OSS Hadoop Use in Big Data Processing

TAKAHASHI Chieko, SERA Naohiko, TSUKUMOTO Kenji, OSAKI Hirotatsu

Abstract

Development of technologies for the processing of "big data" has recently been advanced by network-related enterprises. Apache Hadoop is attracting attention as an OSS that implements storage and distributed processing of petabyte-class big data by means of scaling out based on the above technologies. NEC has conducted the test for the ability of Apache Hadoop for enterprise use and has built systems according to its characteristics. For the sizing of Apache Hadoop that is usually regarded to be difficult, NEC has developed a technology for size prediction by means of simulation. This paper introduces these technologies.

Keywords

big data, distributed processing, scale out, open source, OSS, Hadoop, simulation, sizing, CASSI

1. Introduction

"Big data" and "Open-source software (OSS)" are two of the keywords attracting attention recently in the IT field. NEC has been conducting the technical test for the ability of Apache Hadoop for enterprise use(hereinafter "Hadoop"), which is one of the OSSes attracting attention in the field of big data. One purpose for this verification is to clarify differences from the traditional technologies related to the distributed processing platform such as the DWH (Data Warehouse) and big data and to define their comparative advantages.

In this paper, we first outline the fields of application of Hadoop in section 2 and describe case study of Hadoop applications and then go on to discuss points to be considered in the actual system construction in section 3. Finally, we introduce a simulator that has made it possible to demonstrate the Hadoop performance to get the best performance.

2. Hadoop, Its Real Ability

Hadoop is a framework developed as an OSS based on papers published in 2004 by Google Inc. that deal with the "MapReduce" distributed processing and the "Google File System." It is attracting worldwide attention as a realistic means of processing big data.

2.1 Design Concept of Hadoop

In order to implement large-scale distributed processing, Hadoop has been developed under the following design concept.

- **Linear performance expansion by scaling out**
  Hadoop is designed as a scale-out type processing system that can expand its performance by means of parallel distributed processing using multiple nodes. In general, a scale-out type processing system provides the linear improvement of performance and price following a server extension and enables control of performance limitations and sharp cost increases (**Fig. 1**).
  However, in reality, communications traffic between the increasing numbers of servers cause a bottleneck so that the extended server numbers eventually reach a limit. Meanwhile, the Hadoop framework can remove the bottleneck by distribution of the communications between servers, reduction of the load of the master server and decentralization. It therefore enables construction of a larger cluster (group of servers).

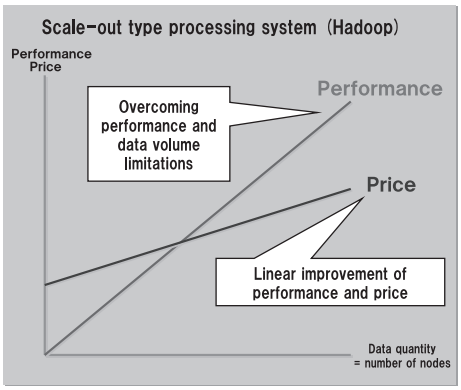

Fig. 1 Characteristics of a scale-out type processing system.

- **A framework that provides distributed processing**
  To process big data efficiently with a scale-out type server cluster, a distributed parallel processing architecture that takes advantage of the number of servers is effective. The Hadoop framework uses a processing system called MapReduce in the development so that optimum distributed processing is automatically executed in the server cluster. This processing program is distributed automatically among the Hadoop slave servers and the results of the distributed processing are later aggregated.
- **Unstructured data processing capability**
  Compared to the processing programs handling structured data such as the RDB (Relational Database), the Hadoop as a distributed processing framework can execute various programs on the stored data. It is therefore suitable for conversion and analyses of unstructured data, for example in the mining of documents and logs and for the conversion of image data formats.
- **Fault resistance enabling use of a large number of common servers**
  When a server cluster is configured with thousands of servers, failures of servers and disk devices should not be considered as "exceptions" but as normally occurring events from the viewpoint of probability. Hadoop creates replications automatically when storing data and, in case of an issue with a server, data block or assigned processing program, the processing is assigned to the data replication in another server. This strategy allows it to avoid system shutdown due to a fault in a server or disk device. Data lost by a server or disk device failure is restored because normal replication is independently created in a normal server (**Fig. 2**).
- **Throughput oriented**
  The design concept of general RDB is to aim at securing low latency by maintaining consistency with the use of complicated mechanisms such as lock and queue. On the other hand, Hadoop does not update data in order to eliminate complicated mechanisms but always presupposes to write new data (Write Once Read Many). As large numbers of slave servers are concurrently storing input and output data, throughputs can be secured for inputs and outputs of big data.

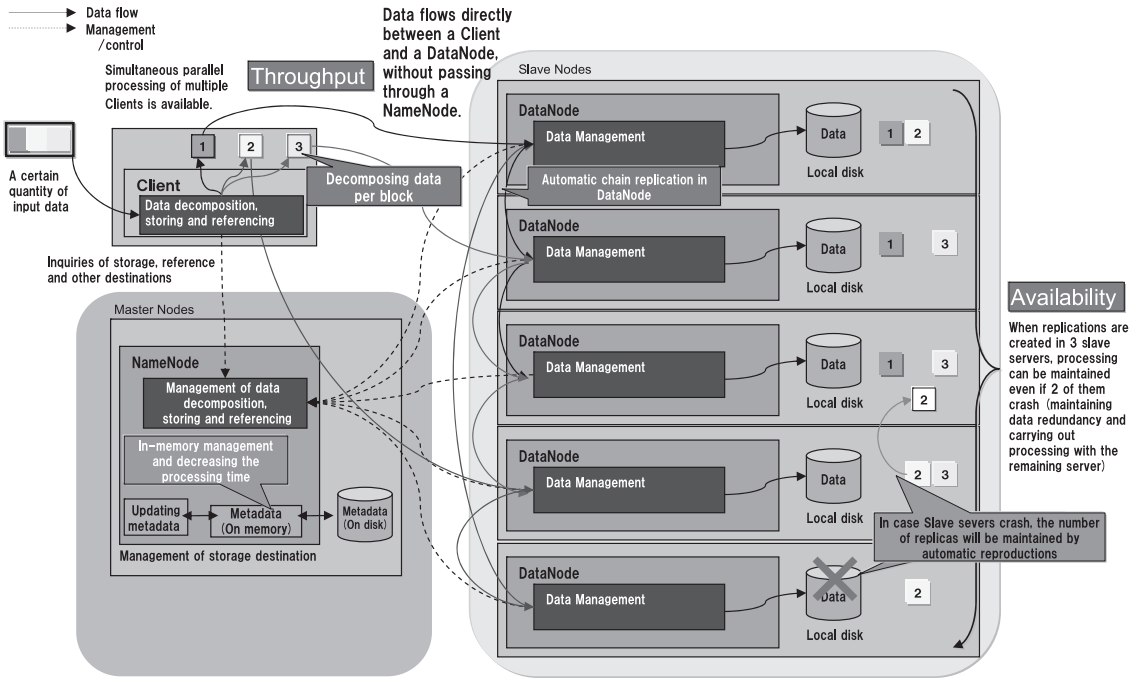

Fig. 2 Data replication and auto recovery.

2.2 Fields of Applications of Hadoop

Hadoop has started to be used mainly in the following three fields.

- **DWH enhancement**
  Hadoop can be applied in DWH processing for access log analysis of a website, data mining, risk analysis, etc.
  In this field, there have been attempts at improvement of the analysis accuracy by performing big data storage, data processing and processing target data filtering at the prior stage of the existing DWH. This procedure was aimed at making the existing DWH solutions capable of handling big data that could not previously be handled.
- **Big data handling**
  Hadoop may also be applied in the processing of big data that was not able to be handled by traditional systems. Coupled with the advancement of life log and M2M (Machine-to-Machine) technologies, Hadoop has started to be applied to big data that has been hard to be analyzed with existing technologies, such as the sensor data and the worldwide credit card usage record data.
- **Mission-critical batch processing**
  The part of batch processing that has previously been executed by the mainframe can be executed by the scale-out type system.
  Applying the Hadoop system to the processing with which expansion of the data quantity or reduction of the processing time is not possible due to the limitations of scale-up type performance will enable data expansion and batch time reduction as the scale-out type processing.

3. Hadoop Solutions for Implementing Large–Scale Distributed Processing Systems

In countries outside Japan, Hadoop has already been marketed in solutions for implementing large-scale distributed processing systems. However, the circumstances in Japan have now changed and Hadoop is recently shifting from the stage of trial use to one of active use in commercial-purpose solutions.

3.1 Examples of Applications of Hadoop Solutions

- **Large-scale log search system**
  The log data such as access and communication records of networked businesses and communication carriers are dynamically generated every day and every minute, to eventually reach huge amounts. It has been difficult for the traditional systems to save and search such big data. Applying Hadoop to a system of this type has succeeded in dealing with the data quantity increase thanks to the scale-out architecture and in reducing the search time significantly thanks to the distributed processing (MapReduce). As the data quantity continues to increase after introduction of Hadoop, the system increases the scale periodically by adding more machines.
- **Sensor data aggregation system**
  This system installs sensors in locations all over Japan, collects the sensor data from them and analyzes it in order to generate useful data and make it available quickly. The system includes a column type database called the HBase on Hadoop to enable simultaneous implementation both of a scalable database and large-scale distributed processing using MapReduce, which used to be difficult with the traditional RDB.
  As the Hadoop/HBase combination has high flexibility, it can deal flexibly with an increase of sensors, addition of collected data types and addition of analysis viewpoints without reconfiguring the data structure.
- **Integrated authentication management system**
  Part of the processing of the integrated authentication management system that has previously been managed with the RDB can be executed by the Hadoop/HBase combination.
  Since the RDB presupposes that the data is structured, it is often not adept at changes such as the addition of user information. However, HBase is a column-type database and is resilient against changes of user information. The installation of the database on the Hadoop allows the MapReduce to perform analysis/extraction of user information and simultaneous modifications following a system reorganization/transfer.

3.2 Actual Cases of Hadoop Solution Building

In this subsection, we introduce the points to be considered or those that are actually improved in the process of building Hadoop solutions.

- **Countermeasures against single points of failure**
  In general, Hadoop takes the redundancy of a large number of machines called the slave servers into consideration so that it can deal with faults including hardware failures without interrupting the service. However, as the master server managing the entire Hadoop system is a

single point of failure, it means that a mechanism for providing the master server with redundancy is necessary. At NEC, we can provide the CLUSTERPRO clustering software to built a system capable of continuing service by switching the master server if it should go down (the latest version of Hadoop incorporates improvements for redundancy of the master server).

- **Concept of operation monitoring**
  The fact that Hadoop uses multiple machines means that the potential of faults including hardware failures is high. It is therefore necessary to monitor the information constantly in order to confirm if individual servers are working normally, how the overall resource consumption trend of the system is, and so on, and to adopt any necessary measures early on. The monitored items include the general resource information on the OS as well as those output uniquely by Hadoop.
- **Utilization of the auto building tool**
  When the number of machines reaches some tens of units and they are built manually, extension of the building period and increase of the probability of mistakes may lead to issues. Therefore, auto building mechanisms such as the OSS Puppet and Chef are used.
- **Tuning**
  The parameter tuning enables more effective operations depending on the characteristics of data handled by the system and on other system requirements. Examples include the tuning for extracting the performance by maximum use of hardware resources and the tuning for assisting stable operation of the entire system by reducing influences on jobs.

4. Hadoop System Performance Evaluation Technology Based on Simulation and Adjustment of Measurements

4.1 Background and Introduction

One of the features of Hadoop is the capability of improving the performance linearly in accordance with the number of servers. As described in subsection 3.2, the appropriate tuning of Hadoop is necessary to obtain the maximum benefit from its features. Nevertheless, the settings of Hadoop are complicated and the number of servers is so large that it is very difficult to obtain the appropriate settings from working out on papers. In addition, huge costs and length of time are required to verify a complicated large-scale Hadoop OSS system in the real world environment.

To solve the above issue, we built a Hadoop system simulator for performance evaluation using the CASSI (Computer Aided System model based SI environment). This has made it possible to obtain the settings efficiently that can achieve the maximum performance from Hadoop.

4.2 System Performance Evaluation Technology Based on Simulation and Adjustment of Measurements of CASSI

In general, a simulation for system performance evaluation involves issues of simulation accuracy and processing time. For the CASSI, we conducted R&D of the system performance evaluation technology based on simulation and adjustment of measurements and succeeded in improving the simulation accuracy. This was done by adjusting the theoretical values that were calculated based on the performance model with the actual measurements measured in the real environment. The simplification of calculations gained from using the actual measurements has also allowed us to reduce the simulation processing time.

4.3 Evaluation by the Hadoop Simulator for System Performance Evaluation

The Hadoop simulator for system performance evaluation can calculate the response time of Hadoop MapReduce, the throughput of map processing, etc., and the resource usage rate of CPU and network, etc.

As shown in the comparison of the simulation time and the actual measurements of the CPU utilization rate during MapReduce processing (**Fig. 3**), the simulation is capable of simulating variations of the CPU utilization rate with high accuracy. The error in the overall response time is as small as around 2.5% (approx. 10 sec.) of the processing time of about 7 minutes.

4.4 Present Achievements and Future Schedule

As described above, our R&D of the system performance evaluation technology based on simulation and adjustment of measurements have enabled easy execution of highly accurate performance evaluation simulations of a complicated large-scale system such as the Hadoop system by working out on papers. The Hadoop simulator for sysytem performance evaluation simulator can be used to perform advanced performance evaluations that have previously been difficult due to

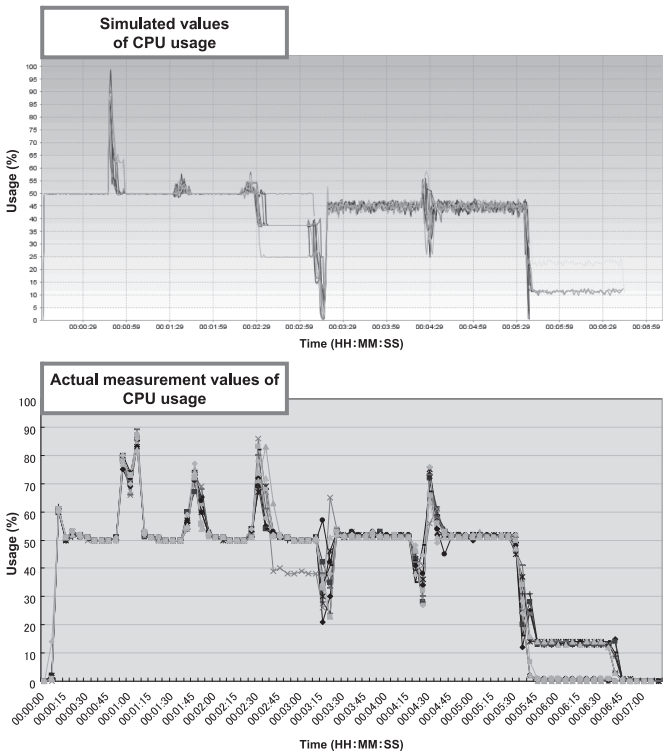

Fig. 3 Comparison of simulated values and actual measurements of the CPU utilization rate.

the restrictions imposed by costs and time. This means checking the performance limit values of a system, easily at a low cost and in a short period.

System performance evaluation technology based on simulation and adjustment of measurements is used not only with Hadoop but it has already been applied to the web 3-tier system, achieving similarly favorable results to the Hadoop system. In the future, while promoting application of the simulator to various other systems, we will introduce further improvements aimed at enhancement of SI quality and reduction of costs.

5. Conclusion

Although Apache Hadoop has not yet completed production, it is expected to continue to develop within the OSS community. No framework that is similar to Hadoop is found among currently marketed products. By making use of the intellectual property as an open-source product, we will continue verifications in order to promote more applications and to provide simulators that can obtain maximum benefits from Hadoop.

*Apache, Hadoop and HBase are registered trademarks or trademarks of the Apache Software Foundation.

*Google is a registered trademark of Google Inc.

Authors' Profiles

TAKAHASHI Chieko
Group Manager
IT Platform Solutions Division
IT Software Operations Unit

SERA Naohiko
Manager
IT Platform Solutions Division
IT Software Operations Unit

TSUKUMOTO Kenji
Assistant Manager
3rd Software Division
Platform Operations Unit
NEC System Technologies, Ltd.

OSAKI Hirotatsu
Assistant Manager
Emerging SI Technology Development Group
IT Services Business Unit

Information about the NEC Technical Journal

Thank you for reading the paper.
If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Japanese

English

Vol.7 No.2 Big Data

Vol.7 No.2
September, 2012

Special Issue TOP