

## Rakuten Cloud-Native Platform Enables Analytics Services

**Tier-1 operator deploys a scalable Dell Data Lakehouse with Rakuten Cloud-Native Platform that provides a private cloud solution for analytics that is easy for data scientists to use**



**Rakuten Symphony**

Communications services providers (CoSPs) are building out their multi-access edge compute (MEC) service delivery infrastructure to meet customer needs for low-latency use cases such as IoT, private 5G, artificial intelligence (AI)/machine learning (ML) inference, content delivery and analysis of massive datasets. With MEC infrastructure these applications can be run locally without requiring transport to a remote data center.

Deploying cloud services for data analytics workloads at scale is complex and requires deployment and management of hardware and complex software layers. A typical AI cloud solution includes GPU-enabled servers, an operating system foundation and a hypervisor for virtualization and containerization. On top of this foundation are AI and analytics pipeline applications. Instantiating and life cycle management of these software layers and testing them before deployment takes a lot of time which doesn't deliver the scalability and flexibility advantages needed to make cloud-based analytics services successful.

A tier 1 CoSP wanted to avoid this complexity as it was developing its analytics service. The operator turned to Rakuten Symphony, an Intel® Network Builders ecosystem Titanium partner, to develop the private cloud that provides the compute needed for analytics applications infrastructure.

### Tier 1 Operator Delivers Data Analytics at the Edge

As part of its data management services family, the operator created an analytics-as-a-service platform. The goal was to provide agile, scalable and automated delivery of analytics workloads.

It was important for the CoSP to have a standardized solution for the analytics but also to have application pipeline delivery and lifecycle management. The solution that Rakuten Symphony proposed was a data lakehouse that could be delivered as a service.

Data lakehouses are ideal for data analytics because they process diverse data types including structured and semi-structured data that are currently stored in data lakes, along with a range of unstructured data types including sensor data, rich media, geo-spatial data, audio, weather data and other challenging data formats. The data lakehouse can sort through structured and unstructured data in real time in order to provide processed and relevant data for business intelligence, analytics and AI applications.

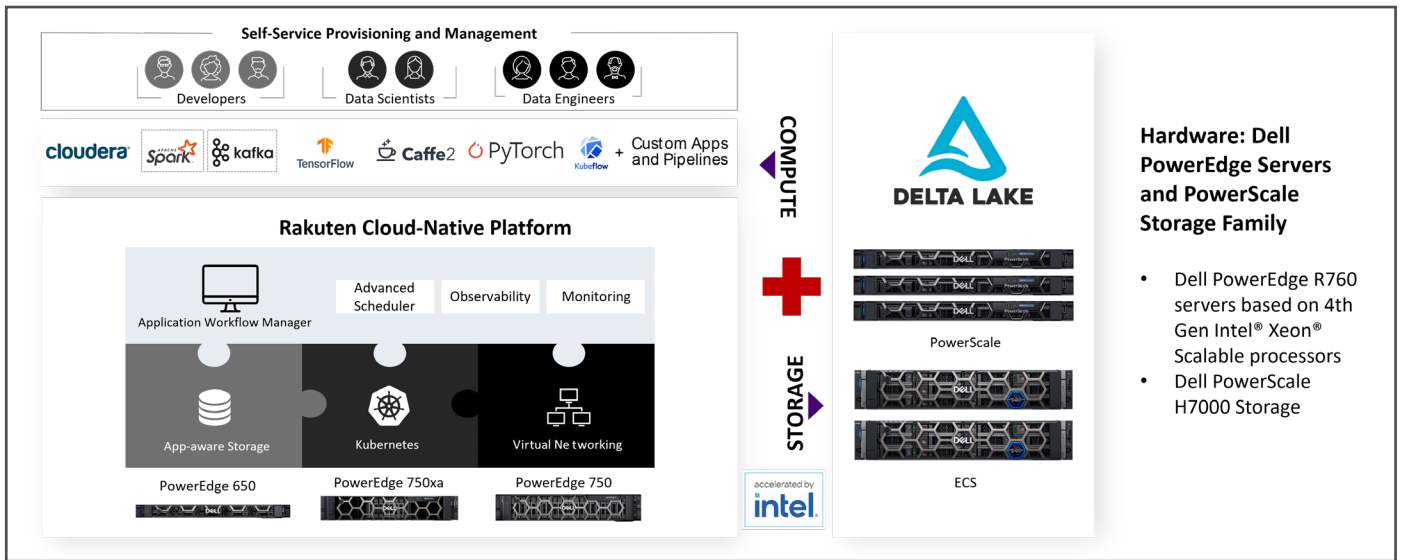


Figure 1. Block diagram of a data lakehouse solution featuring the Rakuten Cloud-Native Platform.

In building out its analytics service, the tier 1 operator needed a technology stack that met the needs of data scientists who will use the service to provide real-time analytics to support their businesses. The foundation of this service includes the following open-source software:

- Delta Lake was the choice for a storage layer that runs on top of an existing data lake and improves its reliability, security, and performance.
- Apache Spark is included in the application stack as a unified analytics engine that data scientists can use to perform exploratory data analysis (EDA) on datasets up to petabytes of data.
- Kubeflow delivers machine learning (ML) and ML system deployment and operations (MLOps) on Kubernetes. Kubeflow is a tool data scientists can use to efficiently manage and scale machine learning workflows.
- TensorFlow allows data scientists to create AI/ ML models. It is a software library used for training and inferring of deep neural networks.

### Analytics as-a-Service for Data Scientists

The next step in making this data lakehouse into a service was working with the tier 1 operator’s IT team to build a cloud infrastructure that was simple to deploy and scalable enough for data scientists, so it didn’t delay or complicate their work. Some of their unique needs included (see Figure 2):

- Self-service rapid workload provisioning
- Simplified application management including QoS
- Performance optimized to achieve desired SLAs
- Highly available platform with data protection

The cloud solution also had to meet the needs of the tier 1 operator’s IT team to simplify their cloud management and provide a flexible, scalable compute and storage platform. Their goal is to provide the scalability needed to enable the data scientists to have an on-demand computing environment that they can easily scale up or down for very compute- and data-intensive services.

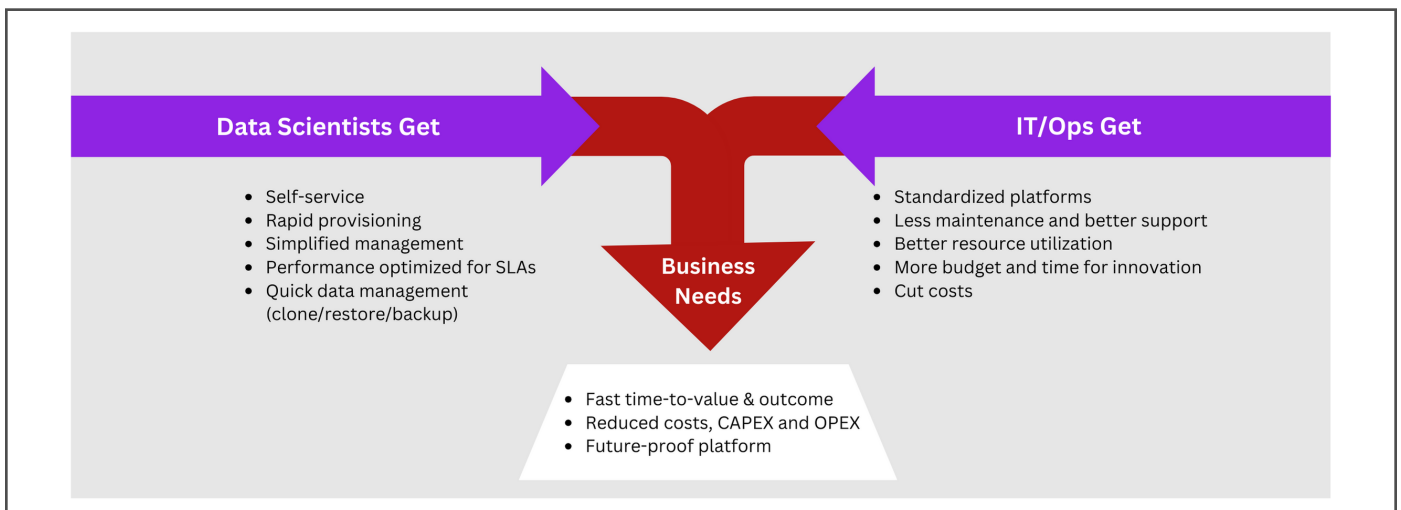


Figure 2. Customer and IT/Ops needs from edge cloud infrastructure.

## Rakuten Cloud-Native Platform Supports Complex Data-Centric Applications

Rakuten Cloud-Native Platform (CNP) was used to deliver a private cloud for data analytics services. Rakuten CNP is an industry-leading Kubernetes platform that is optimized for running storage and network intensive applications such as data lakehouses (see Figure 3). The platform can automate complex applications in minutes and deliver cloud-native agility and scalability for complex data-centric applications – perfect for data scientists setting up their own private cloud.

Rakuten CNP provides an Application Workflow Manager layer on top of Kubernetes that can onboard new applications using a single click interface or by using an API. The software can deploy a complete application pipeline seamlessly.

Rakuten Cloud-Native Storage (CNS) provides app-aware storage that is important for managing the variety of data types that come in a data lake environment. The software creates a local storage abstraction layer on each edge node, making this local storage available to workloads running on other edge nodes.

This software defined storage solution offers advanced application awareness, high availability, data resiliency, and security that enables customers to migrate their stateful workloads to containers seamlessly and transform their edge deployments. Rakuten CNP also has a full slate of storage features including snapshots, clones, replication, backup, data rebalancing, tiering, thin provisioning, encryption, and compression.

## Comprehensive Post-Deployment Management Tools

Once the software is installed and configured, the Application Workflow Manager offers multiple post-deployment management tools including an advanced

scheduler, application observability and monitoring. These tools enable Rakuten CNP to offer a full slate of application-aware storage features including snapshots, clones, replication, backup, data rebalancing, tiering, thin provisioning, encryption, and compression.

### 4th Gen Intel® Xeon® Scalable Processors

4th Gen Intel® Xeon® Scalable processors offer high throughput and low latency and are engineered for on-prem or cloud deployments. The processor family's architecture combines high-performance processor cores with up to eight built-in and workload-specific accelerators for maximum performance efficiency. Integration of accelerators into the processor redefines CPU architecture and provides a more efficient way to achieve higher performance than relying solely on increasing the CPU core count for workload processing.

The integrated accelerators help to reduce a server's power consumption and complexity when compared to a server that uses discrete accelerator cards for additional performance. This power advantage is important to edge servers.

These processors offer up to 80 lanes of PCIe 5.0 connectivity and support Compute Express Link (CXL), a cache-coherent interconnect for processors, memory expansion, and accelerators. With accelerated matrix multiply operations, the 4th Gen Intel Xeon Scalable processors have exceptional AI training and inference performance. Other seamlessly integrated accelerators include Intel® QuickAssist Technology (Intel® QAT) that speeds up data encryption/decryption and compression for faster query throughput for more responsive analytics.

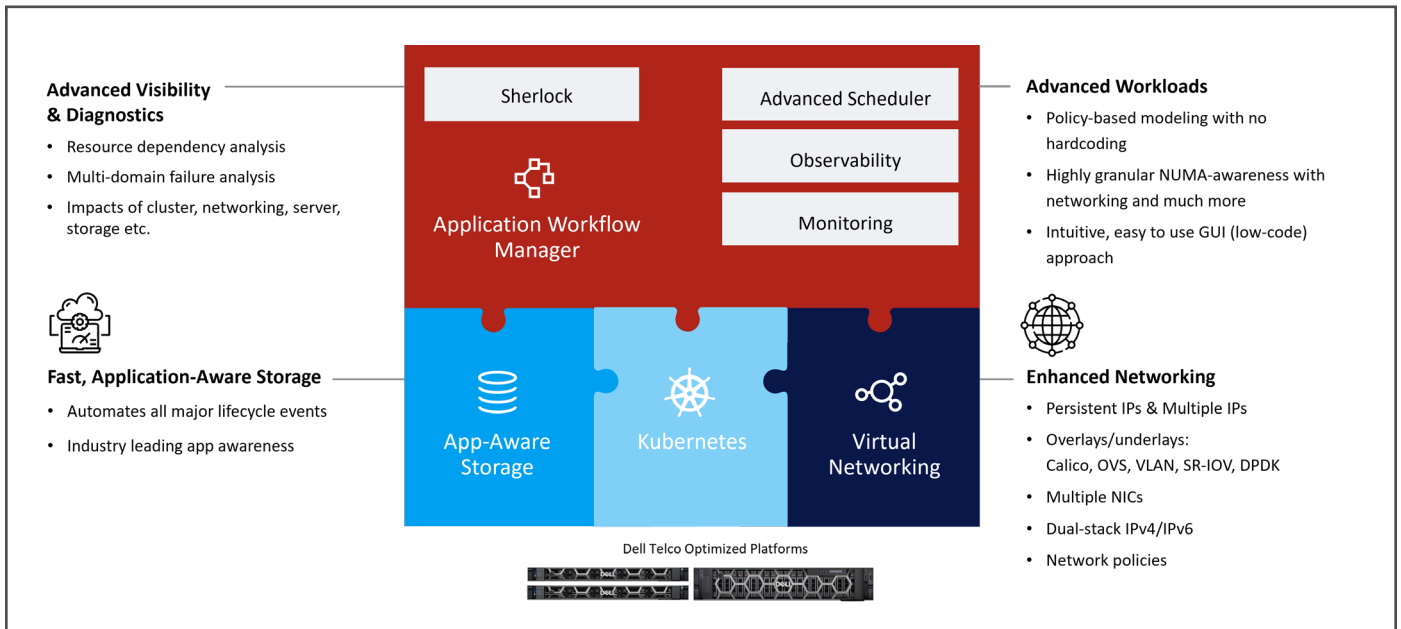
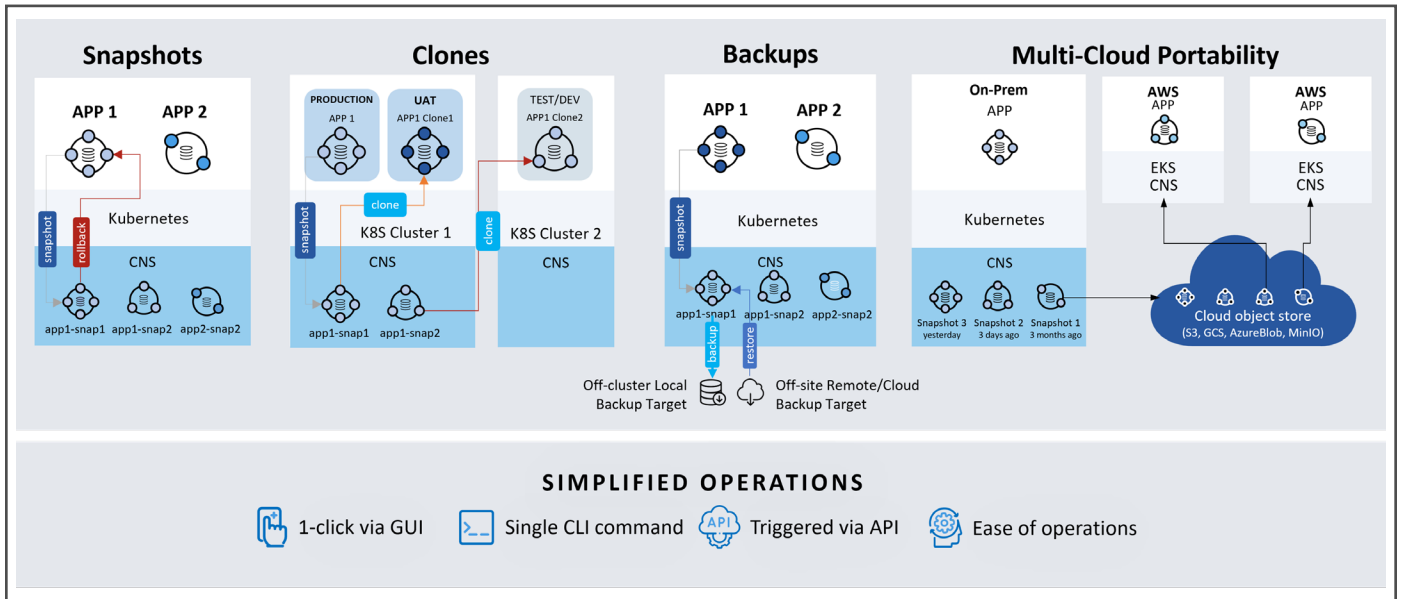


Figure 3. Rakuten Cloud-Native Platform features and functionality.



**Figure 4.** Application-aware storage features offered by Rakuten Cloud-Native Storage include snapshots, clones, backups all with multi-cloud portability.

### Enabling Low Latency and Reduced Data Backhaul Through Edge Cloud

Edge cloud services are essential for meeting the needs of enterprises that are digitizing their operations using IoT, AI/ML, private 5G, and data analytics. These applications benefit from being processed in edge servers for low latency or to reduce data backhaul. This was the case for a tier 1 operator customer that was developing a data analytics service for customers with very large datasets – too large to be transported to a remote data center. To build out its edge compute capabilities, the CoSP chose the Rakuten Cloud-Native Platform optimized on 4th Gen Intel Xeon Scalable processor-based servers. Rakuten CNP builds on Kubernetes offering workflow management, orchestration and application-aware storage. With this CPU-based infrastructure, the tier 1 operator’s data analytics service was delivered on an edge cloud server that met the needs of both customers and internal IT teams.

### Learn More

[Rakuten Cloud-Native Platform](#)

[4th Gen Intel Xeon Scalable processors](#)

[Intel Network Builders Ecosystem](#)



### Notices & Disclaimers

Intel technologies may require enabled hardware, software or service activation.  
 No product or component can be absolutely secure.  
 Your costs and results may vary.  
 Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.