

(July 21<sup>st</sup>, 2014)

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



## Quick Tips for our Insider friends!

Hey Insiders!

This newsletter is coming to you from somewhere off the coast of Kona, HI, where we're just starting our second week of diving on the Kona Aggressor II live-aboard (my 6<sup>th</sup> week on this dive boat). We all passed the PADI Rescue Diver advanced training course two weeks ago, which greatly enhances situational awareness and preparedness for disaster (something that's on our minds in the SQL world too!)

We've released our **December classes in Australia** (Sydney and Canberra) and we're running a **special offer in July only** where our remaining 5-day classes for 2014 (IEPTO1, IEPTO2 in Chicago in October, and the Australia classes) have a **super-early-bird price**, saving a further \$300 off the regular early-bird price. See our [Immersion Event Calendar](#) for details and drill-downs.

Registrations are steadily coming in for the next of our popular SQLIntersection conferences, in Las Vegas this November – see [here](#) for details. And, our already phenomenal line-up is only getting better as Paul White's joining us and we still have a few more surprises to add. Paul's workshop and sessions will be added to the session line-up shortly. Our show will mainly focus on the troubleshooting and performance problems that we know you're facing today but will also highlight best practices in architecture, design, and SQL Server 2014. For a total of 6 full days – you can immerse yourself into great content with speakers that you know will deliver! Also, use the discount code "SQLskills" and you can save \$50 off registration!

The most recent book I've read is Michael Hastings' [\*The Operators: The Wild and Terrifying Inside Story of America's War in Afghanistan\*](#). Hastings is the journalist who wrote the Rolling Stone article that led to General McChrystal's downfall and the book is a much expanded version of that story, with all the background. It's also about the top brass and diplomats running the Afghanistan war and how much of it seemed to be really messed up, including serious friction between the generals and the Obama administration. From reading the book, if everything described is true, I think McChrystal deserved to be fired, regardless of his military prowess. Really interesting stuff, and recommended.

Please [let us know](#) if you liked what you read/saw here and/or have any suggestions for future Quick Tips.

Note: you can get all the prior Insider newsletters [here](#).

## **Paul's Ponderings**

In this newsletter I want to discuss a scenario I've been asked about twice in the last week where the size of a backup did not agree with the expectations of the DBA, and both times involved large, long-running transactions. We'll call the DBA 'Bob', as I promised the people involved that I'd make them anonymous.

Bob was running a script that performed a large ETL operation involving tens of millions of rows of deleted data and new, inserted data within the same transaction (more on that later). Bob knew that the ETL process shouldn't run for more than an hour, so after it had been running for more than 4 hours, he killed the process – which then took a further 3 hours to roll back. It turned out that someone had made a mistake in the script and it was deleting way more data than it should.

The daily differential backup happened that night and Bob expected it to be very small as nothing had changed in the database. Instead it was around 130GB and Bob was left wondering why? If the operation rolled back so that there were no changes in the database then the differential backup should not have backed anything up.

This is incorrect. Although the logical contents of the database were unchanged by the rolled-back ETL process, the physical contents of the database DID change.

Every time a data file page is changed for any reason, the extent that the page is part of is marked as changed in one of the Differential Bitmaps (sometimes known as DIFF\_MAP pages, Differential Change Maps, or DCM pages). This 'mark as changed' operation is irrevocable.

In fact, it's more than irrevocable (if that's grammatically possible!) in that if the change that caused the extent to be marked as changed rolls back, the act of rolling back the change generates MORE changes – i.e. doing the reverse of the original change. The roll-back changes ALSO cause the extents to be marked as changed, as they're changing them too.

The next differential backup happens and all the extents that were touched by the ETL process are marked as changed – physically changed – and so must be backed up.

You might ask why the roll back of a change doesn't unmark the extent as being changed in some way. It's not feasible to do so. Consider the case where there are multiple concurrent transactions that all change pages in the same extent, and then one of the transactions rolls back. It can't unmark the extent as being changed because the extent still contains the changes from the other transactions. Trying to keep track of when it would be possible to unmark each extent as changed would be excessively problematic (essentially involving a very complicated referencing-counting mechanism for each extent).

Furthermore, any time a data file page is changed, the Log Sequence Number in the page header is updated to reflect the LSN of the log record describing the most recent change to the page.

This newer version of the page MUST be backed up so that crash recovery works correctly. Even if the change is rolled back, the roll back involves generating another log record with the reverse change in it (called a COMPENSATION log record), that updates the page's LSN once more. And so this must be baked up as well.

**Call to action:** The next time you're wondering about perceived disparities in backup sizes, work through what's happened at the physical level to the data pages and the log and make sure you understand the algorithms involved. Remember that the more you understand how SQL Server works under the covers, the easier it is to make sense of the behaviors you observe day-to-day.

I'm curious to hear your thoughts on backup size behavior, so please feel free to [drop me a line](#), always treated confidentially, of course.

### **Video Demo**

From Erin: In this week's Insider Video we'll pick up where we left off last time with included columns and nonclustered indexes. We've already covered the internals and structure, now it's time to understand when you might use them and why.

The video is just over 12 minutes long and you can get it:

- In WMV format [here](#).
- In MOV format [here](#).

You can get the demo code [here](#).

Enjoy!

### **SQLskills Offerings**

We've released all the classes for the remainder of 2014, including two in Australia in December. Please know that all of our classes will run and their dates will not change. Please plan accordingly. We expect to release the first half of our 2015 schedule around September.

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

- [Letter to your boss explaining why SQLskills training is worthwhile](#)
- [Community blog posts about our classes](#)
- [Immersion Event FAQ](#)

### **2014 Immersion Events**

Chicago, IL

- October 6-8, 2014: **IE0**: Immersion Event for the Accidental/Junior DBA
- October 9-10, 2014: **IEHW**: Immersion Event on SQL Server Hardware
- October 6-10, 2014: **IEPTO1**: Immersion Event on Performance Tuning and Optimization – Part 1 (formerly IE1)
- October 13-17, 2014: **IEPTO2**: Immersion Event on Performance Tuning and Optimization – Part 2 (formerly IE2)

Sydney, NSW, Australia

- December 8-12, 2014: **IEPTO1**: Immersion Event on Performance Tuning and Optimization – Part 1 (formerly IE1)

Canberra, ACT, Australia

- December 15-19, 2014: **IEPTO1**: Immersion Event on Performance Tuning and Optimization – Part 1 (formerly IE1)

See [here](#) for the main Immersion Event Calendar page that allows you to drill through to each class for more details and registration links.

### **Fall SQLintersection**

This year our Fall SQLintersection conference will be the week of November 10<sup>th</sup> in Las Vegas. See [here](#) for details. Don't forget to use the discount code "SQLskills" (without the quotes and it isn't case-sensitive) and you can save \$50 off registration! We hope to see you there!

### **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](mailto:Paul@SQLskills.com) and [Kimberly@SQLskills.com](mailto:Kimberly@SQLskills.com)