CASE STUDY
Intel® Xeon® Processor E5 Series
Traffic Management
Big Data Analytics

Improving traffic management with big data analytics

Hangzhou Trustway Technology Co. Ltd. significantly improves its transportation management capability using Apache Hadoop on Intel® Xeon® processors

"With Apache Hadoop, we were able to not only store the massive volume of image and video data, but also to enable a large number of users to access the data quickly. Traffic violation data can now be stored for 24 months instead of only three months. And it now takes less than a second to accurately search for the plate number or driving track from the over 2.4 billion records of vehicle data."

Chen Haitao
Hangzhou Trustway Technology Co., Ltd.

A city in Zhejiang, China, has seen its economy develop rapidly in recent years. As a result, local traffic has become much heavier, leading to an increase in vehicle accidents and traffic violations.. The city government needed better ways to monitor and manage local traffic to provide better transportation services to the public. Taking a data-driven approach to the problem, the local traffic department installed more than 1,000 digital monitoring devices in the city's key checkpoints. These devices capture images and video data continuously. The traffic department now faces a terabyte of data each month. The increasing amount of traffic data now poses challenges in the city's ability to effectively manage traffic.

CHALLENGES

- **Enable centralized management of traffic data.** Centralize access to image and video traffic data stored in the data centers of different divisions. Centralize access to traffic management facilities, equipment and application systems.
- **Optimize utilization of massive data.** Store vehicle monitoring data for as long as possible to give information support for departments such as public security, criminal investigation, and economic investigation and front-line police.
- **Improve traffic flow across the city.** Enhance dispatch capability for dealing with various kinds of emergencies and accurately forecast traffic patterns.

SOLUTIONS

- **Deploy unified data center based on Intel Xeon processor E5 series.** Deploy 22 servers running on Intel Xeon processor E5 series and a 198-terabyte storage space for centralized storage for the digital traffic information.
- **Deploy Apache Hadoop.** Use the Hadoop Distributed File System (HDFS) and Apache HBase* to provide permanent storage and seamless expansion of vehicle and traffic violation image data accumulated in the past 24 months. Use Hadoop to retrieve data in real time.
- **Deploy Trustway key vehicle dynamic supervision system.** Taking advantage of the open platform for analytics from Intel, deploy Trustway system for massive-scale data mining and analysis.

TECHNICAL RESULTS

- **Enhanced storage capacity for large data.** Apache Hadoop provided a mass data storage solution with high fault tolerance and throughput, allowing reliable storage for massive information and seamless expansion capacity.
- **Achieved powerful I/O processing function.** The Intel Xeon processor E5 series enhances I/O processing. Now a single server can allow synchronous transmission of a 500KB picture with an average speed of 250 times per second, or asynchronous concurrent storage of 2,000 times.
- **Provided high-performance HBase database.** Apache Hadoop enabled complex data queries in the vehicle monitoring system. It now takes less than a second to accurately search for plate numbers or the driving record of a vehicle from over 2.4 billion records.

BUSINESS VALUE

- **Improved traffic case detection ability.** With 24 months of traffic violation image information stored in the system, traffic police departments easily retrieve vehicle information such as the color, model, and license plate in real time along with other relevant information such as historical behavior, driving routes, the vehicle's operating company, and the identity of the driver.
- **Improved traffic police supervision of motor vehicles.** Traffic police can now easily retrieve plate numbers and the driving track of a passing vehicle from the over 2.4 billion records in the system.
- **Easy access to relevant vehicle analysis data.** Investigating traffic cases that require complex inquiries, such as data from multiple checkpoints or multiple vehicles, now takes just 10 seconds.

Apache Hadoop transforms the traditional storage and processing architecture for traffic data analysis

Traffic management challenges in a developing city

A city in Zhejiang has connected more than 100 intelligent monitoring checkpoint systems, over 300 checkpoint electronic police, and more than 500 video monitoring systems to manage traffic effectively. All the structured data from collection equipment – such as time, place, and vehicle information – are stored in a data center in the city traffic management division, while the semi-structured data such as pictures and videos are stored in a centralized data center.

The traffic management division uses the data to perform real-time analysis of traffic conditions and statistical analysis of traffic accidents as well as to analyze traffic violations and driver records.

The big data challenge

As traffic data continues to grow, the average monthly data has now reached 10 terabytes. Since data such as pictures and video are stored in different data centers in different divisions, it has become difficult to use. Moreover, some traffic management facilities, equipment, and application systems run in silos, which need to be integrated.

The city kept 12 months of traffic data. But the historical data showed that traffic data has grown 60 percent per year, forcing the duration of stored data to become shorter and shorter. The city needed to extend the storage period of traffic data to support public security staff, criminal investigation teams, economic investigations, and front-line police. Vehicle traffic data is often key evidence that helps identify individuals involved in legal cases.

As the city continues to develop, the scale of traffic monitoring operations has also grown. The data collected through monitoring equipment needs reliable storage. Also, with the development of new technologies and the upgrade in the electronic police checkpoint system to high-definition video images, image size is larger than before, demanding better storage performance.

Another problem the city faced was that traffic data could not be fully utilized. The manual data query and statistical analysis had become less efficient with the growing amount of data. Since these data also have potential value for the city's traffic management, the city hopes to transform the video monitoring applications, using the data as preventive traffic information that can boost public security and transportation management.

Dealing with big data using Hadoop

To solve its big data challenge, the city government worked with Hangzhou Trustway Technology Co., Ltd. (Trustway). Based on the specific requirements, Trustway deployed a key vehicles dynamic monitoring system based on Apache Hadoop and Intel Xeon processor E5 series. The solution is connected with traffic monitoring equipment such as the city's checkpoints, video monitoring, traffic flow detection, signal systems, and devices.

This solution provides the city with a big data storage system with high throughput and fault tolerance. The servers based on Intel Xeon processors E5 family deliver powerful I/O. The data center has 22 servers with Intel Xeon processors E5 series and 198 terabytes of storage space. Tests showed that a single server can deliver a full high-definition image of 500KB on average of 250 times per second. In asynchronous transmission mode, the solution supports concurrent access at an average speed of 2,000 times per second. In effect, even with multiple vehicles passing through all the monitoring stations in the city, the solution can deliver real-time and reliable data storage.

Apache Hadoop provides a fault-tolerant platform for data, replicating and distributing the data across the Hadoop cluster. When the data is read, it is automatically verified and, if a verification error is found, the operation is repeated. Intel made a number of enhancements and optimizations to the open-source Hadoop framework to enable higher performance and reliability. For example, the system's I/O performance is enhanced by optimizations to the HDFS data distribution and read/retrieve algorithm.

Using Apache Hadoop, the city's vehicle data can be stored permanently, and the historical data of images related to traffic violations can be stored for 24 months. The system also allows seamless capacity expansion by adding nodes to a live cluster without shutting down the system. This meets the city's requirement of expanding the system capacity to store additional data at its convenience.

Using Apache Hadoop, Trustway's key vehicle dynamic monitoring system helps the city process and analyze traffic data intelligently and efficiently. The solution helps the city obtain full value from the data by enabling vehicle track analysis, fake plate number analysis, and vehicle searching and control.

LESSONS LEARNED

- Apache Hadoop, running on industry-standard hardware, can enhance data storage reliability and throughput performance. It is an ideal storage solution for the semi-structured image and video data in the traffic field.
- Intel Xeon processor E5 series provides powerful CPU and I/O capabilities to deal with the challenges of parallel read-write by the data-intensive collection devices.
- The Hadoop platform is well suited to handle the data storage, mining, and analysis of image and video data. by traffic systems at scale.

Apache Hadoop provides a distributed computing framework — MapReduce* — which allows the processing of structured and semi-structured data. With MapReduce, Trustway's key vehicle dynamic monitoring system was able to provide data mining and retrieval of the video and vehicle images as well as high-speed analysis of vehicle information from traffic data. Trustway's testing showed that it took only about 10 seconds to carry out a collision analysis based on 2.4 billion plates stored on the platform.

The Apache HBase datastore solved the theoretical and practical limitations of a relational database. Tests showed that it took less than one second to obtain a driver's record by querying the vehicle's plate number from 2.4 billion records.

In addition, Hive* data warehouse technology was used to efficiently find associated information such as driver, vehicle, road monitoring data, accident, violation, and public security theft monitoring, making complex computations flexible and simple. The data analysis can be used for effective decision-making.

Looking ahead, Trustway will continue to work with Intel on large data analysis, using updated technology to provide more intelligent traffic management solutions.

Find a solution that's right for your organization. Contact your Intel representative, visit Intel's Business Success Stories for IT Managers (**www.intel.com/itcasestudies**) or explore the Intel.com IT Center (**www.intel.com/itcenter**).

This document and the information given are for the convenience of Intel's customer base and are provided "AS IS" WITH NO WARRANTIES WHATSOEVER, EXPRESS OR IMPLIED, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NONINFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS. Receipt or possession of this document does not grant any license to any of the intellectual property described, displayed, or contained herein. Intel® products are not intended for use in medical, lifesaving, life-sustaining, critical control, or safety systems, or in nuclear facility applications.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

© 2013, Intel Corporation. All rights reserved. Intel, the Intel logo, Intel Xeon, and Intel Xeon Inside are trademarks of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

0313/SHA/PMG/XX/PDF 328805-001EN