

Fujitsu develops distributed and parallel complex event processing technology that rapidly adjusts big data load fluctua

December 16 2011



Figure 1: Structure of Complex Event Processing

Fujitsu today announced that, in an industry first, they have developed complex event processing technology designed for use with cloud technology that employs distributed and parallel processing. This enables rapid adjustment to fluctuations in data loads when processing massive amounts of heterogeneous time series data, now popularly known as "big data."

Today there is an ever-increasing amount of large-volume, heterogeneous time series data, such as sensor data and human [location data](#). To analyze this big data at high speed in order to put it to use, complex event processing technology has been developed. With complex event processing, high-speed processing of fluctuating levels of large-scale time series data is required, but up until now it has been difficult to

adjust to load [fluctuations](#) without pausing the processing operation.

In the development announced today, however, applying distributed and [parallel processing](#) technology to complex event processing enables ever-greater granularity in processing. And with dynamic distribution during execution and very high speeds, immediate adjustments in load fluctuations were achieved without the need to pause processing. As a result, a [throughput](#) function of five million events per second was achieved, and an unprecedented level of time series data was able to be continuously analyzed in real time. This is one of the technologies that will be put to use to support human-centric computing, which will provide precisely targeted services anywhere.

This research was supported in part by the Ministry of Economy, Trade and Industry's Project for the Program to develop and demonstrate basic technology for next-generation high-reliability, energy-saving IT equipment for fiscal 2010 and fiscal 2011.

Complex event processing is a method for mining big data in real time to extract useful information. Methods of processing a large volume of data using a database, for which the data first need to be stored on some medium, such as a disk, are therefore not suitable for real-time processing. By contrast, as shown in Figure 1, with complex event processing, multiple inputs of time series data are processed in memory based upon pre-defined rules (known as "queries"), enabling extremely fast processing compared to methods database storage.

In recent years, the amount of large-volume, heterogeneous data, as represented by such time series data as sensor data and human location data, continues to grow at an explosive pace. There is a strong demand to take this type of "big data" and efficiently extract valuable data that can be put to immediate use in delivering services, such as various navigation services.

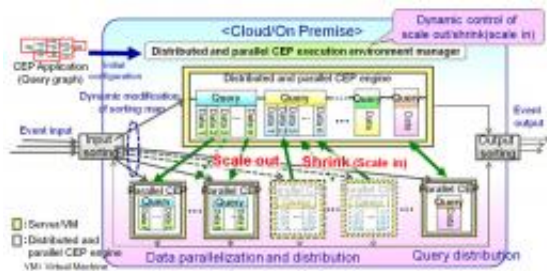


Figure 2: Basic Structure of Distributed and Parallel Complex Event Processing

Complex event processing technology has been used to quickly analyze time series data and put the results to use. However, until now it has been technologically challenging to continuously process incoming time series data in real time, while simultaneously performing high-speed transfers across servers of in-memory data that is still being analyzed and processed.

As a result, in the case of significant fluctuations in time series data due to seasonal and time variations or the occurrence of natural disasters and accidents, there has been a need to either estimate and configure in advance the resources required to accommodate peak load times, or alternatively, to pause the complex event processing and make modifications to the system configuration.

The newly developed complex event processing technology is able to dynamically apply distributed and parallel processing technology to address load variations. This technology enhances the granularity of processing and can transfer this finer processing across servers, as well as be configured based on technology that selects the optimum candidate processing to be transferred. As a result, time series data load fluctuations are adapted to swiftly, and together with entire event

processing that can dynamically scale out or shrink (scale in) across servers, the processing speeds are also raised. Features of the newly developed technology are as follows.

1. Distributed and parallel processing technology-based dynamic load distribution for complex event processing

As shown in Figure 2, unit-based management of complex event processing takes each query and the data parallelization of each query, and then refines it into smaller parts. As methods of dynamic load distribution, both data parallel distribution and query distribution are realized. This makes possible dynamic distribution processing in real time that can adjust to load fluctuations within a limited range of resources. In addition, in the case of processing simple queries, high-throughput of 5 million events per second was achieved.

2. Technology that optimally selects migration processing candidates depending on the status of load fluctuations

Based on the speed of load fluctuations, the properties of each event or query, or the status of the processing load, Fujitsu developed technology that enables the effective distribution of loads and that migrates processing tasks for which the impact of the migration is lowest. As an example, queries that are closely related are, to the extent possible, allocated in the same server, enabling the maintenance of high-speed processing.



Figure 3: Reference Application of Distributed and Parallel Complex Event Processing Technology

The new technology, which will support the creation of human-centric computing in which fine-grained services are offered in all areas of daily life, makes it possible when processing big data to maintain high processing performance and accommodate fluctuations in processing loads without pausing the processing operations. As a result, as illustrated in Figure 3, it will be possible to develop services that provide non-stop analysis—in real time and using only the necessary resources—of the large-scale time series data by leveraging in-house facility (on-premise) systems. In addition, given the flexibility to flexibly alter configurations, the need for strict estimates of the resources is eliminated and will be accessible for even small-scale operations.

[Fujitsu](#) aims to bring products employing the new technology to market during fiscal 2012. The company will proceed with deploying the technology in a wide range of applications, such as its middleware products and Convergence Service Platform, which enables the utilization of valuable data generated through the process of collecting, accumulating and analyzing large volumes of [sensor data](#).

Provided by Fujitsu

Citation: Fujitsu develops distributed and parallel complex event processing technology that rapidly adjusts big data load fluctua (2011, December 16) retrieved 27 April 2024 from <https://phys.org/news/2011-12-fujitsu-parallel-complex-event-technology.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private

study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.