



Interoperability vs. Performance

Can you have the best of both?



Dr. Kim Kuen Tang
Sr. kdb+ Solution Architect

KX STAC M3 Benchmarks - Restrictors



KX is the database in SUT with fastest mean response time in:

- 24 out of 24 Kanaga (3 queries, large data: 60 TB) tests
- 15 out of 17 Actuco (7 queries, medium data: 3.5 TB) tests



Interoperability – functionality and performance

Native SQL Support

- **Unlock** your KX data estate with ANSI SQL compliant SQL interface
- Access **advanced q functionality** through SQL interface while within **1% performance** of q*
- **Integrate seamlessly** using pgwire



PyKX

- Truly **Python first** approach to kdb+
- Run **q analytics anywhere** **Python** will run with **officially supported** interface to kdb+
- Quicker query and conversion to Pandas
- More expansive **type conversions between q and Python**
- Query in **qsql** or **SQL** or write **advanced q analytics** through **context interface**



REST & OpenAPI

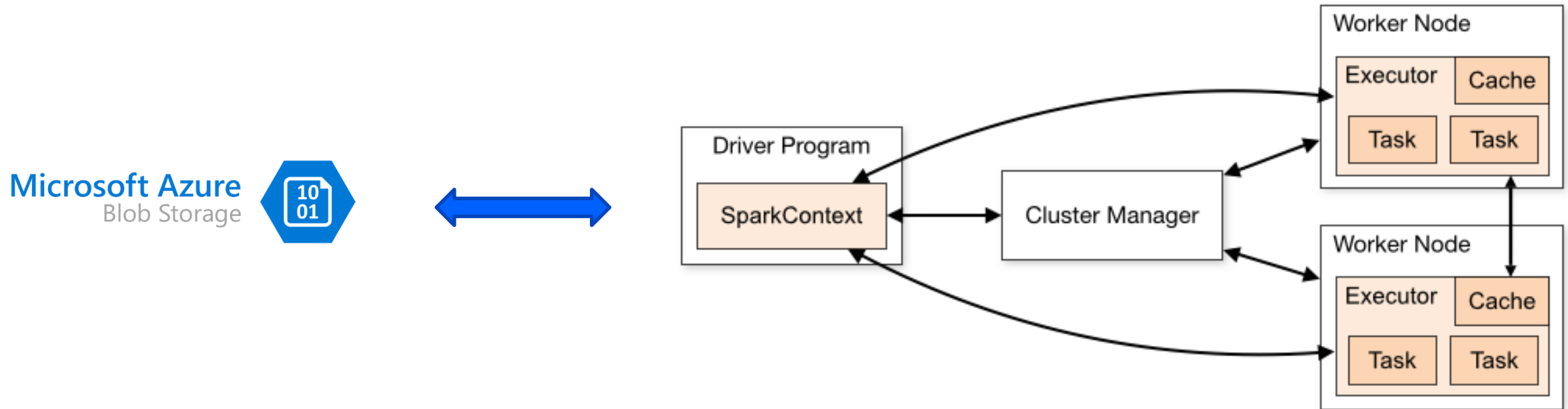
- Kurl is a **REST client** that provides sync and async methods callable from q
- Kurl provides **ease-of-use cloud integration** by registering Azure, Amazon, and Google Cloud Platform authentication information.
- **REST-server** library: expose a RESTful interface to a kdb+ based system



Accelerating Spark workflow with PyKX

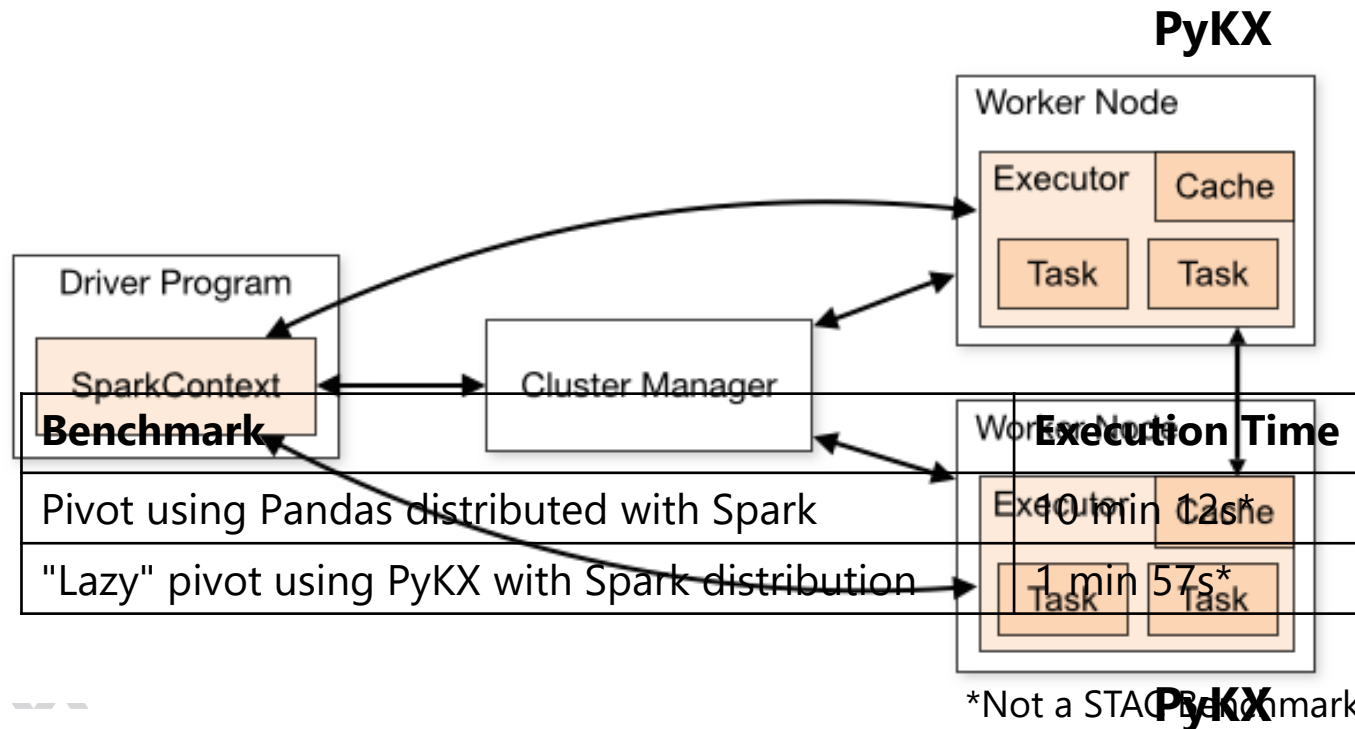
- Open-source NASA aircraft sensor data
- Data stored on Blob Storage in Parquet format
- Perform a large-scale distributed pivot on the dataset

Pivot using Pandas distributed with Spark



Accelerating Spark workflow with PyKX

"Lazy" pivot using PyKX with Spark distribution



Microsoft Azure
Blob Storage

Total number of flights: 180,000

Analysis time saved over pandas: ~ 35hrs

Thank You

For more
information:

Dr. Kim Kuen Tang

ktang@kx.com

Visit www.kx.com

