

Microsoft
WinDays 11

Pod pokroviteljstvom predsjednika
Republike Hrvatske prof. dr. sc. Ive Josipovića



Microsoft
WinDays technology
PROSTORI IDEJA.



SQL Server 2008 R2 kompresija

Josip Šaban, dipl. ing. rač.

Tonči Sviličić, dipl. ing. rač.

SPONZORI

Microsoft
WinDays 11

Pod pokroviteljstvom predsjednika
Republike Hrvatske prof. dr. sc. Ive Josipovića



Microsoft
WinDays technology
PROSTORI IDEJA

SPONZOR KOMUNIKACIJSKIH TEHNOLOGIJA



GLAVNI IT SPONZOR KONFERENCIJE



SPONZOR AUDIO-VIZUALNIH TEHNOLOGIJA



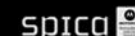
SPONZOR SISTEMSKIH INTEGRACIJA



GLAVNI SPONZOR



SPONZOR TEHNOLOGIJA ZA
OBRADU SLIKA I DOKUMENATA



SPONZOR

STARTNI SPONZOR



MEDIJSKI POKROVITELJI

ORGANIZACIJSKI PARTNER
MICROSOFT WINDAY'S11 KONFERENCIJE



HRVATSKA
UDRUGA
POSLODAVACA

CROATIAN
EMPLOYERS'
ASSOCIATION





Sadržaj predavanja

- Što je kompresija podataka u SQL Serveru
- Kako se upotrebljava
- Vrste kompresije podataka
- Kada koristiti kompresiju
- Sve ostalo...

Preduvjeti

- Poznavanja konceptata kompresije podataka
- Poznavanje rada SQL Server-a



Level 100

Zašto kompresija

- Problem današnjih servera su I/O performanse
 - Osobito ako govorimo o produkcijskim serverima sa tisućama transakcija u sekundi
- Smanjivanjem veličine podataka smanjuje se I/O pritisak ali se povećavaju zahtjevi na CPU
- Korištenje kompresije podataka u SQL Serveru je jednostavno
 - Ali treba znati kako i gdje ju koristiti



Zašto kompresija

- David DeWitt keynote na PASS Summit-u:
 - Od 1985 CPU x 1000 brži
 - Od 1985 Diskovi x 65 (!) brži
 - Veliko povećanje kapaciteta diskova ali propusnost i seek time se nisu toliko povećali što je rezultiralo 150x (relativno) sporijim diskovima

Osnovni zahtjevi

- Moraju biti zadovoljeni osnovni uvjeti
 - Ne smije postojati mogućnost gubitka podataka
 - Mora biti u potpunosti transparentna – ne smije zahtijevati nikakve promjene postojećih aplikacija
- Komprimirati se mogu
 - Cijele tablice koje su spremljene kao stog ili clustered index
 - Cijeli non-clustered indeks

Osnovni zahtjevi

- Cijeli indeksirani view
- Za particionirane tablice i indekse, kompresija se može konfigurirati za svaku particiju
 - Svaka particija objekta može imati svoju postavku kompresije

Ograničenja

- Komprimirati se ne mogu
 - Stogovi sa sparse kolonama
 - File stream podaci ili LOB podaci
 - Tablice sa redovima koji potencijalno prelaze 8.060 byte-ova i koriste row overflow
- Kompresija ne zaobilazi ograničenje limita veličine reda
- Ostalo:
 - <http://msdn.microsoft.com/en-us/library/cc280449.aspx>



Level 200

Dobre i loše strane kompresije

- Potencijalni problemi:
 - Kompresija podataka je dostupna samo u Datacenter/Enterprise verziji baze
 - Kompresija koristi CPU – treba osigurati dovoljno dostupnih CPU mogućnosti da ne dođe do usporenja cijelog sustava
 - Ne komprimiraju se jednako sve tablice i indeksi – treba testirati da se vidi da li korištenje kompresije donosi objektivne dobiti za sustav

Dobre i loše strane kompresije

- Dobiti kompresije (ovisno o podacima):
 - Značajno smanjuje I/O zahtjeve baze sa puno aktivnih transakcija
 - Smanjuje memorijski zapis podataka i (vjerojatno) poboljšava ukupne performanse
 - Više redova na stranici znači da su brži skenovi podataka i COUNT(*) operacije
 - Snapshoti baze su manji i efikasniji
 - Komprimirani podaci znače da je skraćeno vrijeme backup i restore operacija

Dva osnovna tipa kompresije

- Row-level

- Nije stvarna kompresija već se smanjenje prostora postiže korištenjem efikasnijeg načina spremanja podataka za podatke fiksne dužine

- Smanjuje količinu metapodataka koji su potrebni za jedan red podataka
 - Sprema numeričke podatke fiksne dužine kao da su varijabilne
 - Koristi onoliko byte-ova koliko je potrebno da spremi stvarna vrijednost
 - Sprema CHAR tipove podatka kao varijabilne

Dva osnovna tipa kompresije

- Slabije komprimira podatke od page-level kompresije ali i zahtijeva manje CPU-a
- Ako su tablica ili indeks particionirani onda može komprimirati pojedinu particiju
- Korisnik baze i dalje vidi kolone fiksne dužine, ali storage engine u stvari zapisuje podatke kao da su kolone varijabilne dužine

Dva osnovna tipa kompresije

- Page kompresija
 - Automatski uključuje row kompresiju i dodaje prefiks i onda dictionary kompresiju
 - Page kompresija se primjenjuje samo kada je stranica puna i kada SQL Server odluči da će uštedjeti značajnu količinu prostora
 - Prefiks kompresija uključuje slijedeće korake za svaku kolonu
 - Storage engine ispituje sve vrijednosti i izabire najčešći prefiks vrijednosti podataka u koloni

Dva osnovna tipa kompresije

- Najduža stvarna vrijednost koja počinje s tim prefiksom se sprema u CI („compression information structure”)
- Ako je prefiks prisutan na početku podatka, ubacuje se broj koji pokazuje koliko prefiks ima znakova, dok se ostatak podatkovnog niza ne mijenja
 - Ukoliko zapis na početku nema prefiks onda se u komprimirano zapisu zapisuje „0” što povećava dužinu zapisa
- Prefiks kompresija radi na nivou bitova, stoga jednako radi i za stringove i za brojeve

Dva osnovna tipa kompresije

- Nakon što završi prefiks kompresija, nastupa dictionary kompresija – svaka vrijednost se skenira i sve zajedničke vrijednosti se zamjenjuju sa tokenom koji je spremljen u CI dijelu stranice
 - Dictionary kompresija se događa na svim kolonama na nivou stranice

Dva osnovna tipa kompresije

Pre-fix

Dictionary

Page Header

Page Header		
aaabbb	aaaabb	abcd
aaabcc	abbbb	abcd
aaaacc	aaaacc	abbbb

Transparentnost procesa

- Cijeli postupak kompresije se događa na nivou storage engine-a i potpuno je transparentan za vanjske procese
 - To znači da su podaci komprimirani na disku i kao takvi učitani u memoriju
 - Storage engine dekomprimira podatke na njihovom putu od storage engine-a do analizatora upita, ali u buffer se i dalje učitava u komprimiranom obliku
 - Dekompresija se događa u trenutku kad bilo koji drugi sustav zahtijeva podatke iz buffera



Level 300

Transparentnost procesa

- Cijeli postupak kompresije se događa na nivou storage engine-a i potpuno je transparentan za vanjske procese
 - To znači da su podaci komprimirani na disku i kao takvi učitani u memoriju
 - Storage engine dekomprimira podatke na njihovom putu od storage engine-a do analizatora upita, ali u buffer se i dalje učitava u komprimiranom obliku
 - Dekompresija se događa u trenutku kad bilo koji drugi sustav zahtijeva podatke iz buffera

Redoslijed komprimiranja

- Ako je objekt komprimiran bilo row bilo page metodom (koja uključuje row metodu) tada je row kompresija uvijek uključena za svaku stranicu
- Za razliku od toga, kod page kompresije...
 - Storage engine uključuje page kompresiju stranicu po stranicu ukoliko procijeni dobitak za tu stranicu
 - Ukoliko je stranica popunjena a zelimo dodati još jedan red zapisa u nju, tada storage engine testira stranicu za kompresiju

Redoslijed komprimiranja

- Ukoliko se može komprimirati u dovoljnoj mjeri da primi novi red (ili redove) tada se uključuje page kompresija
 - Algoritam provjerava da li nova komprimirana stranica može sadržavati najmanje 5 dodatnih redova ili 25% više redova nego trenutna stranica
 - Ako nijedan od ovih kriterija nije ispunjen, komprimirana verzija stranice se odbacuje

Redoslijed komprimiranja

- Svaki novi red koji se ubacuje u stranicu se ubacuje komprimiran
 - Ali...to ne pokreće ponovnu analizu CI-a, prefiksa ili dictionary tokena
 - Ponovno pokretanje kompletne analize je moguće ukoliko se dogodi update velikog broja zapisa, a točno kada ovisi o ugrađenom algoritmu koji ovisi...
 - O broju update-a stranice
 - O broju redova u stranici
 - O prosječnoj duljini stranice
 - O količini uštedenog prostora

Redoslijed komprimiranja

- Stogovi su rekompimirani ili kod indeks rebuilda ili kod bulk load operacije
 - Kod indeks rebuilda objekta sa page kompresijom, stranica se smatra puna u ovisnosti o postavljenoj fill factor postavci, tako da je i dalje garantiran slobodan prostor

Redoslijed komprimiranja

- U slučaju page split-a, obje stranice naslijeđuju informacije o kompresiji stranice (status, prefikse i tokene) od stare stranice
 - INSERT, UPDATE i DELETE operacije se normalno zapisuju u transakcijski log u row komprimiranom načinu, ali ne u page komprimiranom
 - Iznimka je kada je uključeno logiranje page splitova – kako je to fizička operacija, zapisuju se samo vrijednosti kompresije

Procjena isplativosti kompresije

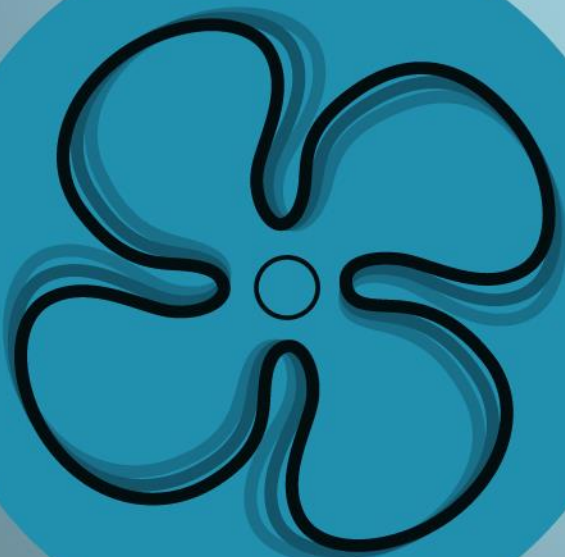
- Da bi znali da li nam se uopće isplati koristiti kompresiju trebamo imati neku procjenu koliko je podatke moguće komprimirati
 - „sp_estimate_data_compression_savings”
 - Procedura radi tako da uzme uzorak od 5% podataka za komprimiranje (algoritam uzima svaku 20-u stranicu, nije slučajni uzorak), kopira ih u „tempDb” bazu i tamo komprimira
 - Detalji njenog funkcioniranja prikazani u demo-u

Trenutne postavke

- „Table properties” i „Index Properties” daju read-only trenutnu situaciju
- Data Compression Wizard
 - Context menu → Storage → Manage Compression
 - On omogućava postavljanje vrste kompresije, analizu isplativosti i kreiranje skripte za kasnije pokretanje
 - Za trenutne postavke postavke nad svim objektima u bazi koristimo metapodatke iz sistemskih tablica (prikazano u demo-u)



DEMO





Level 400

Kako izabrati vrstu kompresije

- `sp_estimate_data_compression_savings`
- Pojednostavljeni kriteriji procjene:
 - Row kompresija je „jeftina”, u prosjeku oko 10% povećanja CPU iskorištenja – koristiti primarno na OLTP sustavu
 - Page kompresija je „skuplja”, ali bolje komprimira podatke – koristiti primarno na DWH-ovima
- Analizirati karakteristike rada (Profiler) radi odlučivanja o kandidatima za page kompresiju

Kompresija - detalji

- Stvaramo 2 kopije „Employees tablice
 - Bez kompresije i sa row kompresijom
 - Kod spremanja row kompresijom SQL Server drugačije tretira podatke manje od 8 byte-ova („kratke kolone“) i veće od 8 byte-ova („duge kolone“)
 - Zato imamo UPDATE koji „skraćuje“ jednu kolonu tako da bude manja od 8 byte-ova

Kompresija - detalji

- Kada se uključi kompresija redovi se počnu spremati način od standardnog (takozvani „Column descriptor (CD) format”)
 - Svaka kolona sada sadržava metapodatke o sadržanim podacima u redu
 - Generalna struktura CD zapisa – broj bitova u svakoj regiji ovisi o vrsti podataka



Kompresija - detalji

– Header

- Statusni byte – bitovi označavaju da li koristimo kompresiju, versioning, da li sadržava „duge podatke”, ...

– CD regija

- Sastoji se od dva dijela
 - Prvi dio je 1 ili 2 byte-a i daje nam broj „kratkih” kolona
 - Nakon toga se nalazi CD polje koje sadrži članove od 4 bita za opis dužine kolone – koliko su dugačke te posebni status ako je kolona „dugačka”



Kompresija - detalji

– Short data region

- Podaci o veličini se nalaze u CD regiji, ali ako ima 100+ kolona moglo bi biti skupo dohvaćati „udaljene kolone”
 - Za minimizaciju troška kolone su grupirane u clustere od 30 kolona
 - Na početku se nalazi „short data cluster array” čiji je svaki član single byte integer i pokazuje sumu veličina svih podataka u prošlom clusteru – efektivno postaje pointer na prvu kolonu u clusteru



Kompresija - detalji

– Long data region

- Svi podaci duži od 8 byte-ova
- To uključuje složene kolone koje ne sadrže podatke već pokazuju na podatke spremljene negdje drugdje (LOB, row overflow data pointers, ...)
- U ovom slučaju potrebna je ofsset vrijednost koja omogućuje SQL Serveru lociranje svake „duge” vrijednosti
 - Sastoji se od tri dijela: offset polja, cluster polja za „duge” kolone i „dugih” podataka



Kompresija - detalji

– Special information

- Nalaze se tri proizvoljne informacije (prisustvo tih podataka označeno je sa bitovima u prvom byte-u zapisa)
 - Forwarding pointer
 - » Kada stog sadrži forwarding pointer koji pokazuje na novu lokaciju gdje je spremljen originalni red – sadržava 3 byte-a headera i 8 byte-a dug ROW_ID
 - Back pointer
 - » Kada je prisutan forwarding te pokazuje na originalni red, sadrži 8 byte-ova dugačak ROW_ID
 - Versioning info
 - » Kada je red modificiran u Snapshot izolaciji, SQL Server dodaje 14 byte-a versioning informacija



Zaključak

- Page kompresija donosi najviše dobiti za sustave koji imaju puno I/O pristupa i sa tablicama koji se jednom zapišu i često čitaju
 - Skladišta podataka i izvještavanje
- Za OLTP sustave preporuka je uključivanje samo row kompresije za izbjegavanje ponovne izgradnje CI zapisa
 - U tom slučaju dodatna CPU potrošnja je minimalna

Zaključak

- U storage engine-u usporedbe se mogu vršiti na komprimiranim podacima jer interne konverzije mogu pretvarati tipove podatka u svoj komprimirani oblik prije usporedbe sa komprimiranim podacima
- Osim toga, samo kolone koje zahtijeva relacijski engine se moraju dekomprimirati u odnosu na dekompresiju cijelog reda

Microsoft
WinDays 11

Pod pokroviteljstvom predsjednika
Republike Hrvatske prof. dr. sc. Ive Josipovića



Microsoft
WinDays **technology**
PROSTORI IDEJA



ISPUNITE WinDays 11 UPITNIK i osvojite nagrade!

Microsoft
WinDays 11

Pod pokroviteljstvom predsjednika
Republike Hrvatske prof. dr. sc. Ive Josipovića



Microsoft
WinDays technology
PROSTORI IDEJA

Gdje ispuniti upitnik?

1. ispred dvorane nakon predavanja
2. u WinDays Hali
3. na svom računalu
4. na vašem mobilnom uređaju

Kada je izvlačenje?

Svakodnevno izvlačenje
nagrada!

Dobitnici nagrada
bit će objavljeni na
WinDays11 portalu.
Nagrade se preuzimaju
na info pultu.



htc
quietly brilliant
Telefon HTC
Mozart

5x



1x

Hrvatski Telekom
MaxTV besplatna
pretplata na
1. godinu.



5x

lenovo
Miš N30A
BADURINI



20x

EPH
Knjiga Avatar -
James Cameron



1x

NOKIA
Telefon Nokia
E5



3x

maistra
ROVINGI VRHAR
Zen spa



2x

maistra
ROVINGI VRHAR
Klasična masaža



3x



Voucher



10x

maistra
ROVINGI VRHAR
x 2 koktela



20x

Microsoft
Office
Professional 2010



7x

Microsoft
Notebook
Cooling Base

Microsoft
WinDays 11

Pod pokroviteljstvom predsjednika
Republike Hrvatske prof. dr. sc. Ive Josipovića



Microsoft
WinDays **technology**
PROSTORI IDEJA

Hvala!