An Introduction to HTAP

Wenliang ZHANG
https://zedware.github.io

Outline

- What is HTAP
- Why HTAP
- Different types of HTAP
- Design space
- Conclusion

What is HTAP

- OLTP
    - Single row insert/update
    - Key lookups (point query [with index])
    - High concurrency
    - High throughput
    - Low response time

What is HTAP

- OLAP
  - Complex queries with lots of joins
  - Scans lots of data
  - Low concurrency etc.
  - Ad hoc queries

What is HTAP

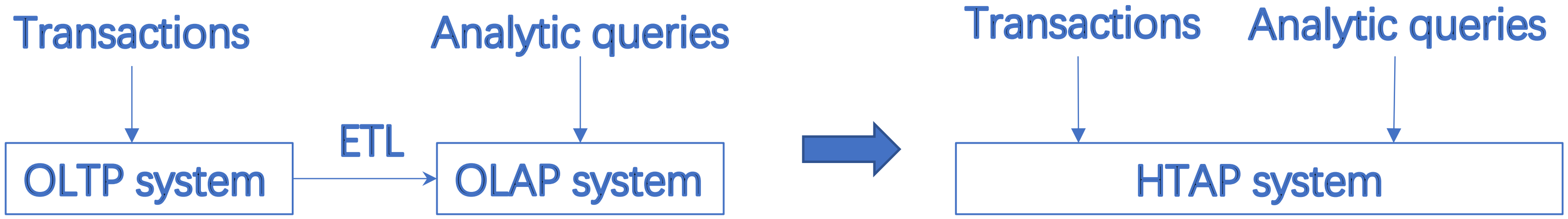
- Hybrid transactional and analytical processing
  - A single system for both transactions as well as reporting queries
  - Loosely defined vs. strictly defined
    - Most systems are solutions, using the loosely definition
- Gartner definition
  - An HTAP system can execute both I/U/D statements as well as analytical queries over all data, including the recently ingested in this transaction, within the same transaction

What is translytical?

- Forrester definition
  - A unified and integrated data platform that supports multi-workloads such as transactional, operational, and analytical simultaneously in real time, leveraging in-memory capabilities including support for SSD, flash, and DRAM and ensures full transactional integrity and data consistency.

Why HTAP

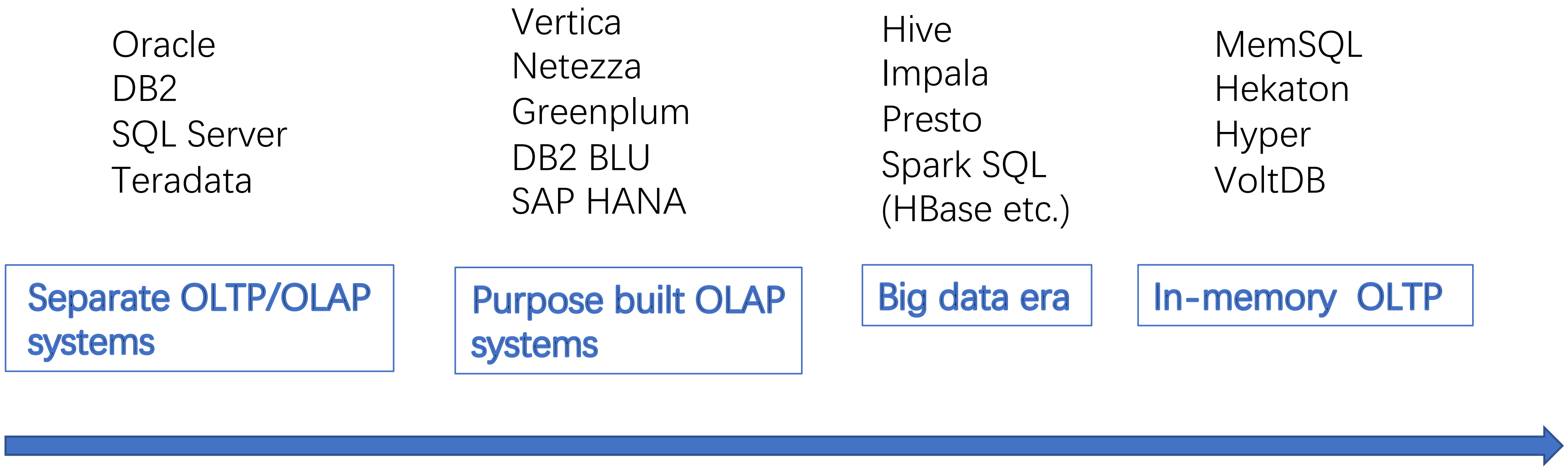
- Realtime analysis of transactional data
  - Analysis on fresh/real-time data
  - Increase business agility
  - Reduce business risks
- Traditional solutions



- Comparison with master/slave

Why HTAP

- The Evolution of products



Oracle
DB2
SQL Server
Teradata

Separate OLTP/OLAP systems

Vertica
Netezza
Greenplum
DB2 BLU
SAP HANA

Purpose built OLAP systems

Hive
Impala
Presto
Spark SQL
(HBase etc.)

Big data era

MemSQL
Hekaton
Hyper
VoltDB

In-memory OLTP

Different types of HTAP

- Most are loosely coupled systems
  - esp. the competitors in China
  - From pure Java OLAP to HTAP
  - From PostgreSQL OLAP to HTAP
  - From MySQL OLTP to HTAP

Different types of HTAP

- Loosely coupled systems
  - Separated OLTP and OLAP engines
- Tightly coupled systems
  - Separate engines but integrated
- A single system
  - A single engine for both workload

Design space

- Single engine vs. separated engines
  - For data ingestion and analytic queries
- Single storage option vs. different options
  - For transactional data and historical data
- Data format
  - Row format, columnar format or mixed format
- Data recency
- Transactional semantics
- Index support

Design space

- Data format leads to big difference
  - Row, columnar, mixed, or even dynamically organized
- Transactional semantic is the biggest difference
  - Play tricks with ACID
  - Single row vs. multiple row transactions
  - Weak TX consistency or use the misleading consistency of CAP
- Index support is hard for some systems
  - Traditional B+ index etc.
  - Secondary index
  - Local index vs. global index

Conclusion

- HTAP is strictly defined by Gartner
- Wildly used in the loosely speaking in industry
- Convergence of SQL database and Big Data
- The major use case is real-time data analysis