

# Framtíð gagnagrunna í nútíma samfélagi



Tómas Helgi Jóhannsson  
Gagnagrunnstjóri  
Reiknistofa bankanna

# Hvað er gagnagrunnur ?

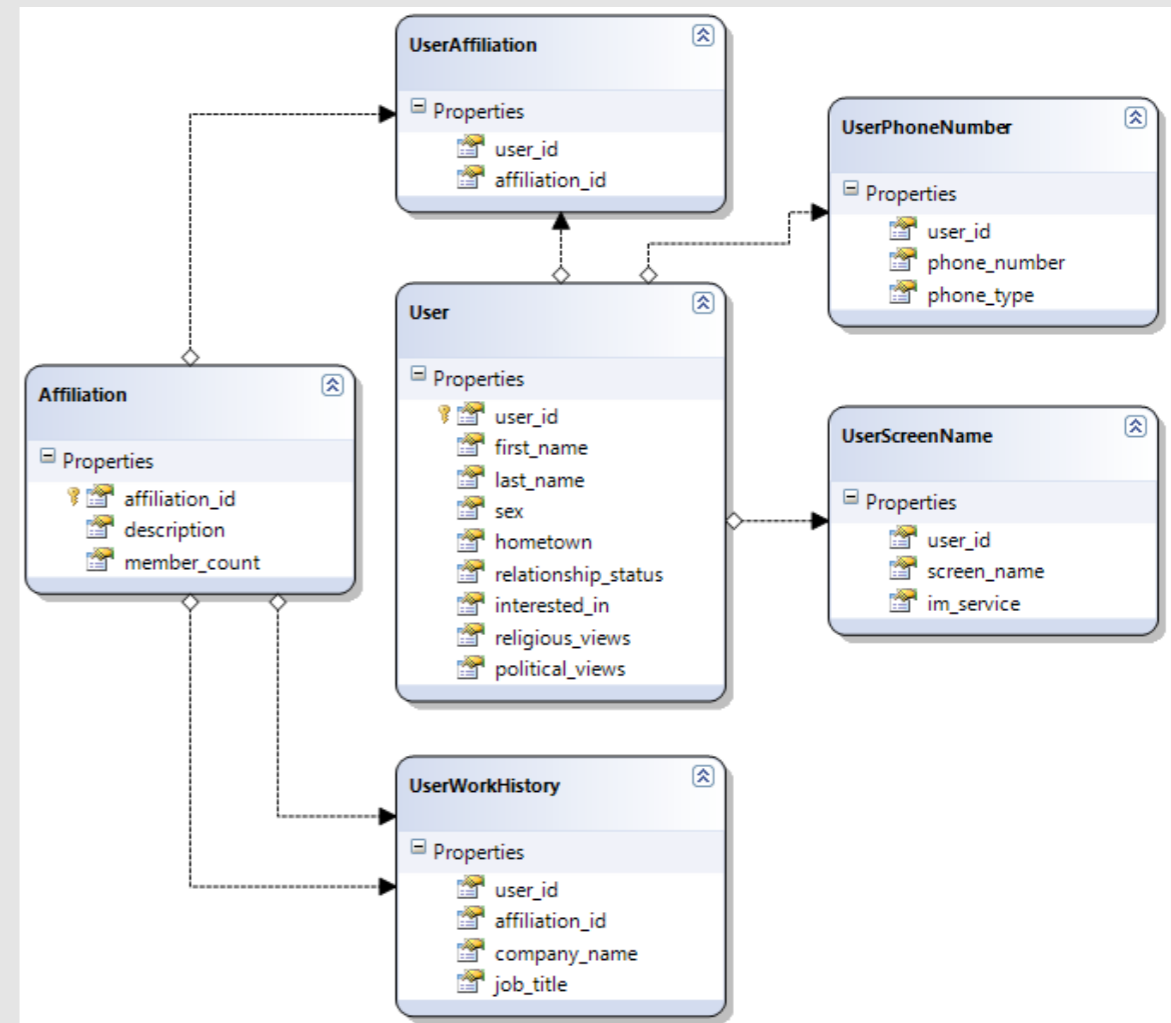
- Safn gagna sem
  - vensluð eru á einn eða annan hátt
  - mótuð eftir titeknu gagnalíkani og gagnahögun.
- Þróun gagnagrunna hefur tekið umtalsvert stökk á undanförunum áratugum.
- Fjölbreyttar gagnagrunnsútfærslur sem henta tilteknum verkefnum og vinnslum komið fram og orðið vinsælar.

# Tegundir fyrirspurna

- SQL fyrirspurnarmálið langalgengasta
- Transactional (OLTP)
  - Stuttar og hnitmiðaðar fyrirspurnir t.d. í banka og verslunarkerfum.
- Analytical (OLAP)
  - Tölfræðifyrirspurnir sem vinna á miklu gagnamagni
- Fyrirspurnir geta verið einfaldar og uppí það að vera mjög flóknar allt eftir fjölda taflna sem tengdar eru saman í fyrirspurninni.

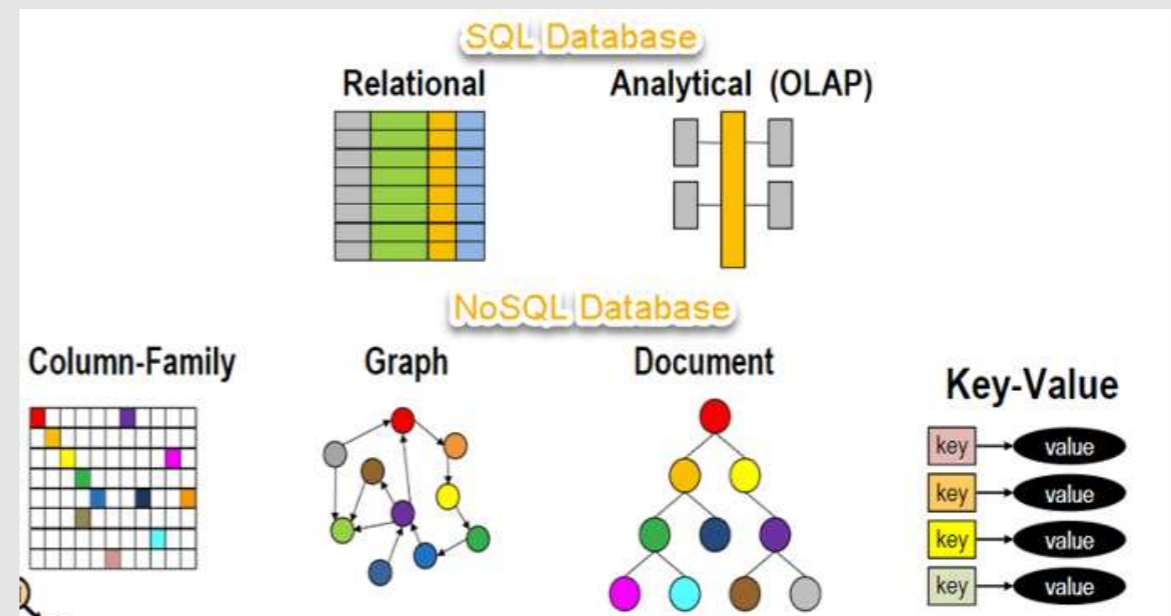
# Töflugagnagrunnar

- Hannað af Edgar F Codd 1960 og hefur verið í stöðugri þróun síðan.
- Langalgengasta form gagnagrunna
- Langmest notað í verslunar og fjármálakerfum.
- Geymir langstærsta hluta þess gagnamagnins sem til verður á heimsvísu.
- Upphaflega einungis row-level högun en síðari ár einnig á colum-level högun
- Seint að aldrei hverfa af sjónarsviðinu



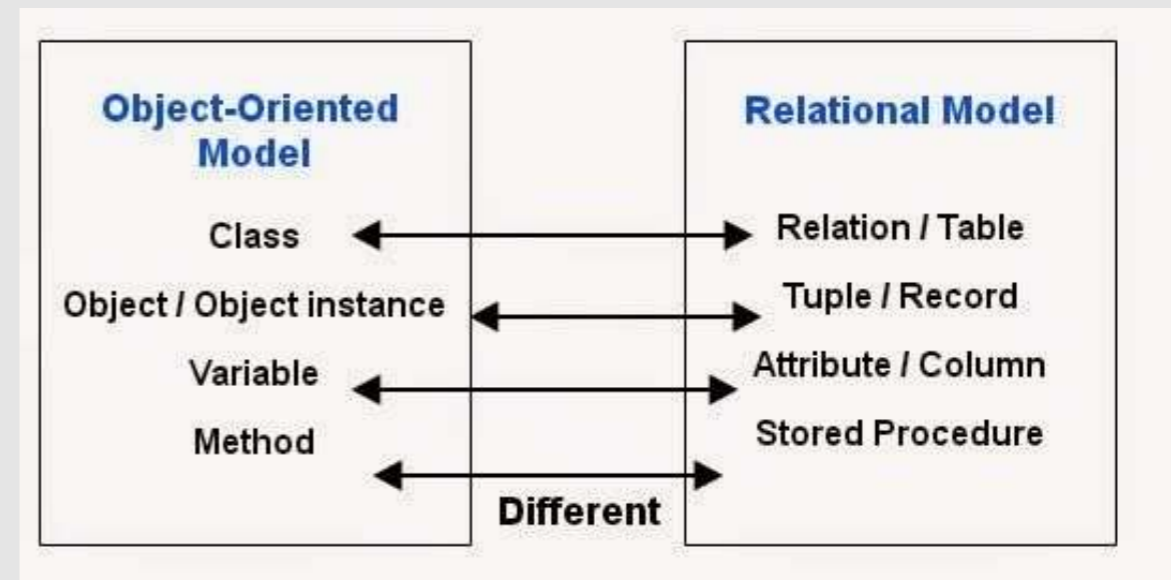
# Skemalausir gagnagrunnar (NoSQL)

- Hafa í raun verið til síðan 1960 en NoSQL
- Einföld gagna-líkön/-högun
- Auðvelt að skala “horizontaly”
- Sveiganlegri og hægt að tengja auðveldlega beint við forritunarmál.
- Ákveðnar aðgerðir hraðvirkari en í töflugrunnum
- Hefur meiri áhrif á áreiðanleika gagna ef ekki er rétt með farið heldur en töflugagnagrunnur.



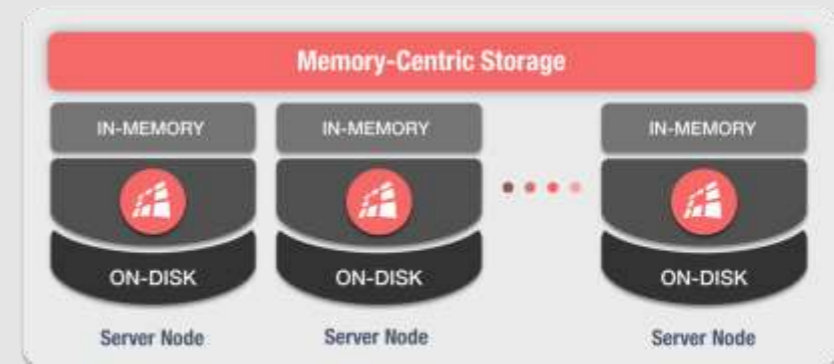
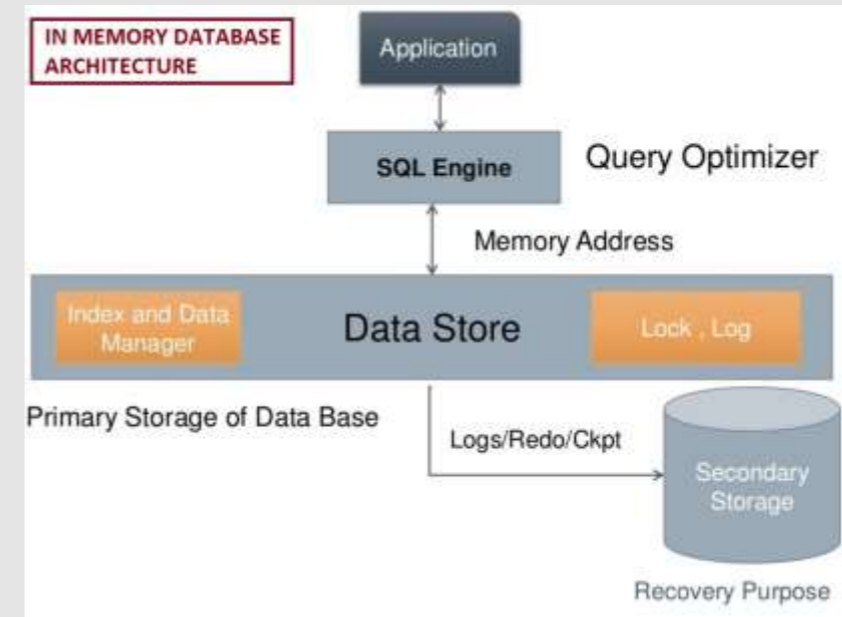
# Hlutbundnir gagnagrunnar

- Komu fram á sjónarsviðið 1990
- Náin tengsl við hlutbundin forritunarmál á borð við C++ og Java
- Gögnin vistuð “as-is” í gagnagrunni, þ.e.a.s. engin umbreyting á formi í minni forritis yfir í gagnagrunninn.
- Flókin indexering þegar flækjustig eykst



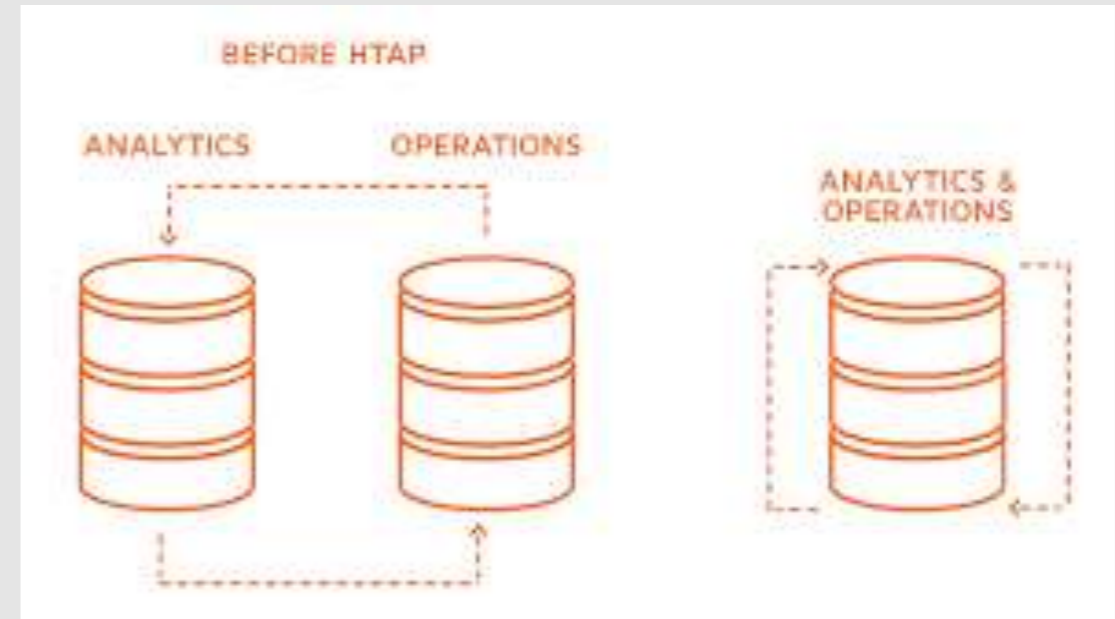
# InMemory gagnagrunnar

- Gögn geymd í minni vélar í stað disks eins og í hefðbundnum grunnum.
- Margfalt hraðari leit og vinnsla gagna.
- Dæmi eru um að vinnslur hafi styst úr mörgum klukkustundum í fáeinar sekúndur.
- Gögn glatast þegar rafmagn fer.
- Framtíðarhögun með samspili við hefðbundna grunna.



# Hybrid gagnagrunnar - HTAP

- Hybrid Transactional Analytica Processing – HTAP
- Vinnur jafn með OLTP og OLAP fyrirspurnir
- Töflur annaðhvort á row-level eða column-level formi eða hvoru tveggja.
- Column level form best í að summera upp dálka sbr. Excel.
- Heilinn í grunninum (optimizer) ákveður hvort hann notar row eða column level töflur eða hvoru tveggja.

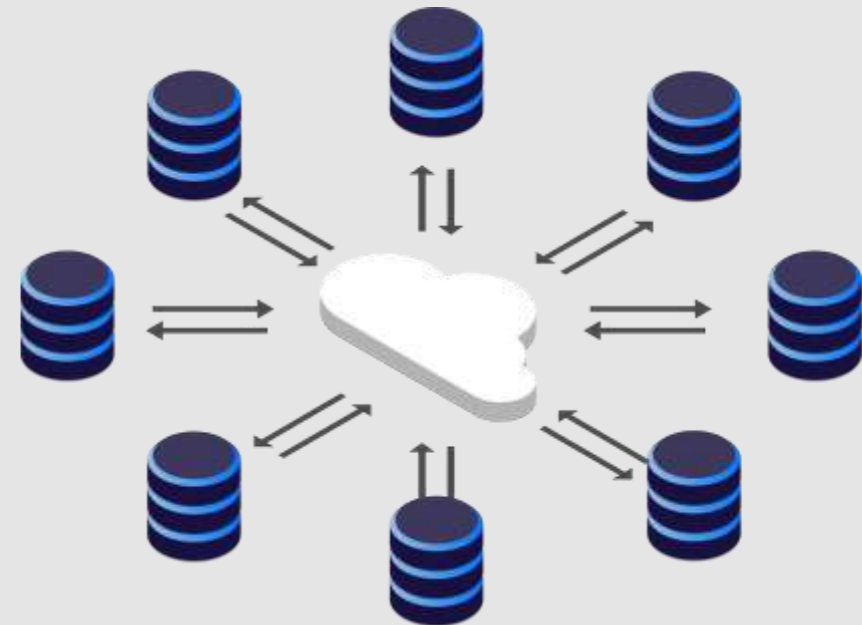


- Dæmi um að flóknar OLAP fyrirspurnir hafi farið úr mörgum mínútum/klukkustundum niður undir sekúndu/ í tíu sekúndur.

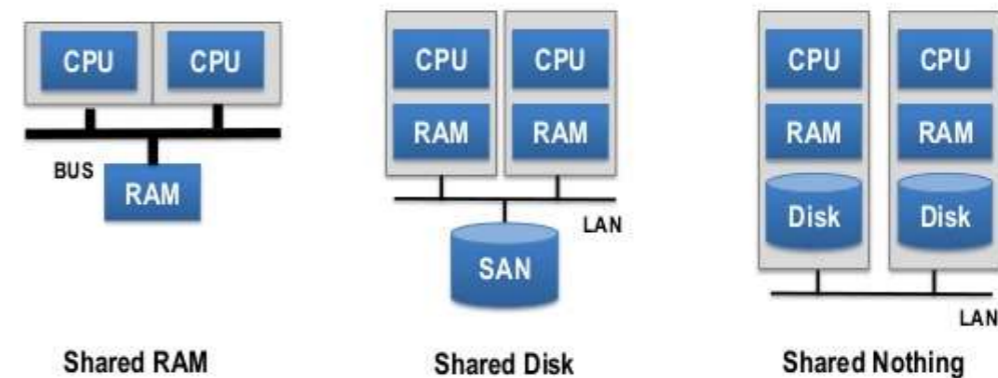


# Cluster gagnagrunnar

- Samanstendur af þremur eða fleiri eintökum af sama gagnagrunni.
- Dæmi um slíka grunna:
  - Oracle RAC
  - Db2 PureScale
  - MariaDB Galera Cluster
  - SQL Server Cluster
- Gögn dreifð milli þriggja eða fleiri véla á sameignlegu diskapooli.
- Sömu gögn dreifð og aðgengileg milli þriggja eða fleiri véla.
- Álag dreifist milli véla.

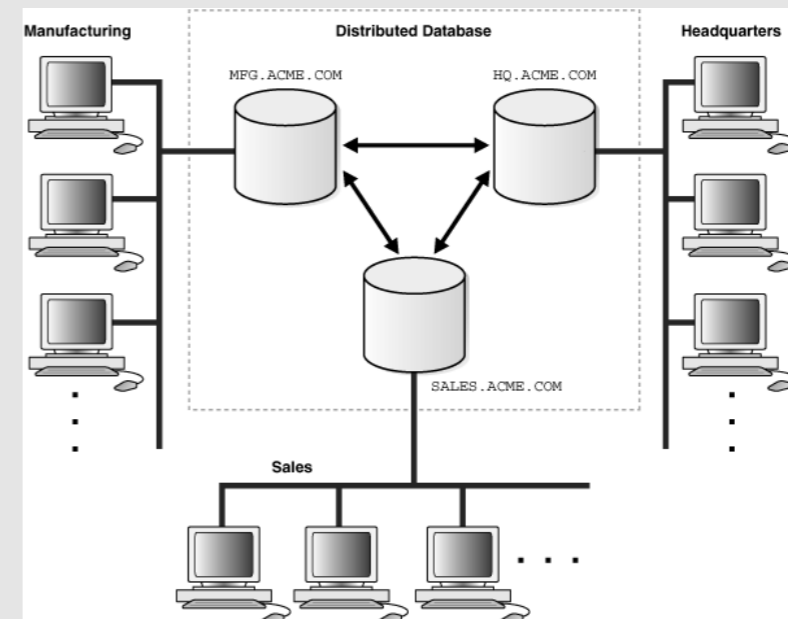
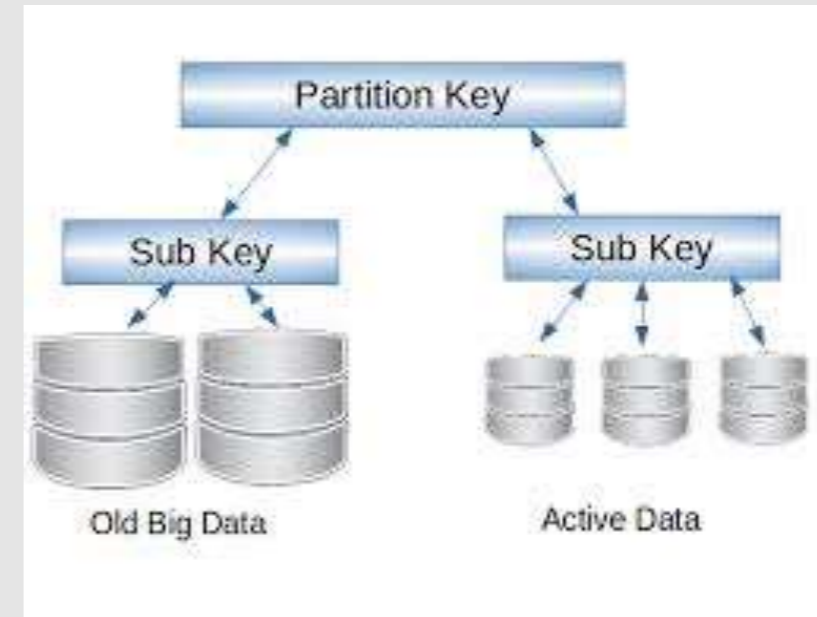


## Shared Nothing Architecture



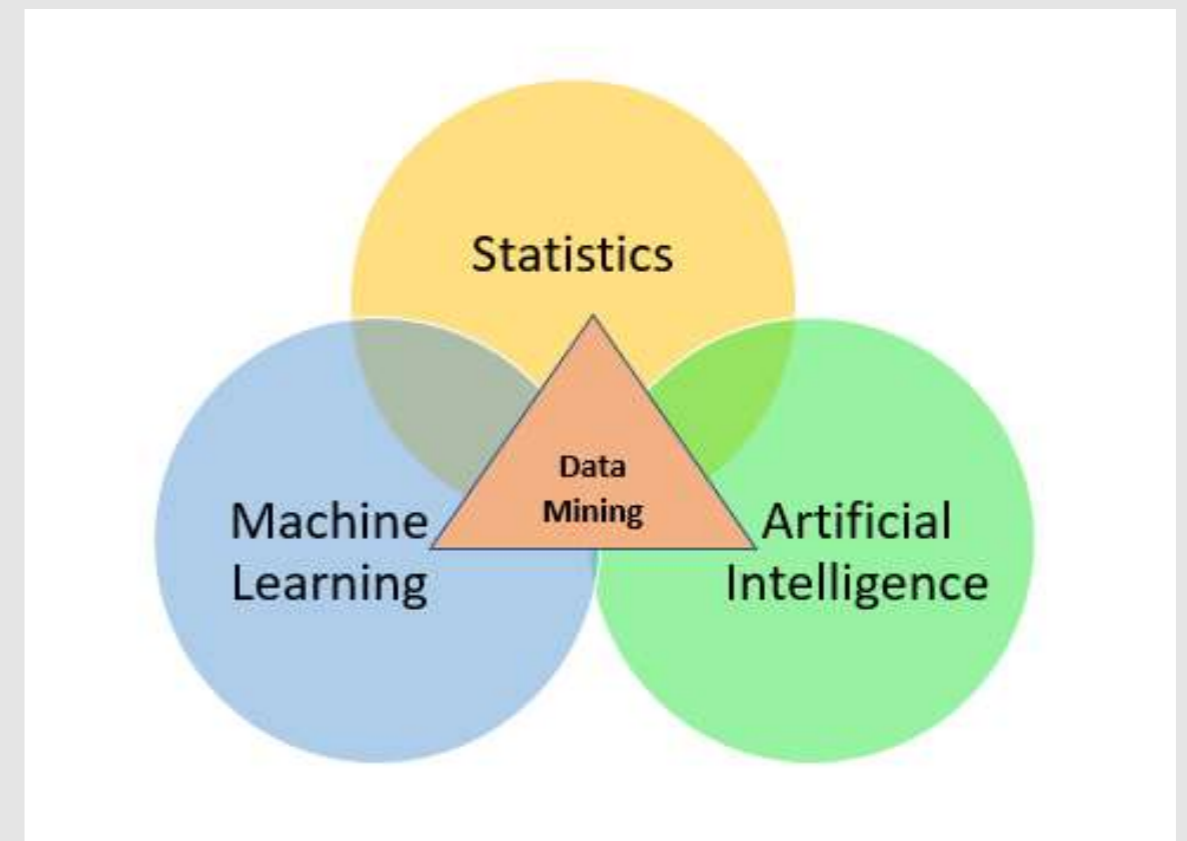
# Dreifing gagna

- Gagnamagn hefur aukist margfalt undanfarna áratugi og engin endir á því.
- Internetið og snjalltæki þar fremstu áhrifavaldarnir.
- Tímafrekar fyrirspurnir ef röng högun og/eða indexun er ábótavant.
- Table partitioning gefur möguleika á að sleppa að leita í gögnum sem ekki falla að leitaraskilyrðum.
- Töflum eða töflupörtum hægt að dreifa milli véla til að dreifa álagi og auka vinnsluhraða margfalt.



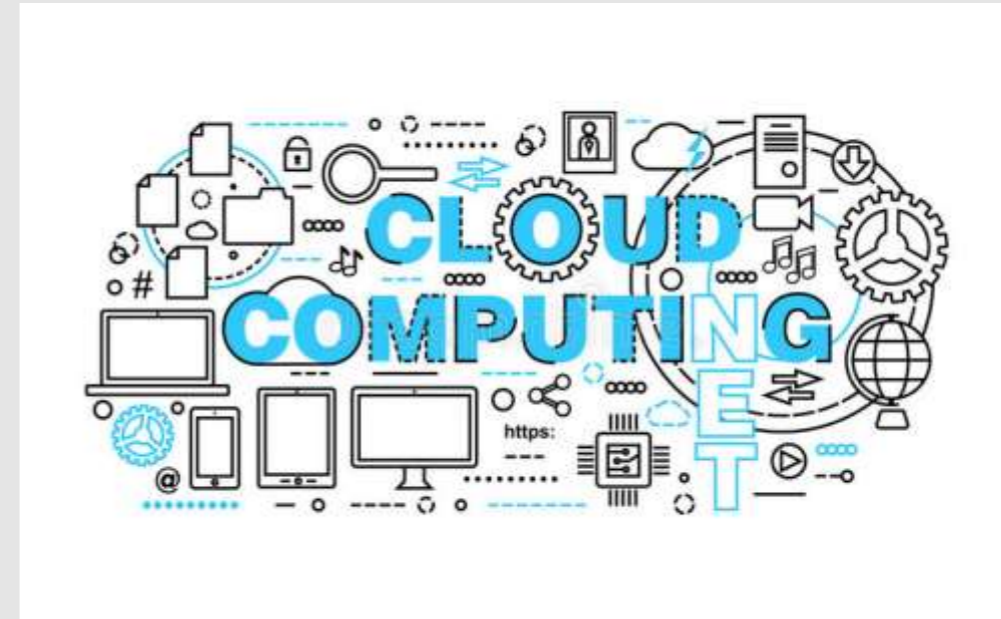
# Gerivigreind (AI) í gagnagrunnum.

- Heilinn í gagnagrunnum (optimizer) ruglast og velur ranga leið að gögnum þ.s. mikið gagnamagn og flækjustig þeirra og fyrirspurnar er hátt.
- Gervigreind og vélnám notað til að styrkja heilann til að takast á við slíkar aðstæður.
- IBM og Oracle komnir lengst í að tengja AI og ML við sína gagnagrunna.



# Gagnagrunnar framtíðarinnar

- Munu innihalda gríðarlegt gagnamagn
- HTAP Cluster grunnar verða áberandi og styðja fjölbreytt form gagna.
- InMemory grunnar fá aukið vægi.
- Gervigreind og vélnám koma sterkt inn sem viðbót við gagnagrunnsheilann (Optimizer)





**Takk fyrir!**