

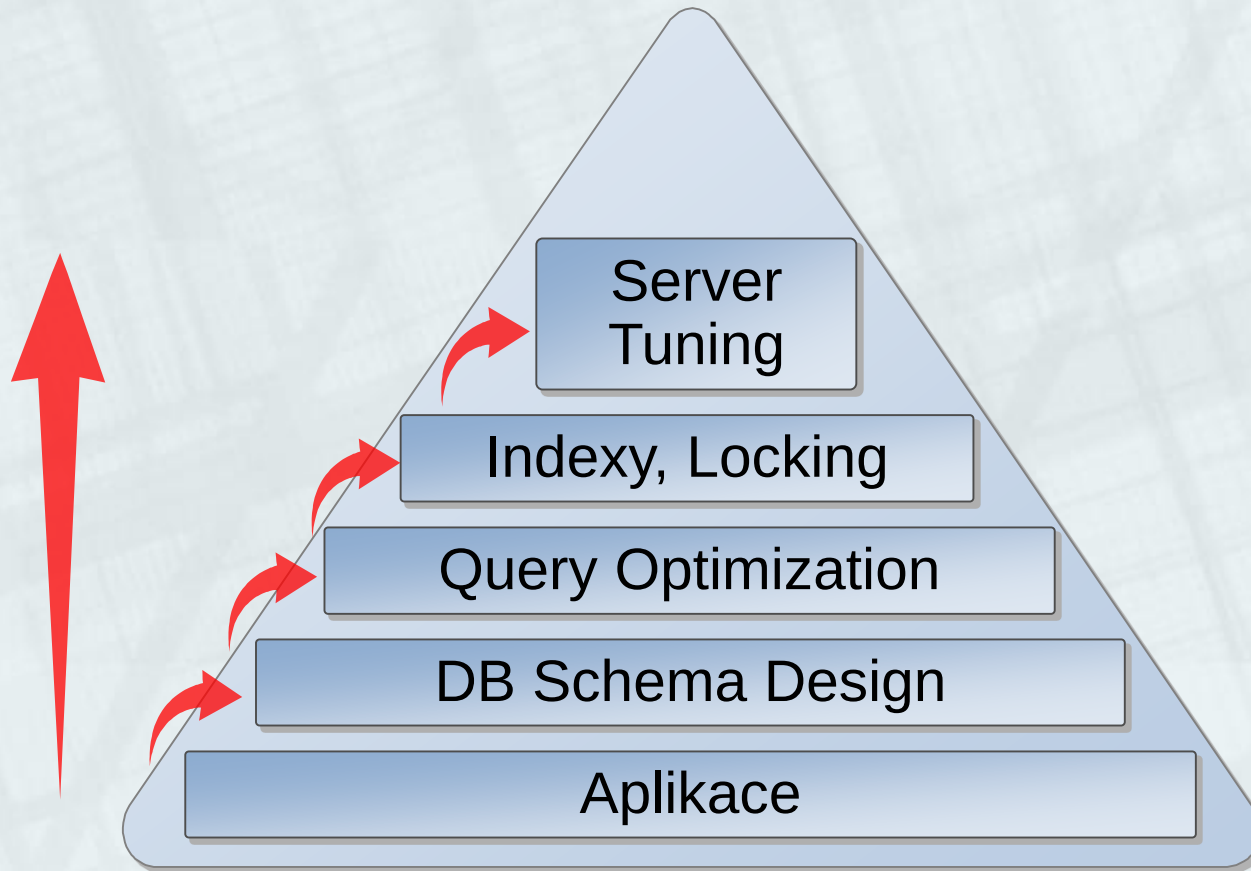
DevCon
Praha

2009

Optimalizace SQL dotazů

*Robert Haken, HAVIT, s.r.o.,
Microsoft MVP ASP.NET, MCT*

Schéma optimalizace



Aplikace

- *scénář použití (single-/multi-user, atp.)*
- *objem přenášených dat*
- *cachování výsledků*
- *SQL Server Profiler*

Schéma databáze

- *"přirozená" normalizace*
- *přiměřeně atomické atributy*
- *bez repeating att./groups*
- *správné závislosti na klíči*

Výchozí schéma

- *umělý jednoduchý PK typu int, clustered*
- *PK auto-increment pokud není tabulka read-only*
- *výjimka pouze M:N vazební tabulky, kde je PK složený ze dvou FK*
- *pouze primitivní FK vůči PK*

Výchozí schéma – pro/proti

- + *umělý PK – nemění se*
- + *monotónní růst PK – clustered index*
- + *int PK – rychlé operace (JOIN atp.)*
- *distribuované systémy – GUID*
- *umělý PK nehlídá přirozené duplicity v datech*

Variace schématu

- *atributová databáze*
- *znázornění dědičnosti objektů – pozor na asociace!*
- *denormalizace – pouze, pokud požadovaného výkonu nejsme schopni dosáhnout normalizovaným sch.*

Schema killers

- *chybné datové typy (varchar místo datetime, atp.)*
- *chybějící reference FK, magic numbers*
- *vícenásobné sloupce (Tel1, Tel2, ...)*
- *konvence pojmenování*
- *chybná normalizace*
- *hromadné číselníky?*

Query Logical Flow

1. *FROM & JOIN*
2. *WHERE*
3. *SELECT*
4. *GROUP BY & agregace*
5. *HAVING*
6. *ORDER BY*

Query Performance

- *query cost*
- *execution plan*
- *execution plan caching*

Query performance killers

- *row-based subqueries/functions*
- *crazy JOINS (přes LIKE, atp.)*
- *zbytečný DISTINCT*
- *CURSORY, nevhodné trigger*
- *univerzální dotazy – mnoho podmínek*
- *prezentační logika (formátování, atp.)*

Indexy

- *B-tree, (non-)clustered index*
- *index seek, range scan, full scan*
- *pokrývající indexy, included cols*
- *záleží na pořadí sloupců*
- *statistiky, estimated rows < 20%*
- *nelze optimalizovat izolovaně*

Indexy - výchozí doporučení

- *PK int clustered index*
- *primitivní indexy pro FK*
- *indexy pro unique constraints sloupce*
- *přidat do indexů sloupec Deleted*
- *soustředit se na klíčové dotazy*

Index killers

- *clustered index sloupce na konci non-clustered jsou zbytečné, jsou tam automaticky*
- *clustered index na sloupci, který se často mění (též monotónnost)*
- *široké clustered indexy – NC je používají jako klíče k datům*

Index killers II.

- *nevyužívané duplikátní indexy*
- *neudržované indexy (fragmentace)*

Server Tuning

- *vše mimo grafické karty :-)*
- *velikost paměti > aktivní data*
- *rozdělení disků – data, log, indexy*
- *multicore CPU*
- *síťová karta*