SOLUTION BRIEF
Intel® Xeon® Processor E5-2600 v2 Product Family
Intel® Solid-State Drive DC S3500 Series
Big Data
Cloud Computing
Entertainment/Media

Look Inside.™

Improving news quality and editing efficiency with big data

Sugon's XData* big data solution, based on Intel® Xeon® processor E5- family and Intel Solid-State Drives, provides unified management and real-time analysis and retrieval of media data sources

中科曙光 Sugon

"For one of China's largest news media agencies, we deployed the XData* big data platform based on the Intel® Xeon® processor E5-2600 v2 product family, allowing storage consistency of index data from multiple media resource libraries and scattered multimedia resources. Utilizing Intel® Solid-State Drives (Intel® SSDs) DC S3500 Series allowed XData to deliver near-real-time query performance."

Cao Zhennan
General Manager, Solution Center
Sugon

The emergence of new media such as social networks continues to impact the traditional media industry in China. Even with its long development history and mature ecosystem, the industry has found it challenging to keep up in this fast-changing environment. As a leader in cloud computing and big data solutions, Sugon helps the industry keep pace with new media developments through its XData* big data solution, allowing them to make full use of rich content, graphics, audio and video resources, and other assets. With XData, Sugon enables the industry to provide excellent content for end users.

CHALLENGES

- **Unify data management.** Enable centralized management of scattered multimedia resources to allow converged storage of both structured and unstructured data.
- **Simplify IT architecture.** Develop a standard information resource service system to lay a strong foundation for meeting the demands of evolving business types and future applications.
- **Improve search precision and efficiency.** Strengthen historical data analysis and associative data processing through an improved text retrieval process and develop different indices for various content types such as text, graphics, audio, and video**.**

SOLUTION

- **Build a unified big data resource center.** Deploy a resource center that can handle large amounts of data using 80 servers based on the Intel® Xeon® processor E5-2600 v2 product family to allow real-time and unified management of both historical and future data.
- **Deploy XData big data analytics platform.** Adopt XData-DRAC* parallel database and xData-Hadoop* distributed system architecture to allow high-speed reading and writing, analysis, and computation of massive amounts of structured and unstructured data**.**
- **Add Intel® SSDs.** Use Intel SSD DC S3500 Series as a high-speed cache memory to enable near-real-time queries and retrievals, statistics, and analysis of data resources.

IMPACT

- **Met media industry's business demands.** The new solution allowed the centralized storage and management of massive data to meet the media industry's demands for unified data storage and business search requests of reporters and editors.
- **Increased user engagement and page access depth.** Being able to efficiently classify users into different groups through cluster analysis helps increase user engagement and page access while reducing the bounce rate and improving business.
- **Enhanced real-time query and response.** When tested with 2 billion pieces of news data including text, photos and videos, the XData big data analytics platform showed optimized performance, allowing simple, fuzzy, and multi-condition queries to receive real-time response.

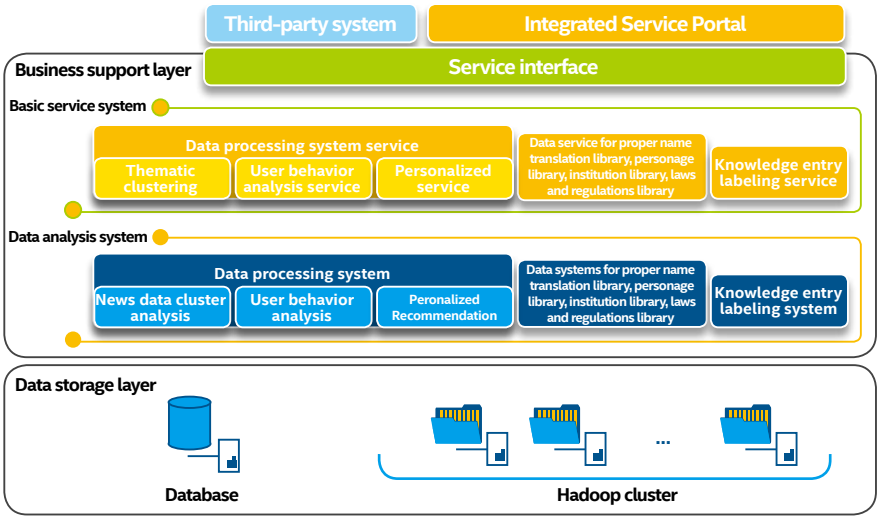
Explosive growth of data raises storage and management issues

Radio and TV media enterprises use various management platforms and file management application systems to enable digitalization of traditional media. Digital resources—like terabits of news releases and petabits of graphics, audios and videos—are usually stored in multiple scattered servers and under different departments inside media companies. As the number of news releases and graphics, audio, and video resources grows, finding data storage and management solutions to keep up with this massive explosion of data is a major challenge.

Value of media data not maximized

Historical data gathered from readers also faces the challenge of insufficient

Centralized storage, analysis, and real-time retrieval of massive amounts of digital data



Architecture of Sugon's XData big data analytics platform based on Intel Xeon processor E5-2600 v2 product family and Intel SSD DC S3500 Series

LESSONS LEARNED

- XData big data analysis platform based on XData-Hadoop* software strengthens data storage reliability and throughput performance, serving as an ideal storage solution for massive amounts of unstructured data such as graphics, audio, and videos.
- Intel Xeon processor E5-2600 v2 product family, combined with Intel SSD DC S3500 Series, enables high-speed, real-time big data processing and retrieval.

utilization due to challenges of retrieving information from unstructured resources such as graphic, audio, and video. News media data suffered from insufficient utilization. Thus, its value was not fully realized. On the other hand, consumers need information in their areas of interest without interruption from irrelevant information. The media industry is challenged to meet consumer needs by using the latest technology to maximize the utilization of media resources to serve the needs of consumers.

Unifying data storage and management

Sugon deployed its XData big data analytics platform at a national news agency in China using XData-DRAC and XData-Hadoop* software to allow centralized storage and management of structured and unstructured data from various sources.

"The news agency has more than 50 million news releases, several hundred million photos, and over 80,000 hours of audio and videos, for which we deployed 80 servers based on Intel Xeon processor E5-2600 v2 product family, with each server equipped with 10 3TB SATA hard disks," explained Zhennan. "The Intel Xeon processor E5-2600 v2 product family has a new platform architecture that supports up to 12 cores, enabling a higher level of concurrent processing. This effectively handled the workload for a data-intensive big data analytics platform."

Improving data retrieval

The XData big data analytics platform also improved accuracy of data retrieval for the news agency. As the platform adopts a feature word extraction algorithm, in combination with text data mining technology, it can extract feature words from news releases and identify implicit similarity among news. It enables reporters and editors to search for existing news data from a single data store and enhances news editing efficiency. This improves users' news browsing. Said Zhennan, "To support the high-speed data analysis and retrieval, we equipped each server with two units of 240GB Intel® SSD DC S3500 Series serving as high-speed cache memory."

Increasing user viscosity and page access depth

The XData platform can also classify users into different groups through cluster analyses, and analyze their composition, geographical distribution, positioning, and user habits. This has increased user viscosity and page access depth, reducing the bounce rate and allowing customers to figure out the causes for user conversion or loss, which boosts the company's market operations.

"Without the excellent computing capability of the Intel Xeon processor E5-2600 v2 product family and the high-performance data cache capability of Intel SSDs, analysis and mining of massive data would not be possible," said Zhennan.

Sugon expects to continue its close collaboration with Intel, relying on its technology expertise including its advantages in big data computing, to meet the increasingly complicated big data computing needs of the media industry.

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**).

0914/SHA/PMG/XX/PDF