Powered By

Bluehole (https://www.bluehole.net)

- We are the developer of 3D MMORPG, TERA.
- We are using Tajo in TERA client log analytics system. Tajo made our analysis works simple through its direct JSON data format support.
- For more information on our Tajo adoption story, take a look at our slide: http://www.slideshare.net/zenos2408/aws-tajo

Database Lab., Korea University (http://dbserver.korea.ac.kr)

- Tajo was originally developed at Database Lab. Korea Univ.
- They have 32 Tajo cluster nodes.

Encored Technologies (http://encoredtech.com)

- The company provides real-time and cumulative electric power for residential and business usage, via NILM (Non-Intrusive Load Monitoring).
- A Tajo cluster is running with 30 nodes, and the system is increasingly expanding.
- The cluster is constructed on AWS EC2, and on-premise environment; it is becoming more dependent on Tajo.
- To archive the purpose of research and data analysis for per-device usage, more than one hundred million records are processed by Tajo on a daily basis.

Gruter (http://gruter.com)

- Gruter is a Hadoop-based infrastructure company which builds big data platforms and provides technical services for the data-driven enterprise market.
- We provide a social network analysis service.
- We have about 32 cluster nodes.
- We use Tajo for ETL and ad-hoc queries on collected social network data sets.
- We process hundreds of giga bytes per day.

HYUNDAI U&I (http://www.hyundai-uni.com/eng/index.jsp)

- We use 20 Tajo cluster nodes.
- We process 4TB (400,000,000 events) per day.

LINEWALKS Inc. (http://www.linewalks.com)

- LINEWALKS provides visual data discovery services for advanced analysis of large-scale data sets of various fields including Medicine and Finance.
- We use Tajo to analyze medical data sets provided by HIRA (Health Insurance Review and Assessment service).
- We use a Tajo cluster of 10 nodes on HDFS for performing ETL and complex data analysis tasks.

Locket (http://www.getlocket.com/)

- They use 10 nodes on Amazon EC2.

LOEN Entertainment (http://www.iloen.com)

- We provide MelOn, the largest online music service platform in South Korea, with more than 28 million users.
- We run Tajo clusters with 50+ nodes on HDFS. In our production workloads, Tajo is 1.5 ~ 3x faster than old Hive setup, using less cluster resources.
- We use Tajo with HBase for analyzing purchase history for target marketing.
- Also we are using Tajo in our data insight platform for ETL and ad-hoc queries, running 1300 SQL queries on a lot of tables up to 8TB.
- We plan to adopt Tajo for our batch jobs in other legacy production cluster.

SK telecom (http://sktelecom.com/en/)

- SK telecom is the largest mobile carrier in South Korea.
- In overall, Tajo is used to analysis logs collected from our cellular network.
- We replaced some commercial data warehouse system by Tajo.
- We use Tajo for ETL and OLAP workloads.
- We use 2 nodes for Tajo Master and 51 nodes for Tajo workers.
- For only ETL, we process over 120 SQL queries that process over 4TB per day.
- We also submit more than 500 different queries through some integrated OLAP tool per day.

TicketMonster (http://www.ticketmonster.co.kr)

- TicketMonster is the fastest growing e-commerce company in South Korea.
- Currently, Tajo is used to analysis logs collected from the service.
- We use a Tajo cluster of 50 nodes to process about 100 SQL queries over 5TB per day