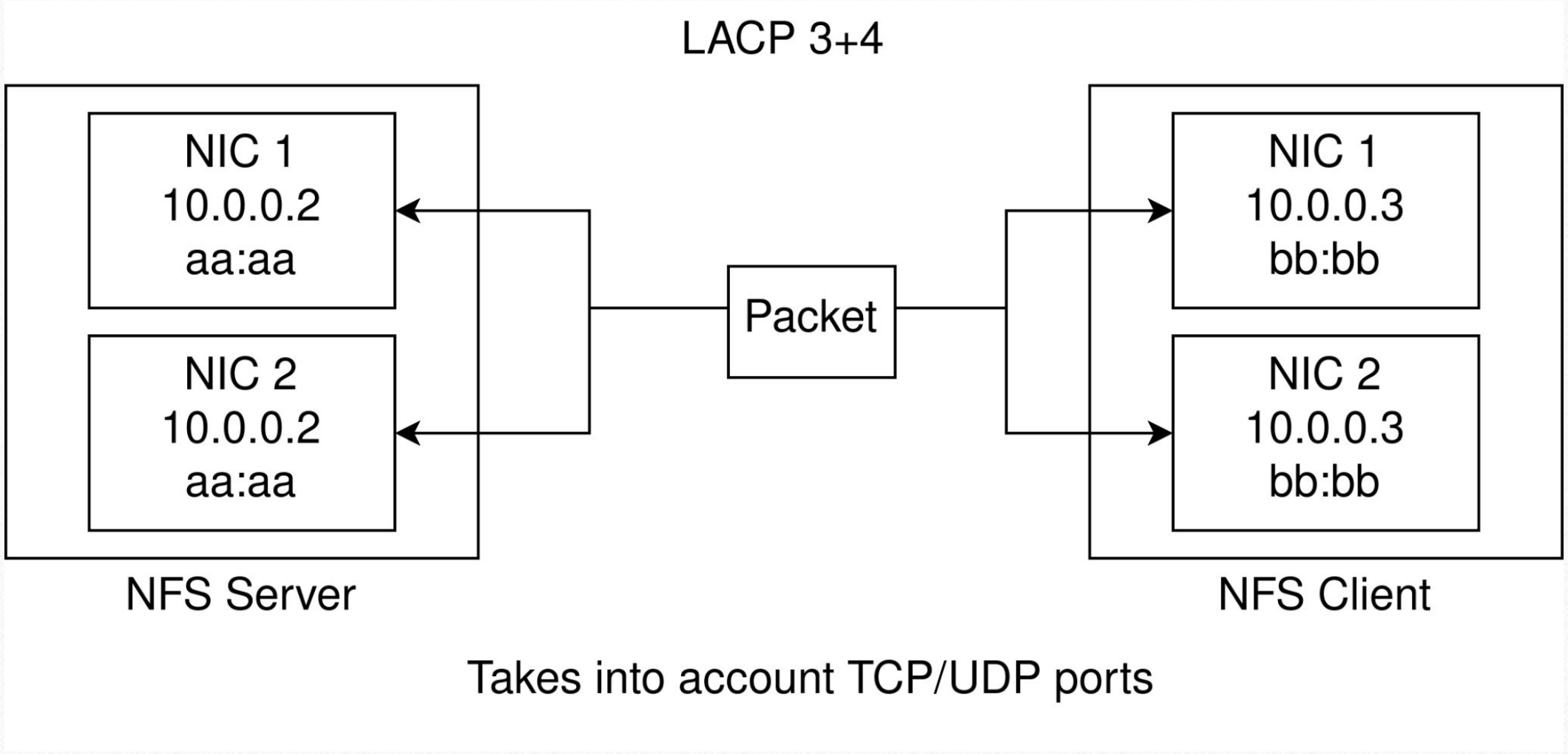
Maximizing Network Bandwidth

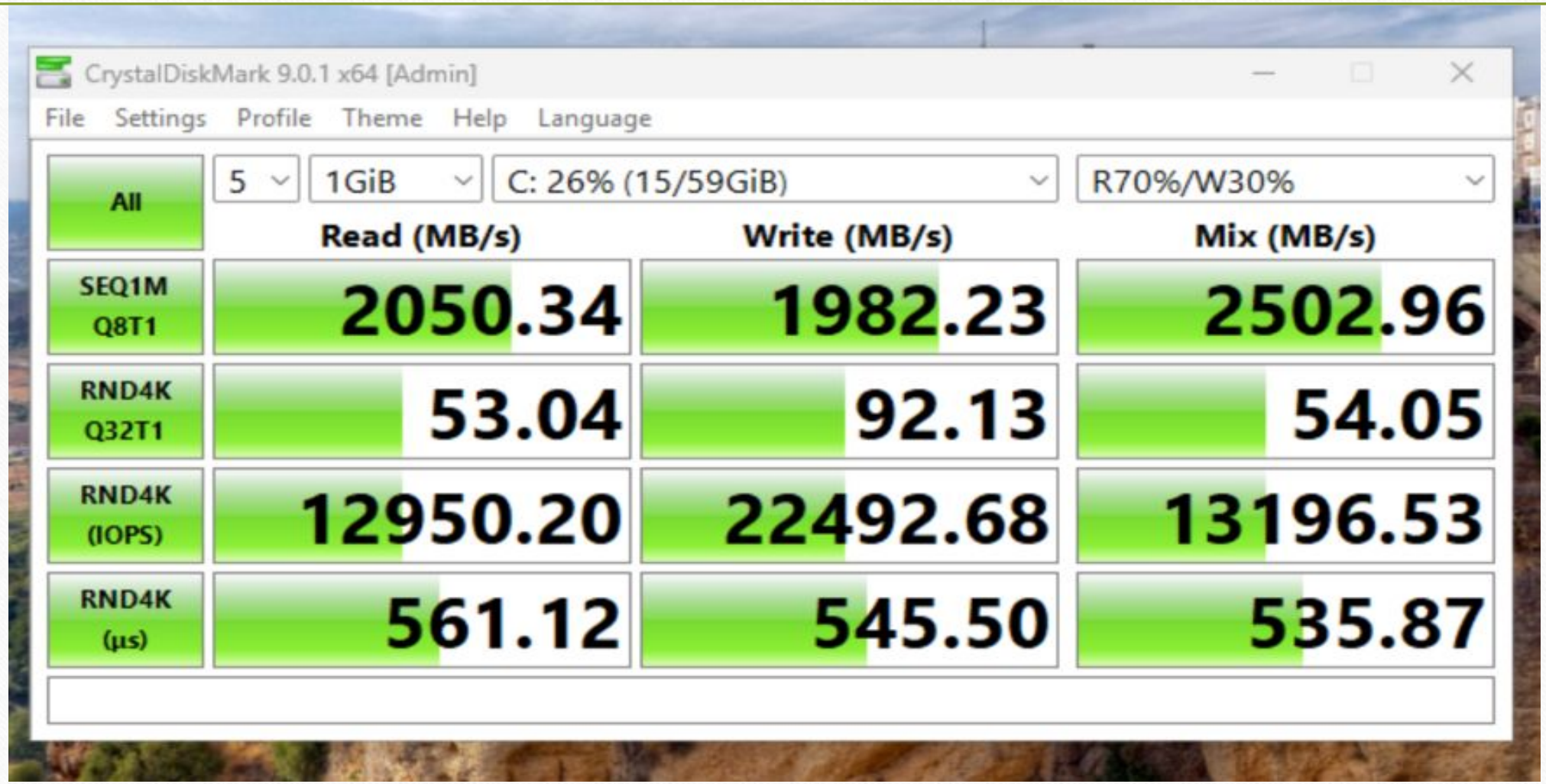
- I bought all the ports – **I want to use all the ports!**
- 40gbps max duplex throughput
  - BUT a single port can only do 10gbps...



Intel X550-T4


LACP 2+3
NIC 1
10.0.0.2
aa:aa
NIC 2
10.0.0.2
aa:aa
Packet
Packet
NIC 1
10.0.0.3
bb:bb
NIC 2
10.0.0.3
bb:bb
NFS Server
NFS Client
Only cares about MAC + IP Addresses


LACP 3+4
NIC 1
10.0.0.2
aa:aa
NIC 2
10.0.0.2
aa:aa
Packet
NIC 1
10.0.0.3
bb:bb
NIC 2
10.0.0.3
bb:bb
NFS Server
NFS Client
Takes into account TCP/UDP ports


CrystalDiskMark 9.0.1 x64 [Admin]
File Settings Profile Theme Help Language
All
5
1GiB
C: 26% (15/59GiB)
R70%/W30%
Read (MB/s)
Write (MB/s)
Mix (MB/s)
SEQ1M Q8T1
2050.34
1982.23
2502.96
RND4K Q32T1
53.04
92.13
54.05
RND4K (IOPS)
12950.20
22492.68
13196.53
RND4K (µs)
561.12
545.50
535.87

Better, but...

The Final Touches

- Increase packet size
  - MTU = 9000
- Latency: pushing the hardware to it's limits
  - Disabling hardware offloads yields a 500% performance increase in random IOPS

```
  GNU nano 8.4      /etc/network/interfaces
auto eno3
iface eno3 inet manual
        mtu 9000
        offload-tso off
        offload-ufo off
        offload-gso off
        offload-gro off
        offload-lro off
        offload-rxhash off
```

CrystalDiskMark 9.0.1 x64 [Admin]

File Settings Profile Theme Help Language

All | 5 | 1GiB | C: 30% (17/59GiB) | R70%/W30%

	Read (MB/s)	Write (MB/s)	Mix (MB/s)
SEQ1M Q8T1	2254.96	2106.04	2713.10
RND4K Q32T1	244.47	254.24	239.47
RND4K (IOPS)	59685.55	62069.34	58465.58
RND4K (µs)	222.83	429.74	309.82

WRCCDC Backend (Peak Performance + Mix, Sync Write)

A lot of optimizations

Not enough time

Closing

Shout out to the WCComps Volunteers

Shout out for hours of setup work this year

- Dex (stepping up)
- Toku
- Alqemy
- Beesod
- Payton Erickson
- Sean Richardson
- Dessa Xaviera
- Jimmy
- Niterabbit

- And many others!

WCCOMPS INFO

- https://wccomps.org
- Feel free to join us and volunteer!

Questions?

wasabi

- Social media:
  - infosec.exchange/@spicywasabi
  - x.com/spiceywasabi
  - bsky.app/profile/spicywasabi

dex

- Contact Info
  - Discord: @me_dex
  - Signal: @dex.123
  - Email: dex@fur.lol
  - Matrix: dex@fur.lol