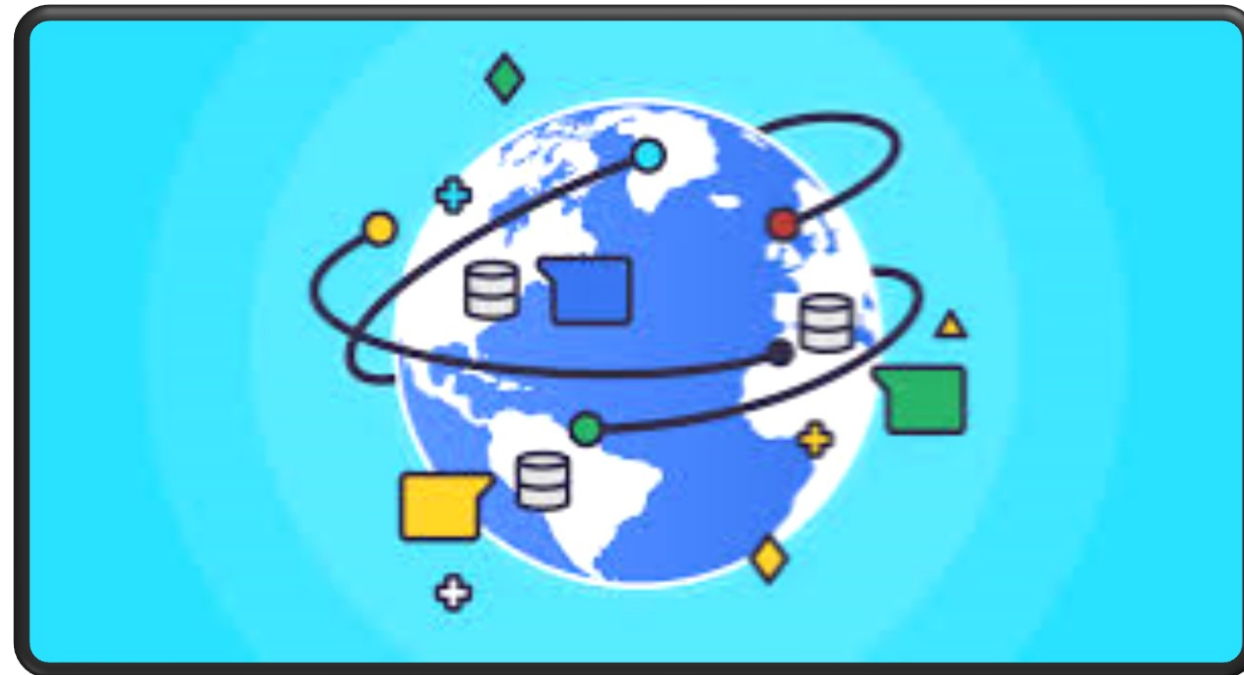


Bases de dades distribuïdes



Objectius



- Identifica les bases de dades distribuïdes: utilitat, diferències, avantatges i inconvenients, distribució de les dades, arquitectura, seguretat i recuperació.
- Identifica les bases de dades centralitzades i les bases de dades distribuïdes: utilitat, diferències, avantatges i inconvenients.

Bases de dades distribuïdes

- Un dels sectors informàtics on més s'està evolucionant darrerament, tot integrant el desenvolupament tecnològic amb la innovació metodològica, és el relatiu als sistemes distribuïts d'informació.
- Per molt dimensionat que tinguem un servidor (memòria, CPUs, disc durs) si anem afegint usuaris arribarà un moment que per **limitacions físiques** no podrà donar servei.
- En entorns d'internet on no podem controlar el nombre d'usuaris que es connectaran al nostre sistema es molt important poder **distribuir la càrrega**.



Bases de dades distribuïdes

- El primer problema que tenim que resoldre és com distribuïrem la nostra base de dades. Imaginem que tenim una taula 'R' existeixen dues formes de distribuir-la:
 - **Rèplica:** El sistema guarda varies còpies de les dades en diferents ubicacions. (redundància intencionada).
 - **Fragmentació:** Consisteix en dividir les dades en diferents trossos de manera que a partir dels fragments, podem recomposar-los.
 - Horitzontal: Dividim les taules en conjunts de registres complets.
 - Vertical: Dividim les taules en els seus camps corresponents.

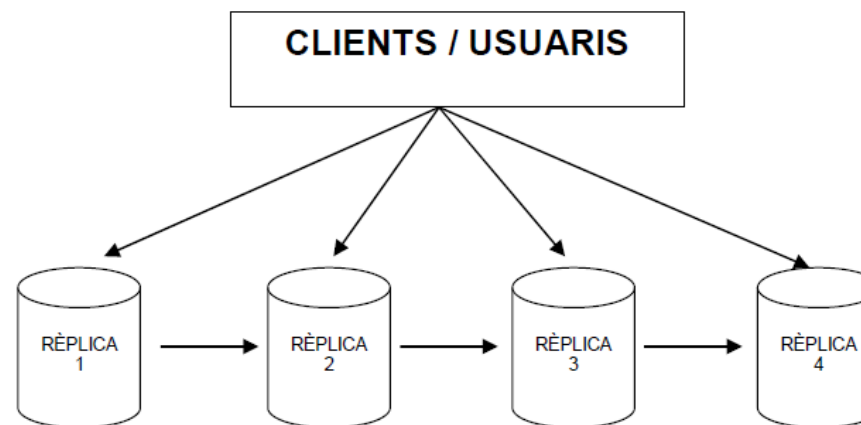
Bases de dades distribuïdes - Rèplica

Avantatges

- Facilitarà molt les operacions de lectura, ja que disposarem de les mateixes dades en diferents localitzacions. Això comportarà que augmenti molt la disponibilitat de les base de dades.
- Repartiment de càrrega: Diferents executant consultes complexes

Inconvenients

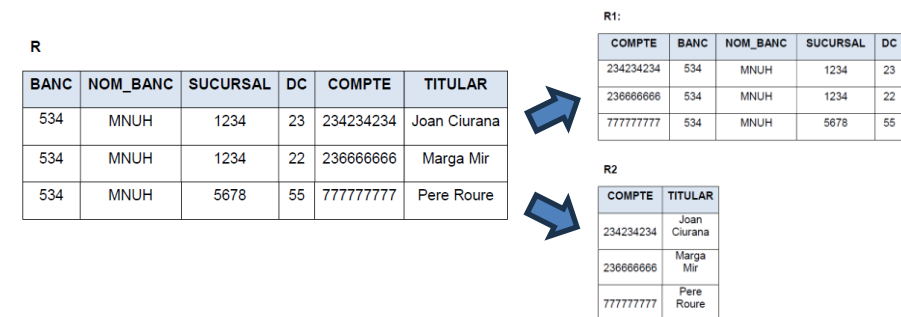
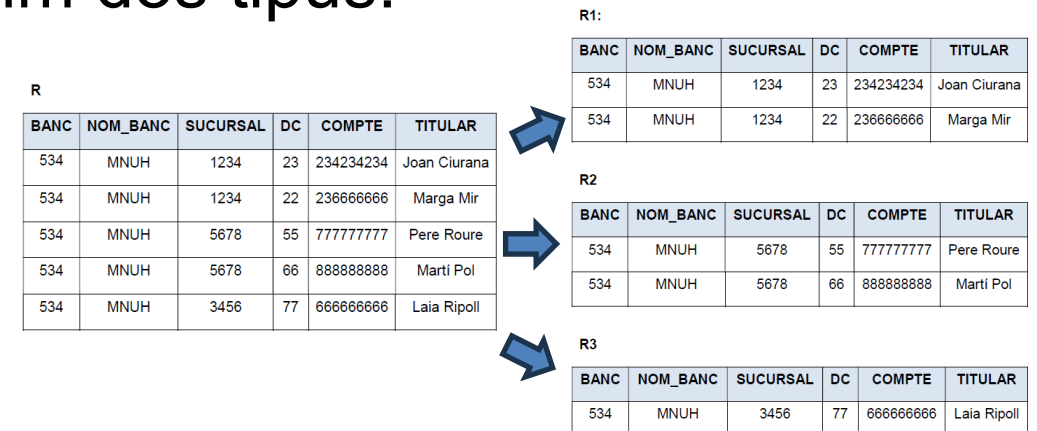
- Operacions d'actualització complexes.
- Cada vegada que haguem de modificar una dades, hem de realitzar-la a totes les ubicacions al mateix temps, ja que sinó perdríem consistència de els dades.



Bases de dades distribuïdes - Fragmentació

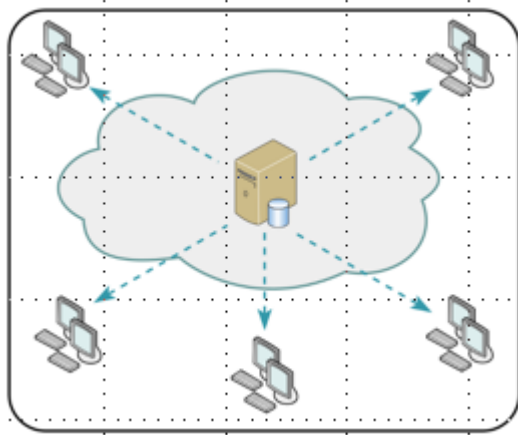
- Fragmentar una relació (taula) R, consisteix en dividir-la en diferents relacions: R1, R2, R3,...Rn. De manera que a partir d'aquests fragments, podem recomposar la relació original R. Tenim dos tipus:

- Fragmentació horitzontal:** Les taules les dividim en conjunts de tuples completes. Per poder crear una fragmentació horitzontal utilitzarem l'operador de selecció o delta (δ) sobre la relació original en el model relacional.
- Fragmentació vertical:** Si dividim les taules per un conjunt de camps. Per crear aquesta fragmentació vertical utilitzarem l'operador de projecció π (π). Si volem recomposar la relació utilitzarem l'operador de join natural (\bowtie).

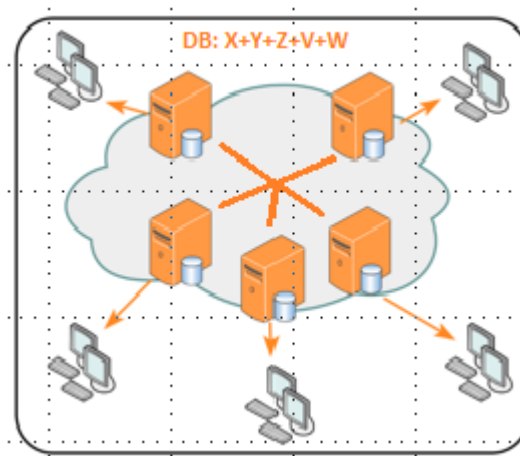


Exemples de distribució

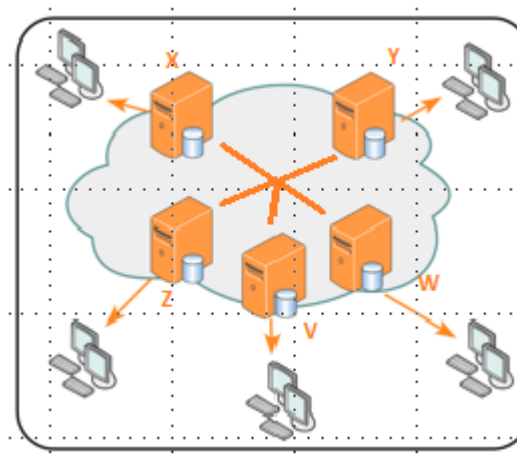
BD Centralitzada



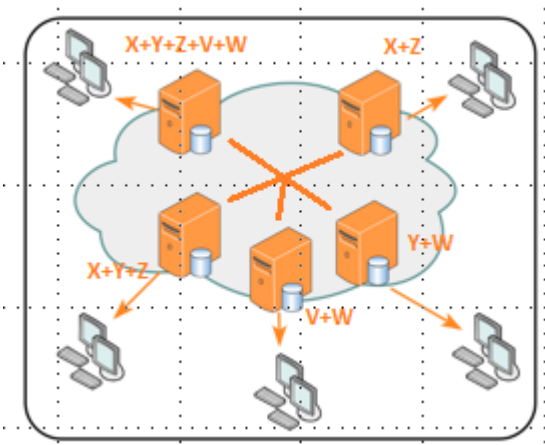
BD Multiplicada



BD Particionada



BD amb node principal o seleccionat



Avantatges

Simplicitat en l'administració i en la generació de consultes o modificació de dades

La base de dades està a tots i cadascun dels nodes. Si apareix un problema en qualsevol dels nodes els altres seguirien funcionant sense problemes.

Ocupa el mateix espai que la base de dades de forma centralitzada. Les operacions d'actualització són molt senzilles, només hem d'actualitzar les dades en un sol node

Les actualitzacions només es realitzarien en un sol node al principal i les consultes en els nodes replicats.

Inconvenients

Problemes de rendiment en el cas de molts usuaris encara que el tinguem molt dimensionat. Problemes de integritat de dades en el cas de un desastre

Necessitem 4 vegades l'espai normal d'emmagatzematge. Les operacions d'esborrat, inserció i modificació de les dades l'haurem de replica en tots els nodes, generant possibles problemes de concurrència i inconsistència.

Si aparegués algun problema en algun dels nodes, no podríem accedir a part de les dades. Les consultes de selecció impliquen accedir a varis nodes

Si cau el node principal haurem de tenir un mecanisme perquè les dades o bé es repliquin o bé construir un nou node principal en base a les seves rèpliques

Pregunta

- Perquè creieu que van aparèixer les CDN?



WEBGRAFIA

- Batini, C.; Ceri, S.; Navathe, S.B. (1992). Conceptual Database Design: An Entity-Relationship Approach. Reading, Massachusetts: Addison Wesley.
- Teorey, T.J. (1999). Database Modeling & Design. The Fundamental Principles (3a ed.). San Francisco: Morgan Kaufmann Publishers, Inc.