Introduction Case Study: Gurunavi, Inc.
Gurunavi accumulates the Bigdata of food with MapR for making a contribution to Japanese food culture

Gurunavi is best known as Japan’s largest gourmet and restaurant guide, originally as a web site and now also through its popular mobile app. The company, founded in 1989, launched its “Gurunavi” restaurant search website in 1996. At present, with operations focused at its headquarters in Tokyo, Gurunavi has a total of 19 offices throughout Japan. They have expanded their business to Shanghai, Singapore, Hong Kong and Taiwan. Their mission is to continue to spread information in its role as a “supporter of restaurants.”

The company adopted Apache Hadoop in 2010 in order to cope with the data generated from their increasing user base of their successful restaurant guide. However, due to the problems relating to the difficulty and costs of operating open source products, Gurunavi decided to migrate to MapR, mainly based on MapR’s higher reliability and superior performance to open-source Apache Hadoop.

Gurunavi is extremely well known throughout Japan for being a restaurant search site. Their restaurant guide has approximately 500,000 with 57 million unique users and a total of 13.97 million registered users.

Yasutomo Tamafune, who is affiliated with the Business Solutions Group at Gurunavi Inc., had the following to say regarding the problems with Apache Hadoop.

“Our mission is to analyze the enormous amount of data that has been accumulated on Gurunavi, and to plan and develop products which match the needs of our users. Recently we have been working on collaborations that are based on the knowledge obtained from big data, such as consulting with food manufacturers and developing new products. We introduced Apache Hadoop in 2010 as the platform for analyzing this big data. However there were two growing issues with the platform: The first was concerns over single point of failure (SPOF), and the other was the increase in burden encountered in the operation of Hadoop-related open source products.”

Apache Hadoop requires complex operational knowledge to mitigate its well known NameNode SPOF issue. In addition, the operational costs and issues related to open source were also a major issue for Gurunavi’s Business Solutions Group. Arai of the Business Solutions Group provided us with the following explanation of this problem.

“When an open source failure occurred, we would basically have to investigate the problem ourselves, and take the necessary measures such as applying patches and/or upgrading to a newer version. This work took up a significant amount of time and labor, and meant that we were no longer able to dedicate sufficient resources to the collection and analysis of big data, which was our original mission. From a business perspective, this was a big problem.”
Open source technology progresses very fast, there are always new functions and patches to look into and verify. We did not think it made good business sense to assign valuable staff to this work. From the moment we started test use of Apache Hadoop we wanted to find a knowledgeable company to provide us with enterprise-grade support, but we could not find one at that time. In addition, we thought that perhaps adopting a commercial platform which had been developed for enterprise use would be better than open source products. Those reasons led us to investigate commercial Hadoop with enterprise support, recalls Tamafune on the background to the selection of MapR.

The Business Solutions Group looked into the products of three companies, including MapR.

"The decisive factors for the selection of MapR were the ease of use of its file system, as well as its reliability and ability to be used in real time. As it has an NFS mount compatible open standard interface, it was easy to place files and accumulate data. As we planned to collect data from multiple departments, not just our own, we determined that the MapR file system could be used easily in any department. In addition, we also paid significant attention to the functionality that allowed for written data to immediately be used in real time. In addition, given the data protection features such as snapshots, many jobs aimed primarily at fault tolerance could be made redundant. We gave high marks for these reliability related features.", explained Tamafune on comparing functions.

Open source technology progresses very fast, there are always new functions and patches to look into and verify. We did not think it made good business sense to assign valuable staff to this work. From the moment we started test use of Apache Hadoop we wanted to find a knowledgeable company to provide us with enterprise-grade support, but we could not find one at that time. In addition, we thought that perhaps adopting a commercial platform which had been developed for enterprise use would be better than open source products. Those reasons led us to investigate commercial Hadoop with enterprise support, recalls Tamafune on the background to the selection of MapR.

The Business Solutions Group looked into the products of three companies, including MapR.

"The decisive factors for the selection of MapR were the ease of use of its file system, as well as its reliability and ability to be used in real time. As it has an NFS mount compatible open standard interface, it was easy to place files and accumulate data. As we planned to collect data from multiple departments, not just our own, we determined that the MapR file system could be used easily in any department. In addition, we also paid significant attention to the functionality that allowed for written data to immediately be used in real time. In addition, given the data protection features such as snapshots, many jobs aimed primarily at fault tolerance could be made redundant. We gave high marks for these reliability related features.", explained Tamafune on comparing functions.

In addition to highly evaluating MapR’s solution in terms of functionality, Arai was also impressed with the support system provided by a Japanese company.

"We also paid note of MapR’s commitment to open source. Rather than just distributing the product, MapR demonstrate a strong commitment to innovation where it matters most of its enterprise customers, such as constant improvements to performance, manageability or developing a native file system that can directly access storage hardware. We were really impressed by this."
In the end, Gurunavi selected MapR based on scoring highly on a combination of higher performance, ease of use, quality of support and commitment to technical innovation.

“By using MapR we only needed to use one third the number of Apache Hadoop servers as we had done in the past. In addition, we are now able to do aggregation of action history in 30 minutes instead of 3 hours. In terms of performance, we have calculated that we are now working about six times faster.

In addition, the volume of new batch processing is increasing every day. We feel that our current cluster can handle the increase with room to spare. The adoption of MapR has allowed us to build a platform in which big data can be processed highly efficiently”, said Tamafune on the results of the introduction.

In addition to the large improvement in costs and processing performance, Ishiage had the following to say regarding how the introduction of MapR improved things on the operational side even more than expected.

“The other day, we ran some tests in our test environment. There were quite a lot of problems with critical OSS failures, however MapR maintenance immediately sent us information on a modified version. In the past we would have had to examine the open source product ourselves, however thanks to using the MapR commercial platform, we were able to receive the relevant support. Combined with the quick response, this left us very satisfied.”

“At present, we are accumulating around 20GB of data every day. The accumulated data is processed in batch jobs overnight. The analyzed data is provided to our registered restaurants in the form of graphs such as on access analysis logs, etc. For restaurants, such access logs are a really important management indicator. In addition, our Business Solutions Group is analyzing trends in food from the big data, and using this as knowledge for the planning of new products and consulting work”, explained Arai on the company’s use after the introduction of MapR.

Gurunavi analyses the big data accumulated in MapR and announces an annual “dish of the year” (http://gri.gnavi.co.jp/dishoftheyear/).

“In the future, we would like for all Gurunavi data to be accumulated on our MapR cluster, and to be able to analyze the data in real time. The mission of our company is to enrich the Japanese diet, and to spread Japanese food culture throughout the world. In order to achieve this, the presence of MapR as a platform for big data relating to food is going to continue to increase in importance”, said Tamafune on the future outlook.