Digging Deep into Disk Diagnoses

David Paper
Sr. Escalations Manager & Technical Smokejumper
Splunk

splunk> .conf19

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Splunk Community Rocks!

This means you

.conf19

splunk>

The 3 Whys

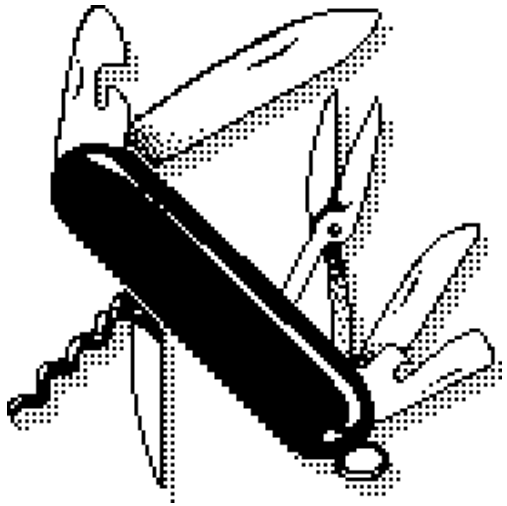
- Why are you here?
  - You ~~want~~ need a fast data platform, storage performance may be hindering it
- Why am I here?
  - Goal in life is to learn from mistakes/problems of others, I want you to avoid being the others
- Why me?
  - Smokejumper: I seeing storage issues with customers too often

Who am I?

- Recovering Ops-aholic
- Details guy in a holistic approach gig
- Share clue, do more interesting things
- Previous .conf speaker: 2013, 2014, 2016, 2019
- Slack & IRC #splunk: cerby

Disk Deep Dive

- Intro - Done!
- Problem definition
- Investigation
- Go Deeper
- SmartStore
- A little Q&A
- Next Steps



- Questions? Bring it
- Discuss the merits? Hold ’em until the end, maybe hallway time

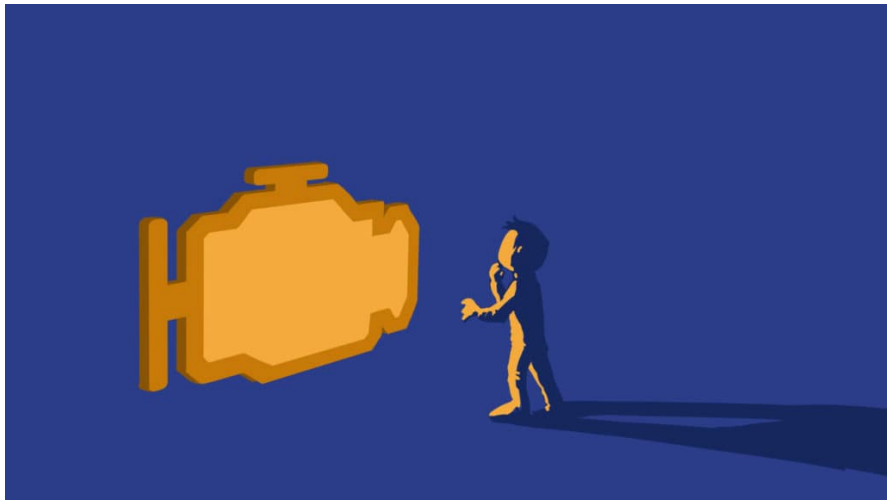
I'm givin' her all she's got, Captain

- Busy Splunk is resource intensive
- Performant environments have sufficient
  - CPU cores
  - RAM ~~swap~~
  - Storage capacity & performance (IOPS?)
- What needs to be fast (we'll define fast later)?
  - Dispatch
  - Hot/warm [cachemanager for SmartStore]
  - Cold [fast enough]

If only there were a Check Engine Light

Tangible issues with sluggish storage

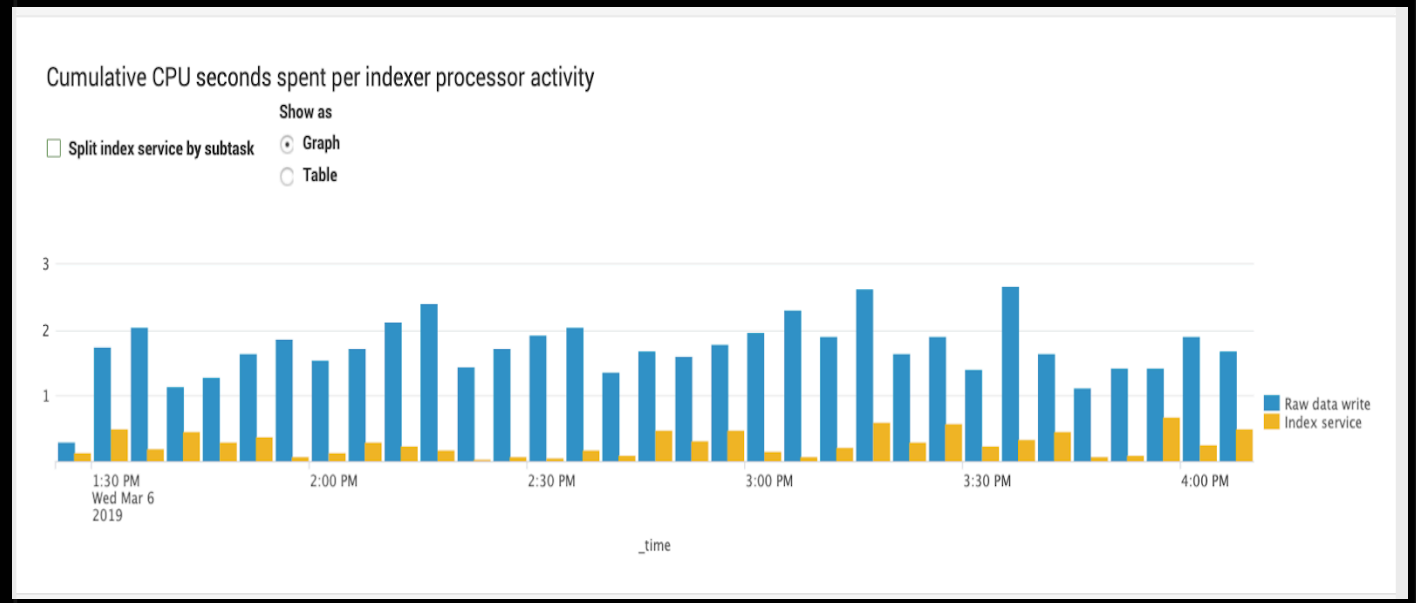
- SH -> IDX bundle rep failures, timeouts, or warning messages
- Ingestion queues filling, Index queue first
- Forwarders begin to queue
- Cluster data replication issues w/in SHC or IDXC
- Rolling restarts take a long time to recover full RF/SF



What can Splunk tell us?

Monitoring Console

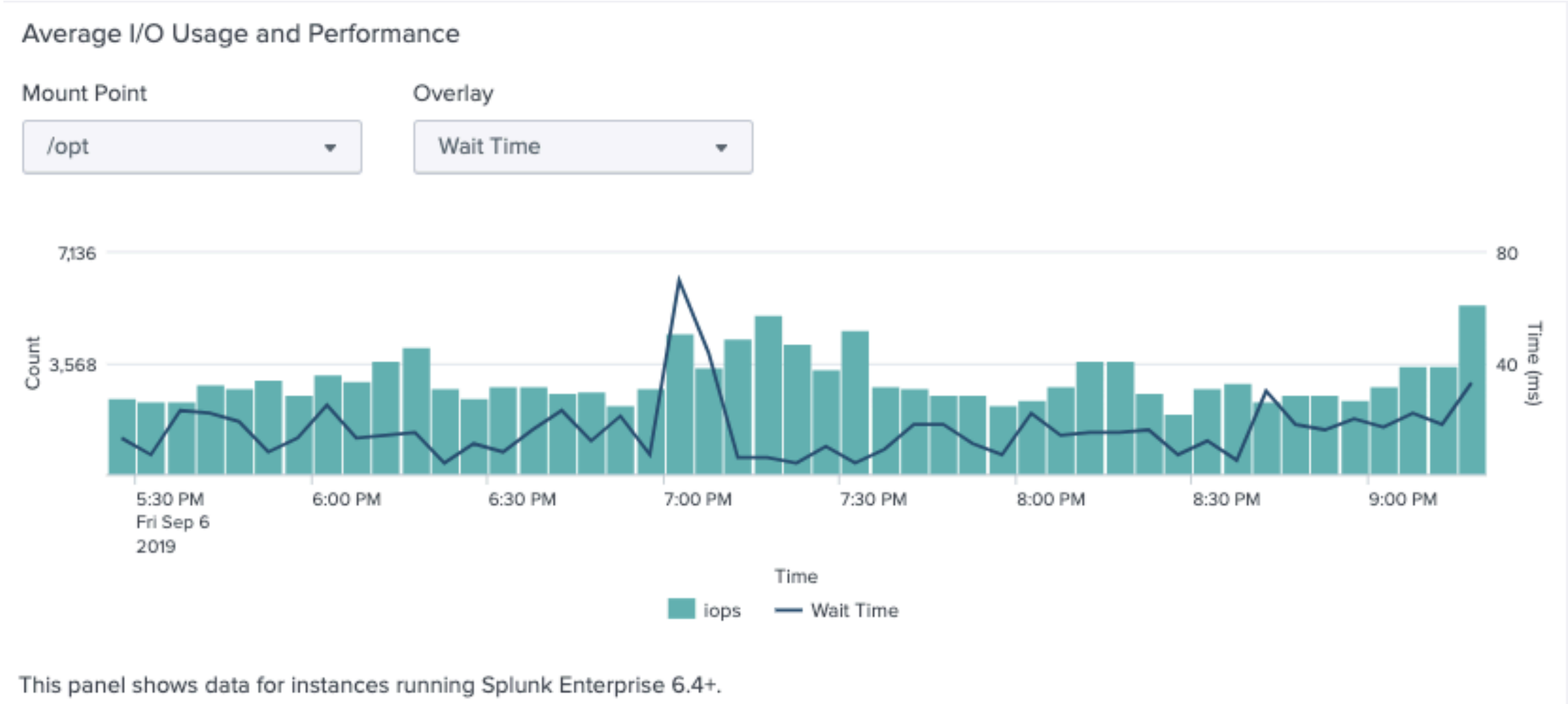
MC -> Indexing -> Indexing Performance -> Indexing Performance: Instance



- "Cumulative CPU Seconds Spent per Indexer Process Activity" panel
- If Raw Data Writes > Indexing Service, indicator storage is struggling to keep up
  - Ingestion queues may not be backed up…yet

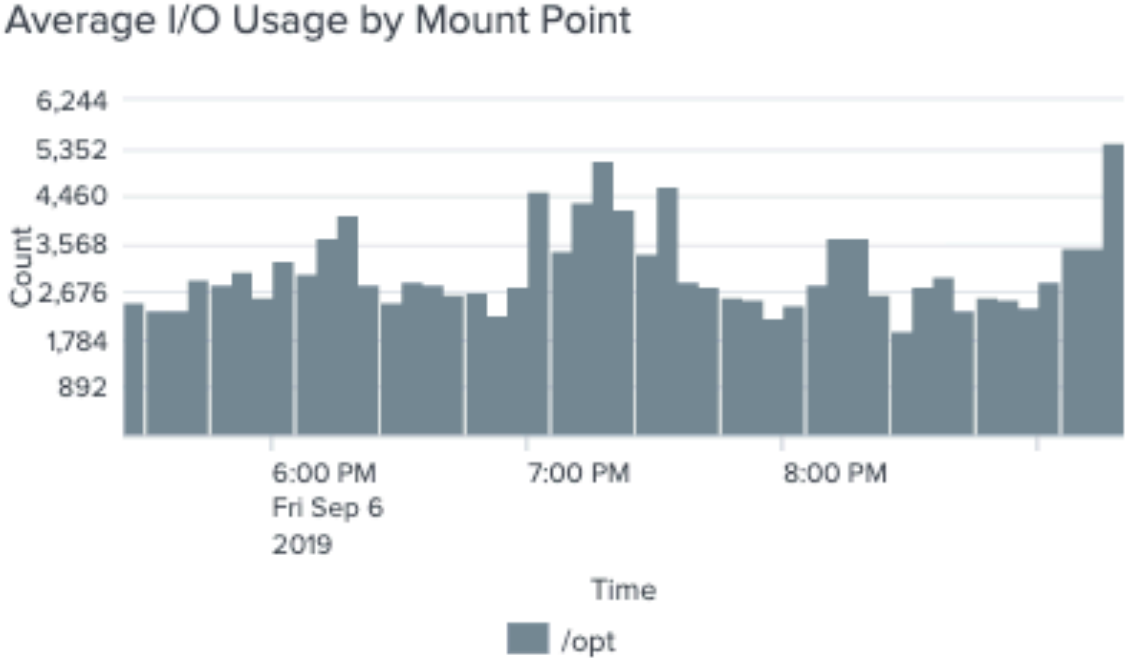
Anything else from MC?

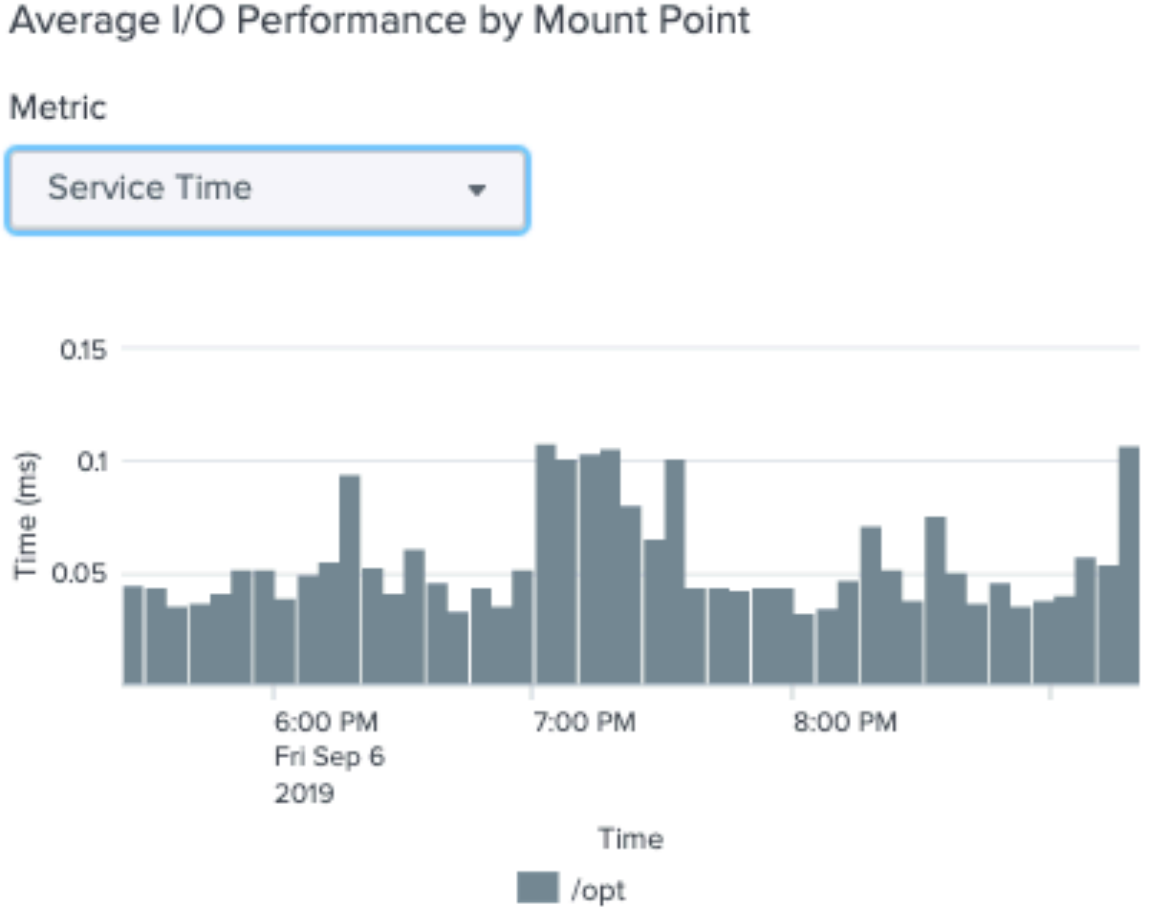
MC -> Resource Usage -> Resource Usage Machine



More from MC?

MC -> Resource Usage -> Resource Usage Machine





splunk> .conf19

What am I looking at on the previous slides?

MC -> Resource Usage -> Resource Usage Machine*

Look for

- IOPS performance that has unexpectedly hit a ceiling or values that should be higher based on historical utilization of the host (effort to suss out)
- Wait times > 10 - 15ms on average
- Spikes may be OK, if no recovery, not OK

*Powered by _introspection data at 60s intervals

Wait, what

Whatcha talkin bout Willis?

- 60s interval is a long time
  - Too long

- iostat (sysstat package in RedHat) to the rescue: `iostat -zx 1`
  - Zero out devices not active (easier to read)
  - Extended format
  - Update every 1 second
- Note: Ignore the first set of results

This is IO deep dive -- Show me

```
Device:         rrqm/s   wrqm/s     r/s     w/s    rkB/s    wkB/s avgrq-sz avgqu-sz   await r_await w_await  svctm  %util
xvda              0.00    26.00    0.00   54.00     0.00  1376.00    50.96     0.46    8.48    0.00    8.48   0.22   1.20
xvdb              0.00     0.00    0.00    6.00     0.00   116.00    38.67     0.01    1.00    0.00    1.00   0.17   0.10
dm-0              0.00     0.00    0.00   80.00     0.00  1376.00    34.40     0.67    8.41    0.00    8.41   0.15   1.20
dm-1              0.00     0.00    0.00    6.00     0.00   116.00    38.67     0.01    1.00    0.00    1.00   0.17   0.10

avg-cpu:  %user   %nice %system %iowait  %steal   %idle
           1.92    0.00    7.27    0.00    0.08   90.73

Device:         rrqm/s   wrqm/s     r/s     w/s    rkB/s    wkB/s avgrq-sz avgqu-sz   await r_await w_await  svctm  %util
xvda              0.00     0.00    0.00    5.00     0.00    49.50    19.80     0.00    0.60    0.00    0.60   0.60   0.30
dm-0              0.00     0.00    0.00    3.00     0.00    49.50    33.00     0.00    1.00    0.00    1.00   1.00   0.30

avg-cpu:  %user   %nice %system %iowait  %steal   %idle
           2.01    0.00    7.54    0.00    0.00   90.45

Device:         rrqm/s   wrqm/s     r/s     w/s    rkB/s    wkB/s avgrq-sz avgqu-sz   await r_await w_await  svctm  %util
xvda              0.00     0.00    0.00   33.00     0.00   208.00    12.61     0.08    2.42    0.00    2.42   0.12   0.40
dm-3              0.00     0.00    0.00   33.00     0.00   208.00    12.61     0.08    2.45    0.00    2.45   0.12   0.40
```

Hold on to your seats!

iostats part 1

util%

- % of time spent handling >= 1 request; not great for SSD/RAID
  - Ideal: 0-10%
  - Good: 10-50%, with spikes higher, but returning to “normal” within 1-2 seconds
  - Bad: 50-100%, with sustained at 100% for many seconds at a time before dropping, returning to 100 within a few seconds consistently

iostats part deux

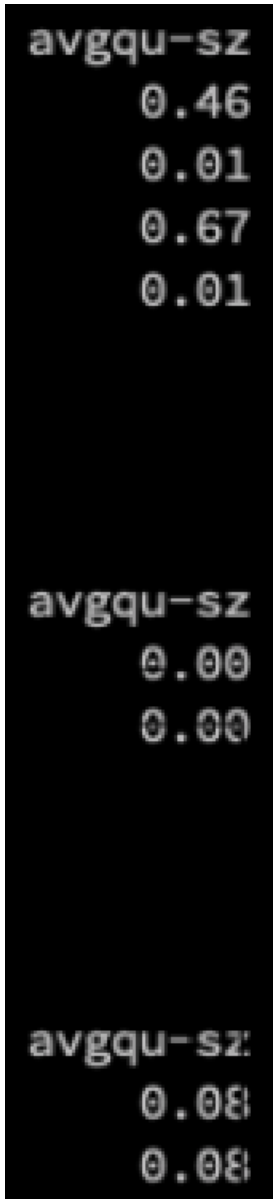
r_await & w_await

```
r_await w_await
   0.00    8.48
   0.00    1.00
   0.00    8.41
   0.00    1.00



r_await w_await
   0.00    0.60
   0.00    1.00



r_await w_await
   0.00    2.42
   0.00    2.45
```

- Disk wait times for read (“r_await”) and write (“w_await”): average time from when a request is put in the queue to when the request is completed
  - Ideal: 0-5 ms
  - Good: 5-15 ms, with spikes higher, but returning to “normal” within 1-2 seconds
  - Bad: 15-300+ ms, higher values sustained for a short number of seconds, and then dropping down, but rarely seeing single digits

iostats part III

avgqu_sz



- Disk wait queues issued to the device (“avgqu-sz”)
  - Ideal: 0-5
  - Good: 5-15, with spikes higher, but returning to “normal” within 1-2 seconds
  - Bad: 15-300+, higher values sustained for a short number of seconds, and then dropping down, but rarely seeing single digits

Note: avgrq-sz is not the same as avgqu-sz, and not in scope here

Resist

- Ideal vs. good vs. bad results are generally easy to discern per environment over time

- All indexers may not be created equal, resist assuming so
  - Compare ingestion rates, indexing queue fill rates, streaming bucket replication and search bundle replication errors: help find better/worse hardware configs
  - May need to look beyond Splunk: Why is storage performing better on some than others

Resist the easy comforts of complacency.

What Is SmartStore?

1. Separates compute from storage
2. Retention != disks attached to indexers anymore; (NFS = Not For Splunk)
3. Designed for Cloud, on prem friendly
4. Storage cost containment
5. Great for long tail retention

- Index data reads based primarily on search w/ smattering of replication
  - Backups
- Writes based on ingestion and replication
- Predictable patterns
- Outage/Cluster recovery drives spikes

AKA Before SmartStore

SmartStore con't

- Decoupled compute from storage = disks are Someone Else's Problem, right? Nope!
- One filesystem to rule them all
- Hot & cached buckets coexist
- Ingestion + replication + downloads from S3 = Writes now highly bursty
  - Up to 8 concurrent bucket downloads per indexer

More SmartStore

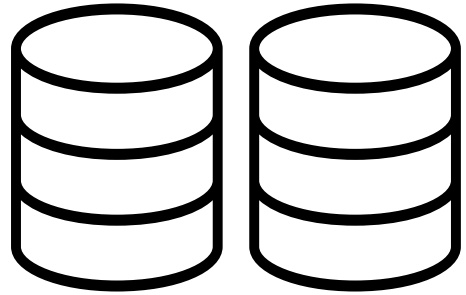
- Min 100MB/sec (1Gb), performant 800MB/s (10Gb) connectivity
  - Concurrently between every indexer and object store
- Indexer RAID Options
  - RAID0/RAID10 can better absorb tsunami of writes without parity write penalty
  - RAID5/RAID6 are creaky, painful write penalty*
- Default SF/RF=3 mitigates disk failure concerns for RAID0
  - SF=RF in SmartStore

*https://blog.storagecraft.com/raid-performance/

Q & … maybe A?

Local disk

- Are the disks installed those the customer thinks they have?*
  - Model numbers -> Disk Speed (5400 RPM vs 15k RPM vs SSD) -> Google -> IOPS
- If spinning rust, are there enough spindles?
- What RAID config is in place?
  - Parity drive failure rebuild stress remaining spindles; disks likely same age
- If using hardware RAID, are there vendor advisories for performance issues with controller firmware?
- Are host, RAID controller and other firmware versions consistent across all indexers?
  - Are they consistent between good and bad indexers?
- Are iLO/DRAC syslogs being splunked for disk messages?

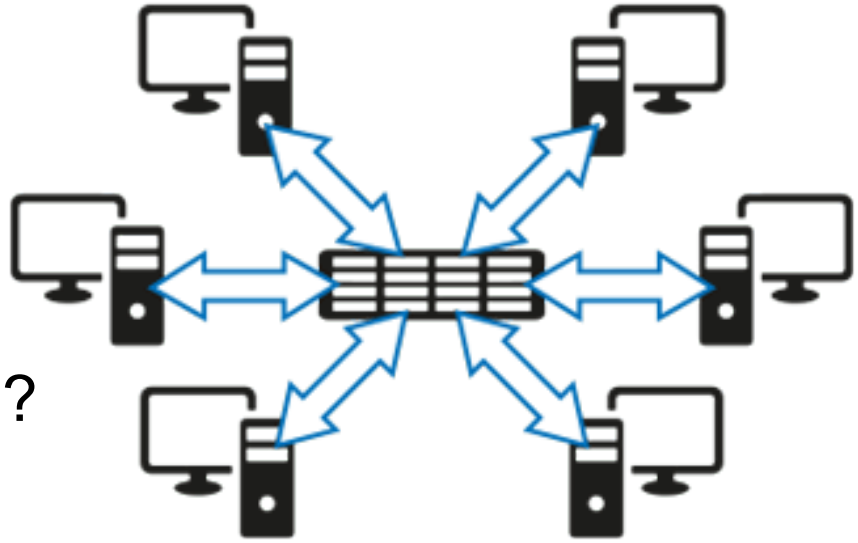


* Indexer Bad Day: 4*7200 RPM in RAID6 = 150 IOPS usable

More Q & A

Shared storage or Virtualized

- Is there enough bandwidth (bytes, packets per second) available all the way to the disk? Do the math.
- Individual SAN ports in the path saturated?
- SAN ports logging errors (ex: buffer exhaustion)?
- Do the disk arrays (trays) have capacity to provide IOPS to indexers mounting LUNs carved from those disks? RAID# at SAN important
- Are the disk array controllers having caching issues?
  - Are there enough controllers?
- Is data pinned to a non-performant tier in the SAN?
- Are there vendor advisories for performance issues with firmware or BIOS settings on storage components?



3 Takeaways

- Performant storage is foundational for Splunk success
  - May have been that way at time of deployment, is it still that way today?
- SmartStore changes the story
  - Fast networks, changing write patterns, creaky storage, oh my!
- Ask questions in your journey to environmental consistency
  - You don’t have to be in this alone

Call To Action

Add MC storage review to regular performance monitoring of your environment.

Daily. Weekly. Monthly.

Be on the lookout for new trends.

This isn’t enough, need more

- Original whitepaper: https://github.com/dpaper-splunk/public/

- Relevant Splunk Community Slack channels
  - #admin
  - #architecture
  - #smartstore
- tags on answers
  - storage
  - architecture
  - administration
  - performance
  - disk

Other Community Talks

- **FNC2051 - I Deleted a Critical Knowledge Object…. Now What? Steve McMaster, Hurricane Labs**
  - **Tues (today!) 4:15pm**
- **FNC2751 - Master joining your datasets without using join. How to build amazing reports across multiple datasets without sacrificing performance. Nick Mealy, Sideview LLC**
  - **Thu 10:30am**
- **FNC2259 - 5 Tips to Better Support Case Resolution. Cary Petterborg, Stage 2 Security**
  - **Thu 11:45am**
- **FNC1549 - Administrators Anonymous: Splunk best practices (and useful tricks) I learned the hard way. Tom Kopchak, Hurrican Labs**
  - **Thu 1:00pm**

Final info

- Find me
  - dpaper@splunk.com
  - @cerby on Splunk usergroups Slack
  - http://splk.it/slack

- Thank you

You could have picked any breakout, and you chose mine. I've appreciated the opportunity to share this info with you, and maybe, just maybe, you won't be the others.

- Please rate me in the .conf19 app!

Q&A

.conf19
splunk>

splunk> .conf19