Speed up Mission-Critical Analytics in the Cloud

Billy Liu, VP of Kyligence, Apache Kylin PMC
Yiming.liu@kyligence.io
About Us

Kyligence = Kylin + Intelligence

• Formed by the creators of Apache Kylin
  • Kylin: Leading open source OLAP for Big Data

• Vision
  • Unleash big data productivity

• Offering
  • Enterprise Kylin and Managed Analytics Service on Cloud

• Funding
  • Redpoint Ventures, Cisco, CBC Capital and Shunwei Capital

• Team
  • Shanghai & Silicon Valley
Apache Kylin: Bring Interactive BI to big data
High Performance and High Concurrency OLAP engine

- **Innovation on data structure**
  - Well-designed cube supports sub-seconds query speed on PB/TB dataset
  - Encoding/Compression/Columnar

- **Lightweight and Scalable Architecture**
  - Distributed Computing by MR/Spark on YARN
  - Storage and Parallel Query on HBase
  - Native on Hadoop

- **ANSI-SQL**
  - JDBC/ODBC/REST API

- **Batch & Streaming**
  - Support Batch and Streaming OLAP in one platform
Extremely Fast Query

KAP Benchmark

(lower is better)

* Based on 4 Nodes, 16 Core CPU, 96 GB Memory per node
Kyligence Enterprise
Automation, Intelligence and Productivity

- High Performance
  - Sub-seconds query speed on massive dataset

- High Concurrency
  - Web-scale OLAP query

- Rich Ecosystem
  - Tableau, PowerBI, MSTR, Qlik
  - Cloudera, Hortonworks, MapR

- Data Sources
  - Hive/SparkSQL/Kafka
  - Cognos/Teradata/Oracle/Vertica/GP

- Automate Everything
  - Auto Cube design based on query pattern intelligently
Hybrid OLAP: Converge SQL on Hadoop and Cube

SQL on Hadoop
- Spark SQL
- Hive
- Impala
- ...

RDBMS

Query Router

SQL

Exploratory analysis
Minute delay

Cube

Mission critical analysis
Sub-second delay

Intelligent speed up
Featured Customers
Trusted by Fortune 500

SAIC Motor
#41 of Fortune 500

Lenovo
#226 of Fortune 500

Lufax Wealth Mgmt
#1 Fintech in China

China Mobile
#47 of Fortune 500

China Pacific Insurance
#252 of Fortune 500

Huatai Securities
#3 Securities in China

Huawei
#83 of Fortune 500

OPPO
#4 Smart Phone Global

GUOTAI JUNAN Securities
#2 Securities in China

1000+ open source adoptions
Success with Partners

BI
- Tableau
- Power BI
- Excel
- MicroStrategy
- Superset

OLAP/Data Mart
- Kyligence

Hadoop
- Cloudera
- MapR
- Hortonworks
- Amazon EMR
- Huawei
- FusionInsight

Cloud
- Microsoft Azure
- AWS
- Google Cloud Platform
- Alibaba Cloud
Case Study: China Pacific Insurance(Group) Co., LTD

Comprehensive Customer Reporting and Analysis

- From legacy DW to Hadoop
- 1000+ Cognos cubes to 2 KAP cubes
- 95% query latency <1s
- 30+ dimensions
- Cost reduction by adopting open source technology
How Big Data Meets Cloud?
Benefits and Challenges for Analytics in the Cloud

**Benefits**
- Scalability & Elasticity
- Continuous Available
- Resiliency & Redundancy
- Global Deployment
- Cost Efficiency
- Access anytime & anywhere
- Agility

**SME**

**Large Companies**
- Security & Privacy
- Mission-critical DW migration
- Low performance at web scale
- System Optimization
- High cost for data-intensive application
- Gap between BI and Big Data
- Skills shortage
- Data movement between RDBMS and Big Data

**Challenges**
Kyligence Cloud Solution

Transform big data analytics to cloud smoothly

- **One Click Provision**
  - Brings users fully deployed KAP and Hadoop stack in minutes

- **Dynamic Resizing**
  - Enables users to extend or shrink computing resource dynamically for on-demand workload

- **Native on Cloud Managed Hadoop**
  - Integrates deeply with cloud native data source, storage, and services on cloud managed Hadoop
Managed services into your space

- Cluster Mgmt
- Cloud Adaptor
- Kyligence Console
- AWS EMR
- Azure HDI
- Cloud Native Storage
- AWS S3
- Azure Blob Storage
- Task Node/Work Node
- Encrypted Key
- VPC
- Edge Node
- Kyligence Analytics Platform
- User Space
- Analyst

© Kyligence Inc. 2018
Integration with Azure

- **OMS**
- **AAD**
- **Admin**
- **Monitor**
- **Scheduler**

**Ingestion**
- **ERP**
- **Devices**
- **Online**
- **Logs**

**Processing**
- **Database**
- **Blob Storage**
- **Data Lake**
- **Event Hub**

**Data Service**
- **KAP Service**
- **Sub-second Response**

**Analysis**
- **Power BI**
- **Excel**
- **Machine Learning**
- **Cognitive Service**

- **Cube**
- **Blob Storage/Data Lake**

© Kyligence Inc. 2018
Cost-Effective for Analytics Workload

Allocate resource as demand (AWS as example)

- **Cost effective by separating**
  YARN workload
  - High-throughput, temporary resource for cubing job
  - Stable, long-live resource for analytics services
  - Allocate resource as demand

- **Enable multiple regions deployment for near access**

---

© Kyligence Inc. 2018
Built-in DevOps tools
Automate diagnostic and optimization services for Kyligence Services

- **Dashboard**
  - Quick insight of Kyligence instance

- **Diagnostics**
  - Figures out system’s bottleneck, issue and exceptions

- **Optimization**
  - Bring suggestion to turn system to be best

- **Knowledge Base**
  - Rich knowledge base from daily support tasks
Benefits with Kyligence Cloud

Speed up Mission-Critical Analytics in the Cloud

• Accelerate big data project go to market from months to days
• Migrate the offline OLAP to be scalable and flexible cloud solution
• Empower Big Data as services globally
• High Performance, High Concurrency, High Productivity
• Seamless Integration with your existing Cloud and BI
• Lower TCO
Demos
Contact Us

info@kyligence.io | http://kyligence.io | @Kyligence