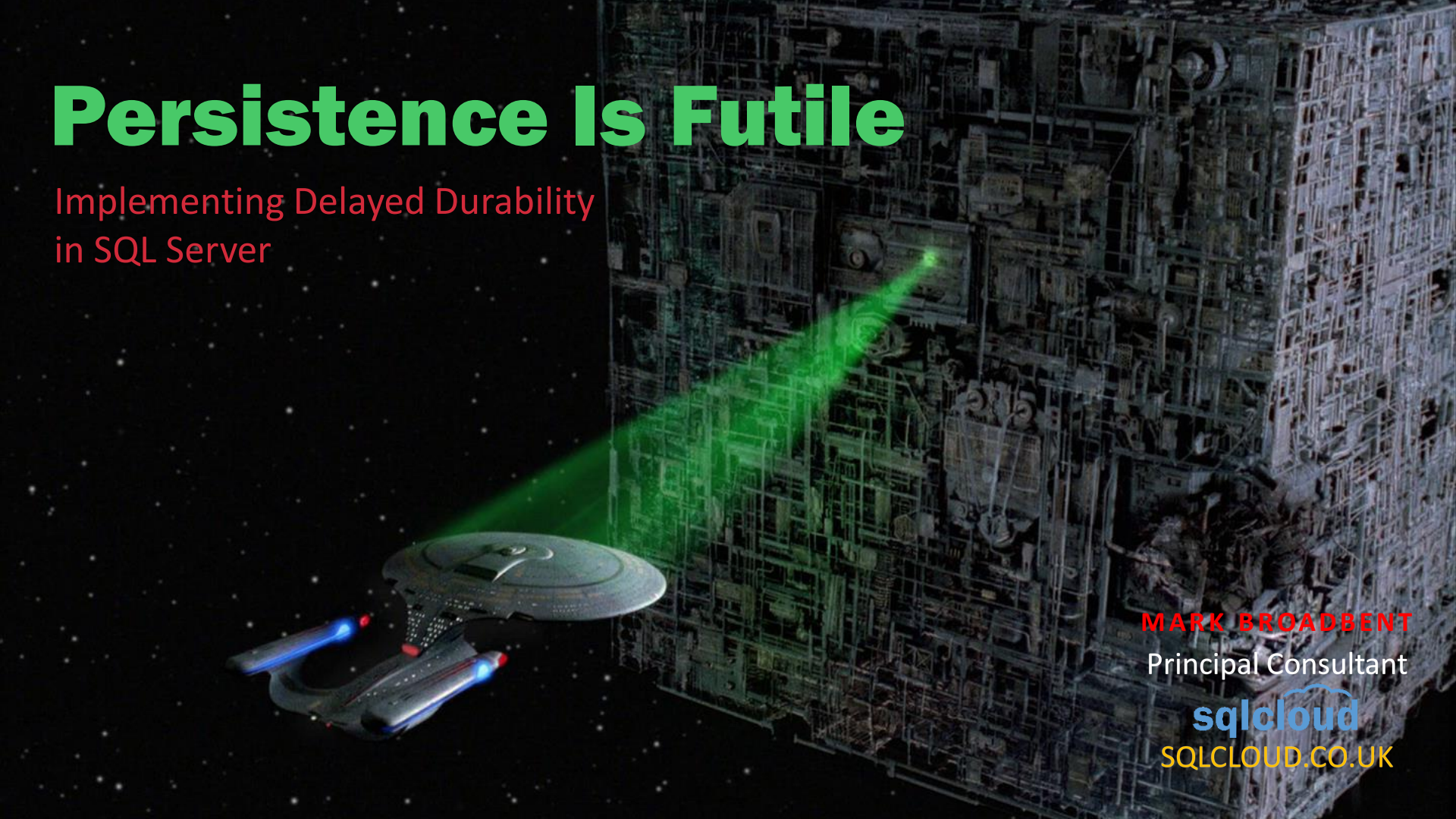
Persistence Is Futile

Implementing Delayed Durability in SQL Server



MARK BROADBENT

Principal Consultant

sqlcloud

SQLCLOUD.CO.UK

Contact…

mark.broadbent@sqlcambs.org.uk

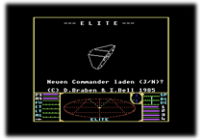

@retracement

tenbulls.co.uk

Guilty pleasures…

SERVERLESS

Likes…



Badges…

Microsoft CERTIFIED
Master: SQL Server

MVP Microsoft® Most Valuable Professional

Microsoft Community Contributor AWARD

Microsoft CERTIFIED Solutions Expert Data Platform

Microsoft CERTIFIED Solutions Expert Data Management and Analytics

PASS

Community…

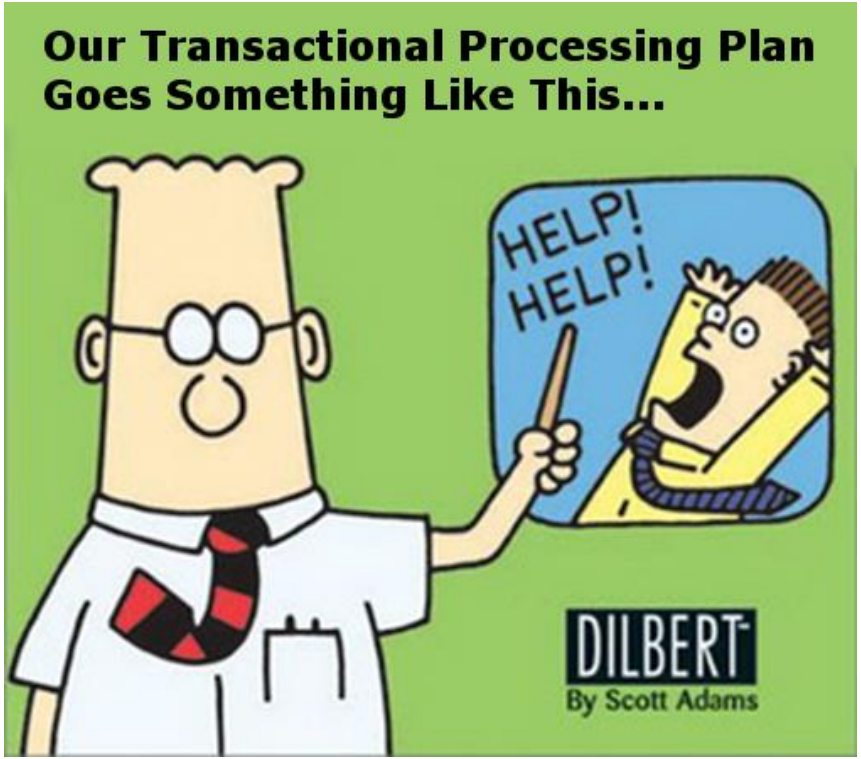
PASS HYBRID VIRTUAL CHAPTER

SQLEA

SQL SATURDAY CAMBRIDGE

SHAREPOINT SATURDAY CAMBRIDGE

Agenda



1 Transactions & Breaking ACID Properties

We'll also need to explain In-Memory OLTP in this context!

2 Log Architecture

3 Delayed Durability Implementation On-disk/ IMOLTP

4 Limitations

5 The best part of the presentation… Gin O'Clock

Requirements

- From SQL 2014 and upward
- On all editions (including Express)
- For SQL on Windows OR SQL on Linux
- In any database recovery mode
  - FULL/ BULK_LOGGED/ SIMPLE

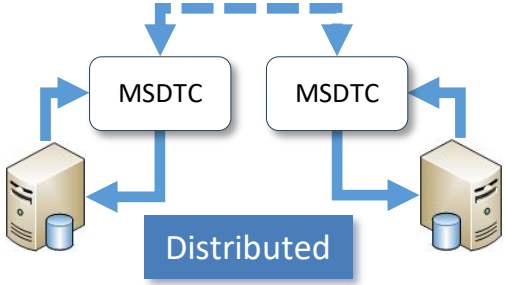
A C I D

Transactions can be…

Explicit

```
BEGIN TRAN
…
COMMIT TRAN
```

```
BEGIN DISTRIBUTED TRAN
…
COMMIT TRAN
```



Batch Scope (MARS)

```
UPDATE …
INSERT …
DELETE …
GO
```

(there is also…)

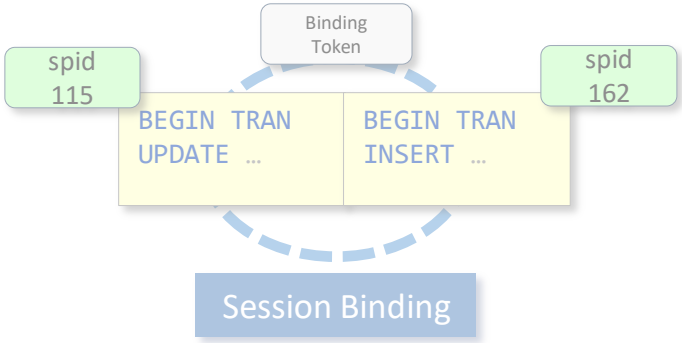
Auto-Commit

```
INSERT …
DELETE …
```

Implicit (evil muthas!)

```
INSERT …
DELETE …
COMMIT TRAN
```

"BEGIN TRAN"



Transactions

```
BEGIN TRAN transaction_1 WITH MARK 'restorepoint'
    BEGIN TRAN
       --do something
    COMMIT TRAN
    SAVE TRAN savepoint
    BEGIN TRAN
       --do something else
    COMMIT
    IF {something_wrong} THEN ROLLBACK TRAN savepoint
COMMIT
```

"Nested Transaction" *1

Savepoint

Outer Transaction

**1 Not to be confused with Atomic Blocks (we'll discuss these later!)*

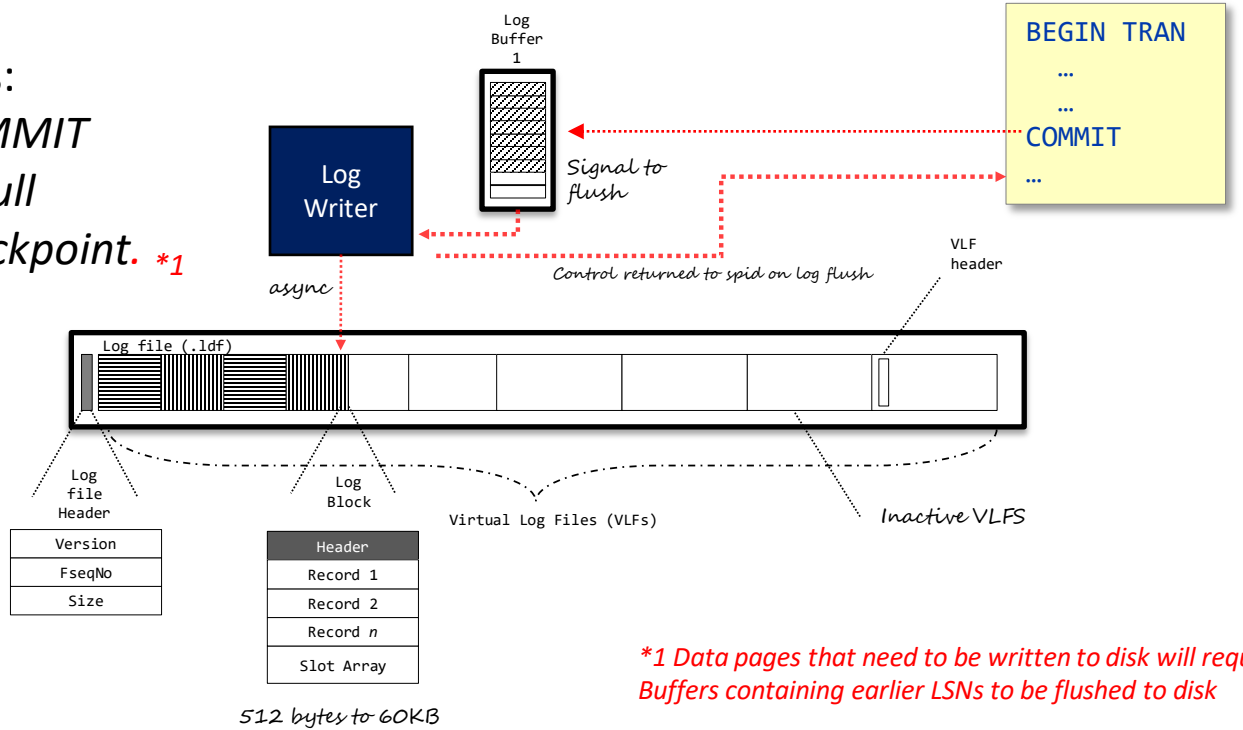
Demo

Transactions and broken ACID properties…

Durable Transactions

Log flushes:

- *On COMMIT*
- *When full*
- *On Checkpoint.* *1

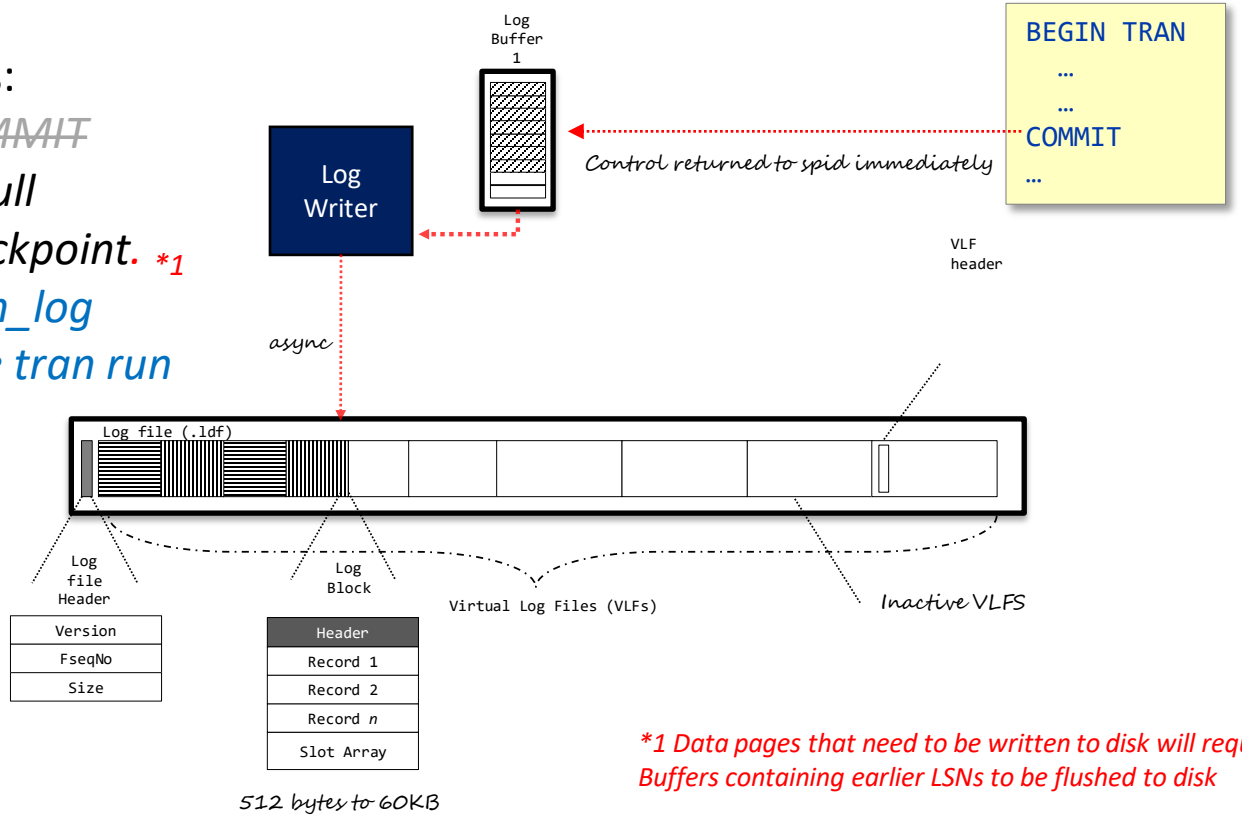


*1 Data pages that need to be written to disk will require Buffers containing earlier LSNs to be flushed to disk

Delayed Durable Transactions

Log flushes:

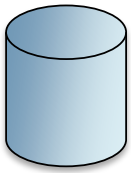
- ~~On COMMIT~~
- *When full*
- *On Checkpoint.* *1
- *sp_flush_log*
- *Durable tran run*



*1 Data pages that need to be written to disk will require Buffers containing earlier LSNs to be flushed to disk

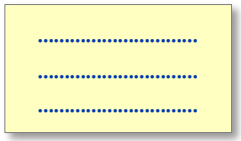
Delayed Durability Hierachy

DATABASE



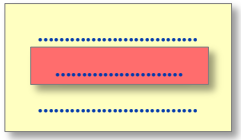
- DISABLED
- ALLOWED
- FORCED

TRANSACTION
(On-disk or IM tables)



- OFF
- ON

ATOMIC BLOCK (IM tables via Native Compiled Stored Procs)



- OFF
- ON

…BUT delayed durability is not guaranteed regardless!

Demo

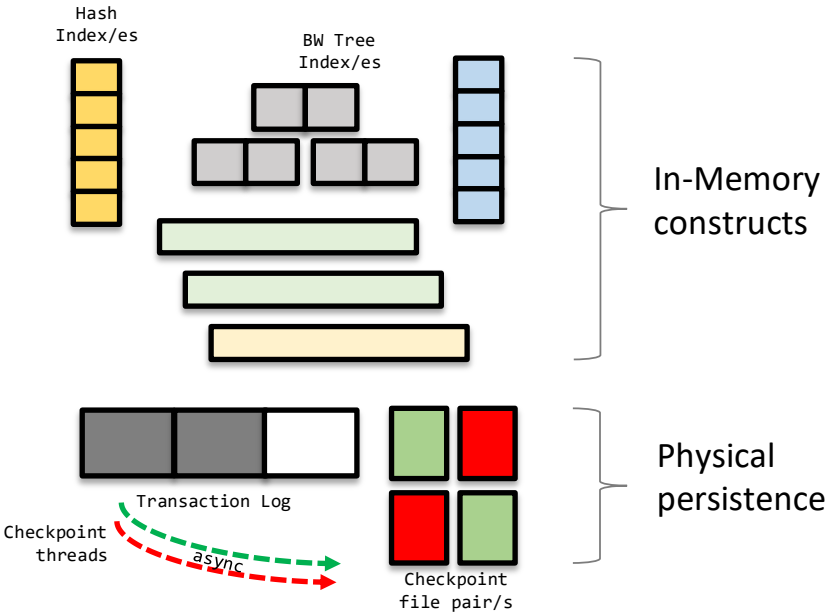
On disk delayed durability

In-Memory OLTP Overview and Log Improvements

- New concurrency model (as of SQL Server 2014)
- All row and index data in memory, all versioning In-Memory
- Lockless and latchless operation
- Physical persistence when SCHEMA_AND_DATA)
  - Thru Checkpoint file pairs + Log

<u>Log Improvements</u>

- SCHEMA_ONLY has no log overhead
- Log block “compression” (less log writes)
- No undo record generation
- Indexes not persisted, rebuilt on start-up
  - …so NO index maintenance logging.



Demo

In-Memory delayed durability

Log Waits and Log Performance Monitoring

- LOGBUFFER wait – time taken to create log record in log buffer
- WRITELOG wait – time taken for log buffer to be flushed to logfile
- Use sys.dm_io_virtual_file_stats to look at file IO, sizes and stalls
- Performance Monitors e.g.
  - Log flushes per second counter of Databases counter set
  - Transactions per second counter of Databases counter set
  - Transactions created per second counter of SQL Server <YYYY> XTP Transactions/ <DB> (instance of object).

Limitations and Special Cases

- Ignored when
  - CDC enabled tables
  - Cross-database or distributed transactions
- Not supported in Transaction Replication
- Reporting and Failover scenarios *1
  - Sync Read-only replicas
  - High Safety Database Mirrors
  - Failover Clustered Instances
  - Log backups and Log shipping targets
  - Hyper-V Quick Migration.

Warning!
Data loss is very possible in these scenarios!

*1 In a nutshell, crash recovery can (and possibly *will*) result in data loss!

Summary

- Delayed Durability attempts to make log flushes to disk more efficient
- Is a compromise of performance over durability
- Force on all trans per DB, a specific trans or on atomic block via Natively Compiled SP
- IMOLTP can remove many bottlenecks
  - But IMOLTP already provides some great logging improvements
  - However Delayed Durability *could* still remove log buffer waits
- It can result in lost data, even under highly available environments.

Thank you for listening!

Email: mark.broadbent@sqlcambs.org.uk

Twitter: retracement

Blog: http://tenbulls.co.uk

Slideshare: http://www.slideshare.net/retracement

Demo: https://github.com/retracement/Persistence_is_Futile