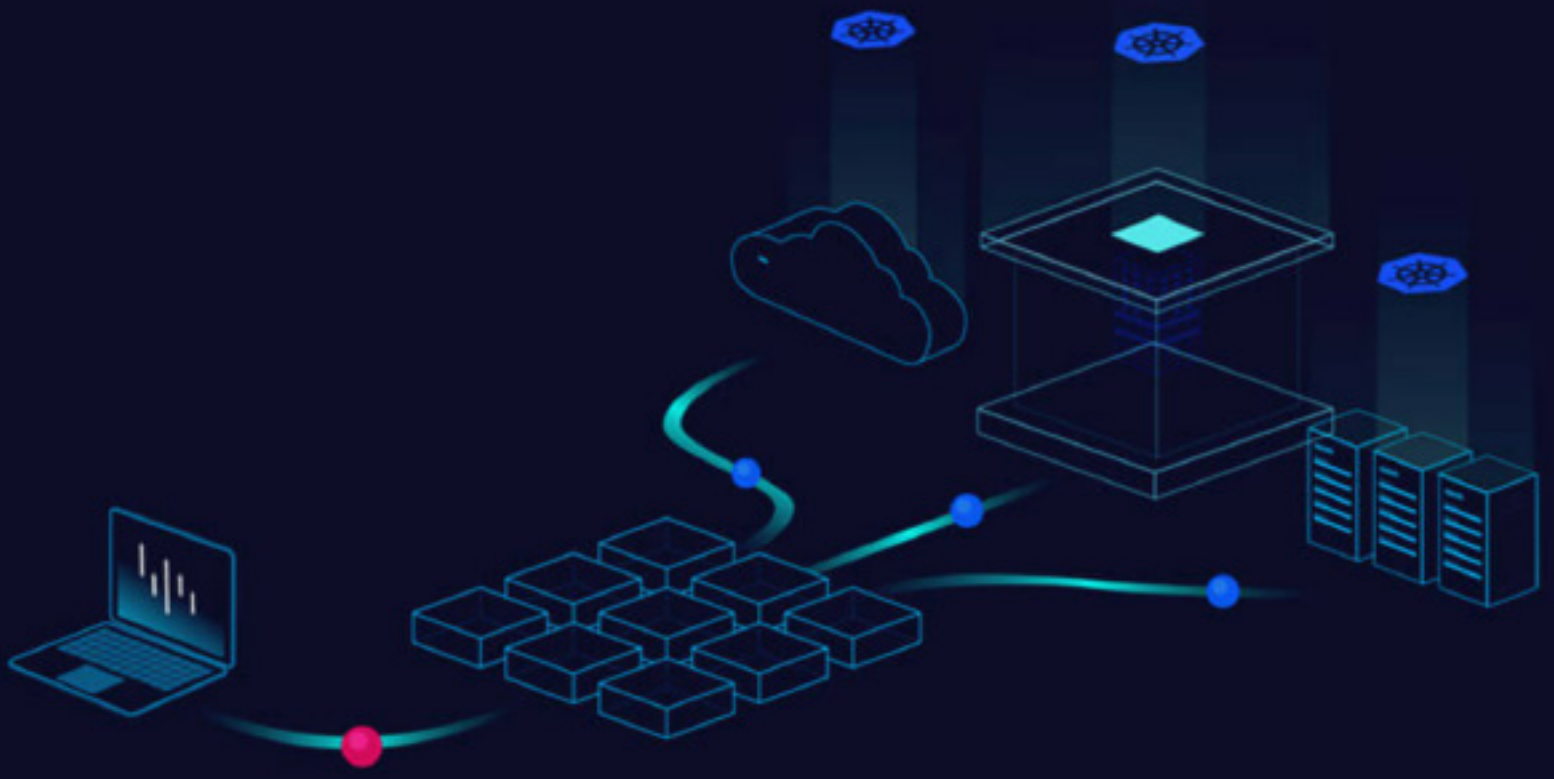
PLATFORM9

White Paper

Enabling the Software-Driven Store:

Solution Architecture for Cloud-Native Infrastructure and Operations



Contents

Introduction

Online and brick-and-mortar retailers understand that the user experience is what drives sales. Consumers want an entertaining, frictionless journey.

They want to be recognized when they visit a page or walk in the door, they want personalized choices and personalized deals, and they want one-click simplicity when they check out or pick up goods. For online retailers, delivering these features typically involves IT and DevOps teams running apps in public clouds or in their own data centers.

However, brick-and-mortar stores are becoming their own mini data centers, hosting sophisticated AI/ML apps that leverage sales data, store-level transactions, purchase histories, and supply-chain and advertising data. For example, when a mobile app reports a customer walking near a store location, perhaps in a mall, it texts a coupon with a discount for an attractive item. When CCTV picks up a license plate entering a store parking lot, the store's server evaluates their recent online browsing history. When trying on an item, digital mirrors running off the local server provide real-time color choices and sale incentives.

Retailers need to be able to deploy and update innovative apps to thousands of geographically distributed physical stores just as easily as they can in their data centers or public clouds.

To take advantage of all these location-based marketing and sales opportunities, retailers need to be able to deploy and update innovative apps to thousands of geographically distributed physical stores just as easily as they can in their data centers or public clouds, and have the ability to do so in a consistent and unified way.

To succeed, they need to:

- Automate infrastructure management and application updates across all of their stores.
- Enable DevOps CI/CD pipelines to bring new capabilities to stores faster.
- Make retail IT work in more public-cloud ways.
- Deploy and centrally manage any type of workload — containers, VMs, or bare metal — across all of their store locations, data centers, and public clouds.

The cost and management challenges of running retail edge apps

Brick-and-mortar retailers face a number of significant challenges in running state-of-the-art apps across data centers, public clouds, and 1000s of stores.

These challenges include:

- Onboarding stores and quickly deploying apps to thousands of servers in distributed store locations.
- Providing a single unified and shared platform for consistent use by multiple users and domains such as business-intelligence analysts, marketing teams, and customers.
- Remotely updating and troubleshooting edge operations and apps and avoiding onsite visits.

What makes meeting these challenges so difficult is that with so many outlets, retailers need to create lean server-per-store and admin-per-store ratios within small facilities. The solution is to manage highly distributed edge IT and DevOps centrally while virtualizing and containerizing the infrastructure and applications running in the stores. Platform9's unique architecture and cloud operations model solves this problem elegantly.

SaaS architecture to manage IT infrastructure across thousands of stores

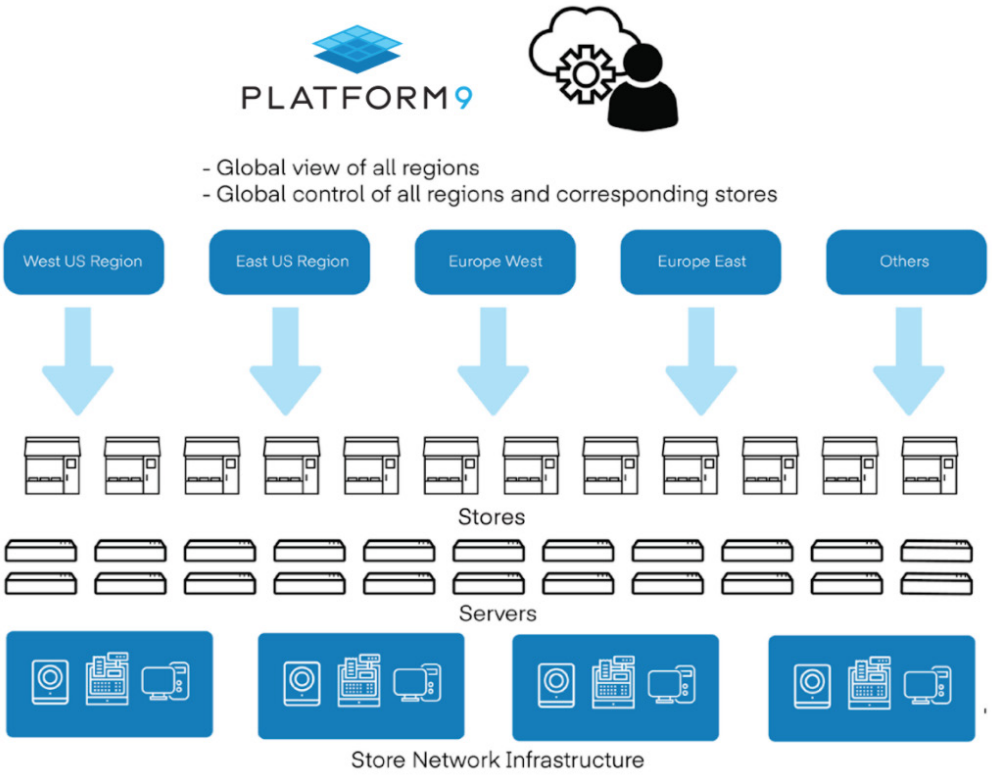
Platform9 has a next-generation, distributed-cloud-to-edge architecture that provides the operational simplicity and ease-of-use of public clouds.

The managed Platform9 service maintains an open environment for open-source, cloud-native stacks including upstream Kubernetes, Prometheus, and OpenStack. Our cloud operations teams deploy and manage distributed infrastructure sites remotely without requiring onsite delivery. A centrally deployed management plane orchestrates the delivery of various capabilities — containers, hypervisors, storage backends, and network backends — to physical infrastructure using automation and operational tools.

Retailers no longer have to worry about the operational burden of up-time/SLA management, upgrades, security patches, and production outages. This approach is fast becoming the standard for distributed cloud-native infrastructure management because of its ability to address large-scale operational complexity and to avoid lock-in.

The example below shows a typical regional and WAN retail architecture. Platform9's centralized management view enables global control of all regions and corresponding stores with a single-pane-of-glass approach.

A centrally deployed management plane orchestrates the delivery of various capabilities — containers, hypervisors, storage backends, network backends — to physical infrastructure using automation and operational tools.

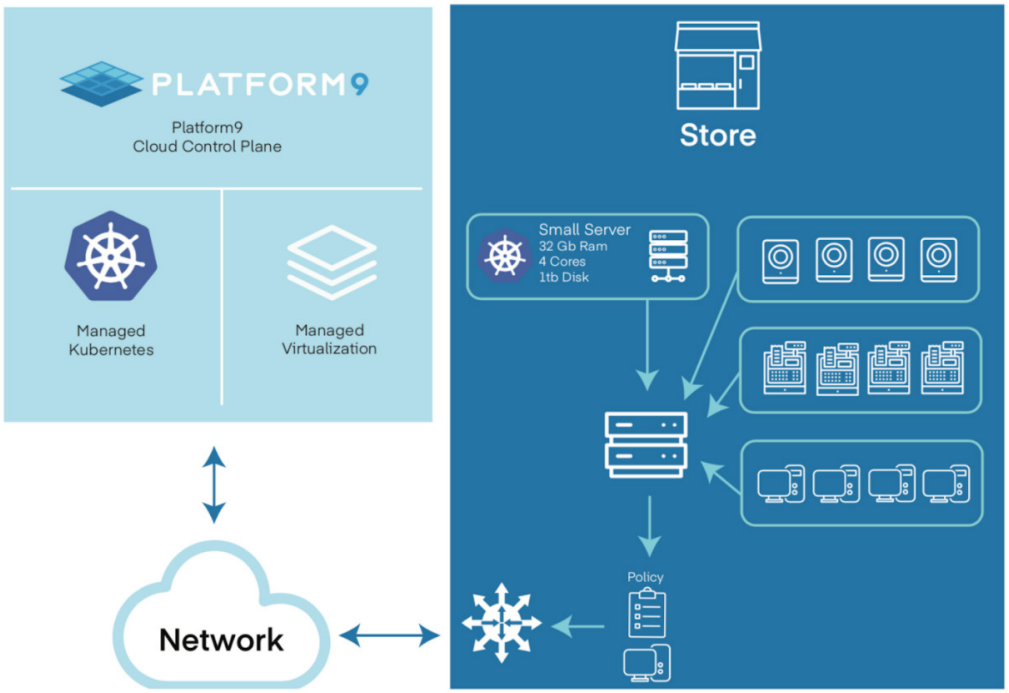


A global chain of smaller outlets such as coffee shops tends to deploy similar architectures to foster networking consistency, lower costs, and smaller IT footprints. A small server (32 GB RAM, 4 cores, 1 TB storage) can serve as a clearinghouse for transactions from all POS terminals, security cameras, and other devices.

In this example, the in-store app is an intelligent order management solution that accepts and schedules customer food and beverage orders (including mobile orders) and coordinates available ovens, coffee machines, and terminal loads with customer arrivals. It also controls waiting times to increase customer satisfaction and speed service.

Stores communicate with a managed Kubernetes instance at their headquarters via an in-store WAN router with a VPN internet connection. Each location's router typically maintains different networking policies, CCTV and POS/IoT devices, and customer access to a hotspot.

Platform9 monitors all infrastructure 24/7/365 with a 99.9% uptime SLA and if a physical component such as a server dies, Platform9 alerts the regional or main office's IT group.



The advantage of this architecture and operational model is that there's no dedicated IT staff at the store location. Platform9 monitors all infrastructure 24/7/365 with a 99.9% uptime SLA and if a physical component such as a server dies, Platform9 alerts the regional or main office's IT group. At the store, someone can take a generic server off the shelf and plug in the network port. The new server bootstraps itself into an OS running Kubernetes. It then connects to clusters in an AWS environment that provision it with all the latest apps. In less than 30 minutes, the store is back online.

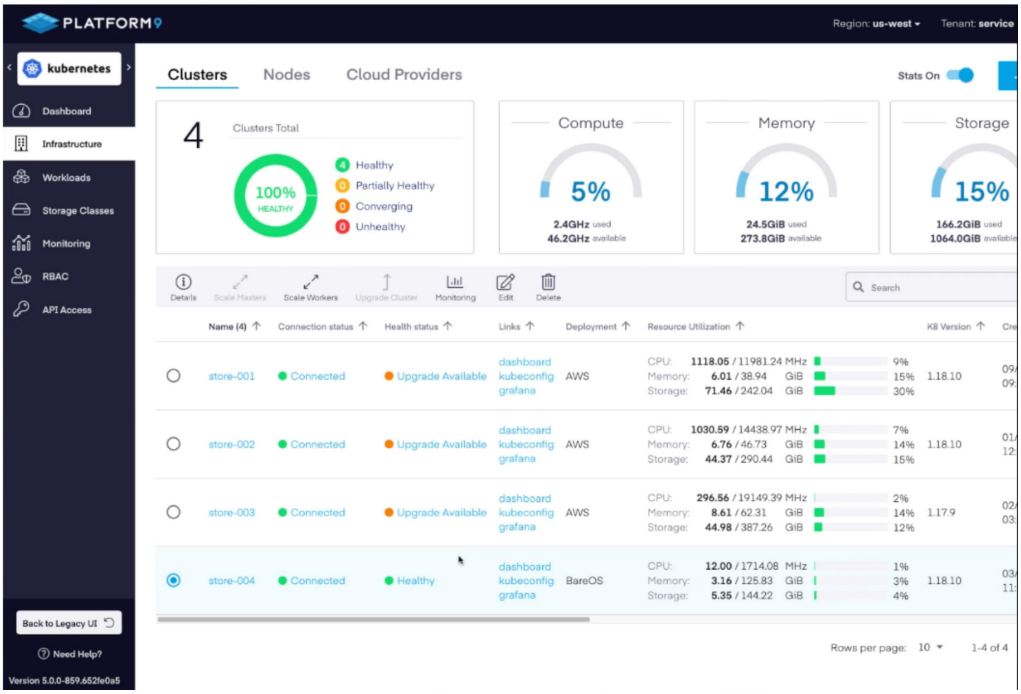
Central management is critical

The Platform9 management plane provides efficient operational automation.

Central management is an important trend in the retail community as they don't want to have resources spread across the country, with engineers flying and driving to deal with issues on-site. Rather, they want self-service for IT and DevOps and to be able to manage all of their edge locations from one data center — monitoring, troubleshooting, provisioning, and upgrading apps and infrastructure.

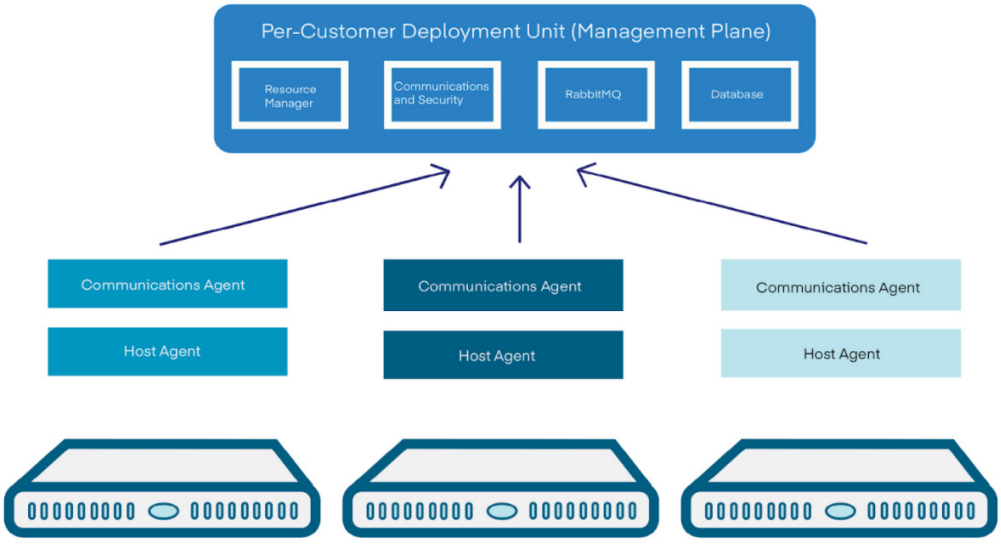
A control or management plane shouldn't dictate the footprint of an app whether it's a monolithically-built three-tier web application or just microservers deployed in a container with several dependencies. You may have a generic server running a bare-metal OS with a container workload on top. You may have VMware or OpenStack running in VMs or Kubernetes running on either VMs or bare metal.

Platform9 supports all these different configurations within one UI and from one location — they should all be first-class citizens of the overall IT platform.



Modern cloud services such as Kubernetes are composite services that are themselves highly distributed and can require complex orchestration.

To make central management of distributed store infrastructure and applications work, the Platform9 management plane provides efficient operational automation. Platform9 deploys a management plane (called the deployment unit) for each region and the associated communication and host agents on the retail store infrastructure. The following diagram shows how this architecture provides all the necessary components to manage and remotely operate 1000s of retail stores.



Automated cloud operations

The management plane has a number of services and components that automate deployment, monitoring, troubleshooting, and upgrades of the specific open-source stacks that are being used. The following topics are important considerations.

"Few organizations have the expertise or want to follow the many associated open-source projects needed to build their own container management software stack."

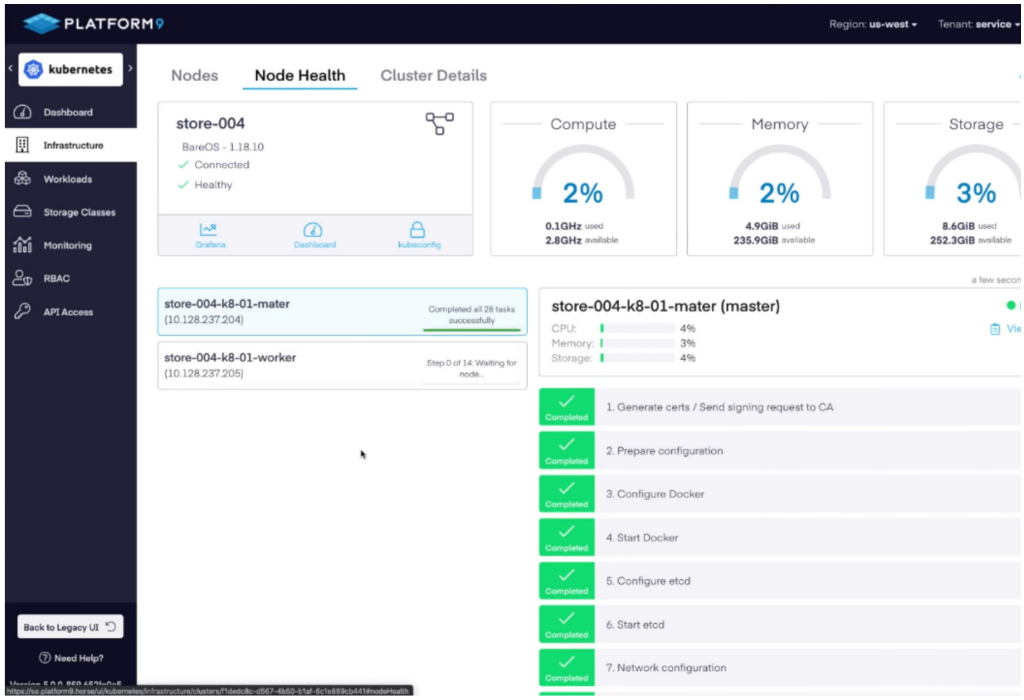
— Gartner, Market Guide for Container Management

Deploying a cloud service like Kubernetes

Most large retail organizations are at some stage of evaluating and implementing Kubernetes to run containerized apps. Without the overhead of replicating different OSs, they can run more apps on the same hardware. Apps are more portable, and because data and configuration are decoupled, containers are by nature high-availability. For these and other reasons such as easy service discovery and centralized configuration, Kubernetes is usually the most appropriate operating system for retail apps.

Modern cloud services such as Kubernetes are composite services that are themselves highly distributed and can require complex orchestration. Deploying control plane components requires addressing redundancy and high-availability capabilities. Most cloud services also require data plane components such as Kubernetes worker node components kubelet and Docker. These worker nodes make sense only in the context of a certain master node. Therefore, the system must support dependency resolution and the integration of data plane components with control plane components.

The Platform9 SaaS management plane makes it simple to centrally deploy Kubernetes services at every retail store by automating all the tedious and complex steps. Importantly, Platform9 uses pure upstream Kubernetes source code straight from the GitHub repo and therefore does not lock customers into its Kubernetes implementation.



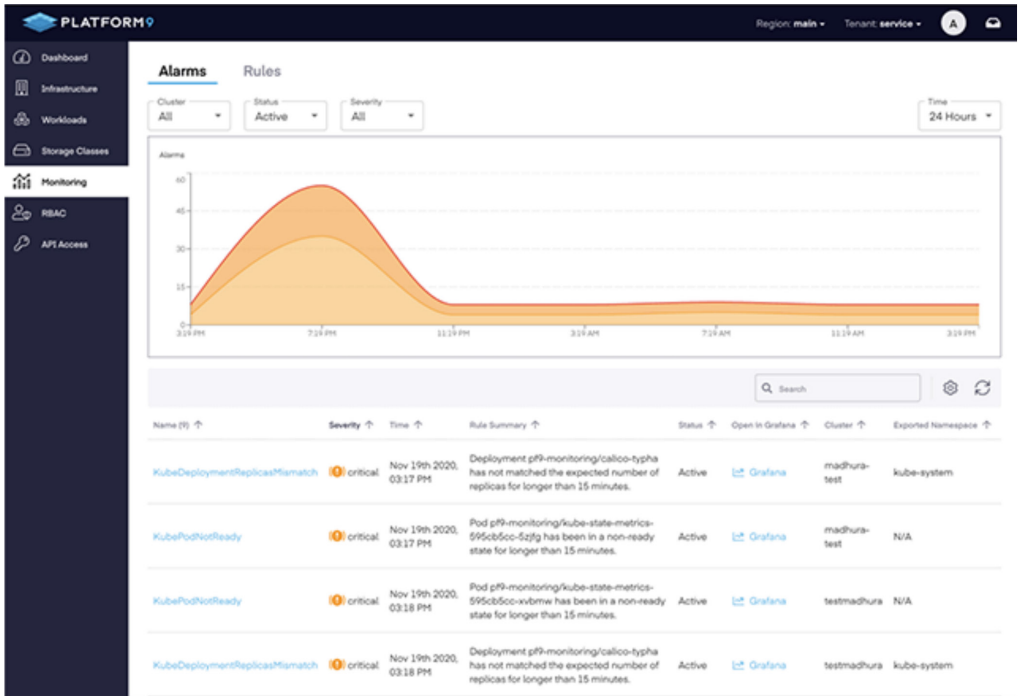
Monitoring a cloud service

Operators need a bird's eye view of all infrastructure activities and be able to quickly drill-down to specific locations. Monitoring and alerts should all be fired off into the central dashboard so the operator can immediately identify what's up and what's down. Should they call the store and have a clerk plug in a server or check the power? In addition, automatic alerts trigger automated processes like reboots or other operations to resolve issues.

Every cloud service and its components can define a set of health metrics which must be monitored on a continuous basis. Since modern cloud services are

highly distributed, a small degradation in certain components can lead to a larger system-wide degradation. To simplify troubleshooting (whether automated or human) and to minimize performance issues, these health probes need to be highly granular.

Platform9 developed a SaaS monitoring system using Prometheus and Cortex where retail infrastructure is remotely monitored in real-time. This reduces the burden on retailers and helps uncover potential problems early. The monitoring system ingests up to 6.1 billion events per day and raises alerts for over 65 different types of events in the distributed retail infrastructure.



Platform9 enables simpler deployments into new locations compared to a manual setup and fully automated deployments after locations have been brought online.

Applying diagnostics to automate/aid troubleshooting

Having good health metrics provides a basis to codify the resolution of common problems via automated runbooks. These runbooks can be built for common problems that occur during normal system operation such as a control plane going offline because of an infrastructure failure.

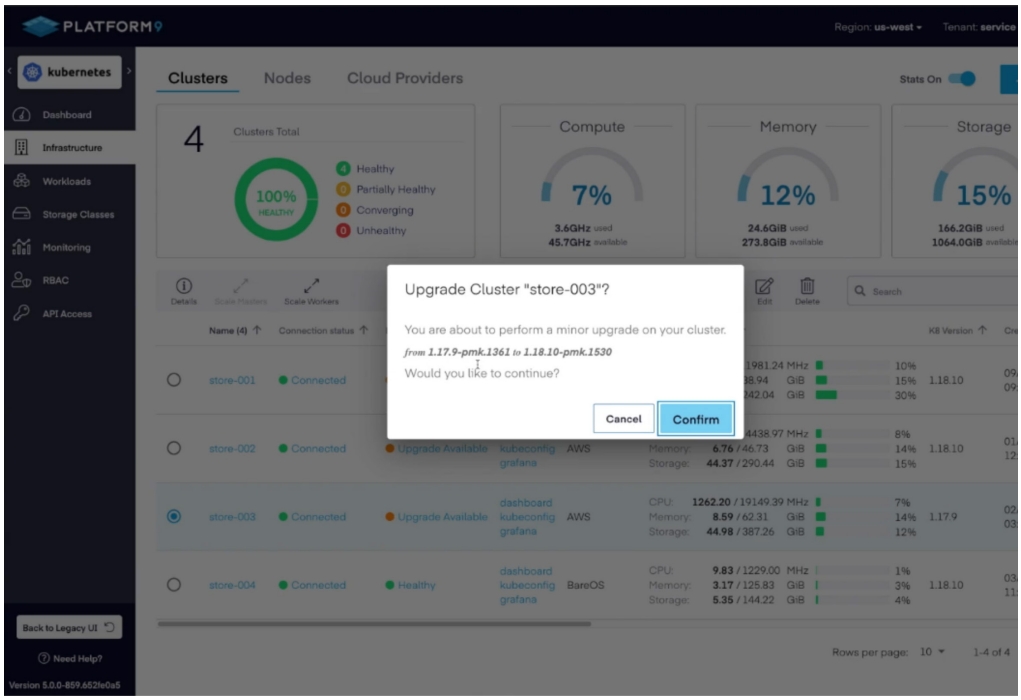
These runbooks are also effective when there are problems with new versions of cloud services or interoperability issues that are found only in the field after a deployment. Since a runbook can operate without requiring a new version of the cloud service, immediate mitigation can take place while developers create a bug-fix or a new version of the cloud service in question. In this way, customers can be operational despite the complex, ever-evolving nature of modern open-source cloud technologies.

Versioning to automate non-invasive, repeatable upgrades

Developers and vendors competing in modern open-source ecosystems are constantly developing new versions both with bug fixes as well as with security and feature enhancements. The management plane makes it easy for customers to stay up to date by fully automating upgrades to new versions.

Developers typically offer these upgrades on a granular basis. For example, upgrading Service A can be independent of upgrading Service B. This makes change control easier for large-scale enterprise deployments. Ideally, they enable self-service to let customers schedule their upgrades when and where they wish.

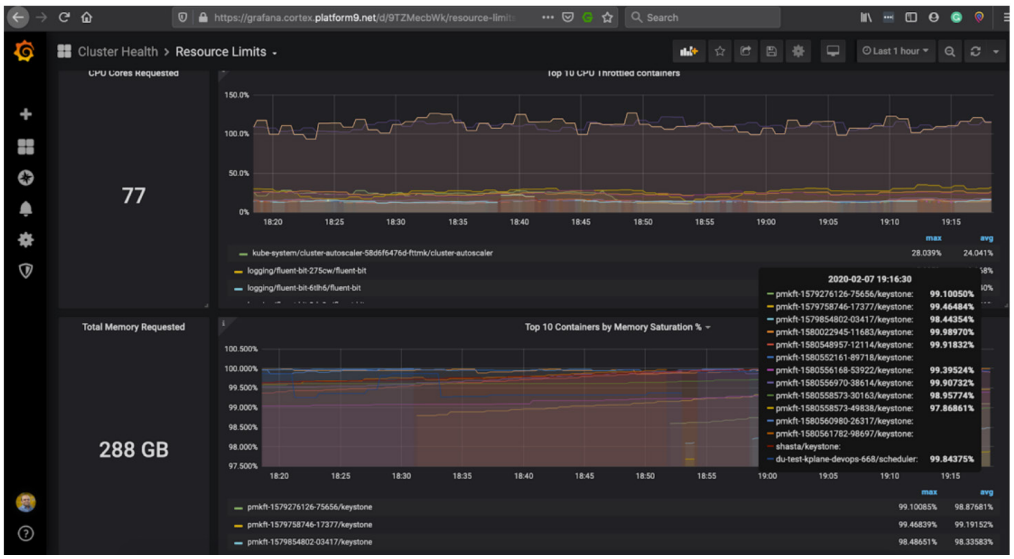
Automated cloud operations can reduce deployment times from months to days. They also optimize health monitoring, runbook-driven resolution of common problems, and streamlined upgrades, with an operational SLA for distributed clouds.



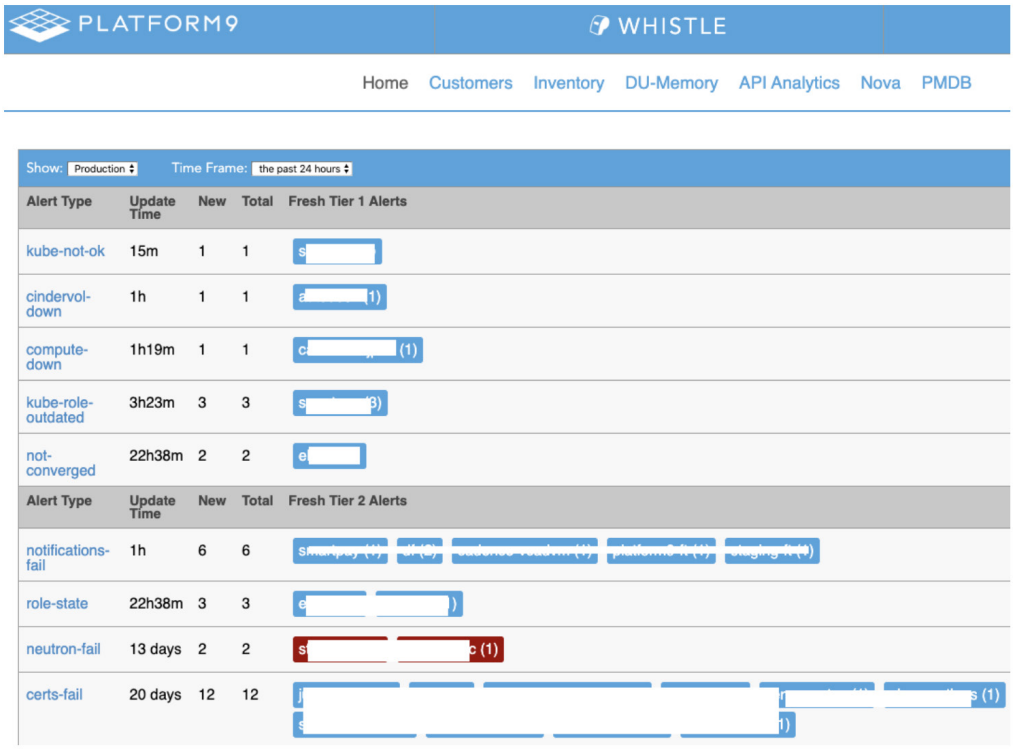
Operational tooling

Retail IT operations typically have resource constraints at remote locations. Small stores mean fewer servers and no onsite IT, necessitating a centrally-operated management plane. There's also a greater likelihood of unpredictable network latency, jitter, and disconnects. Resiliency depends on individual locations being able to reliably keep doing business regardless of connectivity issues.

Platform9 engineers have spent years minimizing the impact of latency as an architectural construct, particularly for environments that depend on a percentage of cellular bandwidth. Platform9 operations teams use a variety of tools for service delivery. The teams monitor and troubleshoot customer installations using metrics gathered in Prometheus and Grafana and efficiently search across large log sets using log aggregation based on the Elasticsearch Logstash Kibana (ELK) framework.



This methodology reduces the number of actionable alerts that Platform9's support teams need to watch, making it easier to quickly identify priority issues as well as the different sites that may be affected by them. The following figure shows how high volumes of raw metrics and alerts can be reduced to a set of actionable alerts and impacted sites (actual customer names are redacted).



Platform9 enables using commodity servers and open-source frameworks and tools across the stack. This means retailers can ship new hardware to onboard new locations without pre-configuring boxes.

Software profiles standardize deployment and configuration

The goal for retailers should be to deploy and maintain apps at edge locations in a standardized manner — and with zero human touch. That means first, setting up cloud-controlled networking with Meraki-like devices. Second, setting up an appropriately configured VLAN. Third, deploying the OS and virtualization layer. Lastly, deploying the apps on top of that environment, using APIs, API gateways, and the like. It's leveraging standardized templates, stores, and configs to get things up-and-running while a GitHub repository stores all the artifacts for every store location.

The fundamental idea is for developers to create apps that are either microservices-based or even traditional three-tier and define them in infrastructure. When the infrastructure knows what apps need, it can provision them using templates in real time. This eliminates the manual work of defining firewalls, clusters to compute, backend databases, and complex interconnects.

The Platform9 management plane uses software profiles (listings of software packages, versions, and dependencies) and configuration profiles (listings of key configuration parameters) to ensure that nodes under management are provisioned consistently with a high degree of repeatability. This provides several benefits:

- Deploying new nodes with confidence since software profiles and configuration profiles are well known and tested and cannot drift.
- Rebuilding failed nodes quickly and reliably.
- Significantly reducing manual work, troubleshooting, and firefighting.

Key machine templates include:

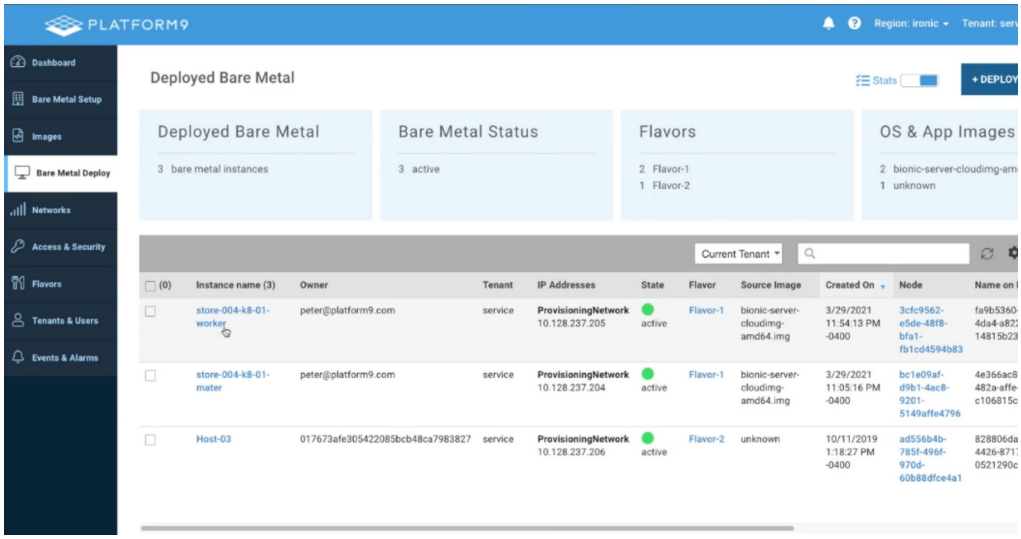
- Hypervisor — openstack-nova, libvirt, and others
- Storage backend — openstack-cinder
- Network node — openstack-neutron
- Storage node — openstack-ceph
- Container-visor — Kubernetes kubelet, kubeproxy, Docker, and others
- Kubernetes master — etcd and others

Platform9 does not lock customers into its Kubernetes implementation: the Platform9 wrapper on top of Kubernetes or the specialized Platform9 API. Our goal is to enable deploying open-source technologies across all your edge locations.

Support for existing store IT infrastructure

Platform9 enables using commodity servers and open-source frameworks and tools across the stack. This means retailers can ship new hardware to onboard new locations without pre-configuring boxes. They also avoid vendor lock-in and enable innovation. Servers should have no dependencies; they can be Intel NUC-type servers, nothing high-end, just readily available. Local hub ports are helpful.

Platform9 can manage these from their bare metal state, putting them into an OS and then automating a bootstrap into a Kubernetes cluster or virtualization. This makes it very easy to manage from one location. Managing from bare metal all the way into virtualization and infrastructure lets you run all kinds of apps in each retail location. You can run monolithic, micro-monolithic, and container-based apps in a simple single-server or dual-server setup.



Platform9 works with a variety of x86 servers and other IT infrastructure that retailers have already invested in for their storage and networking environments. It supports running in centralized core data centers for private clouds and can operate with very low "touch" in hybrid, distributed edge environments. It also supports running on physical machines and existing virtualized environments that use VMware or KVM.

By ensuring consistent management across all of these environments, this architecture delivers the proverbial "single pane of glass" for managing clouds.

The following diagram summarizes the overall approach. Automated cloud operations can reduce deployment times from months to days. They also optimize health monitoring, runbook-driven resolution of common problems, and streamlined upgrades, with an operational SLA for distributed clouds. Before, this was only available via hyperscale public clouds. Now, Platform9 gives retailers the operational simplicity and peace of mind of public cloud computing across 1000s of stores, data centers, and public cloud properties.

Platform9 gives retailers the operational simplicity and peace of mind of public cloud computing across 1000s of stores, data centers, and public cloud properties.

The benefits of this approach include:

- Quick delivery of innovative applications to address fast-changing customer needs and desires.
- Fewer expensive hardware investments and refreshes by using existing footprint to roll-out new applications.
- Reduced operational costs with automated and centralized remote management capabilities.

Conclusion

The brick-and-mortar store is the new battleground for retailers. They need to deliver outstanding in-store shopping experiences that drive new revenues while maintaining operational efficiencies.

Retail environments with top customer experiences will require AI and smart systems that leverage complex cloud infrastructure. And with distributed, resource-constrained outlets, a central management plane that can automate and simplify deployments and upgrades and provide single-pane-of-glass visibility into cloud infrastructure is essential.

With Platform9, retailers can modernize their stores without having to rip and replace their existing hardware. They can run their existing legacy application while rolling our new innovative cloud-native AI applications using Kubernetes as the underlying cloud fabric. Platform9 simplifies the retail IT modernization journey, greatly accelerates the time to get stores up and running from months to days, and reduces operational costs by 90% as a result of hyper-automation at all layers of the store infrastructure stack.

Request a free trial or book a demo — info@platform9.com