Case Study

Intel® Xeon® Processor
Intelligent Storage

Intel®-based Large Object Store Reduces Disk Capacity Requirements and Associated Costs by 43 Percent

To meet expanding demand in Japan for high-capacity cloud archiving, backup, and content delivery, Sakura Internet is deploying a large object store using storage software from Amplidata and Intel® Xeon® processor-based hardware from Quanta QCT.

"Using Amplidata's AmpliStor software, the required storage capacity is approximately 1.7x the original volume of data—43 percent less than the 3x capacity required for triple replication, while providing an equal or greater degree of data durability. For massive amounts of data, this efficiency translates into huge cost savings."

–Katsuyuki Ikeda
Director of New Business Development
Sakura Internet

Summary

Challenge	Due to fast-growing demand from Japanese companies for cloud archiving, backup, and content delivery services, Sakura needed an efficient, extremely scalable solution that provides a high level of data durability.
Solution	Sakura Internet deployed a large object store based on a reference architecture developed by Intel, Amplidata, and Quanta. In testing, the scale-out storage system transferred data up to 3x faster than a competing cloud service.[1] The highly dense object store requires 43 percent less disk capacity for the same volume of data than traditional data-storage methods, with corresponding savings in infrastructure, floor space, and power and cooling costs. [2,3]

Delivering Object Storage Services in the Cloud in Japan

As adoption of cloud computing soars in Japan, Osaka-based Internet service provider Sakura Internet is delivering a growing range of cloud services aimed primarily at small-to-medium size businesses (SMBs) and software developers. Increasingly, Sakura's customers require cloud storage services that support archiving and backup for large amounts of unstructured and infrequently accessed "cold" data. Demand for disaster recovery protection has also increased since the devastating earthquake and tsunami that struck Japan in 2011, says Sakura Internet's Director of New Business Development, Katsuyuki Ikeda. With a lack of object storage services in Japan, the only options available to Sakura Internet customers were services hosted overseas, which suffered from performance and latency issues, according to Ikeda-san.

To provide cost-effective storage capacity to Japanese customers, the company has built a highly efficient cloud computing data center—the first of its kind in Japan and one of the most energy efficient in the world.[4] Sakura's Ishikari data center, strategically located in northernmost prefecture Hokkaido, gives Sakura a strong position in the highly competitive cloud services market.

When planning a large object store that could accommodate petabytes of unstructured data, Sakura began investigating technologies that could meet requirements for efficiency, and scalability. To accelerate time to market, Sakura took a highly proactive approach, searching worldwide for the best technology rather than waiting for local distributors to carry products and translate the associated documentation and specifications.

Erasure Code for Large Object Stores

Sakura Internet's extensive research resulted in the selection of a scale-out large object store solution using Intel Xeon processor-based servers from Quanta QCT and an object-based, software-defined storage solution with erasure code technology from Amplidata. The solution is based on a reference architecture jointly developed by Intel, Quanta, and Amplidata specifically to address the need for high-density, massive scalability, and cost-effective storage with high levels of durability.[5]

Sakura found that Amplidata's AmpliStor* software with erasure code offered much greater storage efficiency and cost savings compared to traditional data-protection schemes such as RAID and triple replication (see the sidebar "Erasure Code for Big Data"). The solution offered high levels of performance, scalability, and data durability. Another advantage was that it was optimized for the newest industry-standard Intel® processors.

The three companies worked together closely to meet Sakura's requirements. Intel Japan helped coordinate the process, which included months of proof-of-concept testing—first at Quanta's labs and then at a Sakura Internet data center, with Quanta engineers on site to help test and optimize performance and durability, and then to integrate the solution into Sakura Internet's environment. "Intel, Quanta, and Amplidata have a collaborative working style, and because each company has offices in the region, it is easy to schedule and conduct effective meetings with local teams. For ambitious projects like ours, this type of support is very important," says Ikeda-san.

Initially, the service is being delivered from the Ishikari cloud data center. As demand grows, the scale-out architecture will enable Sakura to quickly expand the initial 1.3 PB of object storage, while adding disaster recovery protection by spreading data across multiple data center locations.

Figure 1. Sakura Internet large object store configuration.

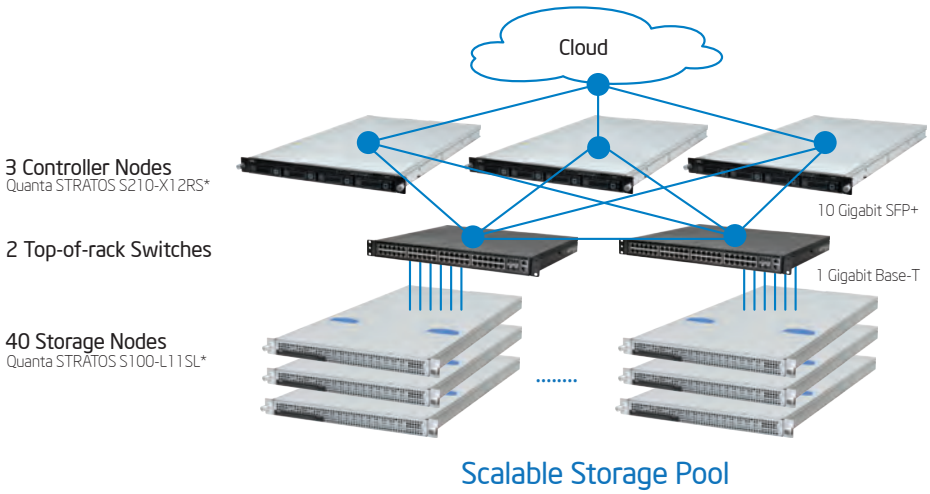

Table 1. Hardware Specifications

	Description	Quantity	
Quanta STRATOS S210-X12RS*			
CPU	▪ Intel® Xeon® Processor E5-2640 (15 MB Cache, 2.50 GHz, 7.20 GT/s Intel® QPI)	2	
RAM	▪ 4 GB ECC DIMM 1600 MHz	8	32 GB per controller
Network	▪ Intel® X520-DA2 Ethernet Converged Network Adapter ▪ Quanta 10 gigabit SFP+* dual port mezzanine card	1 1	10 GbE
Solid-state Drives	▪ Intel® Solid-State Drive 520 Series, 240 GB	2	
Hard Disk Drives	▪ 3.5" 2-TB SATA3, 7200 RPM	2	
Quanta STRATOS S100-L11SL*			
CPU	▪ Intel® Xeon® Processor E3-1220L (3 MB Cache, 2.20 GHz)	1	
RAM	▪ 8 GB ECC DIMM 1600 MHz	2	
Hard Disk Drives	▪ 3.5" 3-TB SATA3 7200 RPM	12	36 TB per node

Solution Delivers Higher Efficiency, Performance, and Data Protection

In preliminary testing, Sakura's scale-out storage system delivered extremely high performance, transferring data up to 3x as quickly as a competing overseas cloud service.[6] Tests also showed that the object store provides data throughput of nearly 3 GB per second, essentially saturating data-center network connections.[7]

AmpliStor's patented BitSpread* erasure coding algorithm enables the system to provide a very high level of data protection much more efficiently than traditional approaches for safeguarding data. Compared with triple replication, Sakura's object store requires 43 percent less disk storage, with corresponding savings in infrastructure, floor space, power, and cooling costs.

Sakura's production system is shown in Figure 1 and Table 1. It consists of a combination of three controller nodes, 40 storage nodes, and two top-of-rack switches, with a total capacity of 1.3 PB. The load-balanced controller nodes are Quanta STRATOS S210-X12RS* servers based on Intel® Xeon® processor E5-2640, with two 240-GB Intel® Solid-State Drive 520 Series for metadata and data caching and 10 Gigabit Intel® Ethernet Converged Network Adapters. This server platform is designed to maximize compute performance while minimizing power consumption and optimized for the demands of the densest data center environments. The storage nodes are Quanta STRATOS S100-L11SL* servers, based on Intel® Xeon® processor E3-1220L, with 12 3-TB SATA hard drives. This ultra-dense 1U storage server is designed to meet the scale-out storage needs of diverse data center applications. The controller nodes run AmpliStor's BitSpread rateless erasure coding algorithm, which encodes and distributes the objects to be stored on the storage nodes. The controllers also decode and assemble objects that are retrieved from the storage nodes.

Efficiency

A key benefit of the solution is higher storage efficiency compared with traditional data-protection schemes such as RAID and triple replication. For example, Sakura Internet ultimately plans to spread the object store across three data centers, protecting data even if a disaster incapacitates an entire facility. "Using Amplidata's AmpliStor software, the required storage capacity is approximately 1.7x the original volume of data—43 percent less than the 3x capacity required for triple replication, while providing an equal or greater degree of data durability. For massive amounts of data, this efficiency translates into huge cost savings," according to Ikeda-san. "With a legacy [triple replication] approach, if you want to store 1 PB of data, you need at least 3 PB of disk storage." With competing systems he considered, nearly 130 servers were required to deliver the same total capacity.

To keep the storage system healthy and optimized without the need for manual intervention, Amplidata's other key patented technology, BitDynamics*, provides automated out-of-band storage management functions such as capacity management, the ability to add storage nodes on the fly, continuous data integrity verification, self-monitoring, and automatic data healing, scrubbing, and garbage collection.

Erasure Code for Big Data

The massive growth of unstructured data is driving a need for large object stores that can provide extremely high levels of performance and scalability. Amplidata AmpliStor* storage software is designed to address these challenges using erasure coding, an alternative to traditional data-protection schemes such as RAID. AmpliStor's BitSpread* rateless erasure coding algorithm mathematically transforms data and spreads it across multiple storage nodes in such a way that information can be recovered in the event of multiple drive failures, node failure, or other data loss. The object store also eliminates the overhead of a traditional file system. While erasure coding algorithms are compute-intensive, Intel® Xeon® processors deliver the required performance and scalability. The technology is attractive for many applications. For example, many sports organizations and film and music companies incur the expense of maintaining vast media archives on tape; by transferring them to object stores and making them available online, the archives are much more widely accessible—and become revenue-generating assets.

Additionally, the AmpliStor stack is fully abstracted from the underlying hardware—meaning that AmpliStor is truly object-based, software-defined storage. Abstraction enables the durability policy to be changed on the fly, without having to move or touch data, which is a huge effort in a RAID system. This enables Sakura to best optimize durability, performance, and capacity over time.

Performance

AmpliStor's BitSpread erasure coding builds more intelligence into the storage system, taking advantage of the performance and rich instructions provided by the Intel processors to achieve benefits such as reduced capacity, higher data throughput, and greater scalability. "Our BitSpread algorithm has been optimized for Intel CPUs," says Wim De Wispelaere, Amplidata's chief technology officer. For example, Amplidata's software extensively uses the Intel® Streaming SIMD Extensions (SSE) 4.2 instruction set and takes advantage of the multiple cores in Intel Xeon processors. BitSpread and optimization for Intel CPUs allows system performance to scale linearly, squeezing all the performance out of each controller node. Depending on the number of controllers, throughput of tens of gigabytes per second is possible.

As a result, Sakura's large object store achieves extremely high performance. In testing, the system provided more than 3x read performance and 2.4x write performance with 1 MB objects, compared to a cloud object storage system from a leading provider. In another test, a three-controller configuration delivered aggregate throughput of more than 2,900 MB per second, essentially saturating the system's 10 Gb Ethernet connection. "Our storage system can ingest and retrieve data faster than the typical Internet connections in data centers," De Wispelaere says.

Sakura Internet's High-efficiency Cloud Data Center

Sakura Internet's Ishikari data center is one of the first in Japan to be designed specifically for large-scale cloud computing. The company built the data center in response to fast-growing customer demand and the recognition that an extremely efficient data center was needed to compete with global cloud service providers who are providing cloud services in Japan from advanced data centers in other countries.

Sakura Internet located the data center in the city of Ishikari, in northern Japan, partly because the cool climate enables the use of outside air for "free" data center cooling. By greatly reducing the need for air conditioning, this helps the data center achieve very high levels of energy efficiency, measured by a low power usage effectiveness (PUE) ratio of 1.11. Sakura Internet further increases energy efficiency by using high-voltage direct current (HVDC) rather than alternating current within the data center. The data center uses a modular building design that can be expanded incrementally, taking advantage of the latest technology at each stage and ultimately accommodating as many as 600,000 servers.

Data Protection

With AmpliStor, Sakura Internet can adjust the level of protection provided through a policy-based approach, up to 15 nines of availability. For example, the company can set a policy that enables the object store to be distributed across multiple data centers, with the ability to survive the loss of a complete data center plus an additional drive. BitDynamics enables the system to be self-monitoring and self-healing; in the event of a disk failure or data loss, the system repairs data in parallel even before a service intervention occurs, and without the lengthy, performance-impacting rebuilds required by RAID systems. Subsequently, service interventions like drive replacements can be scheduled on a periodic basis, such as monthly.

Conclusion

Sakura Internet expects that the large object store based on Intel Xeon processor-based hardware from Quanta QCT and storage software from Amplidata will provide a highly scalable, cost-effective solution that meets rapidly growing demand for big data services such as archiving, content delivery, and backup. As demand grows, the flexible, scale-out design will enable Sakura Internet to quickly expand capacity and add multisite disaster recovery protection.

For more information about large object stores using erasure code, visit www.intel.com/go/storage, www.quantaqct.com, and www.amplidata.com.

Quanta™ Optimize Your Datacenter

AMPLIDATA

SAKURA Internet

---

[1] Sakura Internet internal testing.

[2] Sakura Internet and Amplidata cost data.

[3] The TCO or other cost reduction scenarios described in this document are intended to enable you to get a better understanding of how the purchase of a given Intel product, combined with a number of situation-specific variables, might affect your future cost and savings. Circumstances will vary, and there may be unaccounted-for costs related to the use and deployment of a given product. Nothing in this document should be interpreted as either a promise of or contract for a given level of costs.

[4] See http://ishikari.sakura.ad.jp/index_eng.html.

[5] "Intel® Cloud Builders Guide: Scale-out Storage Infrastructure with Amplidata and Quanta" www.intelcloudbuilders.com/index.php?option=com_productsearch&view=displaysearch&&lang=&filter_companyname=1350&filter_category=&filter_efficiency=&filter_region=&company_name=Quanta&Itemid=181

[6] Sakura Internet internal testing.

[7] "Amplistor Delivers eXtreme Performance: Object Storage System Ingests Data at Near Line Speed." DeepStorage, LLC. http://amplidata.com/?wpdmact=process&did=MjMuaG90bGluaw==

Printed in USA 0114/CM/VF/PDF Please Recycle 329526-001US