Technical Case Study

Delivering Faster: Converting Real Estate Property Data into Actionable Information

How RP Data Reduced Microsoft SQL Server-Based Report Processing Times by 70% with NetApp EF-Series Flash Arrays

RP Data is the largest provider of real estate property valuation information, analytics, and risk management services in Australia and New Zealand. Owned by CoreLogic, one of the world's largest data and analytics companies, RP Data takes in large volumes of data every day. Real estate and property data, including images and financial information, is the foundation of its business. Success hinges on how quickly the company can acquire, cleanse, link, and store the information to provide timely reports to its customers, which include real estate and financial services companies as well as homeowners, investors, and government organizations.

RP Data continuously collects and manages complex real estate property data and imagery from numerous disparate sources. It maintains more than 130 data feeds from government agencies, industry professionals, and proprietary relationships, as well as internally generated data. This highly variable and complex data is quickly transformed into relevant and actionable information and analytics, helping the company's subscribers make confident real estate property-related decisions.

Australia's real estate market has been strong in recent years, and RP Data has experienced double-digit year-over-year growth as well as a significant increase in the volume of data it manages. To keep up with this growth and address customer requests for even faster turnarounds for property valuation reports, it was necessary for the company's IT team to perform a review of the underlying technology stack.

Goals for Property Information Database

With more than 100TB of new data coming in every year, RP Data stores an average of five images of nearly every piece of real estate in Australia and New Zealand. The company's property management database, which runs on Microsoft® SQL Server® 2012, is its most critical workload. The database contains more than 500 million property decision points spanning more than 9 million properties. Fast database transfer times are absolutely crucial to providing accurate, relevant, and actionable data to customers faster. As the database grew, storage I/O became a concern.

"We need to make sure that the time to market from when we receive the data to when it gets to our customers is as short as possible," says Simon Perry, general manager of technology at RP Data. "Having highly available, fast storage is critical to achieving that goal."

To overcome performance issues and improve the customer experience, RP Data wanted to implement storage technology that would allow it to:

- Reduce database transfer times.
- Produce more frequent customer reports.
- Provide high availability and rock-solid reliability at the storage layer.
- Partner with a trusted advisor that could bring together technology leaders to help architect an integrated solution.

How RP Data Accelerated Report Processing Using NetApp

Initially, RP Data considered taking the traditional approach of overprovisioning its storage with more partially filled spinning drives to solve the immediate storage I/O problem. However, with space at a premium at the company's two active-active Brisbane data centers, it needed a solution that could provide greater density and long-term scalability. RP Data turned to NetApp.

"NetApp's desire to understand our business has been very different from other technology companies we've worked with, so we trust their recommendations," says Adrian Jansz, head of Information Communication & Technology at RP Data. "Rather than selling us a traditional disk-based solution that would have met our immediate needs, NetApp carefully considered our business requirements and explained how an all-flash storage solution could provide longer-term performance and scalability benefits for our high-I/O workloads."

RP Data engaged NetApp Global Services and NetApp partner Bridge Point Communications to conduct a proof of concept to demonstrate real-world results for a subset of its property management database. Incoming data from three financial institutions is stored on solid-state drives (SSDs) on a NetApp® EF-Series flash array connected to servers by Fibre Channel. The extract, transform, and load (ETL) process runs and then property valuation reports are generated for financial services customers.

"One reason our proof of concept was successful is that NetApp was able to provide the right people for the right jobs," says Jansz. "They have technical people who can talk at a business level and vice versa. That helps us bridge any potential gaps between the business and IT when deploying solutions."

During deployment, the team leveraged the EF540 reference configuration for Microsoft SQL Server 2012. "It was reassuring to base our deployment on a well-thought-out reference configuration," says Jansz. "The recommendations fit our environment like a glove, which gave our technical team a lot of comfort."

The workload used for the proof of concept is now in production, and RP Data's financial services customers already see the benefits of more frequent, timely reports. The full ETL process will be moved into production on NetApp EF-Series flash arrays in the coming months as RP Data deploys additional systems.

"We need to make sure that the time to market from when we receive the data to when it gets to our customers is as short as possible. Having highly available, fast storage is critical to achieving that goal."

Simon Perry
General manager of technology, RP Data

Ultimately, RP Data standardized on NetApp storage for its entire business, based on:

- NetApp's demonstrated market leadership in virtualized storage
- Its superior flash storage technology with options to fit the requirements of variable workloads
- The company's extensive network of technology partners, such as Cisco, VMware, and Microsoft, that RP Data leverages for other infrastructure components
- The ability to achieve true nondisruptive operations by migrating to the NetApp clustered Data ONTAP® operating system, as described later in this paper

"The outcome of our proof of concept with the NetApp EF-Series flash array is a perfect example of why we standardized on NetApp," says Jansz. "NetApp helped us solve what could have been a major business problem. As a result, we improved database performance to the point where we can get different cuts of the data multiple times a day, allowing us to provide more frequent reports to our customers."

Storage Innovations Behind Accelerated Data Processing

NetApp flash storage technologies were instrumental in helping RP Data achieve its business goals for faster data processing. "We were confident that the NetApp EF-Series was the best flash solution for our business," says Jansz. "It's a mature, enterprise-ready flash offering that doesn't sacrifice reliability for speed."

> "With the NetApp EF-Series flash array, we can consider workloads that were just not possible before, allowing us to provide real-time data to customers much more efficiently and effectively."
>
> **Adrian Jansz**
> Head of Information Communication & Technology, RP Data
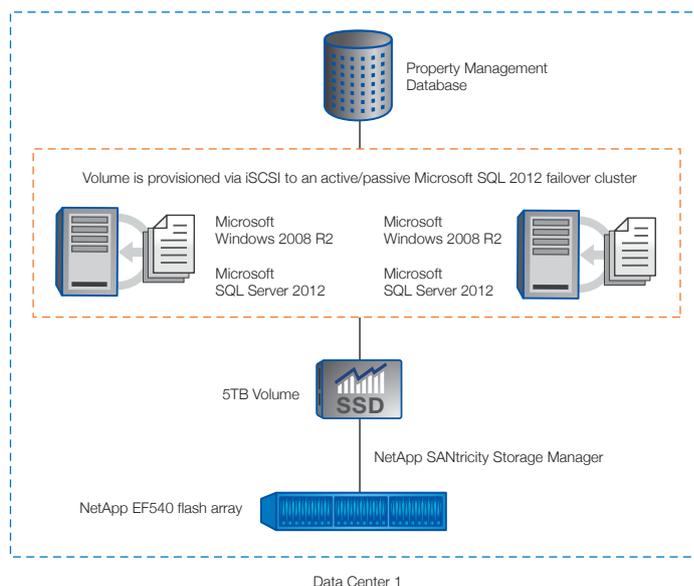
RP Data's NetApp EF-Series flash array configuration for Microsoft SQL Server 2012

Driving Down Latency

Designed specifically for database-driven business operations that demand the highest levels of performance and availability, the NetApp EF-Series flash array delivers the performance of over a thousand 15,000-RPM drives while requiring approximately 5% of the rack space, power, and cooling. As a result, RP Data can provide submillisecond latency and extreme IOPS in a compact, scalable footprint.

"With the NetApp EF-Services flash array, we can give our data team what they've always asked for: more IOPS and faster data processing," says Jansz. "They can run ETL processes throughout the day. Before, they had to do one big build and then splice that out into smaller batches, which takes time and introduces complexity. We're now able to provide updated data multiple times a day instead of weekly for our customers to make business decisions with."

Simplifying Storage Management

RP Data engineers use NetApp SANtricity® Storage Manager software to manage the EF-Series array. Optimized for flash, SANtricity software allows storage administrators to achieve maximum performance and utilization through configuration flexibility and custom performance tuning. All management tasks can be performed while the storage remains online with complete read/write data access, enabling storage administrators to make configuration changes and conduct maintenance without disrupting I/O.

"SANtricity Storage Manager is a great real-time, day-to-day management tool for our administrators," says Jansz. "Dynamic capacity and volume expansion as well as dynamic RAID-level migration make storage management transparent to the business."

Maintaining High Availability

One of the most critical aspects of an enterprise-ready solution is the ability to detect and resolve issues. The NetApp EF-Series flash array offers extensive capturing and monitoring of diagnostic data to provide comprehensive fault isolation and simplify analysis. SSD wear life is proactively tracked, and alerts are issued when the threshold is reached. To allow recovery in the event of an environment issue or component failure, the array includes fully redundant I/O paths with automated failover.

Supporting Business-Critical Applications

Other areas of RP Data's business rely on NetApp FAS3240 storage systems connected with Cisco® switches over SAN protocols, including Fibre Channel and iSCSI. Approximately 70% of the server environment is virtualized with VMware® vSphere®, supporting Linux®- and Windows®-based business applications. NetApp SnapMirror® software replicates changed data blocks between the two Brisbane data centers for disaster recovery purposes.

RP Data serves Windows file shares from the NetApp storage using direct CIFS, eliminating the need to purchase and maintain an additional NAS gateway. "The multiprotocol flexibility that NetApp offers is hugely beneficial," says Jansz. "It removes a whole layer of cost and complexity."

The NetApp AutoSupport™ tool, a web-based remote support diagnostics system, checks the health of the NetApp systems on a continual basis. "We get phone calls from NetApp alerting us to any issues before we're even aware of them," says Jansz.

Offering New Data Services with FlexPod

As RP Data prepared to take on new business challenges, it recognized the need for a converged infrastructure platform to speed application deployment and reduce costs. The company sought a standardized, validated solution that could be easily deployed and duplicated if necessary to respond quickly to customer and market demands. The FlexPod® Express architecture—an integrated computing, networking, and storage solution developed by Cisco and NetApp—was the perfect fit.

"We're deploying FlexPod as the basis of a PCI-compliant rack that will enable us to take advantage of new business opportunities with financial services customers," says Adrian Jansz, head of Information Communication & Technology at RP Data. "We'll be able to use Fibre Channel over Ethernet to converge server and storage traffic, keeping our footprint small and manageable."

IT Benefits and Business Impact

With its property management database benefitting from an all-flash array, RP Data can deliver time-sensitive property information to banks and valuation firms more frequently to enable better-informed decision-making. Property data collected over the weekend—a crucial window for real estate activity—can be processed quickly on Monday morning, and reports can be delivered to customers in time for them to take action on the information before it becomes stale.

"With the NetApp EF-Series flash array, we can consider workloads that were just not possible before, allowing us to provide real-time data to customers much more efficiently and effectively," says Jansz. "The ability to produce valuations on a property faster, or produce multiple batches of valuations on a particular property, will help our customers provide their customers with better experiences."

Because of the NetApp EF-Series, RP Data has been able to:

- Reduce database transfer times by over 70% (50 minutes versus 3 hours).
- Decrease the time needed to process reports by approximately 70% (3.5 hours versus 12.5 hours).
- Enable 99.9999% system availability at the storage layer.

"We've seen in excess of 100,000 IOPS in real-world use cases running the ETL process for our property management database on the NetApp EF flash array," says Jansz. "We've achieved data transfer rates in excess of 1.2 million records per second. That means that processing 10GB of data takes less than a minute."

RP Data's newfound capabilities will help the company innovate and find new ways to assist the real estate market in Australia and New Zealand. "The less time we waste processing data, the more bandwidth our teams have for analytics and product development," says Perry.

What's Next

RP Data is in the process of upgrading its NetApp systems to FAS8040 controllers running the clustered Data ONTAP operating system, giving it the ability to move data nondisruptively between storage controllers during business hours. To accelerate read- and write-heavy workloads that don't require pure flash, RP Data is considering adding SSD shelves to the FAS8040 controllers. It can then use NetApp Flash Pool™ intelligent caching, a feature included with Data ONTAP, to automate storage tiering between SSDs and spinning drives.

"You can't really put a value on the kind of flexibility we get with NetApp. Having the agility to do what we want when we need to is priceless," says Jansz.

About NetApp

Leading organizations worldwide count on NetApp for software, systems and services to manage and store their data. Customers value our teamwork, expertise and passion for helping them succeed now and into the future.

Product List
- NetApp EF540 flash array
- NetApp FAS3240 storage systems
- NetApp Data ONTAP 8.1.3 operating system
- NetApp SANtricity Storage Manager
- NetApp OnCommand® management software
- NetApp Snapshot® and SnapRestore® technologies
- NetApp SnapMirror software
- VMware vSphere 5.5
- IBM and Cisco UCS® servers
- Cisco Nexus® and Cisco Catalyst® switches
- Windows Server® 2012, 2008
- Red Hat, CentOS Linux
- Microsoft SQL Server 2012 database
- PostgreSQL database

Services
- NetApp SupportEdge Premium
- NetApp AutoSupport

NetApp®

www.netapp.com