

The Challenges of Open Source Cloud Database



目录

CONTENT

- 01** | About Oushu
- 02** | Background
- 03** | Cloud Database
- 04** | Apache HAWQ

About Oushu



Founded by Apache HAWQ core team. Apache HAWQ is 1st Apache Database TLP initiated by team in China



Focusing on AI and Big Data. Products: Cloud Database **OushuDB/HAWQ**、**Littleboy AI**、**Lava**



The first Chinese Company selling database to top US companies. Hundreds of enterprise users across the world



Most team members are from **EMC, Oracle, IBM, Teradata, Google, Amazon** et al.



Invested by Sequoia and Red point. Microsoft accelerator 2018



Core members hold dozens of US patents, Work published on top database conference: SIGMOD



Oushu Customers & Partners

Finance



Energy

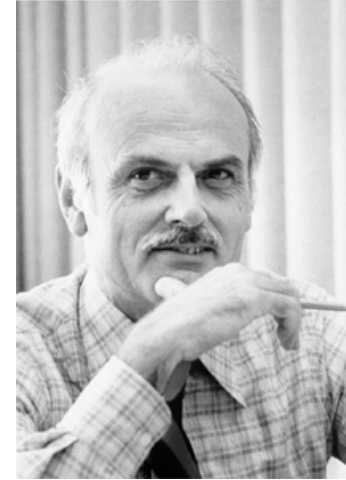


Telcom



Database History : 57 years

- Database: 1962
 - Inverted File Database System
 - System Development Corporation
- Several phases of Database
 - 1960s: Navigational DBMS (network & hierarchical)
 - ✓ Integrated Data Store (IDS)
 - ✓ Information Management System (IMS)
 - 1970s - 1990s: SQL/Relational DBMS
 - ✓ OLTP, Data warehouse
 - 2000s - Present
 - ✓ MPP, Hadoop, NoSQL (XML, KV, Graph, Tree), NewSQL, Cloud Database



Cloud Database

- Public cloud : Database as a service
 - Amazon Redshift (ParAccel MPP)
 - RDS (PostgreSQL, MySQL et al)
 - OushuDB
- Private Cloud: Virtual machine/Docker container

Cloud Database vs Traditional Database

- Same

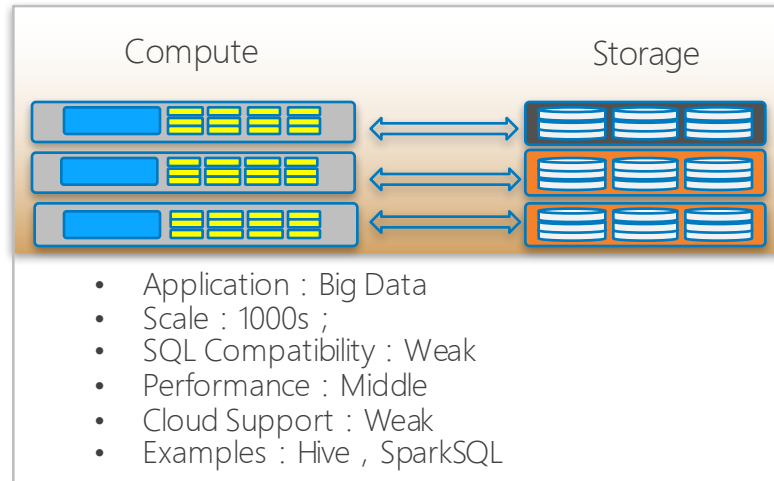
- Data model & Querying Language
- Query optimization & Execution
- Indexes & Storage
- Transaction Processing
- et al

- Difference

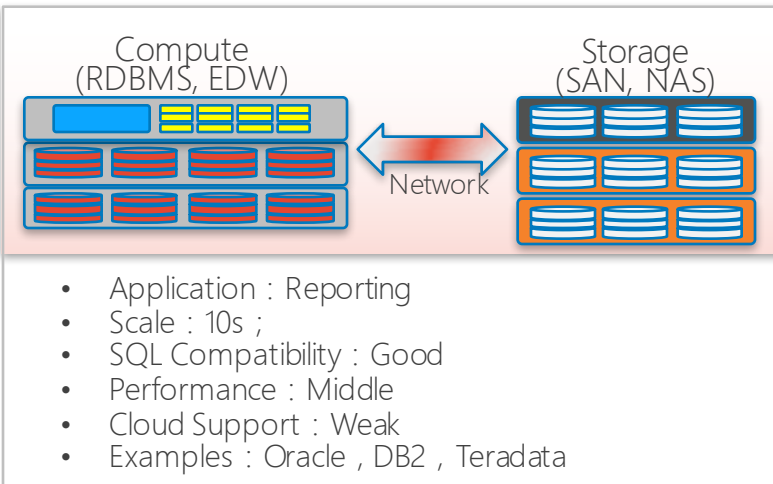
- How users use the database
- Billing
- Running Environment
 - ✓ Virtualization : Resource Management
- Ecosystem
 - ✓ Infrastructure Services : S3 et al.
- Elasticity
- Security

The Evolution Path of Analytical Database

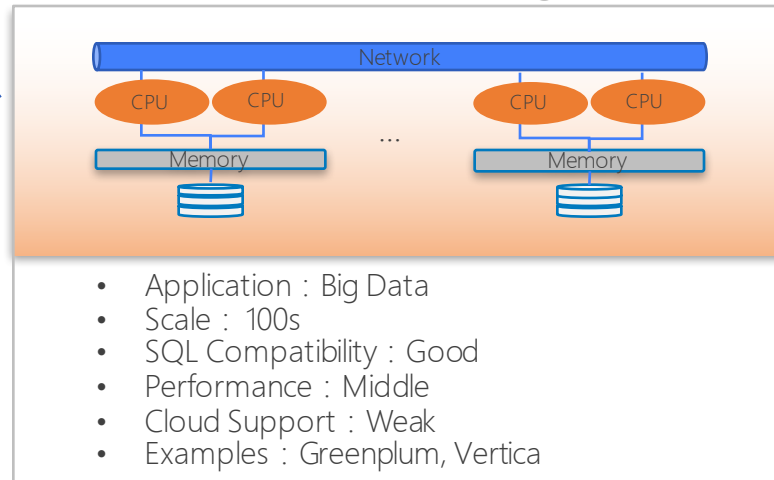
SQL-on-Hadoop Separation of Storage & Compute (2000s)



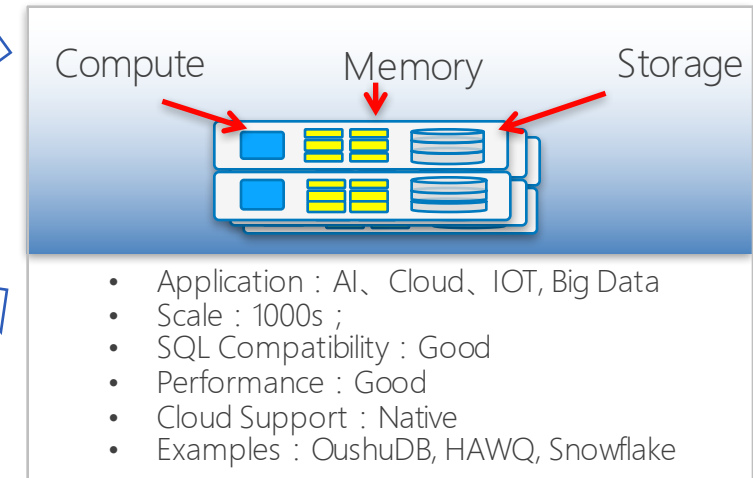
Traditional DW Dedicated Hardware (1980s)



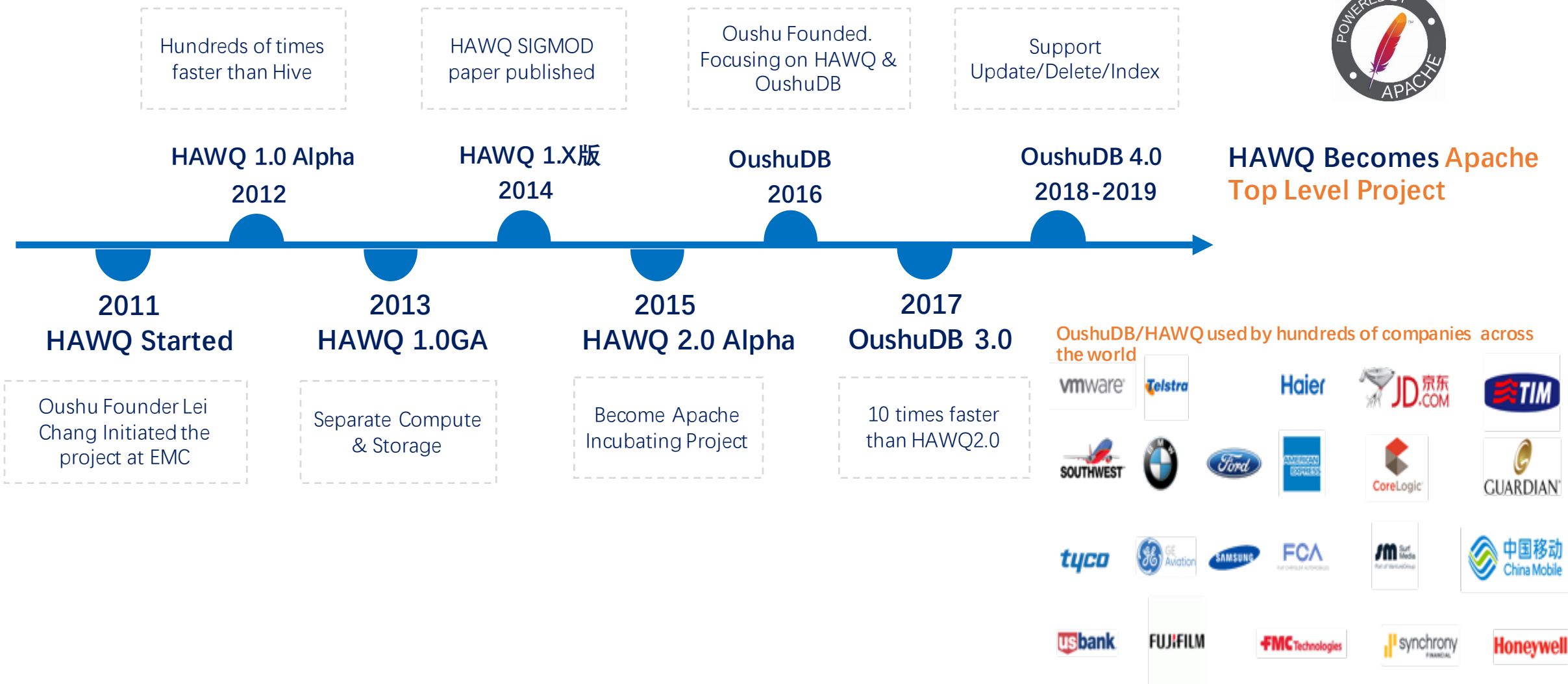
X86 MPP : Share Nothing (2000s)



Cloud Native (about 2015)



HAWQ & OushuDB

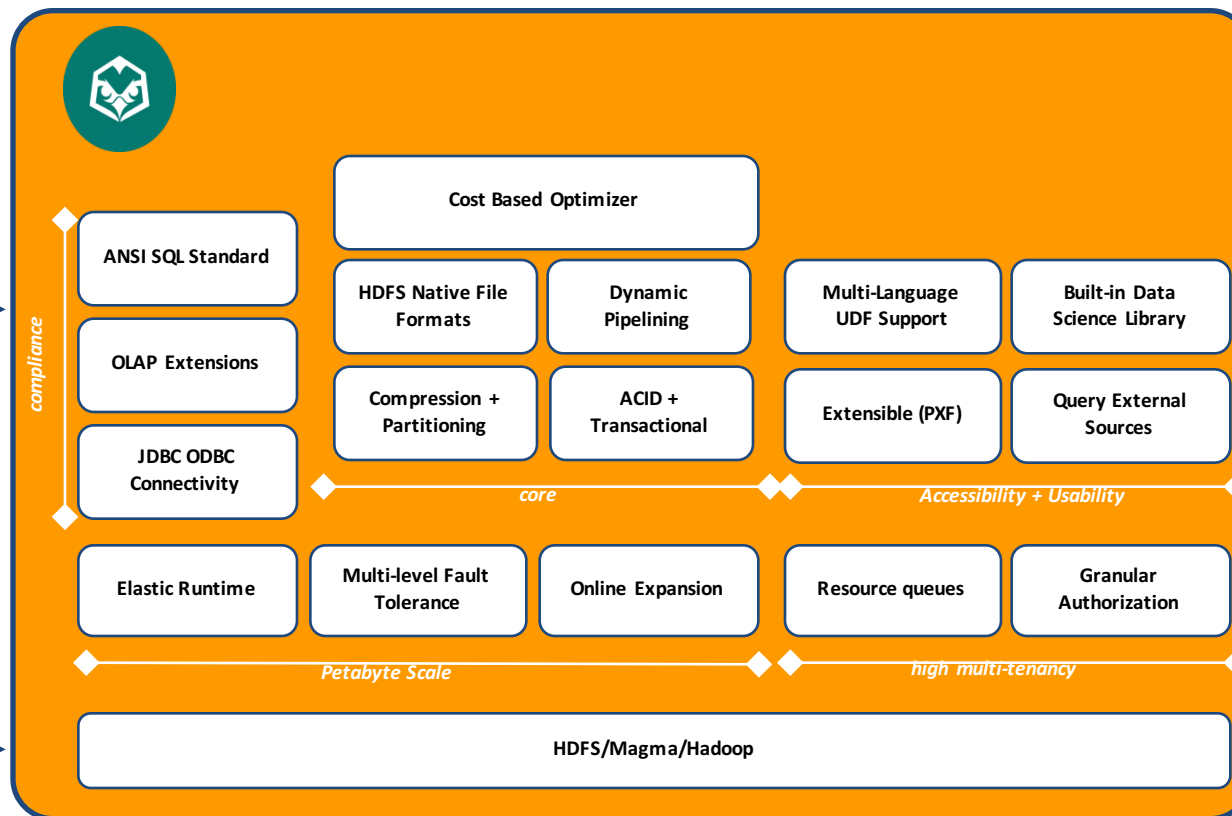


HAWQ Main Features



- Leverage Existing Skills & Tools
- Easily Integrate with Other Tools

- Well Integrated with Hadoop Ecosystem

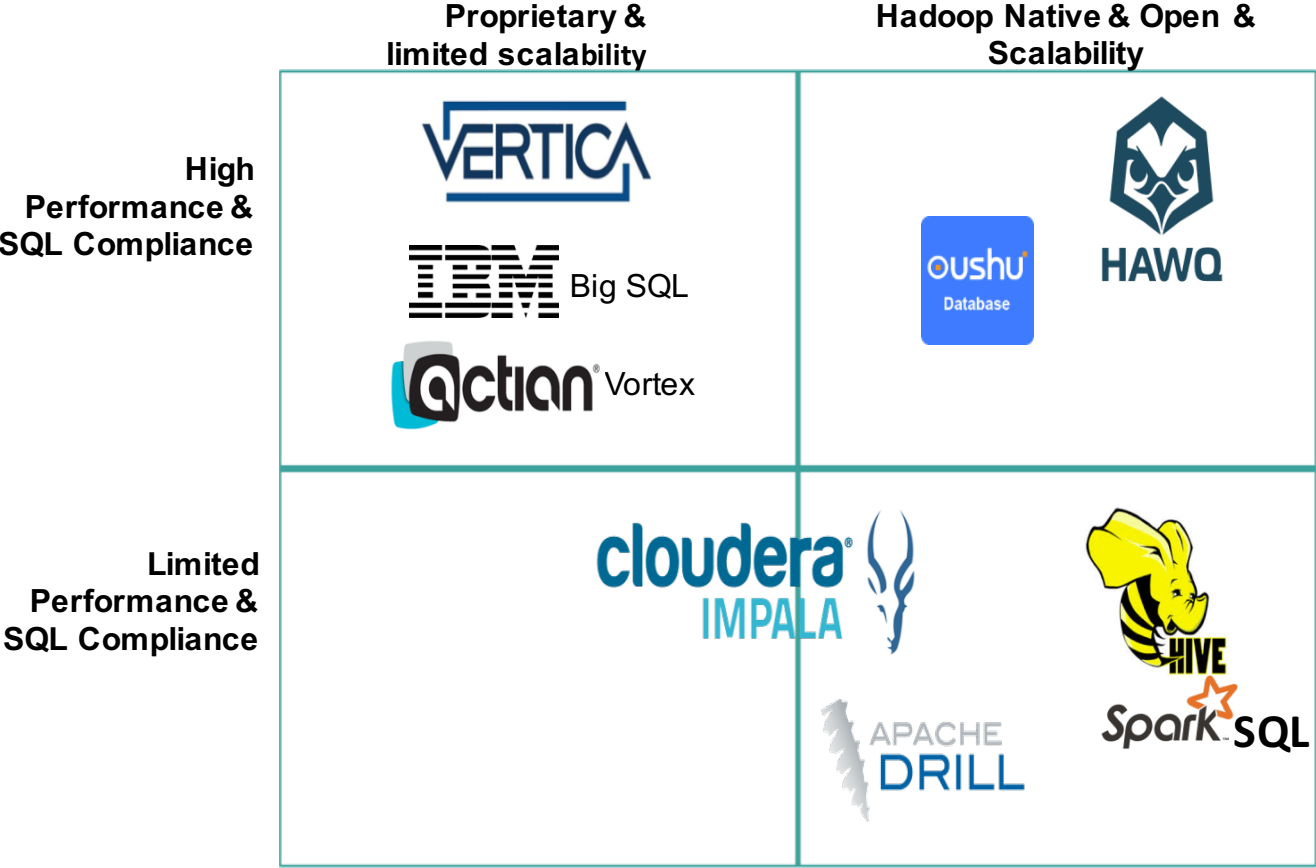


- Discover New Relationships
- Enable Data Science
- Analyze External Sources
- Query All Data Types!



- Manage Multiple Workloads
- Petabyte Scale Analytics
- Sub-second Performance

HAWQ vs Others



Contributing to HAWQ

- Documentation
- Wiki
- Bug reports
- Bug fixes
- Features
- Website: <http://hawq.incubator.apache.org/>
- Wiki: <https://cwiki.apache.org/confluence/display/HAWQ>
- Repo: <https://github.com/apache/hawq>
- JIRA: <https://issues.apache.org/jira/browse/HAWQ>
- Mailing lists: dev/user@hawq.apache.org

Code contribution process

- Start a JIRA
- Fork a github repo: <https://github.com/apache/hawq.git>
- Clone your repo to local
- Add the github repo as “upstream”
- Create a feature branch and commit your code
- Start a pull request for code review



感谢观看



让人类只为兴趣而工作