

Partiționarea datelor pentru o aplicație de tip cloud

~tema 2 DATC~

La ora actuală, datele sunt împărțite în partiții care pot fi gestionate și accesate separat.

Partiționarea datelor are mai multe avantaje:

➤ *Îmbunătățește scalabilitatea*

Când un sistem de baze de date este scalat, acesta va ajunge la o limită hardware fizică. Dacă datele sunt împărțite pe mai multe partiții, fiecare pe un server separat atunci sistemul poate fi scalat aproape pe o perioadă nelimitată.

➤ *Îmbunătățește performanța*

Operațiunile de acces la date pe fiecare partiție au loc pe un volum mai mic de date; operațiunile care afectează mai multe partiții pot rula în paralel.

➤ *Îmbunătățește securitatea*

Se pot separa datele sensibile de cele insensibile în diferite partiții, iar pe cele sensibile se pot aplica diferite controale de securitate.

➤ *Oferă flexibilitate operațională*

Sunt oferite unele oportunități pentru operațiunile de reglaj fin, maximizarea eficienței administrative și minimizarea costurilor; se pot defini diferite strategii de gestionare, monitorizare, backup, restaurare și alte sarcini administrative bazate pe importanța datelor din fiecare partiție.

➤ *Potrivește depozitul de date cu modelul de utilizare*

Fiecărei partiții îi este permis să fie implementată pe un alt tip de depozit de date, pe baza costului și a caracteristicilor încorporate pe care le oferă depozitul de date; de exemplu, datele binare pot fi stocate în depozitul blob, în timp ce datele mai structurate pot fi stocate într-o bază de date.

➤ *Îmbunătățește disponibilitatea*

Separarea datelor pe mai multe servere evită un singur punct de eșec; dacă o instanță cade atunci doar datele din partiția respectivă nu sunt disponibile, operațiunile pe alte partiții putând continua.

Pentru a determina dacă aveți nevoie de o strategie de partiționare și care ar trebui să fie aceasta, trebuie să luați în considerare trei întrebări despre datele dumneavoastră:

- 1) *Volum*: Cât de mult veți stoca în final?
- 2) *Viteză*: Care este rata la care vor crește datele dumneavoastră?
- 3) *Varietate*: Ce tip de date veți păstra?

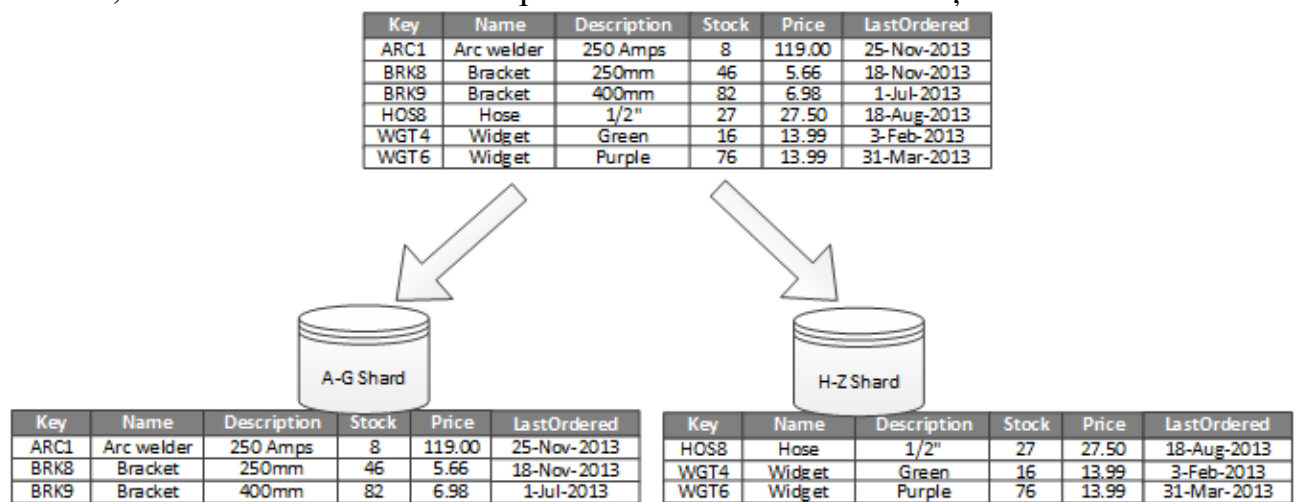
Pentru a vă asigura că nu o să întâmpinați blocaje, trebuie să alegeți cu grijă tipul de partiționare pe care o să îl folosiți astfel încât aplicația să se adapteze eficient pe măsură ce crește.

Există 3 tipuri principale de partiționare a datelor:

- ❖ Partiționare pe orizontală
- ❖ Partiționare pe verticală
- ❖ Partiționare funcțională

Partiționare pe orizontală

- Fiecare partiție este un depozit de date separat, dar toate partițiile au aceeași schemă
- Fiecare partiție este cunoscută ca un shard și deține un subset specific al datelor, cum ar fi toate comenzile pentru un anumit set de clienți

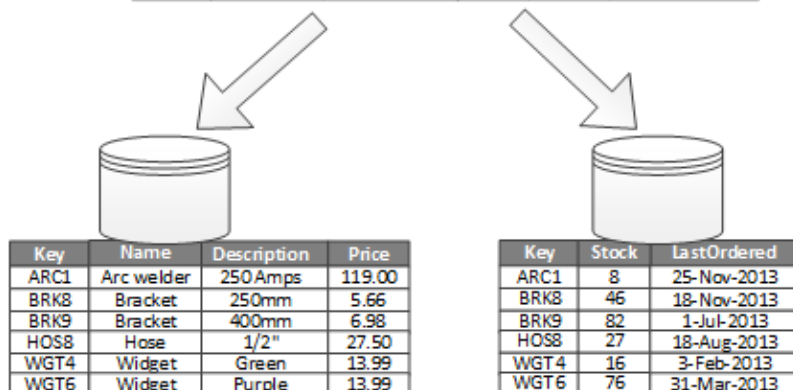


- În exemplul de mai sus, datele despre inventarul produselor sunt împărțite în shards bazate pe cheia produsului; fiecare shard deține datele pentru o gamă continuă de chei shard(A-G și H-Z), organizate în ordine alfabetică
- Cel mai important factor este alegerea unei chei de shard, deoarece poate fi dificilă schimbarea cheii după ce sistemul este în funcțiune
- Shard-urile nu trebuie să aibă aceeași dimensiune, fiind mai importantă echilibrarea numărului de solicitări
- Trebuie evitată crearea unor partiții “fierbinți” care pot afecta performanța și disponibilitatea

Partiționare pe verticală

- ✓ Fiecare partiție deține un subset de câmpuri pentru elementele din depozitul de date; câmpurile sunt împărțite în funcție de modelul de utilizare al acestora; de exemplu, câmpurile frecvent accesate pot fi plasate într-o singură partiție verticală și câmpurile accesate mai puțin frecvent în altul
- ✓ Cea mai obișnuită utilizare pentru partiționarea verticală este reducerea costurilor de I/O și de performanță asociate cu preluarea elementelor accesate frecvent

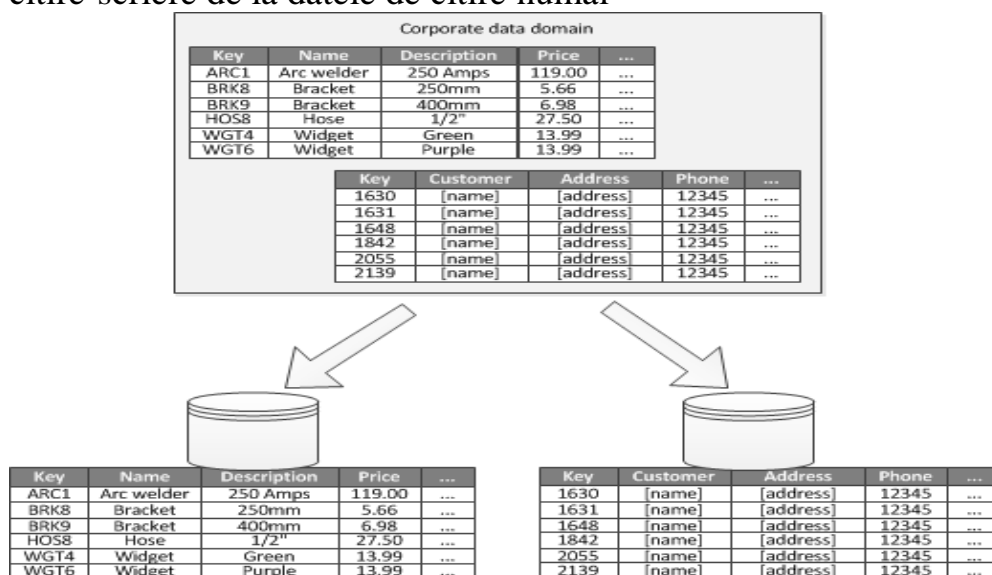
Key	Name	Description	Stock	Price	LastOrdered
ARC1	Arc welder	250 Amps	8	119.00	25-Nov-2013
BRK8	Bracket	250mm	46	5.66	18-Nov-2013
BRK9	Bracket	400mm	82	6.98	1-Jul-2013
HOS8	Hose	1/2"	27	27.50	18-Aug-2013
WGT4	Widget	Green	16	13.99	3-Feb-2013
WGT6	Widget	Purple	76	13.99	31-Mar-2013



- ✓ În exemplul de mai sus, diferite proprietăți ale unui element sunt stocate în diferite partiții; o partiție deține datele care sunt accesate mai frecvent, inclusiv numele produsului, descrierea și prețul; o altă partiție deține datele privind inventarul: numărul stocurilor și data ultimului ordin
- ✓ Datele sensibile pot fi stocate într-o partiție separată la care se adaugă securitate adițională
- ✓ Partiționarea verticală operează la nivelul entității într-un depozit de date

Partiționare funcțională

- datele sunt agregate în funcție de modul în care acestea sunt utilizate de fiecare context limitat din sistem
- de exemplu, un sistem de comerț electronic ar putea să stocheze datele facturii într-o singură diviziune și datele inventarului de produse în altul
- partiționarea funcțională este o modalitate de a îmbunătăți performanțele de izolare și acces la date
- o altă utilizare obișnuită pentru partiționarea funcțională este separarea datelor citire-scriere de la datele de citire numai



- figura de mai sus prezintă o prezentare generală a partiționării funcționale în care datele de inventar sunt separate de datele clientului
- această strategie de partiționare contribuie la reducerea privind accesul la date în diferite părți ale unui sistem

Aceste strategii pot fi combinate și este recomandat să fie luate în considerare atunci când se proiectează o schemă de partiționare.