Weird Stuff I Saw While … Working With Heaps

Really, How Bad Is It to Use Heaps?

Rick Lowe

Data FLowe
Solutions LLC

rick@data-flowe.com

DataFLowe

http://dataflowe.wordpress.com/

Microsoft CERTIFIED Master
SQL Server® 2008

Microsoft CERTIFIED Solutions Master
Charter - Data Platform

What Is a Heap

- Table without a clustered Index
- Doesn’t automatically have a B tree
- Disorganized data

- Table that only supports scans

Demo Break 1

- Creation script
- System health session
- Update demo 1

What's a Deadlock?

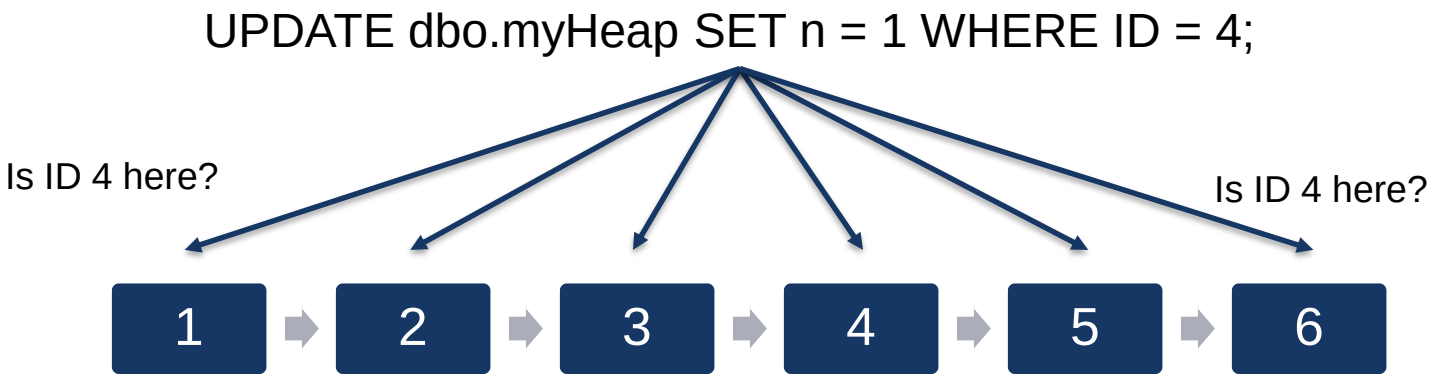
has

has

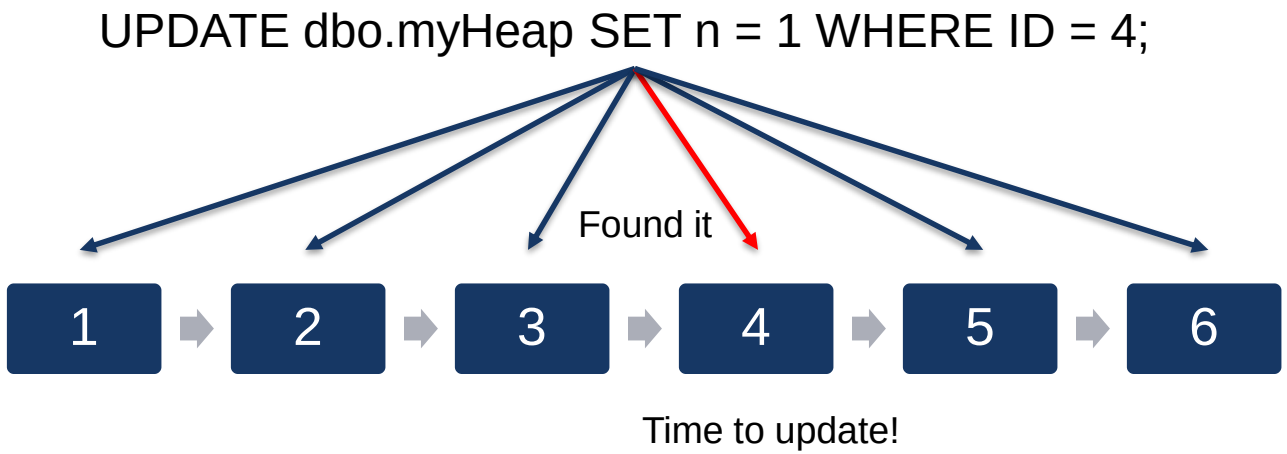
needs

needs

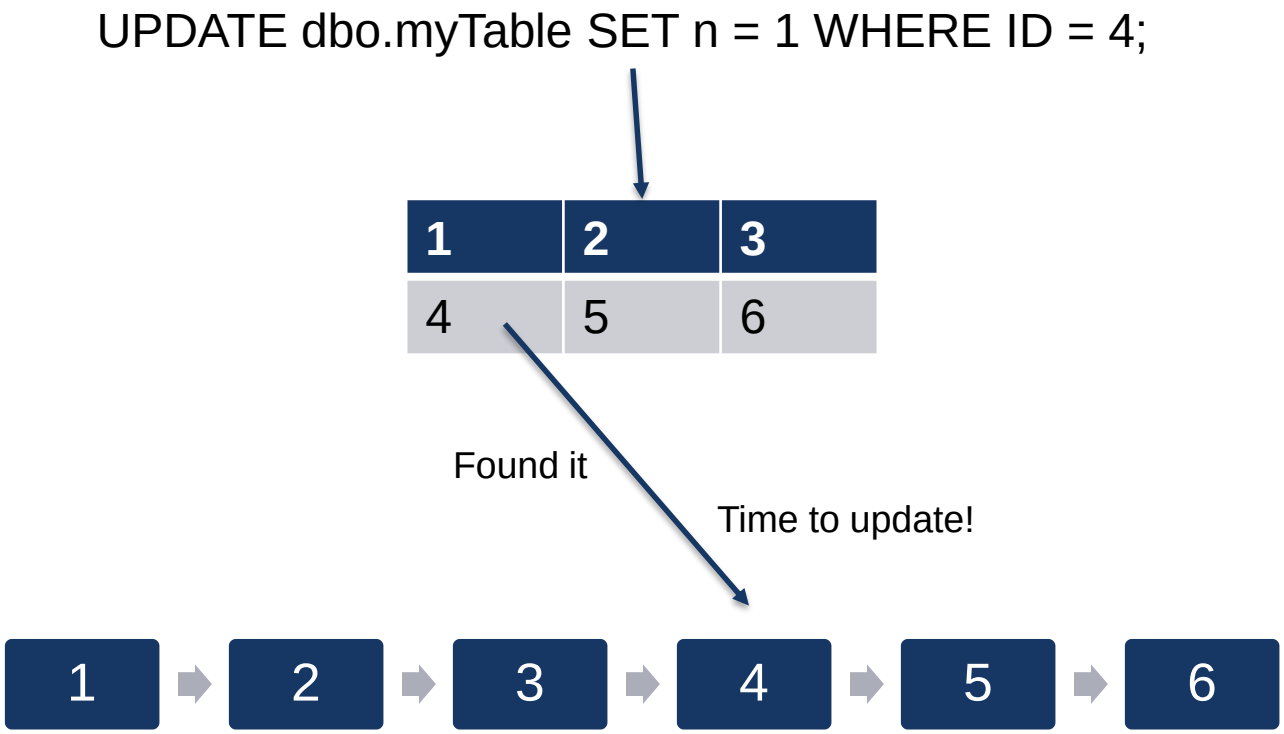
Modifying a Heap



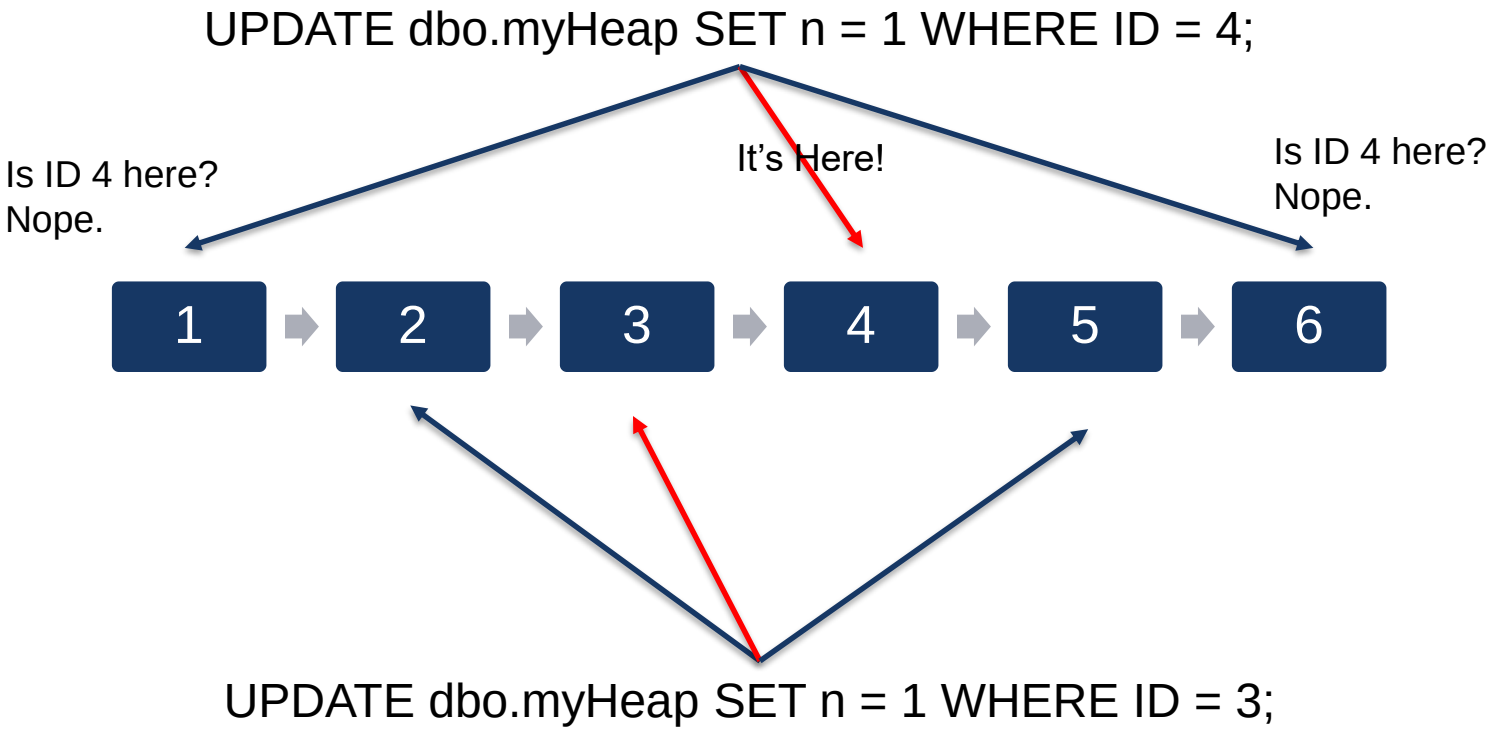
Modifying a Heap (cont'd)



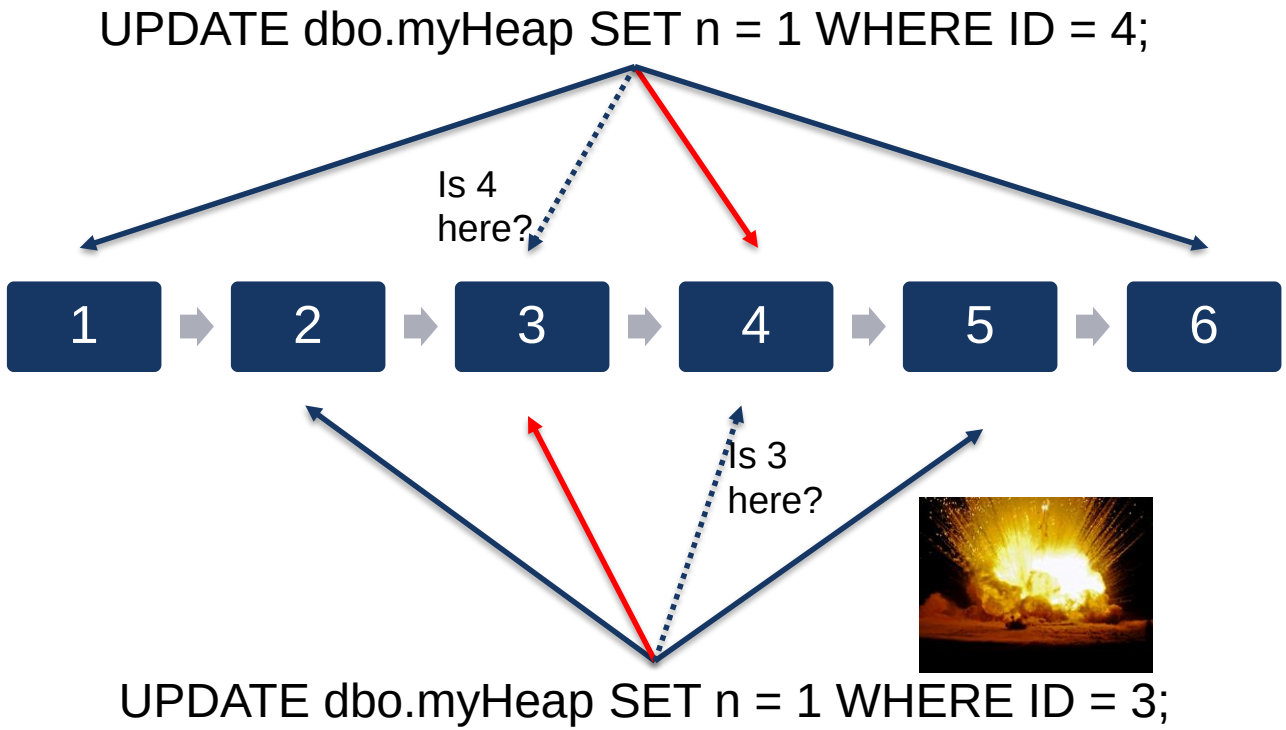
Modifying a Clustered Index (comparison)



Modifying a Heap – Two Connections



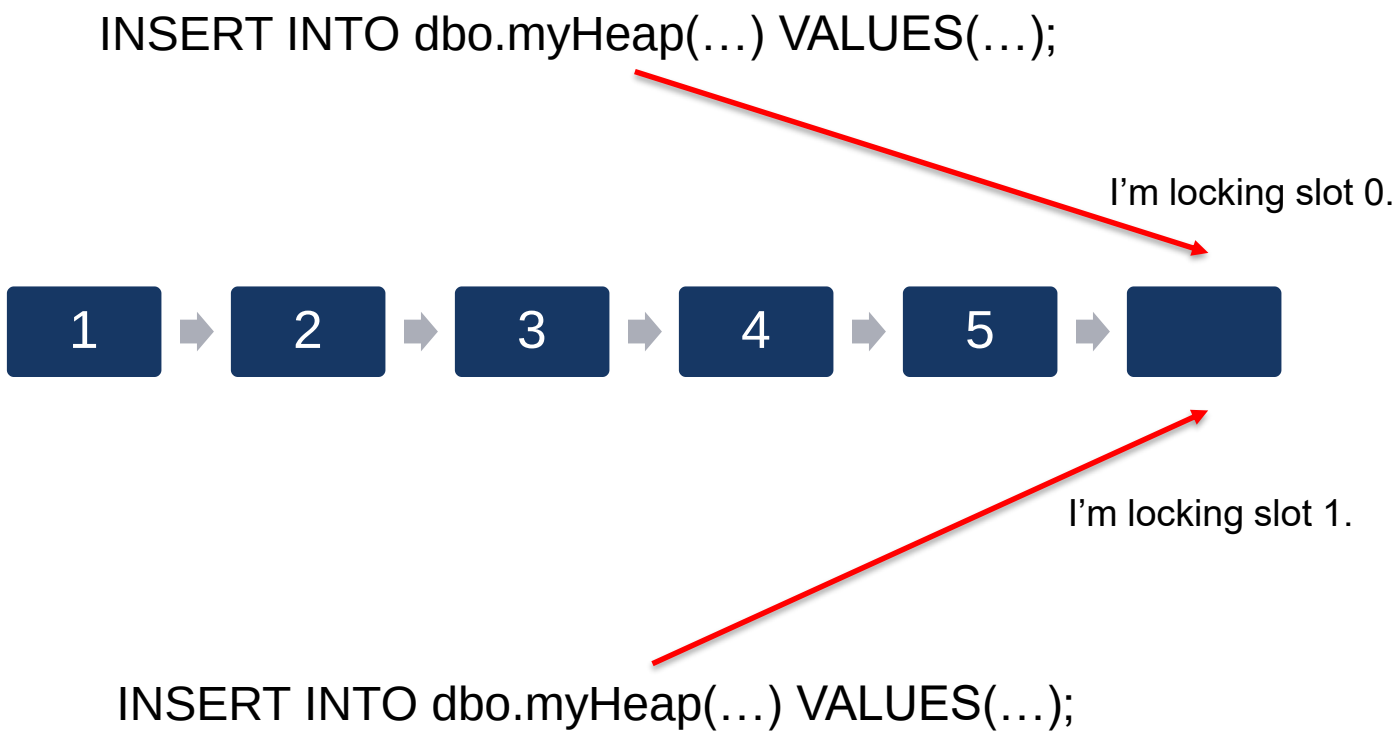
Modifying With a Heap – Two Connections (Contd)



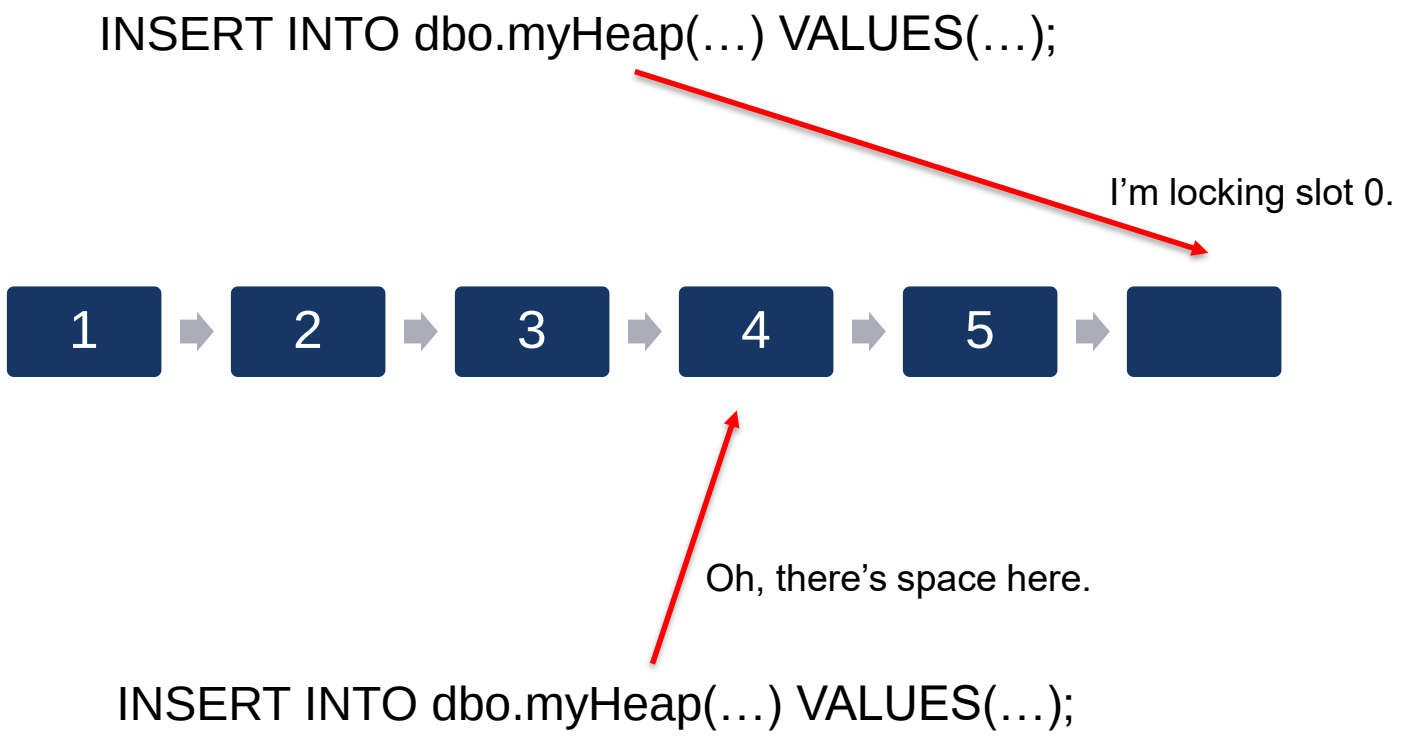
Issues With Updating Heaps

- No seeks (in the absence of NC index)
- Also no guarantee of uniqueness (scan continues even after record is found)
- Many update locks, one exclusive lock
- If multiple connections are doing similar update, connection A may hold an exclusive lock on data that connection B wants to scan and vice versa

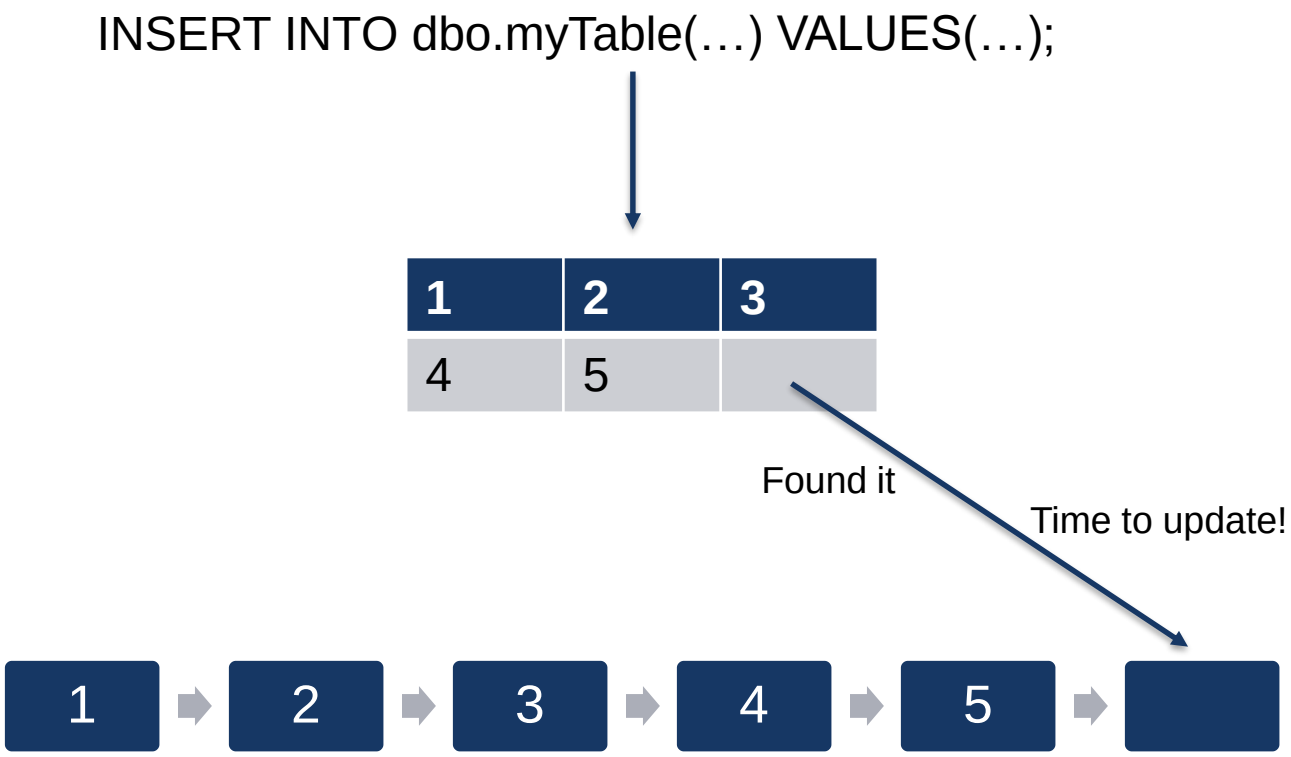
Inserting Into Heaps: Best Case



Inserting Into Heaps: Reused Space



Inserting Into Clustered Index (Comparison)

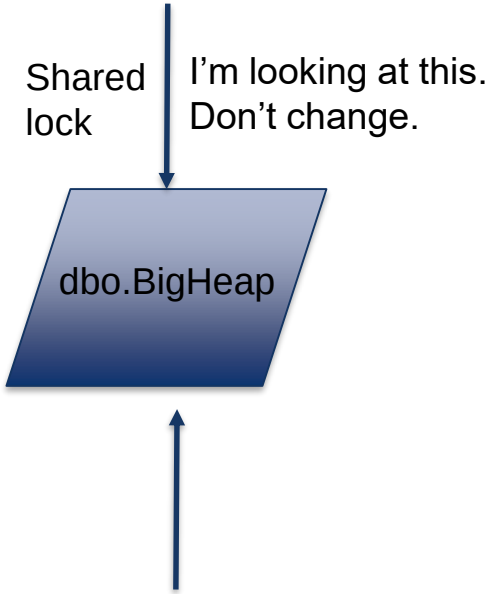


Demo Break 2

- SSIS Demo
- Speed comparison

One Way to Make Insert Fail

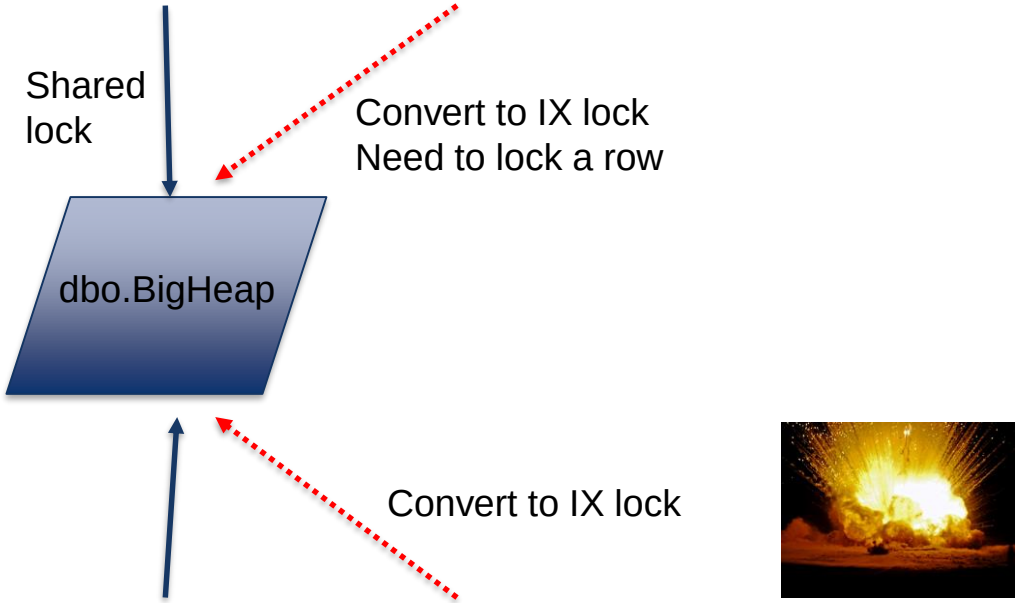
SELECT @n = COUNT(*) FROM dbo.myHeap
WHERE ClientID = @cust WITH(SERIALIZABLE)



SELECT @n = COUNT(*) FROM dbo.myHeap
WHERE ClientID = @cust WITH(SERIALIZABLE)

One Way to Make Insert Fail (cont'd)

INSERT INTO dbo.MyHeap(ID, IsTrial, BigBlobbyData) VALUES(…)



INSERT INTO dbo.MyHeap(ID, IsTrial, BigBlobbyData) VALUES(…)

Issues With Inserting Into Heaps

- No order, so new records could go anywhere
- Reusing space means looking for space
- Mixing SELECTS and INSERTS are more problematic due to table scans
- Many believe inserts into heaps are faster. That’s not always true.
- If table eventually needs a clustered index, how long does it take to convert from heap?

Demo Break 3

- Update demo 2 – heap with NC index

Indexing Cheat Sheet

- Heaps are OK for inserting lots of transitory, unsorted data through a single connection.
- Heaps aren’t ideal for select, update, delete
- Heaps aren’t ideal for multiple connections
- Clustered index inserts may outperform heap inserts when data is ordered
- Usually not difficult to sort data before loading

Minimal Logging

- Can apply to bulk insert to clustered index
- Minimal logging can be restrictive
- Clustered index data pages only minimally logged for empty table
- … But NC index on heap only minimally logged when empty
- … so either way of limited use for partial loads.

Non-Clustered Indexes

- Adding NC indexes or unique constraints to heaps rarely makes sense
- Adding NC index to heap often increases frequency of deadlocks (more things to lock)
- NC indexes also tend to make deadlocks more likely when inserting to clustered index
- Best case for inserts is often a table with clustered index and no NC indexes.

Ghost Cleanup

- http://www.mikefal.net/2012/10/17/a-heap-of-trouble/
- Large, partitioned heap
- Volatile Key
- Ghost cleanup stopped working

Thank You