# SQL Server: Detecting and Correcting Database Corruption

Module 4: Consistency Checking

Paul S. Randal
http://www.SQLskills.com/blogs/paul/
Paul@SQLskills.com

pluralsight
hardcore developer training

# Introduction

- **Although SQL Server can discover corruption, it only works if the corrupt pages happen to be read from disk**
  - There could be parts of the database that are corrupt that are rarely used
- **Regular consistency checking is necessary to help detect corruption**

- **In this module we'll cover:**
  - Consistency checking
  - Backup checksums
  - Frequently asked questions

# Consistency Checks

- **Consistency means that something is as it is supposed to be**
  - E.g. if data page X in the leaf level of a clustered index points to data page Y as being the next page in the index, page Y should exist, should be part of the index, and should have the correct key values on it
- **Consistency checking in the context of SQL Server means running a process that validates the consistency**
  - Usually this means running DBCC CHECKDB
  - We'll discuss DBCC CHECKDB and related commands in the next module
- **Also known as integrity checking**
- **There are some important questions about consistency checking**
  - How often?
  - Which databases?
  - Why not use backup checksums instead?

# Which Databases Should Be Checked?

- **All databases should be consistency checked**
  - Include read-only databases
- **Definitely include system databases**
  - *master*, as it contains information about databases, users, logins
    - Corruptions in it can cause an instance to fail to start
    - *mssqlsystemresource* is checked automatically when master is checked
  - *msdb*, as it contains SQL Agent jobs, schedules, backup history, and more
  - *model*, as its content is copied to create new databases
  - *tempdb*, as corruptions in it can cause an instance to shut down
    - Note that if you use the Maintenance Plan Wizard to create a job to check integrity, *tempdb* is excluded
  - *distribution*, to avoid replication problems

# How To Run Consistency Checks?

- **Three options for running consistency checks:**
  - Manually
  - Automated script, using SQL Agent
  - Automated using a SQL Server Maintenance Plan
    - See the upcoming course *SQL Server: Maintenance Plans* for details
- **We'll cover what DBCC commands to use in the next module**
- **Free, widely used, and easy-to-use set of maintenance scripts available from Ola Hallengren at http://ola.hallengren.com/**

# How Often Should Consistency Checks be Run?

- **It all depends on a combination of:**
  - Stability of I/O subsystem
  - Backup strategy
  - Acceptable downtime if corruption occurs
  - Acceptable data loss if corruption occurs
  - Time window available to take the extra I/O and CPU load
  - What kind of system it is (e.g. production, test, backup)
- **I always like to recommend at least once a week**
- **Examples:**
  - No backups and persistent corruptions
  - VLDB with very low downtime/data loss tolerance

# Can You Ever Get a Guarantee?

- **No matter how often you run consistency checks, you cannot get a guarantee that a database is free of corruption**
- **Here's an analogical story…**
- **Consistency checking is very similar in that:**
  - It does not and cannot read all the database pages at the same time
  - After a page has been read, it may become corrupt on disk
  - The corruption may not be noticed until the next time consistency checks are performed
- **This is why regular consistency checking is necessary**

# How To Consistency Check a VLDB?

- **Consistency checking a multi-terabyte database can take many hours!**
- **Four options for reducing the run time:**
  - Don't run any checks (don't do this!)
  - Use DBCC CHECKDB options to reduce time and resources necessary
    - Discussed in module 5
  - Break up the checks over time using other DBCC commands
    - Discussed in module 5
  - Use another system to run the consistency checks

# Consistency Checking Using Another Server (1)

- **This method removes the consistency checking workload from the production server**
- **Methodology:**
  - Perform full database backup on the production server
  - Restore database on another server
  - Run consistency checks on the restored database
- **What to do if a corruption is found?**
  - Is it the I/O subsystem on this server?
  - Is it the backup file?
  - Is it the database on the production server?
- **If corruption is found, which should be rare, it forces you to run consistency checks on the production server**

# Consistency Checking Using Another Server (2)

- **Some people advocate consistency checking a mirror database or an availability group secondary replica as a way of consistency checking the principal database or primary replica**
- **This is not valid**
- **The various servers are stored on different I/O subsystems and corruptions do not propagate between servers**
  - E.g. ensuring that a mirror database has no corruption does not imply anything about the state of the principal database
- **Consistency checking should ideally be performed on all copies of a database**
  - Especially true for availability groups if a secondary replica is being used to offload full backups from the primary replica

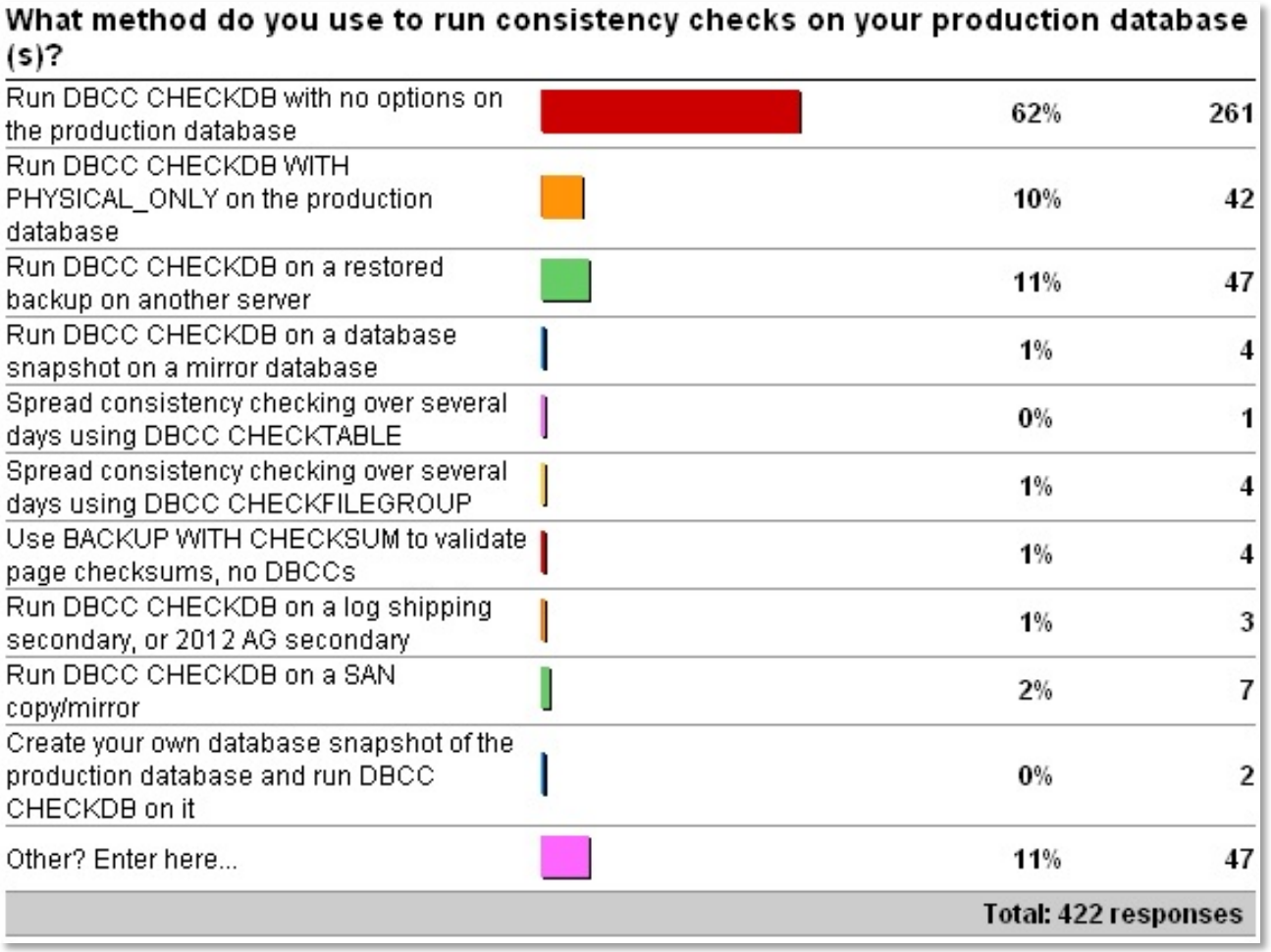
# Backup Checksums

- **Use WITH CHECKSUM when performing a backup**
  - Validates all page checksums in the database as the pages are read
  - Calculates a checksum over the entire backup stream and records it in the backup header
  - Negligible performance impact
  - Gives peace of mind that no I/O errors are present in the database
  - I recommend you always do this
- **Allows the integrity of the backup to be checked using RESTORE VERIFYONLY … WITH CHECKSUM**
  - Gives peace of mind that no I/O errors are present in the backup
  - Backups are susceptible to I/O subsystem corruption just like data files
- **Allows page checksums to be checked during restore operations**

# Backup Checksums vs. Consistency Checks

- **Some people advocate using backup checksums as a replacement for consistency checks**
- **This is not valid**
- **Although backup checksums will find problems caused by the I/O subsystem, what if the page was corrupt before it was written to disk?**
- **For example:**
  - Page in memory gets corrupted by bad memory chip
  - Page is then written to disk by the buffer pool
  - The page will have a valid page checksum, but the contents are corrupt
- **Only consistency checks using DBCC can detect this**

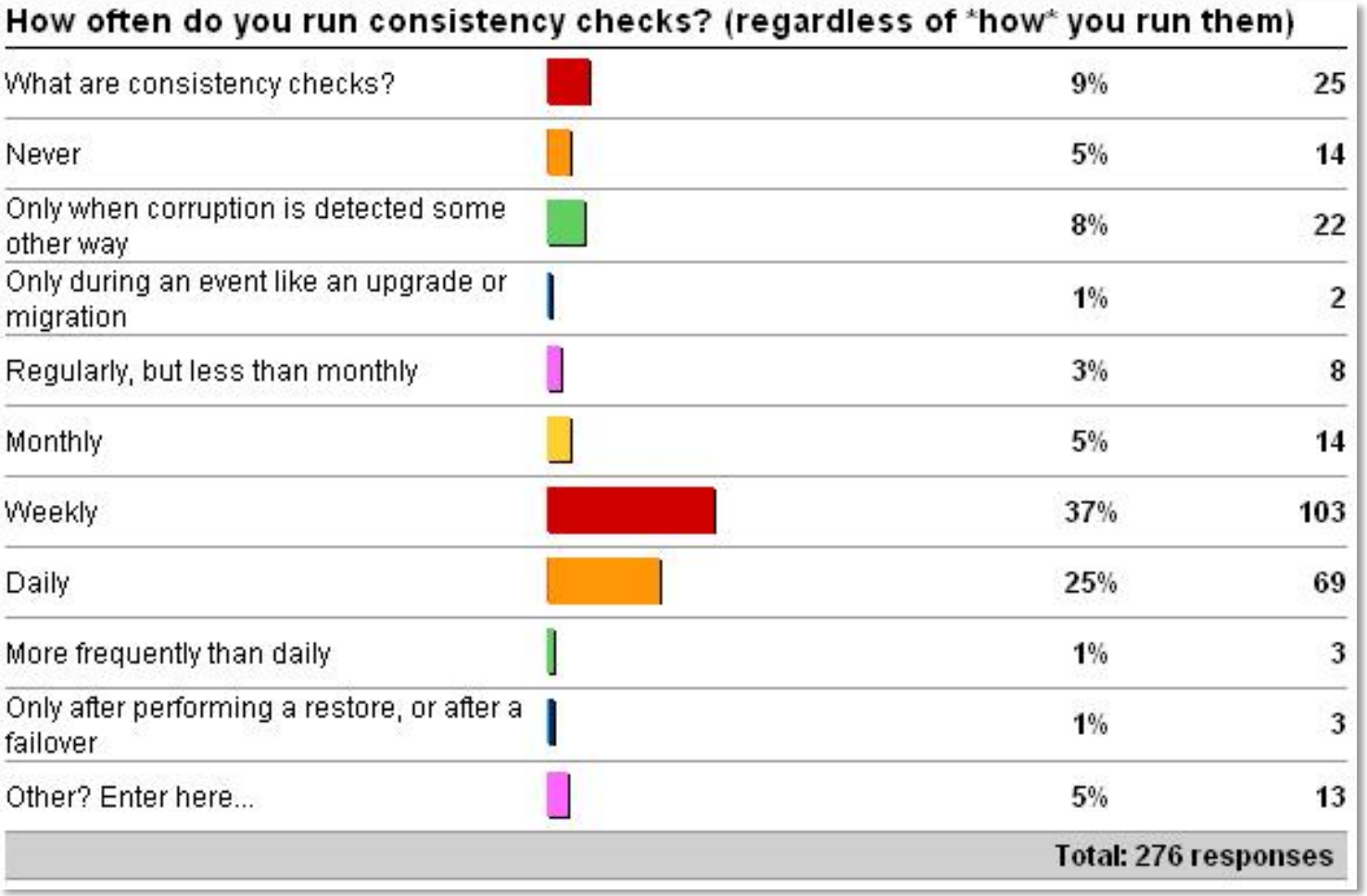
# Consistency Checking Survey (1)



| What method do you use to run consistency checks on your production database (s)? | | |
|---|---|---|
| Run DBCC CHECKDB with no options on the production database | 62% | 261 |
| Run DBCC CHECKDB WITH PHYSICAL_ONLY on the production database | 10% | 42 |
| Run DBCC CHECKDB on a restored backup on another server | 11% | 47 |
| Run DBCC CHECKDB on a database snapshot on a mirror database | 1% | 4 |
| Spread consistency checking over several days using DBCC CHECKTABLE | 0% | 1 |
| Spread consistency checking over several days using DBCC CHECKFILEGROUP | 1% | 4 |
| Use BACKUP WITH CHECKSUM to validate page checksums, no DBCCs | 1% | 4 |
| Run DBCC CHECKDB on a log shipping secondary, or 2012 AG secondary | 1% | 3 |
| Run DBCC CHECKDB on a SAN copy/mirror | 2% | 7 |
| Create your own database snapshot of the production database and run DBCC CHECKDB on it | 0% | 2 |
| Other? Enter here... | 11% | 47 |
| | Total: 422 responses | |

- Source: my blog post at [http://bit.ly/Hq3m77](http://bit.ly/Hq3m77)

# Consistency Checking Survey (2)



| How often do you run consistency checks? (regardless of *how* you run them) | | |
|---|---|---|
| What are consistency checks? | 9% | 25 |
| Never | 5% | 14 |
| Only when corruption is detected some other way | 8% | 22 |
| Only during an event like an upgrade or migration | 1% | 2 |
| Regularly, but less than monthly | 3% | 8 |
| Monthly | 5% | 14 |
| Weekly | 37% | 103 |
| Daily | 25% | 69 |
| More frequently than daily | 1% | 3 |
| Only after performing a restore, or after a failover | 1% | 3 |
| Other? Enter here... | 5% | 13 |
| | Total: 276 responses | |

- **Source: my blog post at [http://bit.ly/1mUuRw](http://bit.ly/1mUuRw)**

# Summary

- **Regular consistency checking is necessary in conjunction with using page checksums**
- **Make use of another server if the production server doesn't have the resources available to consistency check VLDBs**
- **Backups should always use the WITH CHECKSUM option, and this is not an alternative to regular consistency checking**

- **In the next module, we'll discuss:**
  - DBCC CHECKDB and related commands
  - What DBCC CHECKDB does