Development of A Scalable and Flexible Data Logging System Using NoSQL Databases

M. Kago, A. Yamashita

JASRI/SPring-8, Hyogo, Japan

Oct. 8, 2013
Current System

- Relational database management system (RDBMS)
  - Time-series data
  - Stable operation for 16 years

New System

- NoSQL (Not only SQL) database
  *NoSQL is defined as a new type database management system that is non-relational.

New System Features

- **Scaling-out**
  The system can easily grow the performance by adding more low-cost servers.

- **High Reliability**
  There was no single point of failure (noSPOF).

- **Flexible Data Format**
  The system supports various data type such as integers, reals, strings, arrays and maps.

- **Low Latency Access**
  Users can take the latest data in microseconds order.
New System

NoSQL Database

- **Apache Cassandra**
  - Distributed database without SPOF
  - Excellent fit for time-series data
  - Perpetual archive

- **Redis**
  - In-memory key-value store
  - Real time data cache

System Overview

- Event log
- Event log
- Event log

- Relay Servers

Long-term Test

- The system had been inserted 50,000 messages/sec for 3 months.
  
  => No data loss during the test even when the server was forced a shutdown.

High reliability and stability

Poster ID: TUPPC012