(November 6th, 2018)

If you know someone who would benefit from being an Insider, feel free to forward this PDF to them so they can sign up here.

Note: As an Insider, you can read all prior Insider newsletters here.

Quick Tips for our Insider friends!

Hey Insiders,

This newsletter is coming to you from Redmond, where we're putting the finishing touches to our 2019 Q1 live, online class schedule, including a new class on columnstore indexes.

And talking of 2019, **live, in-person classes for Spring 2019 are open for registration** – details below, including a **new IEAzure class expanded to four days!**

SQLskills News

We're going to see a lot of you in the coming weeks, and for the first time in four years, **our own awesome Fall 2018 SQLintersection show doesn't clash with the PASS Summit**. You still have time to join us in sunny Las Vegas as SQLintersection is the first week of December with 18 top-notch speakers delivering 40 sessions and 9 full-day workshops! We're really looking forward to it; our show is smaller than many others so we really have a chance to chat with folks and get your questions answered. And one of our favorite evening events – **SQLafterDark** – is always a huge hit (*and great fun too*)! Check out all the details here.

Live, in-person classes: we have all of our usual classes open for registration for April/May 2019 in Chicago, including:

- IEPTO1: Performance Tuning and Optimization – Part 1
- IEPTO2: Performance Tuning and Optimization – Part 2
- IECAG: Clustering and Availability Groups
- IE0: Junior/Accidental DBA
- IEUpgrade: Upgrading and New Features
- IEPML: Practical Machine Learning
- IEAzure: Azure SQL Database, Azure VMs, and Azure Managed Instance

We're also adding a new IEPowerBI and a new IECS (columnstore) class – details on those in the next few weeks.

All the classes have discounts (ranging from $100-200) for registering in 2018, and you can find all the logistical, registration, and curriculum details by drilling down from the class schedule page here.

Finally, even if you can't join us in person, I've put out a call for **2019 remote user group sessions** and we've done more than 75 this year already! If you'd like one of us to present for your user group, check out my blog post here.

Book Review

The most recent book I've read is Peter Godfrey-Smith's *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. This is a really interesting philosophy/popular science book that uses the main example of cephalopods (octopus, cuttlefish, and squid) to investigate how intelligence and consciousness developed through evolution. How does an animal experience its environment? How and why do they communicate? It's not just about cephalopods and not just about development of the mind. As a complete neophyte in philosophy I found some of the questions he pondered to be fascinating, especially why a creature with such obvious intelligence as an octopus lives such as short lifespan, when having enough neurons to show that intelligence is a very expensive investment for a creature to make - so why grow all that for such a short time? Highly recommended!

The Curious Case of…

This section of the newsletter explains recent problems we've helped with on client systems; they might be something you're experiencing too.

I had an interesting email conversation over the weekend with someone who was having a weird corruption issue. They'd seen a backup fail on a certain page in a large database, run *DBCC CHECKDB*, and decided to fix the corruption by dropping and recreating the nonclustered that the corrupt page was in – then the backup succeeded.

That wasn't the part they thought was weird. Two days later, a backup failed again, on the same page in the database, and a subsequent *DBCC CHECKDB* found that indeed that page was corrupt, but in a different table this time.

Their question to me was how can that possibly happen?

The answer is that the I/O subsystem has a problem with the storage for that offset in the data file. When the original corrupt index was dropped, the corrupt page was deallocated. At some later point that page was allocated again, to a different table, and was successfully written to disk (I know this for sure, because the first thing a data backup does is a checkpoint, and if that succeeded, then there are no dirty pages in memory from before the backup started). The page then failed to be read from disk – a clear I/O subsystem problem.

Bottom line: The smoking gun here was that the same location in the data file became corrupt as part of two different tables and two different times, ruling out (to me) a SQL Server bug or a memory corruption. When there's a corruption like this, I usually suspect the I/O subsystem first.

Paul's Ponderings

One thing that confuses people a lot is NUMA configuration, with complications from hardware, SQL Server's Automatic soft-NUMA, licensing, and which DMV you're querying – and problems can play havoc with performance. We had an interesting case recently which Jonathan wrote up and I've editorialized – enjoy!

A question came to Glenn by email from a former client asking about the NUMA configuration of their SQL Server 2017 instance. The hardware is a 2-socket server using Intel E5-2697A v4 processors @ 2.6 GHz with 16 physical cores each. With hyper-threading enabled, this equates to 64 logical processors in the system.

The question was why when they queried the DMV *sys.dm_os_nodes* did they see four online NUMA nodes with 16 schedulers each instead of two online nodes with 32 schedulers each, which is what their hardware is.

The reason for this is Automatic soft-NUMA in SQL Server 2017, which sub-partitions the hardware NUMA configuration from two nodes to four nodes since there are more than 8 physical cores in each hardware node. So, the resulting configuration is four soft-NUMA scheduler nodes each with 8 physical cores and their associated additional 8 hyper-threaded logical processors assigned to them (i.e. four nodes with 16 schedulers each).

This sub-partitioning of Automatic soft-NUMA shows up in the error log as follows:

> *Automatic soft-NUMA was enabled because SQL Server has detected hardware NUMA nodes with greater than 8 logical processors.*

And this blog post from Microsoft has more details and explanation.

Now, what about the DMV information and how to see the actual hardware configuration?

The *sys.dm_os_nodes* DMV shows the SQLOS scheduler nodes and not the underlying hardware nodes. Scheduler nodes can be created 1:1 for hardware nodes or N:1 under a soft-NUMA configuration. There are two ways to control this:

- From SQL server 2016 onwards, if there are more than 8 physical cores in a hardware NUMA node, then soft-NUMA with sub-partitioning will automatically be used.

- If more fine-grained control is needed, registry keys exist that allow control over the soft-NUMA layout

The *sys.dm_os_memory_nodes* DMV is how to see the actual hardware NUMA layout that is presented to SQLOS when SQL Server starts.

With that conundrum explained, there was still a problem in this system's configuration because the first two nodes showed 16 schedulers online each, and the last two nodes only showed 4 schedulers online in each so the nodes were imbalanced. This kind of imbalance causes problems for task assignment and scheduling. The error log confirmed that there was a licensing model problem for the instance that was limiting it to 40 logical processors:

> *SQL Server detected 2 sockets with 16 cores per socket and 32 logical processors per socket, 64 total logical processors; using 40 logical processors based on SQL Server licensing. This is an informational message; no user action is required.*

In this case, the licensing model used is limited to 20 cores, but since it is a physical server, the hyper-threads on those cores are included which gives a license-limit of 40-logical processors. Had this been a VM, even if it was configured for 64 vCPUs, the hyper-threads are not counted as being assigned to the same core and so the limit would have been 20 logical processors (and we've seen clients with exactly this problem in VMs). Jonathan explores this problem in more detail in this blog post from 2012.

This same problem does carry through the later versions of SQL Server as well. The reason the NUMA layout in *sys.dm_os_nodes* is 16, 16, 4, 4 is that hardware node 0 gets 32 of the 40 logical processors assigned to it, and the second node gets 8. This is because the first 40 schedulers in the system are deemed online, counting from scheduler 0, so the first 32 are in node 0 and the next 8 are in node 1 – hence the imbalance. This imbalance can be corrected, if the 40-processor limit is valid, using *ALTER SERVER CONFIGURATION* so that the Automatic soft-NUMA layout is balanced.

So, what happened that caused the 40 cores limit in this case? It turned out that someone had accidentally downloaded the wrong installation media from the volume licensing site and they had to correct the licensing key to allow all 64 cores to be used in the instance.

Call to Action: A lot of things can affect NUMA configuration, and an incorrect NUMA configuration can really hamper workload performance, so it's always worth checking to make sure the configuration is what you think it should be – and now you know what to look for.

Glenn's Tech Insights

This section of the newsletter highlights recent news and views from the hardware and Windows worlds that we think will be interesting to SQL Server community members.

Microsoft Releases SQL Server 2014 SP3

On October 30, 2018, Microsoft released SQL Server 2014 SP3 (Build 12.0.6024). This Service Pack includes all of the fixes that were previously released in SQL Server 2014 RTM CU1 through SQL Server 2014 SP2 CU13.

This will be the last Service Pack for SQL Server 2014, which is going to fall out of Mainstream support on July 9, 2019. Given that timeline, there will probably be about four more cumulative updates on top of SQL Server 2014 SP3 before it falls out of support.

Microsoft has back-ported even more features from SQL Server 2016 into SQL Server 2014 SP3, which is another good reason to test and deploy this update as soon as possible.

13.7 million IOPS on a Single Windows Server 2019 Hyper-Converged Cluster

At the recent Ignite conference, Microsoft's Cosmos Darwin demonstrated a single, twelve-node, hyper-converged cluster, that was able to deliver 13.7 million 4K read IOPS (with 100% reads), with an average latency of 40µs (microseconds). The cluster was also able to deliver 9.4 million 4K read IOPS (with a 90% read, 10% write ratio). Finally, this cluster was able to deliver 535.86 GB/s of sequential reads with a 2MB block size. The cluster had about 100TB of usable storage.

The HCI cluster used in the demonstration has the following main components:

- 12 x 2U Intel S2600WFT reference servers
- Each node has 384GB (12 x 32GB) of DDR4-2666 RAM
- Each node has two, 28-core Intel Cascade Lake-SP processors
- 5TB of Intel Optane DC Persistent Memory as cache
- 32TB (4 x 8TB) Intel DC P4510 PCIe NVMe flash drives as capacity
- Two Mellanox ConnectX-4 25Gbps NICs

Each server node was running Windows Server 2019 Datacenter, with the Storage Spaces Direct feature installed. The cluster was running 312 Hyper-V virtual machines running Microsoft DiskSpd to generate the I/O workload.

More details about this demonstration are included in this blog post.

#TBT

(Turn Back Time…) This section highlights some older resources we've referred to recently that you may find useful, plus blog posts we've published since the previous newsletter.

Partitioning is the theme for this TBT:

- Kimberly's intro to partitioning post: *SQLskills SQL101: Partitioning*
- Original partitioning whitepaper that Kimberly wrote (which is scenario based so it's still very applicable to all versions): *Partitioned Tables and Indexes in SQL Server 2005*
- Additional whitepaper written with focus on new features rather than scenarios: *Partitioned Table and Index Strategies Using SQL Server 2008*
- Kimberly's recorded partitioning training for the MCM certification
- And, if you're at PASS Summit this week, be sure to join Kimberly's session: ***Architecting for Scalability and Performance in Hybrid Workloads*** on Wednesday afternoon at 3:15 (75 minutes) in room 6B.

Here are a few of the blog posts we've published since the last newsletter:

- Glenn: *Automatic Plan Correction is Enterprise Edition Only Feature in SQL Server 2017*
- Glenn: *New Intel Desktop Processor Families*
- Glenn: *SQL Server Diagnostic Information Queries for October 2018*
- Glenn on SQLPerformance.com: *Three Easy SQL Server Performance Wins*

I hope you find these useful and interesting!

Video Demo

In this week's demo, Jon investigates a scalar user-defined function option to short circuit execution on NULL inputs to eliminate unnecessary function executions during query processing in SQL Server, which can have a profound effect on performance with the right data set.

The video is about 6.5 minutes long and you can get it in WMV format here.

The demo code is here.

Enjoy!

Upcoming SQLskills Events

The last event we have coming up in 2018 is our own SQLintersection conference in Las Vegas during the first week in December – see here for details – we hope you can join us!

Our first set of 2019 live, in-person events has been announced for Chicago in April/May and we'll be adding a couple more in-person classes in a week or so. Just FYI, our first quarter events will all be live, ***online*** classes and we'll be announcing those around late October.

Each and every event has a different focus as well as different benefits – from deep-technical training in our Immersion Events to wide-ranging topics at SQLintersection where you can learn more effectively how to keep moving forward in both your environment and your career! And, of course, one benefit you'll always get from in-person events is networking; we hope to meet/see you at an event soon!

To help your boss understand the importance of focused, technical training, we've also added a few items to help you justify spending your training dollars with us:

- Letter to your boss explaining why SQLskills training is worthwhile
- So why do you want to come to our training? And the winners are…
- Community blog posts about our classes
- Immersion Event FAQ

LIVE, IN-PERSON Immersion Events:

Chicago, IL, April/May 2019

- **IEPTO1**: Immersion Event on Performance Tuning and Optimization – Part 1
  - April 29-May 3
- **IECAG**: Immersion Event on Clustering and Availability Groups
  - April 29-30
- **IEPowerBI** – details coming soon!
  - April 29-30
- **IE0**: Immersion Event for the Junior/Accidental DBA
  - May 1-3
- **IEUpgrade**: Immersion Event on Upgrading SQL Server
  - May 1-3
- **IEPTO2**: Immersion Event on Performance Tuning and Optimization – Part 2
  - May 6-10
- **IEPML**: Immersion Event on Practical Machine Learning
  - May 6-10
- **IEAzure**: Immersion Event on Azure SQL Database, Azure VMs, and Azure Managed Instance
  - May 6-9

Click here for the main Immersion Event Calendar page that allows you to drill through to each class for more details and registration links.

Summary

We hope you've enjoyed this issue - we really enjoy putting these together.

If there is anything else you're interested in, we'd love to hear from you - drop us a line.

Thanks,
Paul and Kimberly

Paul@SQLskills.com and Kimberly@SQLskills.com