

Keskusmuistitietokannat – HyPer OLTP&OLAP-yhdessä

Arto Kärki

Helsingin yliopisto

20.4.2012

Keskusmuistitietokannat – HyPer OLTP&OLAP-yhdessä

Esitys pohjautuu artikkeliin: A. Kemper, T. Neumann "Hyper: A Hybrid OLTP&OLAP Main Memory Database System Based on Virtual Memory Snapshots"

Keskusmuistitietokannat - HyPer

Sisältö

Johdanto: ongelma

Erilaisia ratkaisumalleja

HyPer arkkitehtuuri

Tapahtumien semantiikka ja toivutus

Tehokkuuden arviointia

Yhteenveto

Keskusmuistitietokannat - HyPer

Johdanto: ongelma

Tosiaikainen tapahtumien käsittely, OLTP (OnLine Transaction Processing) ja tosiaikainen tiedonjalostus, OLAP (OnLine Analytic Processing) asettavat erilaisen haasteen tietokanta-arkkitehtuurille.

OLTP:n tulisi mahdollistaa nopea tapahtuminen käsittely, toisaalta OLAP raporttien muodostaminen tarkoittaa monimutkaisten ja samalla hitaiden kyselyiden tekemistä.

Keskusmuistitietokannat - HyPer

Johdanto: ongelma

Nämä erilaiset tavoitteet on pyritty saavuttamaan siten, että tietokanta on jaettu kahdeksi erilliseksi osaksi: OLTP tietokannaksi ja datawarehouse tietokannaksi (OLAP).

Tämä tarkoittaa kuitenkin resurssien haaskausta: tietoa kopioidaan tietokannasta toiseen. Myös tiedon tuoreus kärsii: tietoa siirretään yleensä vain kerran yössä ETL-proseduurilla (extract-transform-load).

Keskusmuistitietokannat - HyPer

Johdanto: ongelma

ETL-proseduurista lyhyesti:

- extract: tietoa poimitaan useasta eri lähteestä (tietokannoista, tiedostoista, jne.).
- transform: tiedon muunto lopullisen tallennuspaikan vaatimaan muotoon (tiedon valintaa, koodien avaaminen, tiedon järjestäminen, tiedon yhdistely, jne.).
- load: tiedon lataaminen datawarehouse-tietokantaan.

Keskusmuistitietokannat - HyPer

Johdanto: ongelma

Tietojen jakamisongelmaa korjaamaan on kehitetty hybridijärjestelmiä jotka pystyvät käsittelemään sekä ajantasaiset tapahtumat että aikaa vievät raporttien muodostamiskyselyt.

Liiketoiminnan kannalta kriittisen OLTP tietokannan koko on rajattu, mikä puoltaa keskusmuistin käyttöä. Esim. suuren kaupallisen yrityksen, Amazonin, vuoden tilausrivit vievät tilaa noin 54 GB. Nämä tilausrivit ovat myynnin tärkein tietomassa.

Keskusmuistitietokannat - HyPer

Sisältö

Johdanto: ongelma

Erilaisia ratkaisumalleja

HyPer arkkitehtuuri

Tapahtumien semantiikka ja toivutus

Tehokkuuden arviointia

Yhteenveto

Keskusmuistitietokannat - HyPer

Erilaisia ratkaisumalleja

Keskusmuistitietokantajärjestelmien kehitys alkoi OLTP tietokannoista. Ensimmäisiä olivat:

- TimesTen (Oracle),
- P*TIME/Transact in Memory (SAP v. 2005) ja
- Solid Information Technologyn kehittämä SolidDB kehitetty Helsingissä (IBM).

Keskusmuistitietokannat - HyPer

Erilaisia ratkaisumalleja

SolidDB:ssä rivitason muistivedokset ovat peilattuja rivejä eivätkä peilattuja sivuja. SolidDB:ssä on havaittu tapahtumien läpimenon lisäykseksi 30% ja keskusmuistin käytön väheneminen. Sivutason peilaus juontaa juurensa relaatiotietokantajärjestelmän varhaiseen kehitykseen.

Keskusmuistitietokannat - HyPer

Erilaisia ratkaisumalleja

Viimeisimmät keskusmuistitietokannat voidaan jakaa kahteen sovellusalueeseen: OLAP ja OLTP:

- MonetDB (OLAP) perustuu sarakemuistiskeemaan,
- TREX perustuu myös sarakemuistiskeemaan (nyk. Business Warehouse Accelerator (SAP)),
- SAP:n hybridijärjestelmä on TREX:n ja P*TIME:n yhdistelmä jossa OLTP päivitykset liitetään ajoittain OLAP TREX tietokannan sarakemallin muistiin.

Keskusmuistitietokannat - HyPer

Erilaisia ratkaisumalleja

H-Storen, OLTP, (nyk. VoltDB) kehittäjät:

- analysoivat erilaisia perinteisen tietokantajärjestelmän käytöstä syntyviä ylimääräisiä kustannuksia (puskurin hallinta, lokitus, lukitus, jne.).
- todistivat keskusmuistitietokannan toimivuuden joka käsittelee tapahtumat peräkkäisjärjestyksessä ilman em. Synkronointikustannuksia. Tällöin tyypillisen liiketapahtuman (esim. tilauksen tai maksun käsittely) kesto 2 – 10 μ s.

Keskusmuistitietokannat - HyPer

Erilaisia ratkaisumalleja

Harizopoulos, Abadi, Madden ja Stonebraker jättivät ensimmäisenä pois tietokantalukot ja sarjallistivat OLTP tapahtumat. Tämä arkkitehtuuri tekee tarpeettomaksi kalliin olioiden ja hakemistojen lukitsemis- ja salpaustarpeen, koska yksittäinen päivitystapahtuma omistaa koko tietokannan, tai oman osionsa tietokannasta.

HyPer noudattaa samaa arkkitehtuuria OLTP:n osalta.

Keskusmuistitietokannat - HyPer

Erilaisia ratkaisumalleja

Muita projekteja:

- Crescando, Zürich: käsittelee kyselyt eränä selaamalla jaksoittain läpi koko datan, eli suorittamalla jatkuvia kyselyjä tietovirtaan,
- Shore-TK-järjestelmä, Lausanne: tavoitteena lukituksen optimointi ja lokituksen tehokkuus käytettäessä uusia moniytimisiä suorittimia ja
- Blink (IBM Smart Analytics Optimizer), OLTP TK-järjestelmän laajentaminen muistitietokannalla OLAP kyselyitä varten.

Keskusmuistitietokannat - HyPer

Sisältö

Johdanto: ongelma

Erilaisia ratkaisumalleja

HyPer arkkitehtuuri

Tapahtumien semantiikka ja toivutus

Tehokkuuden arviointia

Yhteenveto

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

HyPer on uusi RISC-tyylinen tietokantajärjestelmä (RISC = reduced instruction set computer, ideana yksinkertainen ja tehokas suoritin, monimutkaisimmat toiminnot KJ:n tai kääntäjän tasolla).

HyPer ei noudata vanhojen TKHJ:ien kaavaa ja hyödyntää uusimmat laitteistojen ja KJ:n tarjoaman toiminnallisuuden.

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

HyPer arkkitehtuuri on suunniteltu siten että OLTP tapahtumat ja OLAP kyselyt voidaan suorittaa samaan keskusmuistitietokantaan, ilman että kumpikaan siitä häiriintyy.

Perinteisen levyperustaisen tietokantapalvelimen puskuriallas ja sivurakenne on jätetty pois. Tieto sijaitsee melko yksinkertaisessa optimoiduissa keskusmuistin tietorakenteissa virtuaalimuistissa.

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

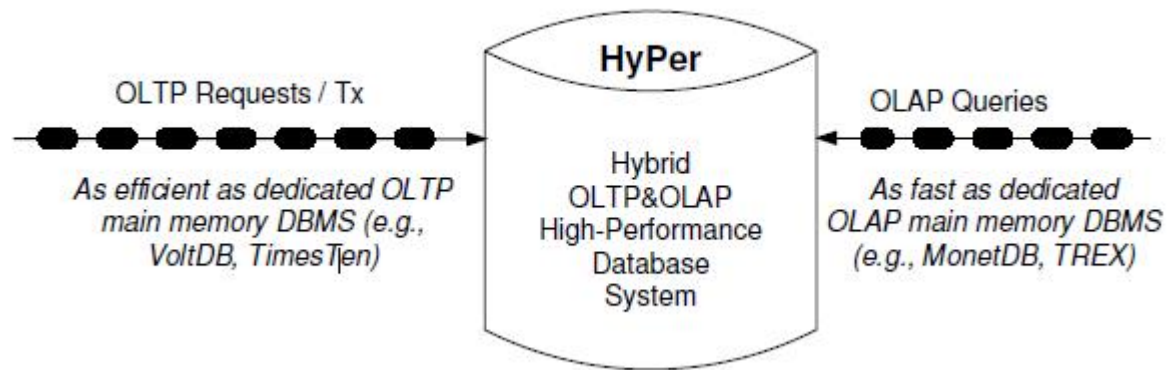
Näin voidaan hyödyntää käyttöjärjestelmän/suoritin tason osoitekäännös täysin ilman mitään lisäviittauksia. HyPer prototyyppi toimii joko sarake- tai rivimoodissa.

Vaikka virtuaalimuisti voi kasvaa suuremmaksi kuin fyysinen keskusmuisti, on tietokannan koko rajoitettu yhtä suureksi kuin on fyysinen keskusmuisti, näin on vältetty KJ:n hallinnoiman virtuaalimuistisivujen sivutukselta.

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

Tavoite:



Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

OLTP käsittelystä.

Koska kaikki tieto on saatavilla keskusmuistista, niin tiedonsiirto-odotusta ei ole. HyPer luottaa yhden säikeen tekniikkaan jonka esitteli Harizopoulos, Abadi, Madden ja Stonebaker, jossa kaikki OLTP tapahtumat suoritetaan peräkkäin. Tämä arkkitehtuuri tekee kalliin olioiden lukitsemisen ja salpaaminen tarpeettomaksi, koska yksi päivitystapahtuma omistaa koko tietokannan.

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

Peräkkäissuorittaminen on mahdollista vain puhtaassa keskusmuistitietokannassa jossa ei ole tarvetta levyoperaatioille.

Keskusmuisti-arkkitehtuurissa tyypillinen liiketapahtuma (eli tilaus tai maksutapahtuma) kestää vain noin 10 μ s. Se tarkoittaa kymmeniä tuhansia tapahtumia sekunnissa, ja se riittää täysin suuren kokoluokan liiketoimintasovellukselle.

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

OLAP muistivedosten hallinnasta.

Jos monimutkaiset OLAP kyselyt sallittaisiin OLTP:n kanssa samaan työkuormaan, niin ne tukkisivat järjestelmän, koska kaikkien seuraavien OLTP tapahtumien tulisi odottaa pitkäkestoisen kyselyn valmistumiseen saakka. Vaikka sellainen OLAP kysely valmistuisi 30 ms aikana, niin kysely lukitsisi järjestelmän niin pitkäksi ajaksi jossa tuhannet OLTP tapahtumat olisivat valmistuneet.

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

HyPer hyödyntää käyttöjärjestelmän toiminnallisuutta ja luo virtuaalisia muistivedoksia, kopioituiksi prosesseiksi.

Unixissa tämä tapahtuu luomalla lapsiprosessi OLTP:stä käyttäen `fork()` järjestelmäkutsua. Jotta tapahtumien eheys säilyy, tulee `fork` suorittaa vain kahden sarjallistetun tapahtuman välissä, ei koskaan kesken tapahtumaa.

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

fork:n käytöstä Unix-ympäristössä lyhyesti:

- luo uuden prosessin ja kopioi sille isäprosessin muistin.
- Isäprosessille fork palauttaa luodun lapsiproessin numeron ja lapsiprosessille numeron nolla.
- Isä- ja lapsiprosessi jakavat samat fyysiset muistisegmentit, koska molempien virtuaaliosoitteet osoittavat samaan fyysiseen keskusmuistiosoitteeseen.

Keskusmuistitietokannat - HyPer

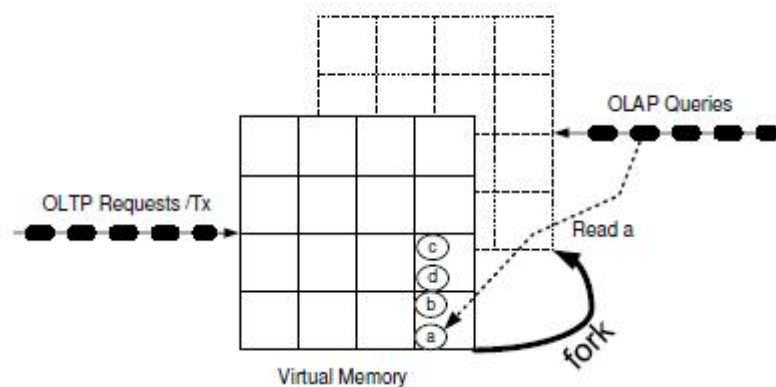
HyPer - arkkitehtuuri

Kopioitu lapsiprosessi sisältää täydellisen kopion isäprosessin osoitetilasta. Fork käskyllä luotua virtuaalimuistivedosta voidaan käyttää OLAP kyselyissä. Muistivedos pysyy täsmälleen sellaisena kuin se ilmeni fork komennon tapahtuessa. KJ ei kuitenkaan fyysisesti kopioi muistisegmenttejä heti. Se toteuttaa laiskaa, tarvehakuista toisinnusta (copy-on-update). Isäprosessi (OLTP) ja lapsiprosessi (OLAP) jakavat samat fyysiset muistisegmentit koska molemmat virtuaaliosoitteet osoittavat samaan fyysiseen keskusmuistiosoitteeseen.

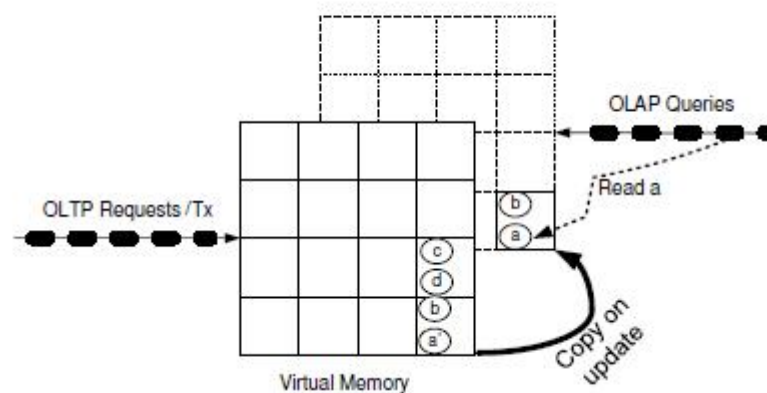
Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

Fork



Uuden muistivedoksen teko



Tarvepohjainen kopiointi –
yhtenäinen muistivedos

Keskusmuistitietokannat - HyPer

HyPer - arkkitehtuuri

Useita OLAP istuntoja:

- kyselyt lukevia, voidaan helposti suorittaa rinnakkain eri säikeissä, sama osoitetila.
- moniydin koneiden käyttö.

Monisäikeinen OLTP prosessointi:

- lukevien OLTP tapahtumien rinnakkaistaminen

Keskusmuistitietokannat - HyPer

Sisältö

Johdanto: ongelma

Erilaisia ratkaisumalleja

HyPer arkkitehtuuri

Tapahdumien semantiikka ja toivutus

Tehokkuuden arviointia

Yhteenveto

Keskusmuistitietokannat - HyPer Tap. semantiikka ja toivutus

OLAP kyselyistuntojen muistivedoseristys tekniikka:

Muistivedoseristämisessä tapahtuma/kysely näkee jatkuvasti tapahtuman tietokantatilaa yhtenäisenä, kuten se oli silloin ennen kuin tapahtuma alkoi. On olemassa erilaisia mahdollisuuksia toteuttaa kyseinen muistivedos, kun tietokannan muutokset tapahtuvat samaan aikaan: peruutus (rollback), versiointi (versioning), peilaus (shadowing) ja virtuaaliset muistivedokset.

Keskusmuistitietokannat - HyPer Tap. semantiikka ja toivutus

Tapahtumayhtenäisestä arkistoinnista:

Virtuaalimuistin muistivedoksia voidaan hyödyntää luomalla varmistusarkistot koko tietokannasta vakaalle tallennusvälineelle. Tyypillisesti arkisto kirjoitetaan nopean varmistusverkon kautta (1 – 10 GB/s) sille omistetulle tallennuspalvelimelle samassa konekeskuksessa. Säilyttääkseen tämän siirtonopeuden, tallennuspalvelimella tulee olla useita (yli 10) levyjä vastaamaan siirtonopeutta.

Keskusmuistitietokannat - HyPer

Tap. semantiikka ja toivutus

OLTP tapahtumien synkronoinnista:

1-säikeisessä OLTP:ssä ei synkronointimekanismia koska OLTP omistaa koko TK:n. Monisäikeisessä moodissa on kaksi tapahtumatyyppiä:

- ositus-rajoitetut tapahtumat voivat lukea ja päivittää tietoa niiden omissa osioissaan ja voivat myös lukea tietoa jaetuista osioista,
- osituksen ylittävät tapahtumat voivat päivittää jaettua tietoa ja hakea (lukea tai päivittää) toisen osion tietoja.

Keskusmuistitietokannat - HyPer Tap. semantiikka ja toivutus

OLTP tapahtumien synkronoinnista:

- HyPer järjestelmä sallii enintään yhden ositus-rajoittuneen tapahtuman per ositus rinnakkain. Siksi ei ole tarvetta millekään lukitukselle tai salpaamiselle koska osituksilla ei ole ylivuotavia tietorakenteita ja jaettu tieto on vain lukukäytössä.
- osituksen ylittävät tapahtumat sitä vastoin tarvitsevat täyden poissulkevuuden, ja ne ovat näin ollen kalliita tapahtumia, koska ne joutuvat odottamaan lukon saantia.

Keskusmuistitietokannat - HyPer Tap. semantiikka ja toivutus

Tapahtumien kestävydestä:

- kaikkien sitoutuneiden tapahtumien vaikutukset tulee palauttaa virheen jälkeen. Tähän HyPer käyttää loogista lokia kirjoittamalla lokiin talletettujen proseduurien parametrit jotka edustavat tapahtumia.
- perinteisessä TK-järjestelmässä looginen lokin kirjoitus on ongelmallista koska järjestelmän kaatumisen jälkeen tietokanta voi olla epäyhtenäisessä tilassa. Näin ei voi käydä HyPer:ssä, koska uudelleenkäynnistys tehdään tapahtumayhtenäisestä arkistosta.

Keskusmuistitietokannat - HyPer

Tap. semantiikka ja toivutus

Tapahtumien atomisuudesta:

- kaikkien epäonnistuneiden tapahtumien aiheuttamat muutokset tulee poistaa tietokannasta.
- vain eksplisiittisesti keskeytetyt tapahtumat tulee huomioida (R1-palautus).
- aktiivisten tapahtumien peruminen (R3) ei ole tarpeen HyPer:ssa, koska tietokanta on ei-pysyvässä muistissa ja looginen lokikirjaus on tehty vain silloin kun tapahtuman onnistunut sitoutuminen on taattu.

Keskusmuistitietokannat - HyPer Tap. semantiikka ja toivutus

Yhtenäisten muistivedosten siivoamisesta:

Peruutuslokiä voidaan käyttää luomaan toimenpide-yhtenäisestä virtuaalimuistivedoksesta joka oli luotu joidenkin tapahtumien ollessa vielä kesken, tapahtumayhtenäinen muistivedos. Tämä on hyödyllistä monisäikeisessä OLTP järjestelmässä koska silloin vältetään täysin tapahtumakäsittelyn tilan varmistaminen (quiesce).

Keskusmuistitietokannat - HyPer Tap. semantiikka ja toivutus

Toipuminen järjestelmävirheestä:

- palautumisprosessi perustuu tietokanta-arkistoon ja lokiin,
- palautus voidaan aloittaa tuoreimmasta täydellisestä arkistosta, joka palautetaan keskusmuistiin. Lokia käytetään aikajärjestyksessä, alkaen arkistoidun muistivedoksen tehneen fork:n ensimmäisestä lokimerkinnästä.

Keskusmuistitietokannat - HyPer Tap. semantiikka ja toivutus

Toipuminen järjestelmävirheestä:

- koska arkisto voidaan palauttaa 10GB/s siirtonopeudella ja loki voidaan levittää 100 000 tapahtuman sekuntivauhdilla, kestää noin 100 GB kokoisen tietokannan palautus vain muutamia minuutteja, jos varmistus arkistoidaan tunneittain,
- jos tämä palautusaika on liian pitkä, voidaan silloin käyttää toisinnettuja HyPer-koneita. Virhetilanteessa tehdään vaihto toimivaan koneeseen hyvin nopeasti.

Keskusmuistitietokannat - HyPer

Sisältö

Johdanto: ongelma

Erilaisia ratkaisumalleja

HyPer arkkitehtuuri

Tapahtumien semantiikka ja toivutus

Tehokkuuden arviointia

Yhteenveto

Keskusmuistitietokannat - HyPer

Tehokkuuden arviointia

HyPer-prototyypin tehokkuusarvio perustuu TPC-CH suorituskykytestiin, joka on yhdistelmä kahdesta standardisoidusta TPC suorituskykytestistä.

TPC-C testi on suunniteltu OLTP tietokantajärjestelmän tehokkuuden arviointiin ja TPC-H OLAP kyselyiden tehokkuuden arviointiin.

Molemmat suorituskykytestit simuloivat kaupan tilausjärjestelmää (tilaus, maksu, toimitus).

Keskusmuistitietokannat - HyPer

Tehokkuuden arviointia

TPC-CH testin tietokantaskeemassa ovat kohteiden kardinaliteetit sekä kardinaliteettien (min, max). Kardinaliteetit vastaavat tietokannan lähtötilaa TPC-C testin alussa ja kasvavat (varsinkin tilaukset ja tilausrivit) testin aikana. Tietokannan lähtötila skaalautuu tavaratalojen, asiakkaiden, tilausten ja tilausrivien lukumäärää lisäämällä. Alkuperäinen TPC-skeema, joka on säilytetty koskemattomana, koostuu yhdeksästä relaatiosta.

Keskusmuistitietokannat - HyPer

Tehokkuuden arviointia

OLAP-kyselyt:

Täydellistä OLTP&OLAP testiä varten TPC-H testin 22 kyselyä on muokattu TPC-CH skeemaa varten. Uudelleenmuotoilulla on varmistettu että kyselyt säilyttävät niiden semantiikan (liiketoiminnan näkökulmasta nähtynä) sekä niiden syntaktisen rakenteen. OLAP kyselyt eivät hyödy tietokannan osituksesta koska ne kaikki vaativat tiedon selaamista yli kaikkien ositusrajojen.

Keskusmuistitietokannat - HyPer

Tehokkuuden arviointia

Erilaisten HyPer kokoonpanojen tehokkuudesta:

Kaikki testit on tehty TPC-C kokoonpanolla, jossa on:

- 12 tavarataloa,
- 360 000 asiakasta,
- 3.6 miljoonaa tilausriviä (1 GB).

HyPer voidaan konfiguroida rivi- tai saraketason varastoksi. OLTP-käsittelyssä ei havaintoja erityisistä eroista tehokkuudessa, mutta OLAP kyselyt nopeutuivat huomattavasti sarakemallin varastoskeemalla. OLTP ja OLAP tehokkuusluvut ilmoitetaan vain sarakemallista varastosta.

Keskusmuistitietokannat - HyPer

Tehokkuuden arviointia

HyPer testi kuten myös MonetDB kyselytesti ajettiin tavallisella palvelimella, seuraavin määrittäyksin:

- Dual Intel X5570 Quad-Core-CPU, 8MB Cache
- 64GB RAM
- 16 300GB SAS-HD (ei benchmark-testissä)
- Linux käyttöjärjestelmä RHEL 5.4
- Hinta: 13.886 euroa (ale hinta yliopistolle)

VoltDB testi on tehty vastaavanlaisella laitteistolla (dual-quad Xeon CPU Dell R610), mutta 6-solmun klusterilla.

Keskusmuistitietokannat - HyPer

Tehokkuuden arviointia

HyPerin yksittäisellä vakiopalvelimella saavutetut tulokset vastaavat VoltDB:n 6-solmun klusterilla saavuttamia tuloksia. VoltDB:n julkaisut osoittavat, nämä tulokset vastaavat parhaita julkaistuja suuren mittakaavan levyperustaisen TK-konfiguraation TPC-C tuloksia. HyPer OLTP:n läpimenot saavutettiin jopa silloin kun yksi, kahdeksan tai kolme rinnakkaista OLAP prosessia teki jatkuvasti OLAP kyselyitä rinnakkain OLTP työkuormaan.

Keskusmuistitietokannat - HyPer

Tehokkuuden arviointia

HyPerin OLAP-kyselyiden vasteaikoja vertailtaessa MonetDB:hen, huomataan, että nämä kaksi kyselymoottoria ovat yhtä tehokkaita.

MonetDB käyttö perustui pelkästään OLAP-kyselyihin.

Keskusmuistitietokannat - HyPer Tehokkuuden arviointia

Query No.	one query session (stream) single threaded OLTP		HyPer configurations 8 query sessions (streams) single threaded OLTP		3 query sessions (streams) 5 OLTP threads		MonetDB no OLTP	VoltDB no OLAP
	OLTP throughput	Query resp. times (ms)	OLTP throughput	Query resp. times (ms)	OLTP throughput	Query resp. times (ms)	1 query stream Query resp. times (ms)	only OLTP results from [18]
Q1		67		71		71	63	
Q2		163		233		212	210	
Q3		66		78		73	75	
Q4		194		257		226	6003	
Q5		1276		1768		1564	5930	
Q6		9		19		17	123	
Q7		1151		1611		1466	1713	
Q8		399		680		593	172	
Q9		206		269		249	208	
Q10		1871		2490		2260	6209	
Q11		33		38		35	35	
Q12		156		195		170	192	
Q13		185		272		229	284	
Q14		122		210		156	722	
Q15		528		1002		792	533	
Q16		1353		1584		1500	3562	
Q17		159		171		168	342	
Q18		108		133		119	2505	
Q19		103		219		183	1698	
Q20		114		230		197	750	
Q21		46		50		50	329	
Q22		7		9		9	141	
	new order: 56961 tps; total: 126576 tps		new order: 29359 tps; total: 65269 tps		new order: 171384 tps; total: 380868 tps			55000 tps on single node; 300000 tps on 6 nodes

Keskusmuistitietokannat - HyPer

Sisältö

Johdanto: ongelma

Erilaisia ratkaisumalleja

HyPer arkkitehtuuri

Tapahtumien semantiikka ja toivutus

Tehokkuuden arviointia

Yhteenveto

Keskusmuistitietokannat - HyPer

Yhteenveto

- OLTP ja OLAP samassa tietokannassa ilman että ne häiritsevät toisiaan,
- muistivedoksen ylläpito ja käsittelyn tehokkuus, joka ilmenee OLTP:n läpimenoaikoina ja OLAP kyselyiden vasteaikoina, saavutetaan laitteiston tukeman tarvepohjaisen toisinnuksen avulla ja samalla ylläpitäen muistivedoksen yhtenäisyys,
- HyPer vaikuttaa lupaavalta arkkitehtuurilta, se voi olla jatkossa todellinen vaihtoehto keskusmuistitietokantojen markkinoilla.