



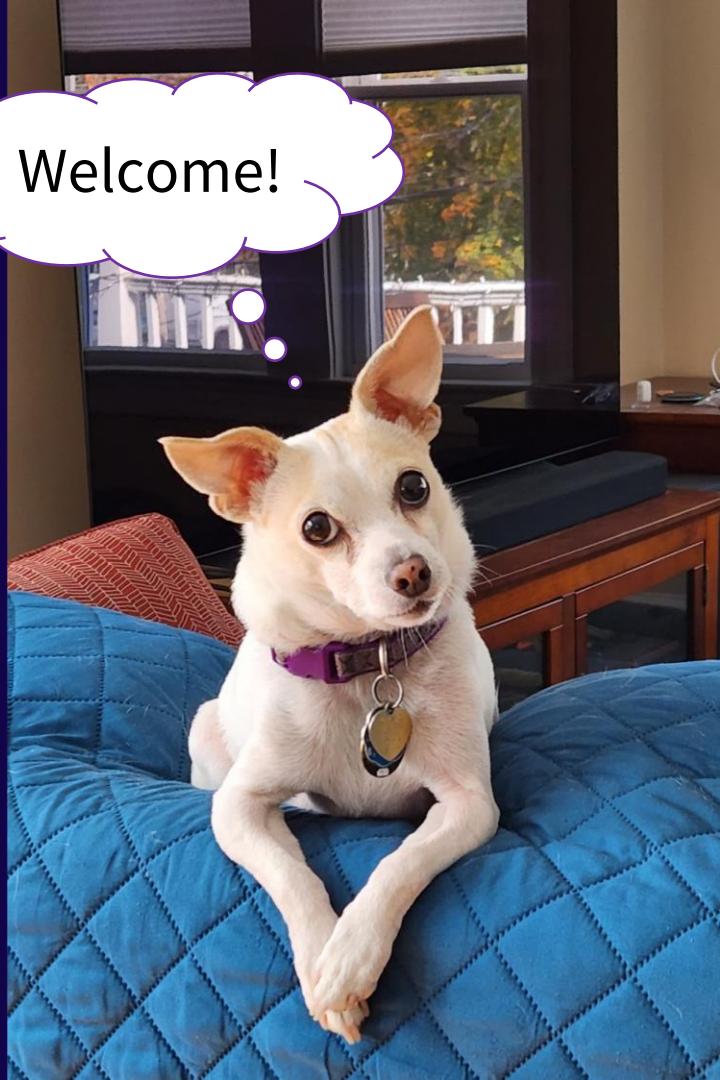
A Practical Deep Dive into I/O for the T-SQL Performance Tuner

Andy Yun

He/him

Field Solutions Architect
Pure Storage

Welcome!





Andy Yun

Field Solution Architect

- SQL Server DBA & DB Dev
- Microsoft Certified Master Server



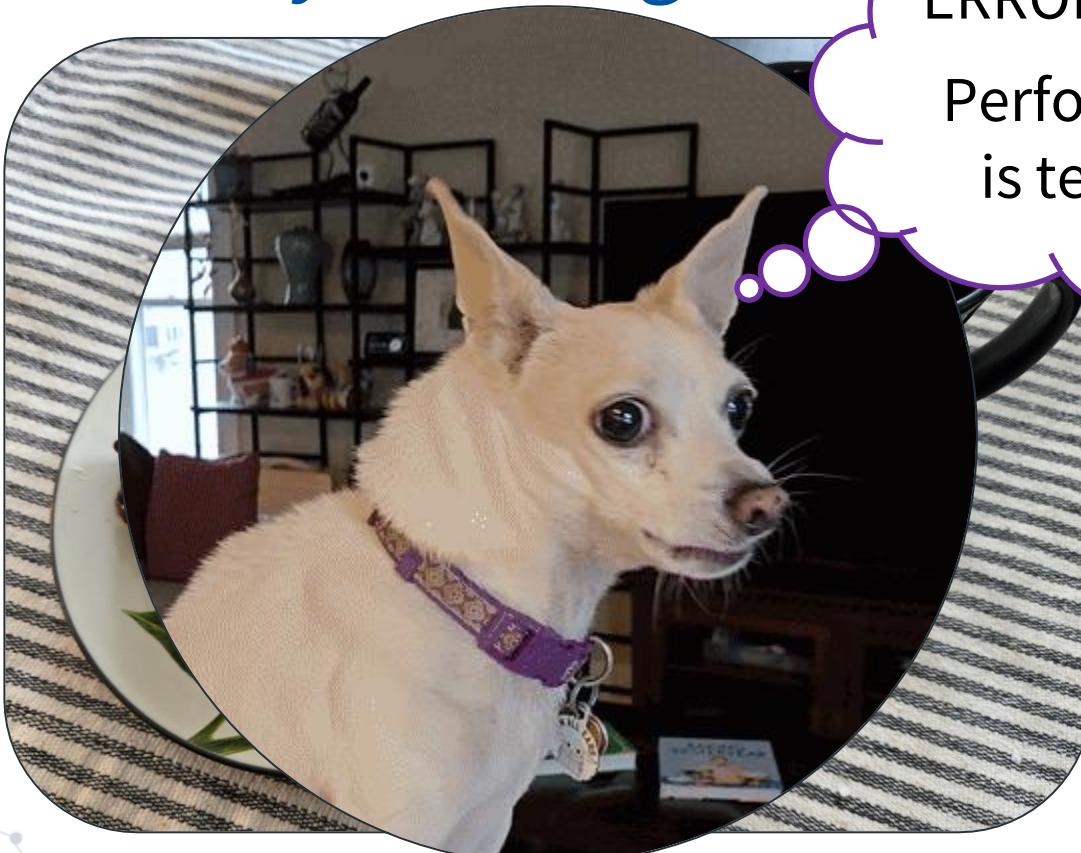
group

on – Director-at-Large



sqlbek@gmail.com
<https://sqlbek.wordpress.com/>
<https://www.github.com/sqlbek/>

Quiet Monday Morning



ERROR ALERT!

Performance
is terrible!

What's Wrong?

Log file summary: No filter applied

Date ▾	Source	Message
10/2/2024 7:06:51 PM	spid41s	SQL Server has encountered 231 occurrence(s) of I/O requests taking longer than 15 seconds
10/2/2024 7:01:38 PM	spid41s	SQL Server has encountered 353 occurrence(s) of I/O requests taking longer than 15 seconds
10/2/2024 6:56:35 PM	spid41s	SQL Server has encountered 702 occurrence(s) of I/O requests taking longer than 15 seconds
10/2/2024 6:51:32 PM	spid41s	SQL Server has encountered 14 occurrence(s) of I/O requests taking longer than 15 seconds
10/2/2024 6:45:43 PM	spid41s	SQL Server has encountered 216 occurrence(s) of I/O requests taking longer than 15 seconds
10/2/2024 6:32:11 PM	spid41s	SQL Server has encountered 240 occurrence(s) of I/O requests taking longer than 15 seconds

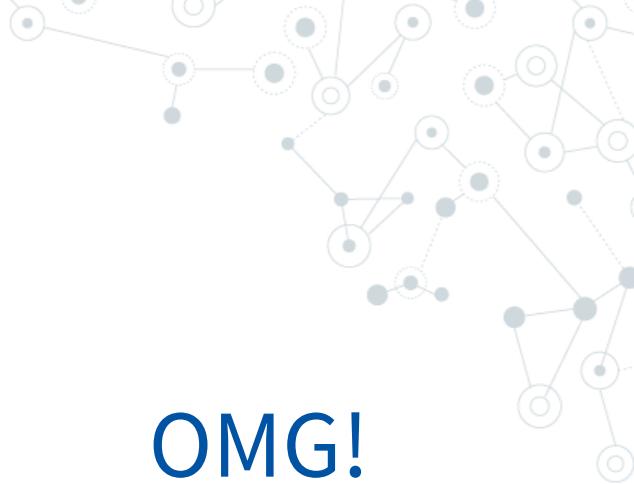
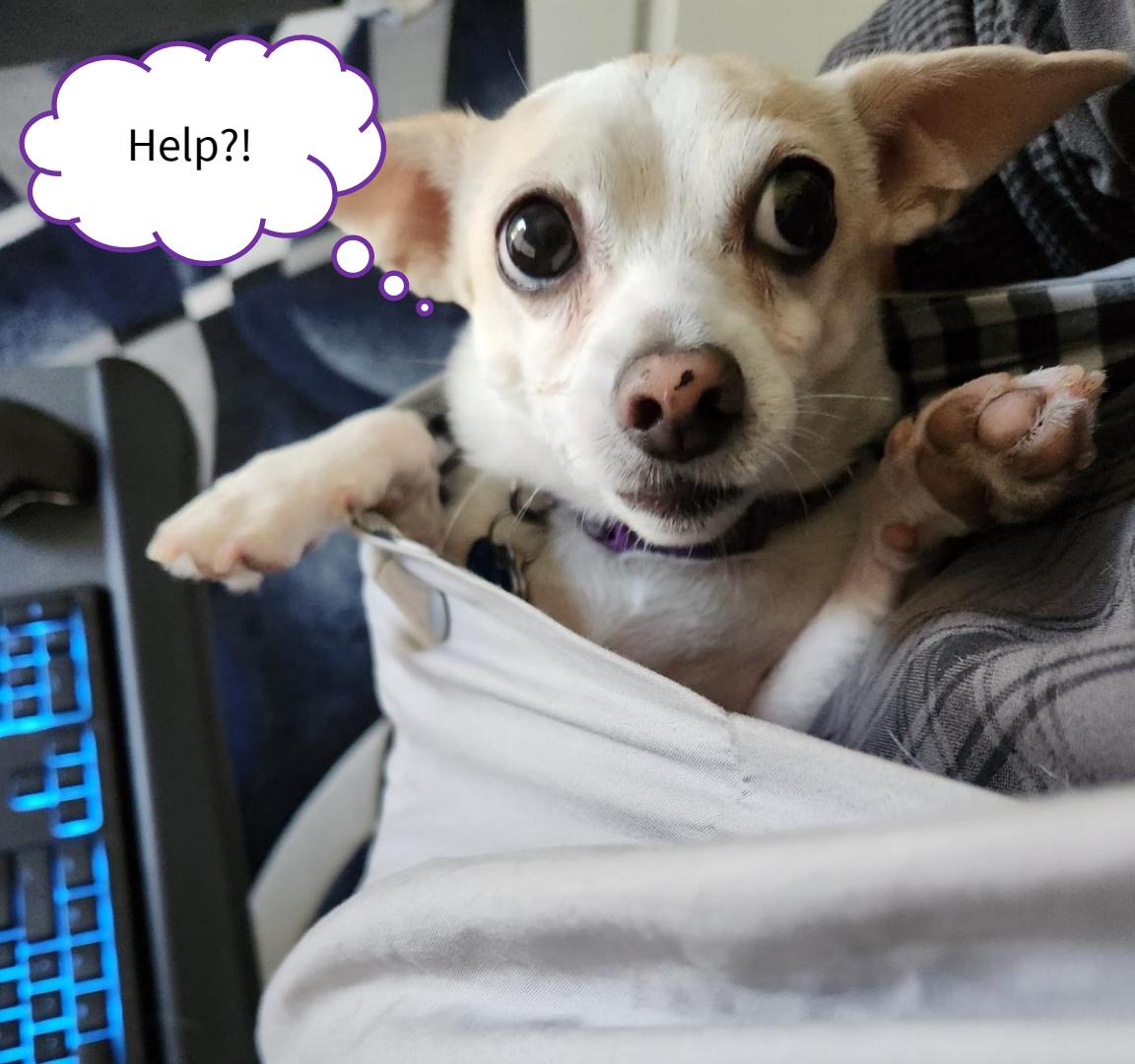
Selected row details:

Date 10/2/2024 6:56:35 PM
Log SQL Server (Current - 10/2/2024 7:01:00 PM)

Source spid41s

Message

SQL Server has encountered 702 occurrence(s) of I/O requests taking longer than 15 seconds to complete on file [M:\Data\CookbookDemo_iSCSI.mdf] in database id 12. The OS file handle is 0x0000000000000015C0. The offset of the latest long I/O is: 0x00000001c3a6000. The duration of the long I/O is: 20021 ms.



I/O is Expansive...

Practical Deep Dive

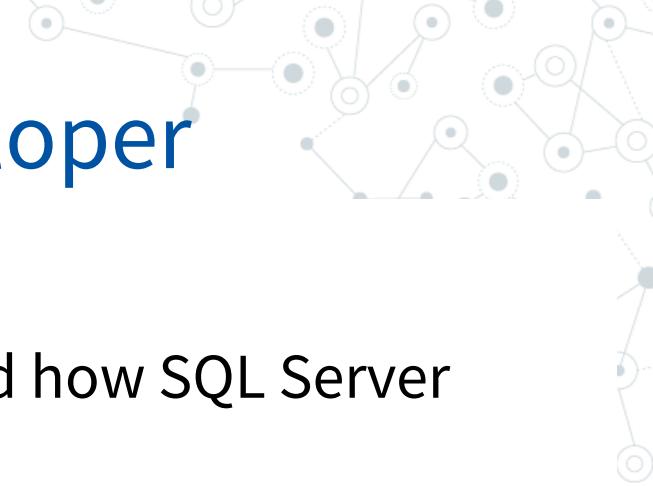
into I/O for the

T-SQL Performance Tuner

T-SQL Developer?



I'm Primarily a T-SQL Developer



“Why should hardware staff be Performance Tuners:

hardware sta

- still have to understand how SQL Server works
- often have to help troubleshoot production
- often have to explain to non-DBAs how SQL Server works
- can more effectively communicate with those who DO deal with hardware



Knowledge Progression

◎ T-SQL Developer

You do not have to
know everything...

◎ T-SQL

But it is ideal
to be FAMILIAR
with many things

◎ SQL



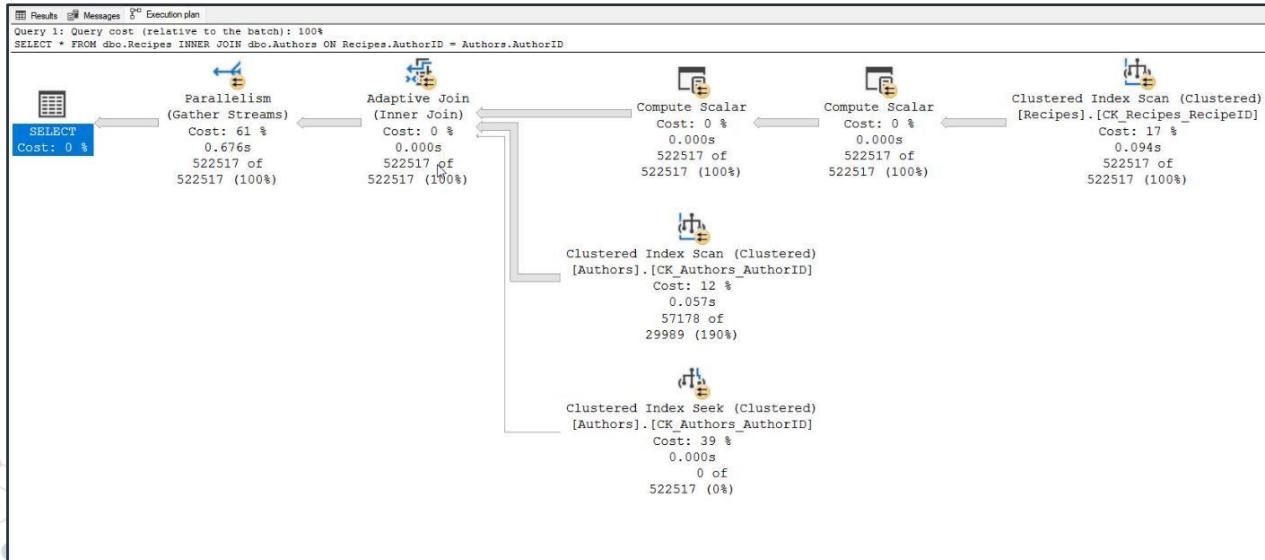
Foundations First

All Starts with a T-SQL Query

```
SELECT *
FROM dbo.Recipes
INNER JOIN dbo.Authors
ON Recipes.AuthorID = Authors.AuthorID
```



Query
Optimizer



TOP CHEF DINNER

MAY 18TH, 2024 | THE WESTBROOK INN

FRIED RED SNAPPER

BANH HOI NOODLES | VIETNAMESE HERBS | SESAME LEAVES | SCALLION OIL | HABANERO NUOC MAM

Jannie Tran | The Black Sheep | Las Vegas, NV

MOLE DE CALACUADE

CHICKEN | TORTILLA | PICKLED ONIONS | SALSA TATEMADA

Marie Mason | Boca | Tucson, AZ

THAI CHAAT

CRISPY RICE | BARBECUE PRAWN & EGGPLANT BHARTHA | SMOKED LAMB RIB

Avising Tharua, Agni & Joyas | Columbus, OH

PIRI PIRI SHORT RIB EN ESCABOLADO

MESQUITE SMOKED PLANTAIN | JOLLOF RICE POWDER | SPRING ONION 'SALTADO' | PICKLED CHARRED OKRA | WILD FENNEL AVOCADO CREAM | GUAVA MISO GLAZE

Nelson German | Alamar & Bobie Mesa | Oakland, CA

SWEET PLANTAIN PAIN PERDU

DIRI AK LET | DARK COCOA DIJON DIJON FEUILLETINE | SWEET PLANTAIN CREMEUX | TABLET NWA

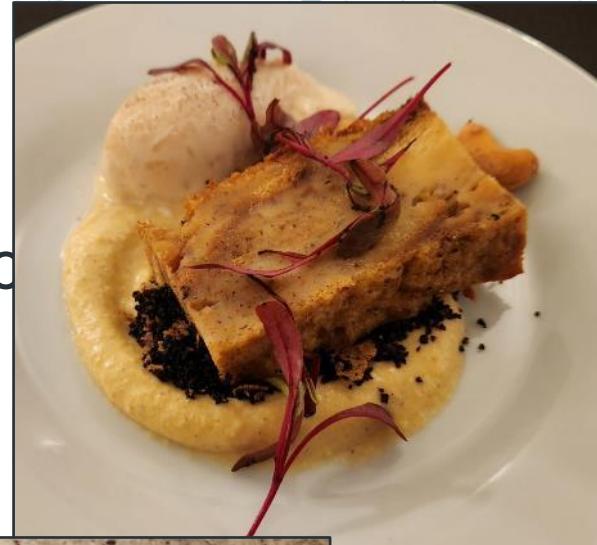
Curtis Vieud | Anseanu, Greenleaf, Pavilion | NH

DEBORAH MELKIN

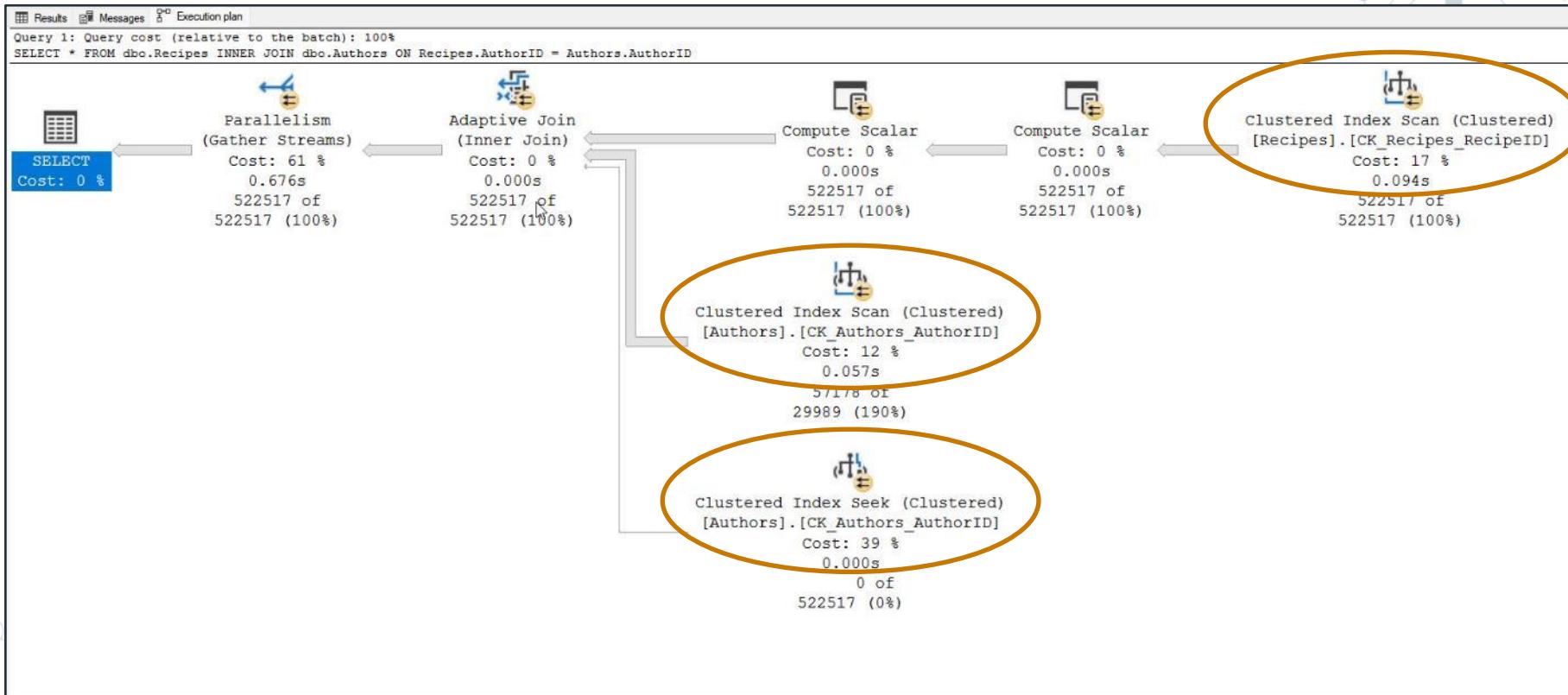
I'm stuck.
Help!

It's Declarative

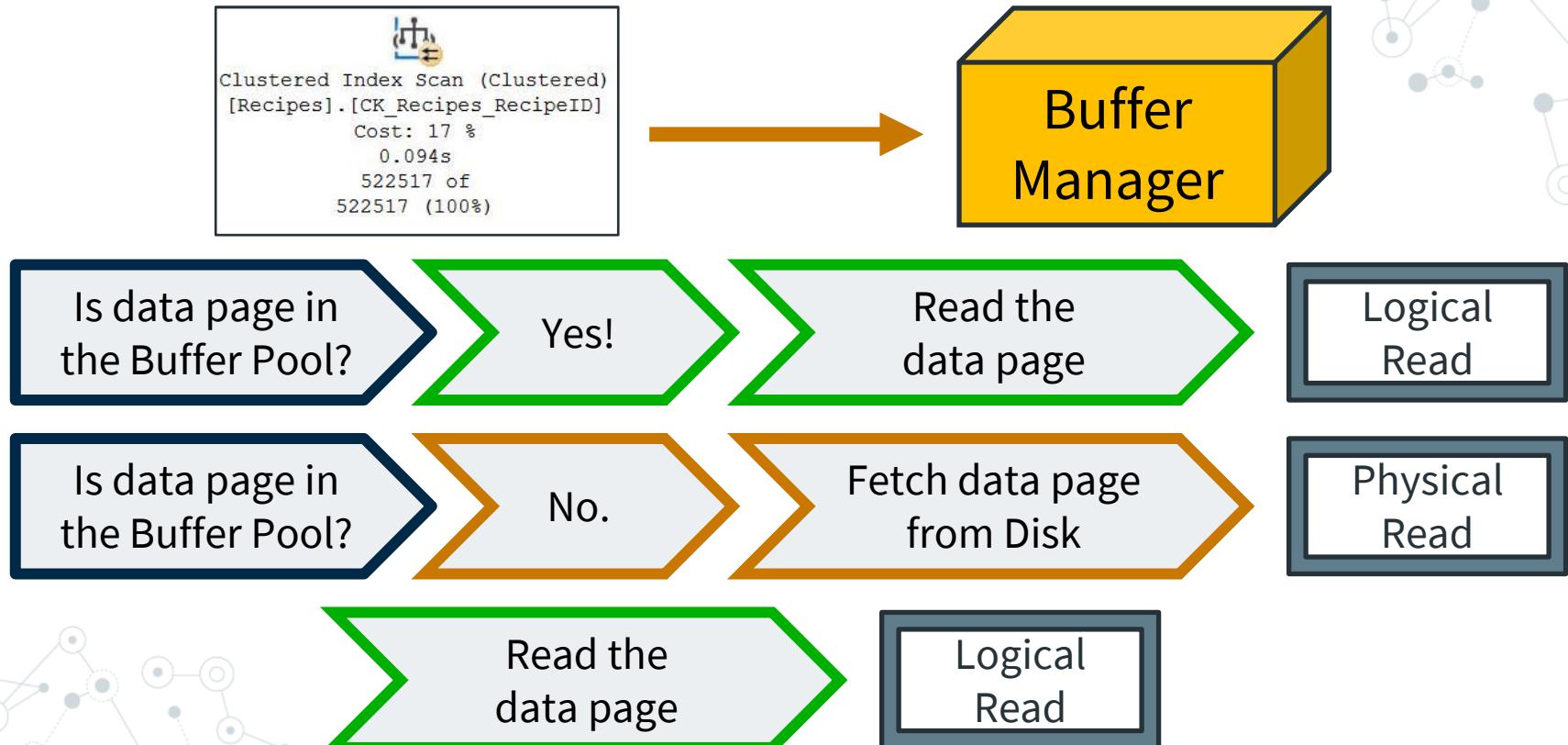
... tell me WHAT you want... but no



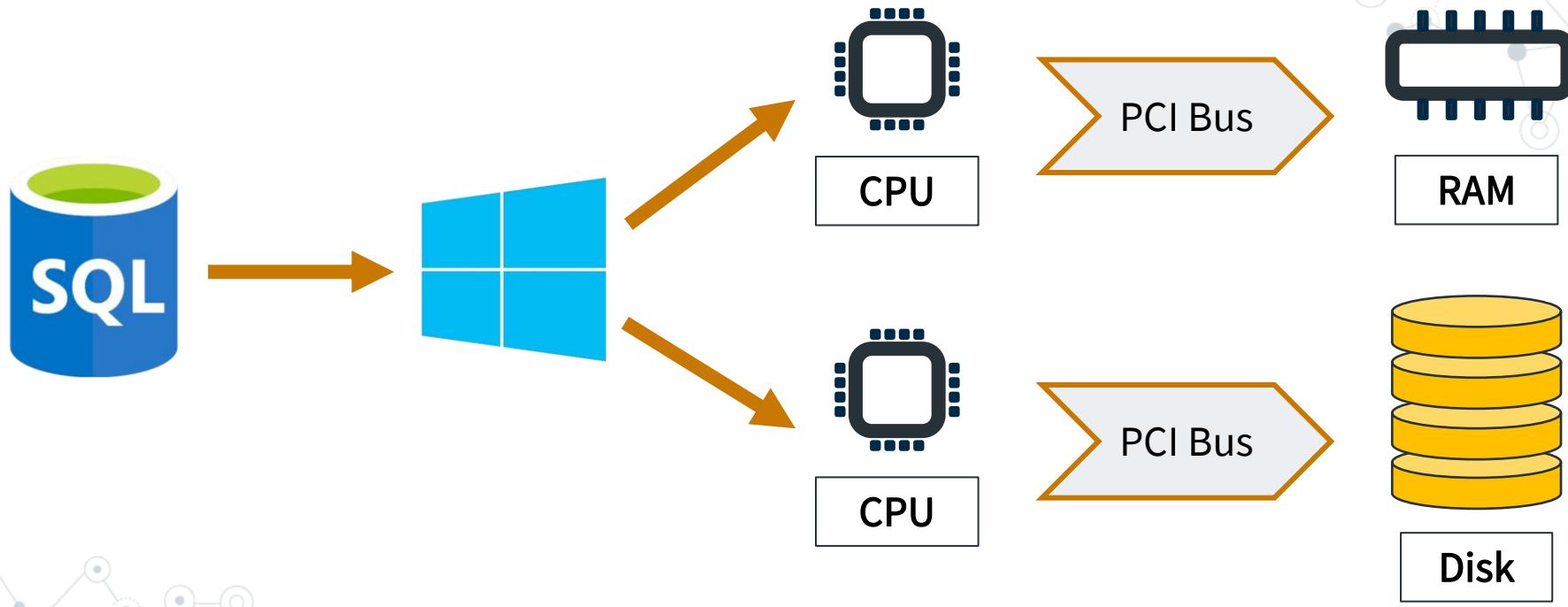
Access Methods in Execution Plan



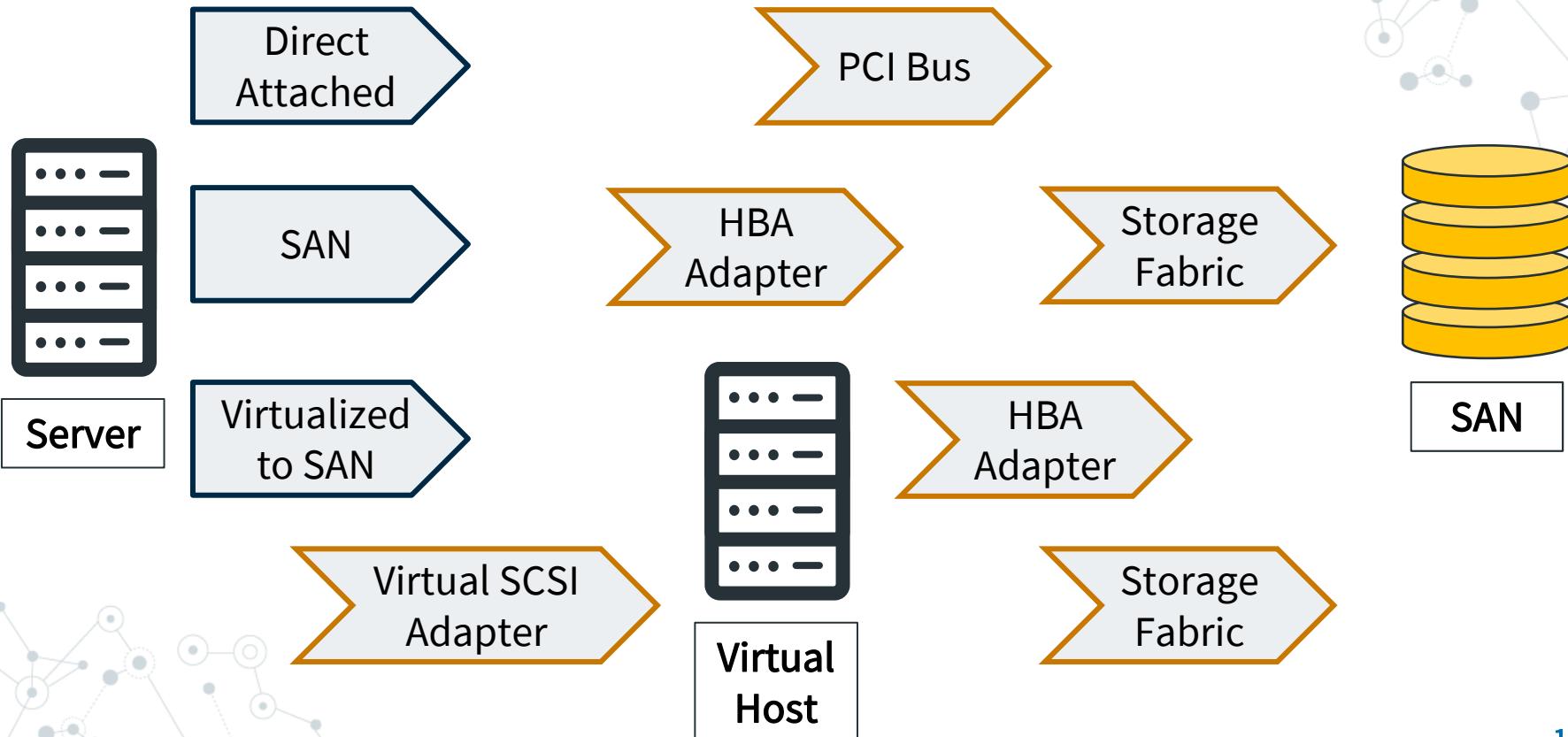
Execution Plan Workflow



Who Dispatches I/O?



Physical Pathway Refresher



Remember

Disk Reading & Writing
has many layers

Access from memory
always faster

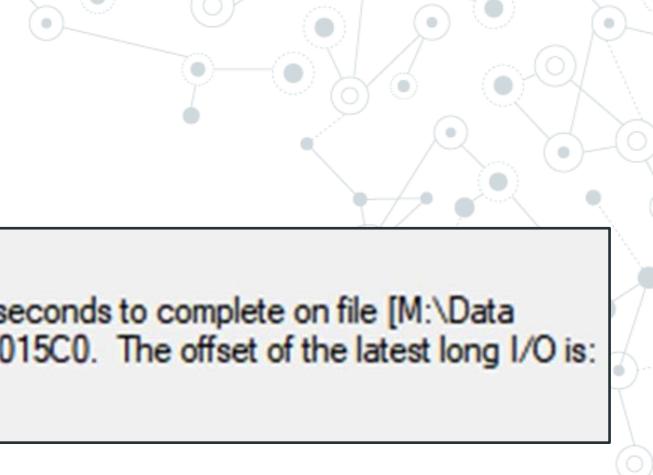


*“I Like Running
Fast Too!”*

Back to the Story



Error 833



Message

SQL Server has encountered 702 occurrence(s) of I/O requests taking longer than 15 seconds to complete on file [M:\Data\CookbookDemo_iSCSI.mdf] in database id 12. The OS file handle is 0x00000000000015C0. The offset of the latest long I/O is: 0x0000001c3a6000. The duration of the long I/O is: 20021 ms.

Microsoft Learn:

... SQL Server has issued a read or write request
to/from disk

... the request has taken longer than 15 seconds
to return.



How is Error 833 even tracked?

- ◎ Has ANY I/O been pending +15 seconds
- ◎ Recorded every 5 seconds
- ◎ Reported every 5 minutes



Aren't there
coverage gaps
here?

Before Digging In...

- ◎ Ask questions
- ◎ Has this ALWAYS been happening?
 - Since when?
 - For how long?
- ◎ Remember:
 - To be mindful of the entire stack
 - To ask EVERYONE

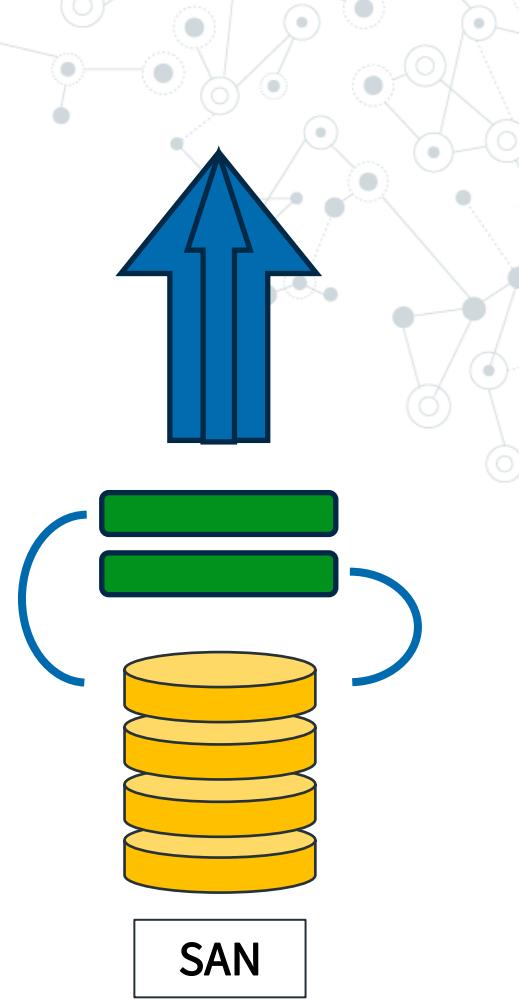
Remember when
that “top of some
thingy” happened?



A Tale from the T

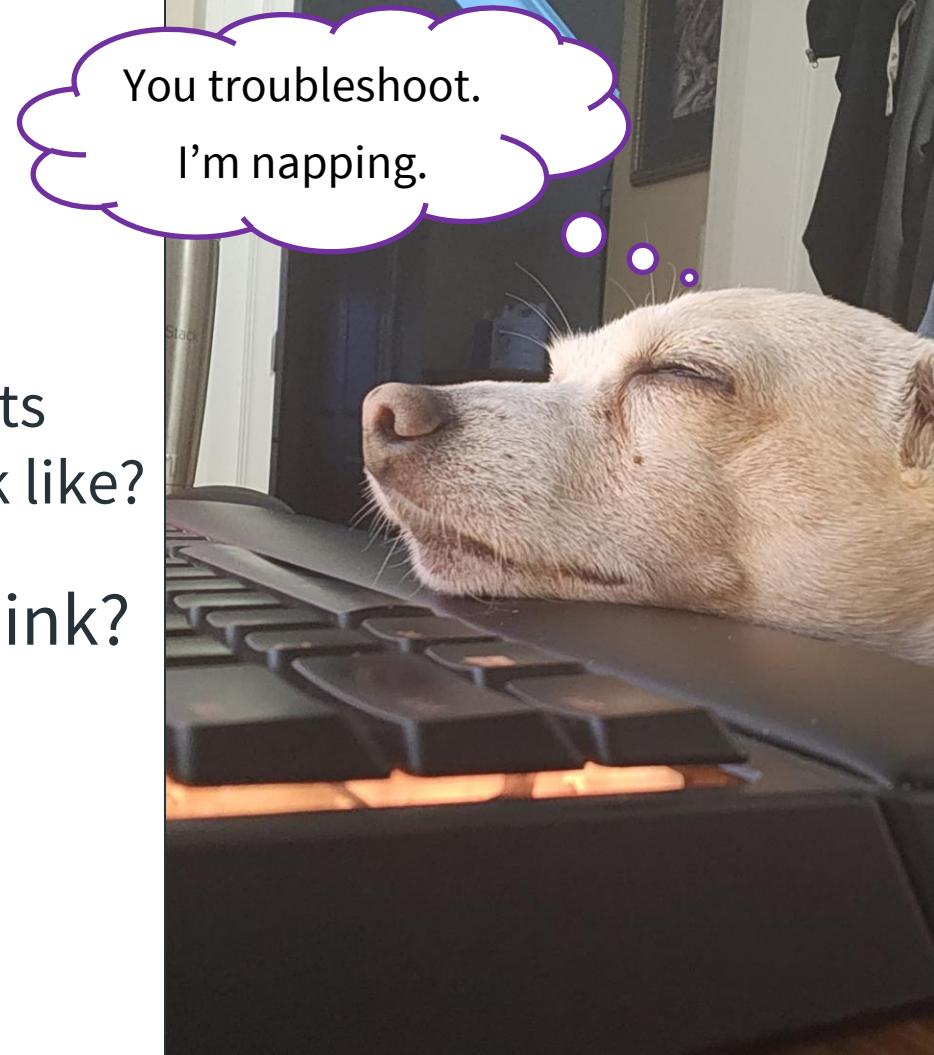
- ◎ Upgraded storage
- ◎ Escalation team:
 - Database Admin
 - Windows Admin
 - Virtualization Admin
 - Storage Admin
- ◎ Key Detail: Storage Fabric changed w. storage upgrade
iSCSI -> Ethernet!

Anyone talk to the Network team?



From the Top

- ◎ I/O Wait Types
- ◎ sys.dm_io_virtual_file_stats
 - What does latency look like?
- ◎ What does Windows think?
 - PerfMon – latency?



Advanc

Administrator: Command Prompt

C:\Windows\system32>FL

Filter

For more information about the study, contact Dr. Michael J. Hwang at (319) 356-4000 or email at mjhwang@uiowa.edu.

WdFilter

WdFilter

IndEiltor

WDFilter

110

100

psychiatry

C:\Windows\system32>

C:\Windows\System32>FLTC Instances	Filter	Volume Name	Altitude	Instance Name	Frame	SprtFtrs	VIStatus
CldFlt	D:		180451	CldFlt	0	0000000f	
CldFlt	C:		180451	CldFlt	0	0000000f	
CldFlt	\Device\HarddiskVolumeShadowCopy5		180451	CldFlt	0	0000000f	
CldFlt	\Device\HarddiskVolumeShadowCopy6		180451	CldFlt	0	0000000f	
FileCrypt	D:		141100	FileCrypt Instance	0	0000000f	
FileCrypt	\Device\HarddiskVolumeShadowCopy6		141100	FileCrypt Instance	0	0000000f	
FileInfo	D:		40500	FileInfo	0	0000000f	
FileInfo	C:		40500	FileInfo	0	0000000f	
FileInfo	G:		40500	FileInfo	0	0000000f	
FileInfo	\Device\HarddiskVolumeShadowCopy5		40500	FileInfo	0	0000000f	
FileInfo	\Device\HarddiskVolumeShadowCopy6		40500	FileInfo	0	0000000f	
FileInfo	\Device\Mup		40500	FileInfo	0	0000000f	
SFPMonitor	C:		385100	SFPMonitor - Top Instance	0	0000000f	
UCPD	D:		385250.5	UCPD - Top Instance	0	0000000f	
UCPD	C:		385250.5	UCPD - Top Instance	0	0000000f	
UCPD	\Device\HarddiskVolumeShadowCopy5		385250.5	UCPD - Top Instance	0	0000000f	
UCPD	\Device\HarddiskVolumeShadowCopy6		385250.5	UCPD - Top Instance	0	0000000f	
UCPD	\Device\Mup		385250.5	UCPD - Top Instance	0	0000000f	
UCPD	G:		385250.5	UCPD - Top Instance	0	0000000f	
WdfFilter	D:		328010	WdfFilter Instance	0	0000000f	
WdfFilter	C:		328010	WdfFilter Instance	0	0000000f	
WdfFilter	\Device\HarddiskVolumeShadowCopy5		328010	WdfFilter Instance	0	0000000f	
WdfFilter	\Device\HarddiskVolumeShadowCopy6		328010	WdfFilter Instance	0	0000000f	
WdfFilter	\Device\Mup		328010	WdfFilter Instance	0	0000000f	
WdfFilter	G:		328010	WdfFilter Instance	0	0000000f	
Wof	D:		40700	Wof Instance	0	0000000f	
Wof	C:		40700	Wof Instance	0	0000000f	
Wof	\Device\HarddiskVolumeShadowCopy5		40700	Wof Instance	0	0000000f	
Wof	\Device\HarddiskVolumeShadowCopy6		40700	Wof Instance	0	0000000f	
bfs	D:		150000	bfs	0	0000000f	
bfs	C:		150000	bfs	0	0000000f	
bfs	\Device\HarddiskVolumeShadowCopy5		150000	bfs	0	0000000f	
bfs	\Device\HarddiskVolumeShadowCopy6		150000	bfs	0	0000000f	
bfs	\Device\Mailslot		150000	bfs	0	0000000f	
bfs	\Device\Mup		150000	bfs	0	0000000f	
bfs	\Device\NamedPipe		150000	bfs	0	0000000f	
bfs	G:		150000	bfs	0	0000000f	
bindflt	C:		409800	bindflt Instance	0	0000000f	
gameflt	D:		189850	gameflt Instance	0	0000000b	
gameflt	C:		189850	gameflt Instance	0	0000000b	
gameflt	\Device\NamedPipe		189850	gameflt Instance	0	0000000b	
luafv	C:		135000	luafv	0	0000000f	
npsvctrig	\Device\NamedPipe		46000	npsvctrig	0	00000008	
wcifis	D:		189900	wcifis Instance	0	0000000f	
wcifis	C:		189900	wcifis Instance	0	0000000f	

Instance Name

WdFilter Instance

WdFilter Instance

IdFilter Instance

WdFilter Instance

Leaf Instances

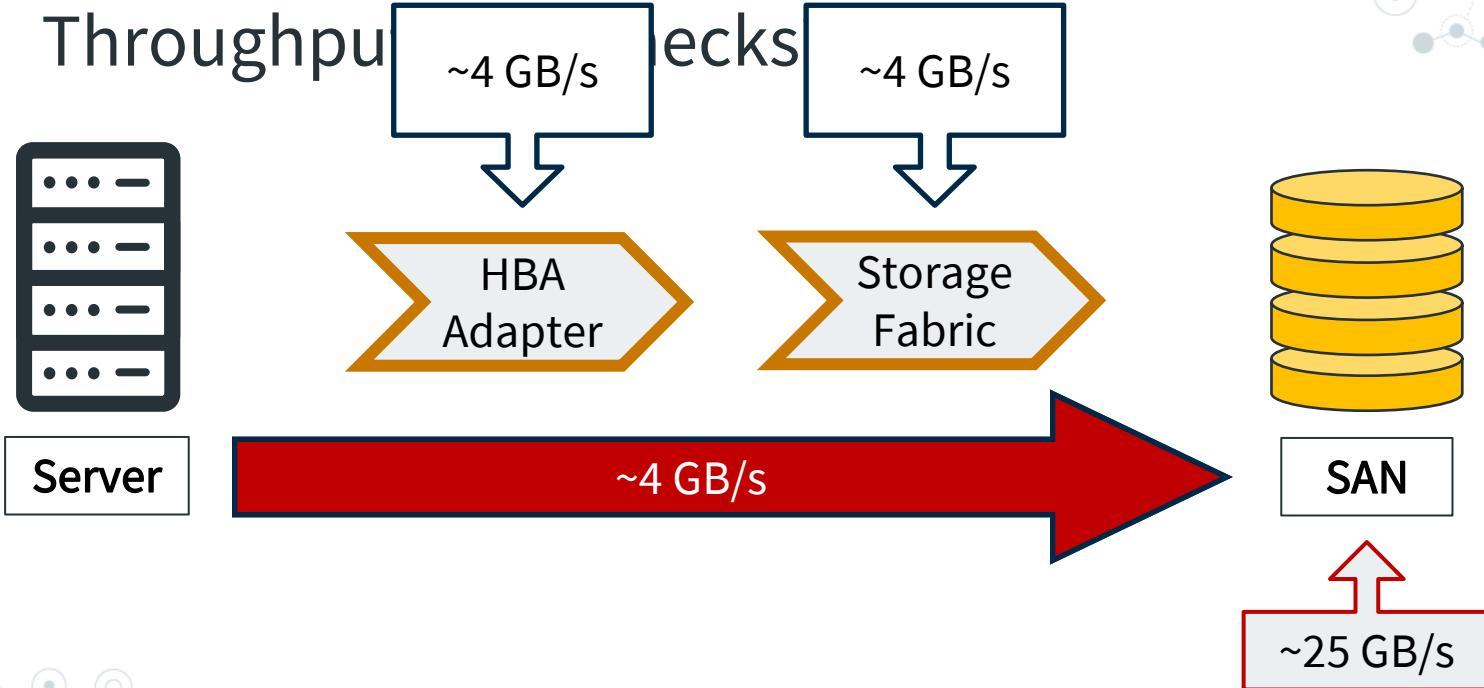
ANSWER

1000



Saturating Your Storage?

- Throughput bottlenecks



Monitor Your Stack

- ◎ Know how to query your queries
- ◎ “Story: the array gets crushed at 5:17AM...”



Queries Don't Get Along

- ◎ Competing

- CPU & I/O

Hey stop,
it's my dinner time!





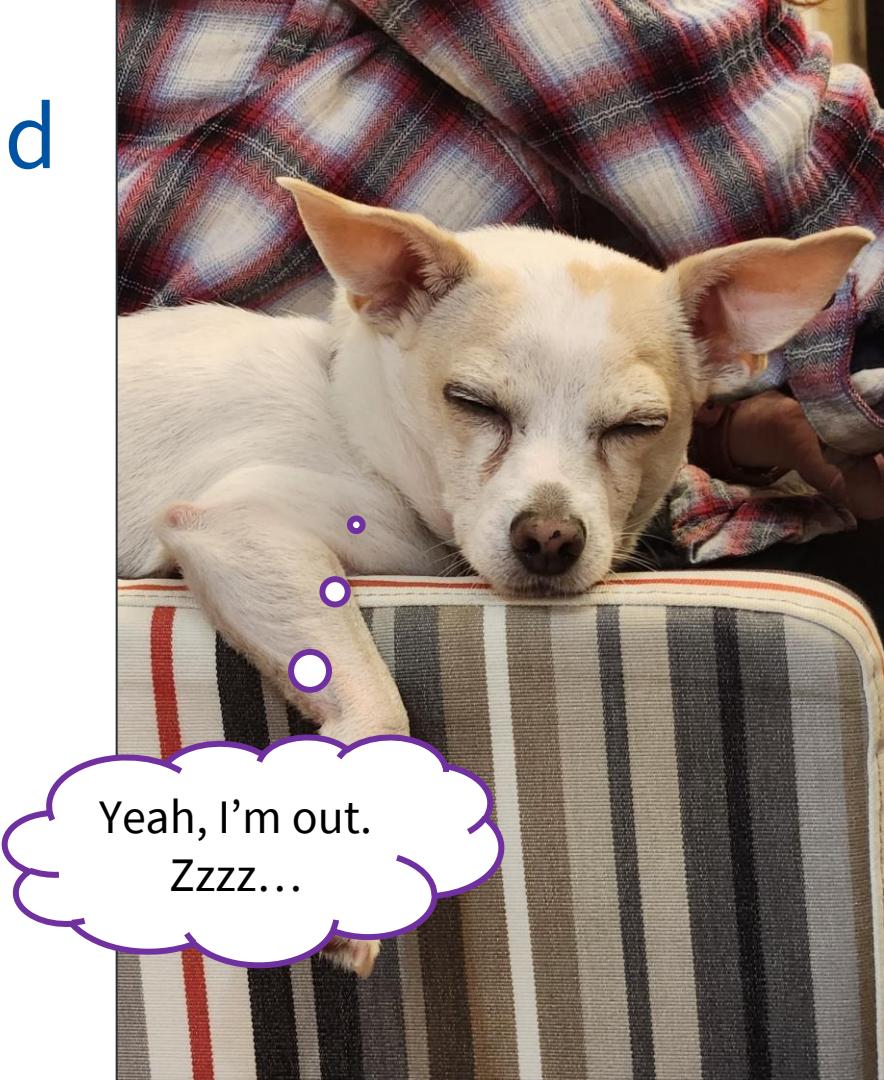
CPU & RAM are Limited

- ◎ CPU orchestrates I/O
- ◎ Memory Grants
- ◎ Spills to TempDB



Fine, It's Not Workload

- ◎ Analyze I/O Path
- ◎ MS Docs: In-Depth page dedicated to Error 833
 - Term “I/O Path” mentioned 11 times!!!



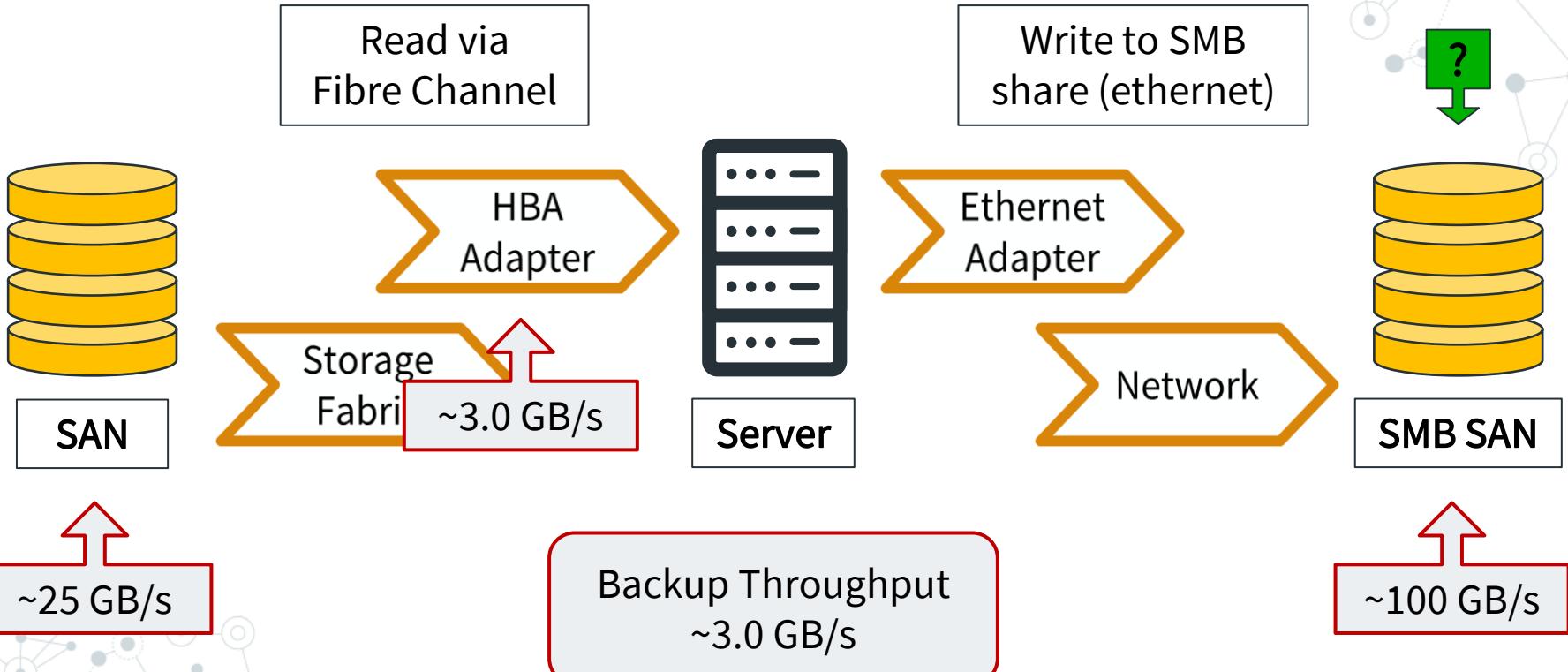
I/O Pathways



- ◎ What is each “element” in pathway capable of?
- ◎ Max possible Throughput?
- ◎ Max possible IOPs?
- ◎ Can you capture/monitor this?



A Tale about Backups



U N
N G!

What gives?!? We spent a ton of money on SUPER FAST STORAGE!!! Why's performance tanking?!?!

FLASH (STORAGE) ARRAY

Monday...

ALERT!!!

I shared some of my treat with you



An Introduction to Queuing

- ◎ Queue contains I/O requests to be processed
- ◎ Queue Depth (Limit) – Max # of I/Os that can be processed simultaneously
- ◎ Queue Length - # of I/Os currently being processed
 - A.k.a. - Pending I/O or I/O being serviced

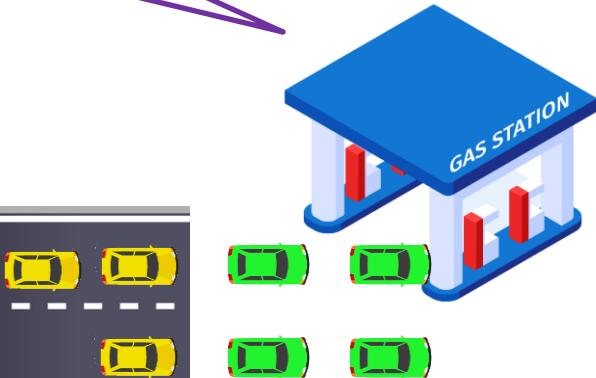
What I Found

Queue is a Line, r

You can't wait here!
Go home and come
back later.

...

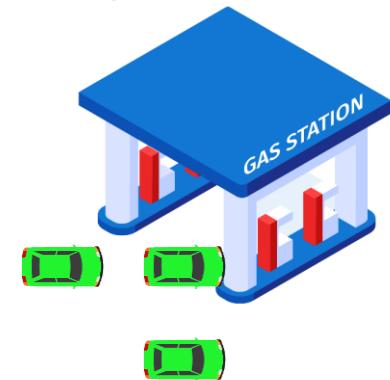
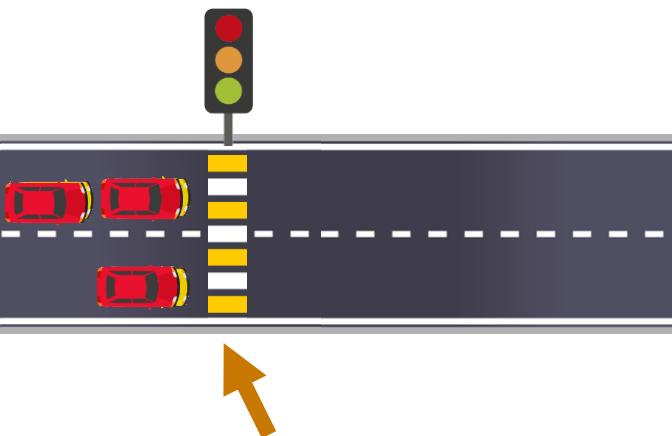
Four (4) Fuel Pumps



Current
Queue Length?

HBA
Queue Depth = 3

Four (4) Fuel Pumps



Current
Queue Length?

Queue Depth

- ◎ “Traffic signals” that buffer the flow of I/O
 - Queues help to not overwhelm lousy legacy storage
- ◎ But what if you have modern, cutting-edge **FAST** storage?
 - Do default values defined “long ago” still make sense?

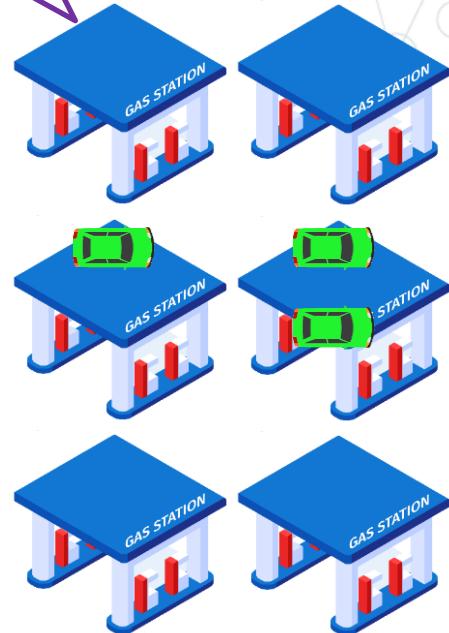
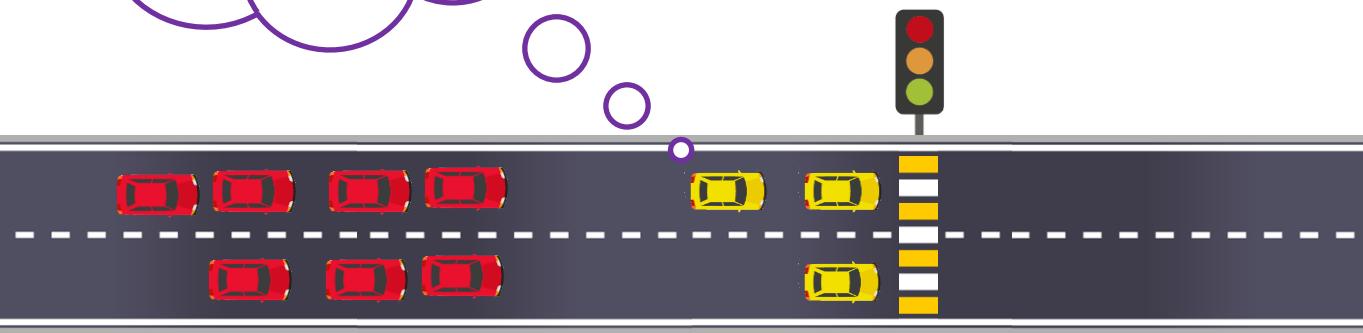
(*cough* cost threshold for parallelism *cough*)

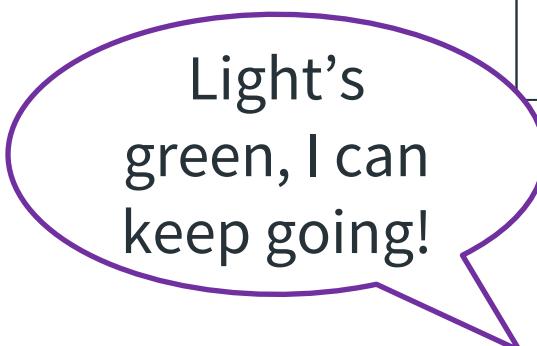
Why hasn't anyone changed the light timings?

HBA
Q.D = 3

Where is everyone?

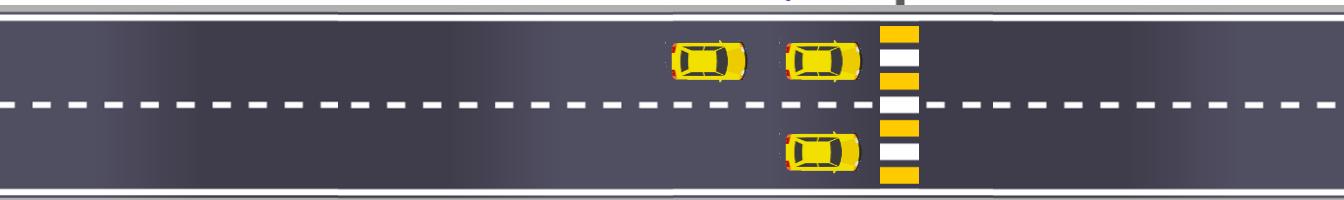
6x gas stations
24 Fuel Pumps



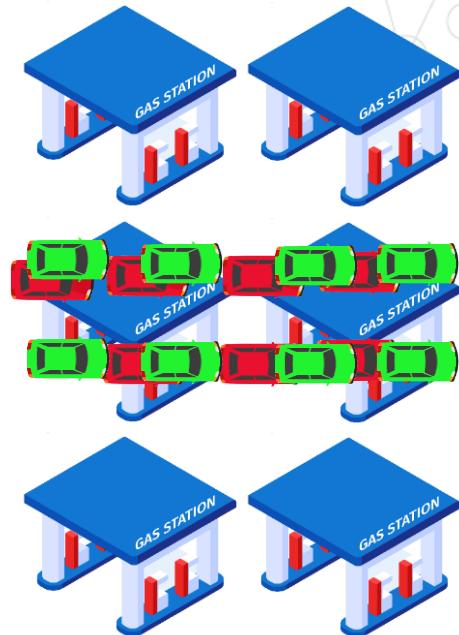


Light's
green, I can
keep going!

HBA
Q.D = 24



6x gas stations
24 Fuel Pumps



Queuing Considerations - VMware

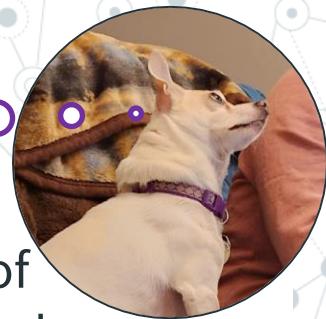
- ◎ Default SCSI Adapter – LSI Adapter
- ◎ Para-Virtual (PVSCSI) Adapter x 4
 - Distribute Volumes/Drives
- ◎ HBA (or Port) Queue Depth defaults

Queuing Considerations - Windows

- ◎ Average Disk Queue Length
- ◎ Current Disk Queue Length
- ◎ Average Disk Sec/Transfer – aka Latency
- ◎ SQL Server uses asynchronous I/O requests

Back to Error 833

So what was
root cause?

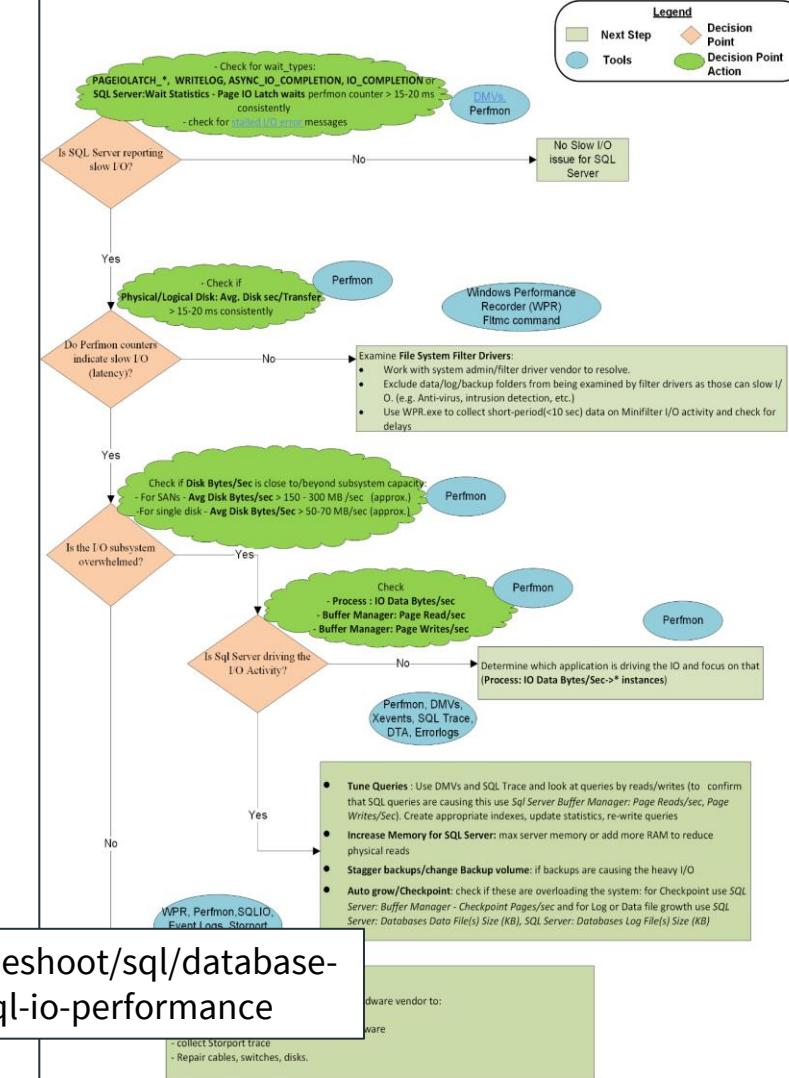


- ◎ The workload exceeds the I/O path capabilities ... because of misconfiguration ... or ... hardware capacity has been reached.
- ◎ The workload exceeds the current system capabilities, such as I/O, CPUs, and HBAs.
- ◎ The I/O path has malfunctioning ... firmware or a driver issue.
- ◎ The I/O path has malfunctioning hardware components.
- ◎ Performance issue at the operating system level.
- ◎ Filter driver intervention in the I/O process or storage path of database files. For example, antivirus program.

<https://learn.microsoft.com/en-us/sql/relational-databases/errors-events/mssqlserver-833-database-engine-error?view=sql-server-ver16#more-information>

Microsoft Methodology

- ◎ “*Troubleshoot slow SQL Server performance caused by I/O issues*”
- ◎ Logical process of elimination...



<https://learn.microsoft.com/en-us/troubleshoot/sql/database-engine/performance/troubleshoot-sql-io-performance>

Let's Talk Tools

- ◎ PerfMon Counters
- ◎ DMV Queries
- ◎ sp_whoisactive



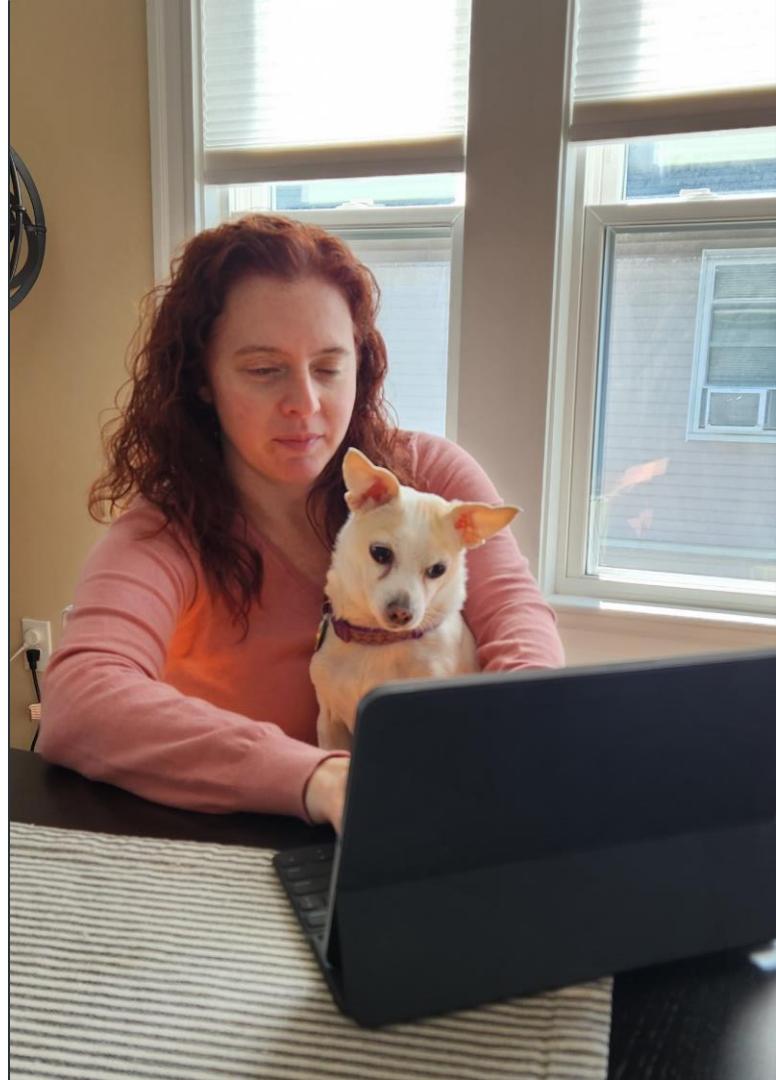
I'm bored.
Give me
something new
to play with!



I/O Tools

- ◎ sp_PressureDetector
 - By Erik Darling
- ◎ Resource Monitor
 - NOT SSMS Activity Monitor
 - Buried in Task Manager
- ◎ Process Monitor (aka ProcMon)
Sysinternals

Demo



The Most Powerful Tool



Remember
that Reddit
story?



Tales from Reddit...

- ◎ “How can load multiple +20mil datasets without overloading it?”
 - Drop & reload using BULK INSERT
- ◎ “Statements are overwhelming the resources and slowing the API down”

- Synonyms
 - Bulk Insert batch size or TABLOCK in a file
 - Co Inc
 - Table compression
 - Table partitioning trickery
- 

Tales from Reddit...

- ◎ How are you bottlenecking? Discrete

Symptoms?

vCores	1	2	4
Max IOPS	320	640	1280
Max Log Rate (MBps)	4.5	9	18

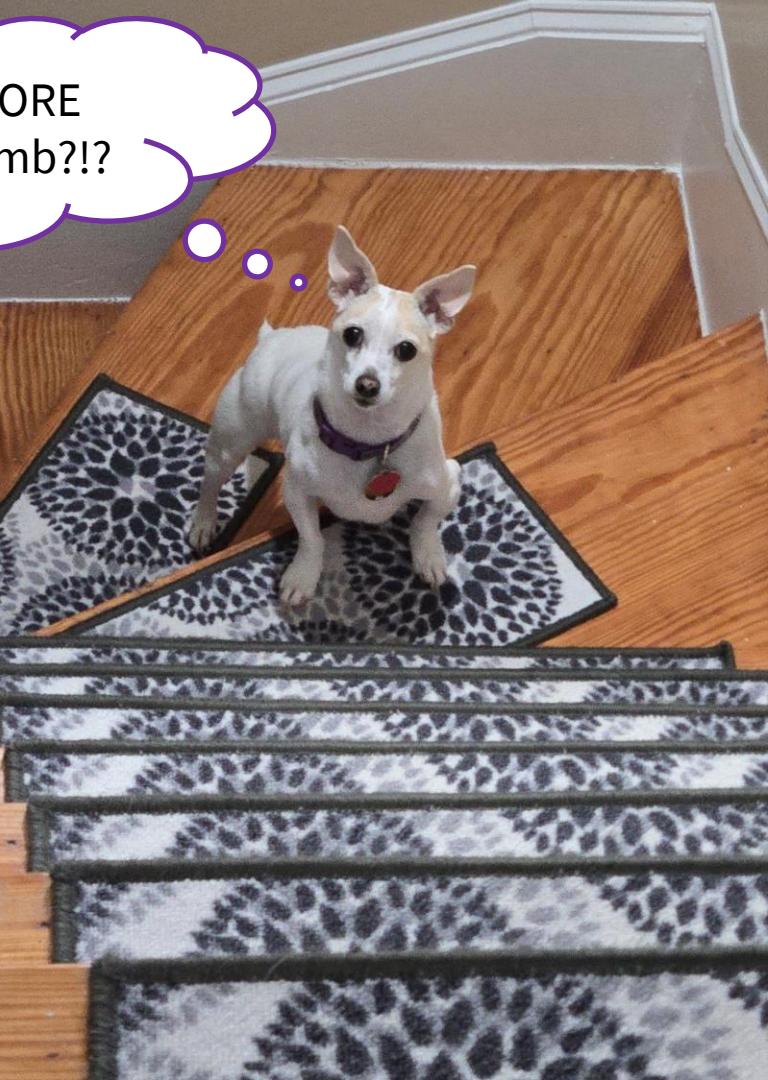
- ◎ Okay, what flavor & service level of Azure SQL?
“Azure SQL DB - Gen5, 1 vCore”

Conclusion

Summary

- ◎ Explored I/O Pathways
- ◎ CPU & RAM contention
- ◎ Throughput bottlenecks
- ◎ Queuing

There's MORE
stairs to climb?!?





This is a Trilogy...

Let's Dive Into SQL Server I/O To Improve T-SQL Performance

<https://youtu.be/fDd4lw6DfqU>



Is Storage the Root Cause of Your Performance Woes... or Not?

<https://youtu.be/kL9UVME6PO4>

Remember

Performance Tuners
do NOT need deep
expertise in everything

Broad familiarity should
be your end goal



*“Why are you doing
so much work?!?”*

Tune your queries!”

Learn More: Resources

How It Works: Bob Dorr's SQL Server I/O Presentation: Bob Dorr

<https://techcommunity.microsoft.com/t5/sql-server-support-blog/how-it-works-bob-dorr-s-sql-server-i-o-presentation/ba-p/316031>

How It Works: Debugging SQL Server Stalled or Stuck I/O Problems - Root Cause: Bob Dorr

<https://techcommunity.microsoft.com/t5/sql-server-support-blog/how-it-works-debugging-sql-server-stalled-or-stuck-i-o-problems/ba-p/315461>

Discussion About SQL Server I/O: Bob Dorr

<https://techcommunity.microsoft.com/t5/sql-server-support-blog/discussion-about-sql-server-i-o/ba-p/316268>

How It Works: How is SQL Server Error 833, 15 Sec I/O Detected: Bob Dorr

<https://techcommunity.microsoft.com/t5/sql-server-blog/how-it-works-how-is-sql-server-error-833-15-sec-i-o-detected/ba-p/3209732>

SQL Server I/O Fundamentals: Microsoft Learn

<https://learn.microsoft.com/en-us/sql relational-databases/sql-server-storage-guide?view=sql-server-ver16>

<https://github.com/sqlbek>

Learn More: Resources

Troubleshoot slow SQL Server performance caused by I/O issues: Microsoft Learn

<https://learn.microsoft.com/en-us/troubleshoot/sql/database-engine/performance/troubleshoot-sql-io-performance>

833 Database Engine Error: Microsoft Learn

<https://learn.microsoft.com/en-us/sql/relational-databases/errors-events/mssqlserver-833-database-engine-error>

Stuck Stalled IO Messages: Amit Banerjee

<https://troubleshootingsql.com/2009/12/30/stuck-stalled-io-messages/>

Outside the Big SAN Box: Identifying Storage and SAN Latency in SQL Server: Kendra Little

<https://kendralittle.com/2016/06/16/outside-the-big-san-box-analyzing-storage-and-san-latency-in-sql-server-dear-sql-dba/>

Analyzing I/O Performance for SQL Server: Glenn Berry

<https://sqlperformance.com/2015/05/io-subsystem/analyzing-io-performance-for-sql-server>

<https://github.com/sqlbek>

Learn More: Resources

VMware I/O queues, "micro-bursting", and multipathing: Chad Sakac

https://virtualgeek.typepad.com/virtual_geek/2009/06/vmware-io-queues-micro-bursting-and-multipathing.html

Storport Queue Management: Microsoft Learn

<https://learn.microsoft.com/en-us/windows-hardware/drivers/storage/storport-queue-management>

You aren't using Resource Monitor enough: Scott Hanselman

<https://www.hanselman.com/blog/you-arent-using-resource-monitor-enough>

sp_PressureDetector: Erik Darling

https://erikdarling.com/sp_pressuredetector/

<https://github.com/sqlbek>

Learn More: Filter Driver Resources

Tools for minifilter development and testing: Microsoft Learn

<https://learn.microsoft.com/en-us/windows-hardware/drivers/ifs/development-and-testing-tools>

Altitude Ranges: Microsoft Learn

<https://learn.microsoft.com/en-us/windows-hardware/drivers/ifs/load-order-groups-and-altitudes-for-minifilter-drivers>

Known allocated Altitudes: Microsoft Learn

<https://learn.microsoft.com/en-us/windows-hardware/drivers/ifs/allocated-altitudes>

Using WPR: Yong Rhee

<https://yongrhee.wordpress.com/2020/11/15/wpt-wprui-capture-high-cpu-disk-i-o-file-registry-networking-mini-filter-private-bytes-virtual-bytes-paged-pool-nonpaged-pool-and-or-application-slowness/>

Analyzing WPR output with WPA: Microsoft Learn

<https://learn.microsoft.com/en-us/windows-hardware/test/wpt/opening-and-analyzing-etl-files-in-wpa>

Windows performance toolkit for performance analysis: Sumit Khedkar

<https://sumit-khedkar.medium.com/windows-performance-toolkit-for-performance-analysis-d172ec442e68>

<https://github.com/sqlbek>

Your feedback is important to us



Evaluate this session at:

www.PASSDataCommunitySummit.com/evaluation

Thank you

Message for the end of the presentation goes here

Andy Yun



sqlbek@gmail.com



<https://github.com/sqlbek>