

TECHNICAL WHITE PAPER

Automating Portworx Data Services on Equinix Metal

Learn how simple it is to configure and run Portworx® on Equinix Metal.

Contents

Introduction	3
How to Use This Guide	3
Portworx Data Services on Equinix Metal	3
Equinix Metal	4
Portworx Data Services	4
Automating Deployment of Portworx Data Services on Equinix Metal	6
Prerequisites for Running the Terraform Module	6
Executing the Terraform Module	7
Benefits of Automating Portworx Data Services on Equinix Metal	7
Validating Apache Kafka on Equinix Metal and AWS Bare Metal	8
Test Environment	8
Kafka Performance Comparison: Test Results	9
Conclusion	12
Appendix	12
About the Authors	13



Introduction

In a rapidly-evolving business landscape, organizations must deal with a variety of challenges to stay competitive. The challenges range from modernizing legacy technology infrastructure to managing high-volume data-driven applications. In addition, as data continues to scale and grow, the demand to create insight and value from that data is higher than ever before. While many organizations believe that transitioning to a cloud platform is the answer, it can lead to multiple challenges, including massive cost overruns and cloud lock-in.

This is where bare metal as-a-service (BMaaS) can make a significant difference. Bare metal as-a-service is a model in which a vendor deploys dedicated physical IT infrastructure or bare metal to the customer's data center with the same on-demand scalability, convenience, and agility of a cloud service.

Portworx Data Services on Equinix Metal is a platform-as-a-service (PaaS) solution on dedicated, bare-metal hardware, physically managed by Equinix Metal. Portworx Data Services enables the deployment of data services on Kubernetes. The Equinix platform is fully interconnected with every major public-cloud provider and features more than 356,000 global interconnects to nearly every telecom and ISP provider. Enterprises can take advantage of PDS on the Equinix Metal platform to obtain cloud-like capabilities without the costs and risks associated with such cloud platforms.

How to Use This Guide

This white paper explains the benefits of automating the deployment of Portworx Enterprise and Portworx Data Services on the Equinix Metal server using a custom-built Terraform module provided by Pure Storage®. In addition, this paper validates the deployment of Kafka using Portworx Data Services on Equinix Metal, and compares the performance of Kafka when it is deployed using Rook/Ceph on Equinix Metal and using Portworx Data Services on AWS bare metal.

The target audience for this document includes, but is not limited to, platform engineers, devops engineers, systems architects, database administrators, storage administrators, and IT professionals.

Portworx Data Services on Equinix Metal

Portworx Data Services on Equinix Metal is a simple, on-demand, hosted, full-stack, (network, storage, compute) deployment. As depicted in Figure 1, the automation of the storage components, server components, and network connectivity—known as down stack automation—is solved by Portworx Data Services on Equinix Metal. After deploying infrastructure within a matter of minutes, customers can quickly install up-stack solutions, whether they are containerized, virtualized, or bare metal applications.



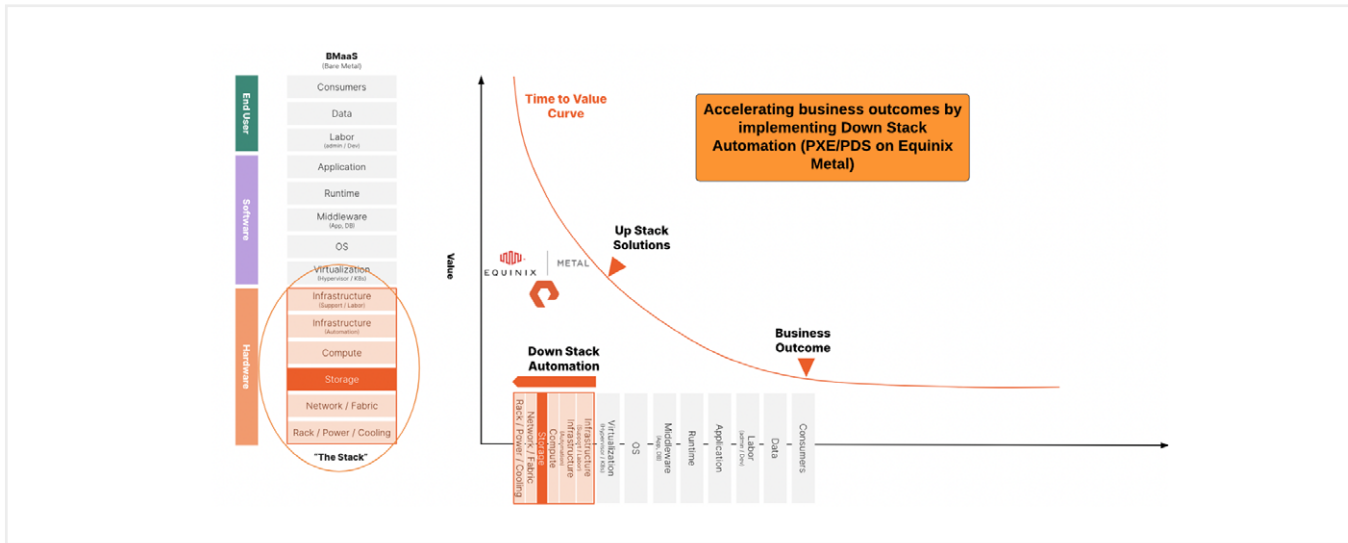


FIGURE 1 Time to value of bare metal as-a-service

Portworx Data Services on Equinix Metal Server has the potential to implement a new business proposition quickly for an organization.

Equinix Metal

Equinix Metal helps enterprises bring together and interconnect hybrid multicloud infrastructures at global scale. Equinix Metal helps digital businesses reduce time to market and reach users with high performance compute that can be deployed through popular developer tools in both on-demand and reserved models.

Equinix helps businesses rapidly deploy as-a-service networking, security and hardware across global data centers, as an alternative to buying, owning, and managing the physical infrastructure.

Portworx Data Services

Portworx Data Services is a database-as-a-service (DBaaS) for Kubernetes. With Portworx Data Services, organizations obtain the benefits of database-as-a-service, without the lock-in of a cloud platform. Portworx Data Services reduces the complexity of deploying and managing data services on top of Kubernetes clusters.

Portworx Data Services has many advantages, including:

- Running on any Kubernetes cluster, in the cloud or on-prem
- Deploying a wide-ranging catalog of data services, with 1-click deployment
- Providing a single pane-of-glass for managing and monitoring all data services and backups
- Supporting the entire stack, from storage through application



Portworx Data Services deploys data services such as Cassandra, Elasticsearch, Kafka, MongoDB Enterprise, PostgreSQL, and MySQL on Kubernetes. These pre-configured templates can be customized to meet the specific requirements of an enterprise.

However, to use the Kubernetes cluster with Portworx Data Services, Portworx Enterprise has to be installed in the environment. Portworx Enterprise provides storage provisioning, specifically tailored for containers and Kubernetes orchestration. Portworx Enterprise enables organizations to manage and scale their containerized applications seamlessly, ensuring high availability, performance, and data resilience. Portworx Enterprise helps manage and operate persistent storage for containerized applications in enterprise settings.

Figure 2 illustrates the validated architecture of Portworx Data Services on Equinix Metal.

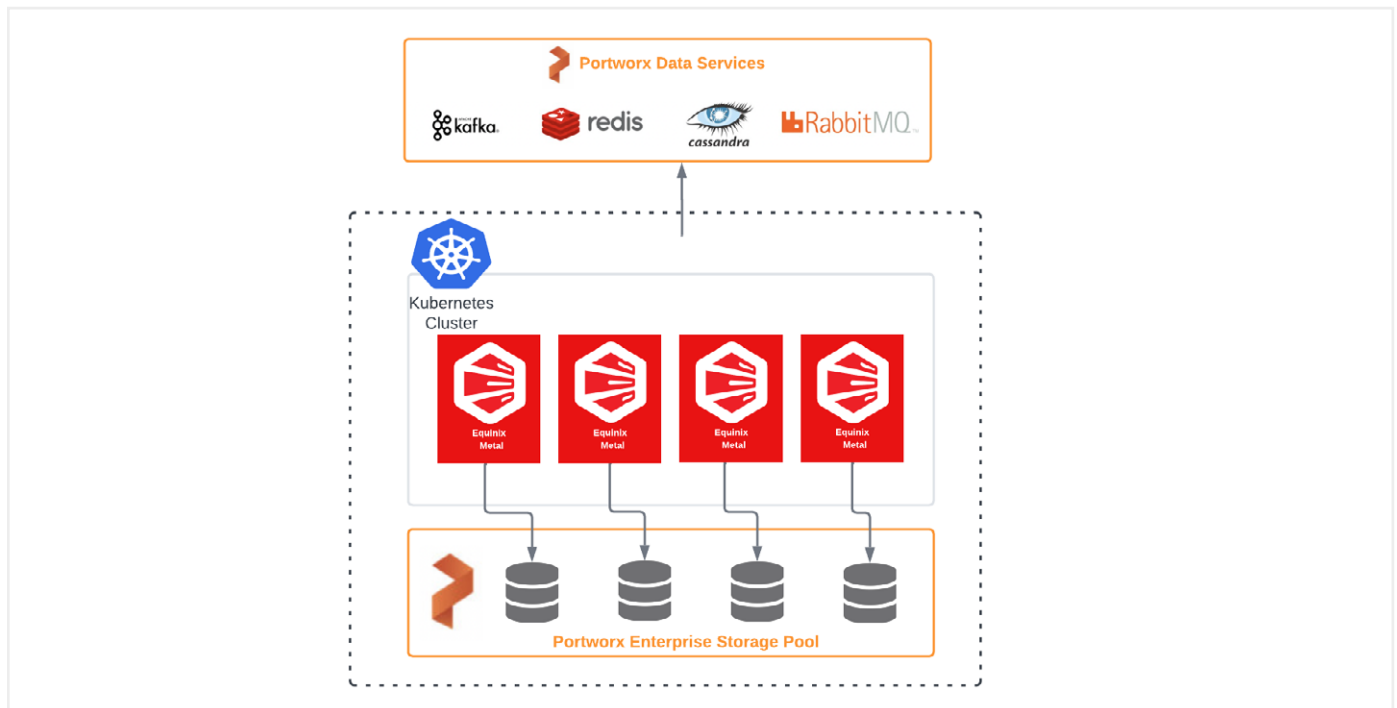


FIGURE 2 Validated architecture of Portworx Data Services on Equinix Metal

Automating Deployment of Portworx Data Services on Equinix Metal

Pure Storage offers a Terraform module that automates the deployment of Portworx Data Services on the Equinix Metal server. Using the Terraform module, you can avoid the complexity of installing and configuring each of the required components manually. Instead, you can specify the deployment criteria for the cluster in a configuration file, and the Terraform module handles the deployment tasks. The module also incorporates the best practices for the Portworx Data Services deployment on Equinix Metal. Figure 3 depicts this workflow.

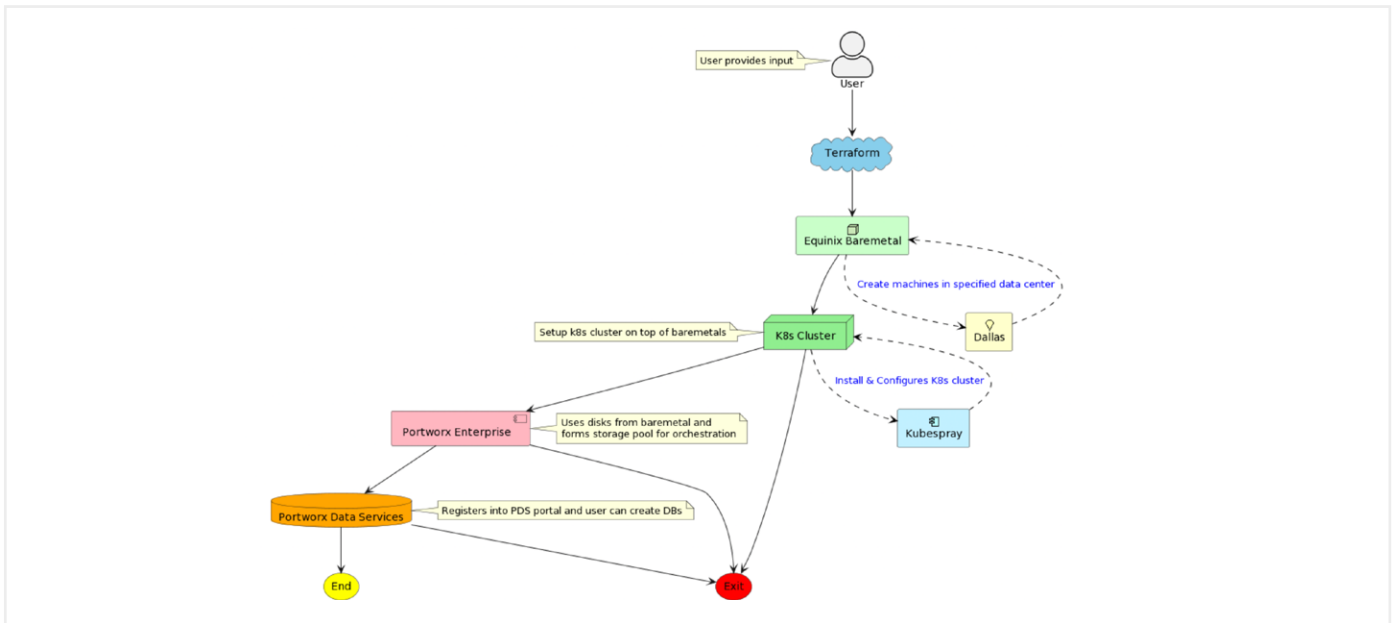


FIGURE 3 Automating the deployment of Portworx Data Services on the Equinix Metal Server using terraform

Prerequisites for Running the Terraform Module

- Deploy a c3.medium.x86 or higher Equinix Metal server.

For current standard plans offered by Equinix, see [Equinix Standard Gen3 Servers](#).

- The Linux host machine that runs the Terraform module must have Git (any version) and Python (version 3.8 or higher) installed.
- Create an [Equinix Metal account](#), set up an [organization](#) through the Equinix dashboard, and then create a [project](#).
- The Equinix Metal server should have at least a minimum of three disks: one for OS, one for KVDB, and one or more disks for Portworx storage pools.
- Create a [Portworx account](#) and then contact the [Portworx sales team](#) for onboarding to Portworx Data Services. You require licenses for both Portworx Enterprise and Portworx Data Services.
- Install [Portworx Data Services](#).



Executing the Terraform Module

When you run the Terraform module, it first provisions the Equinix Metal servers. It then sets up the Kubernetes cluster with Portworx Enterprise as the storage platform. Finally, it integrates the Kubernetes cluster into the Portworx Data Services portal.

After the automatic deployment, the Equinix nodes and their IP addresses are displayed. Sign in to the Portworx Data Services portal and verify whether the Kubernetes cluster is added as a deployment target. Then use the Portworx Data Services portal to deploy the database services.

Configuration (variable.tf) File

You can define deployment criteria in a configuration file called the variable.tf file. Some of the configuration parameters in the variable.tf file contain default values. You can change these default settings.

For example, the Terraform module sets up a 4-node Kubernetes cluster and sets the metro to Dallas (DA) by default.

To change these settings, edit the following parameters in the file:

- **cp_node_count:** Modifies the number of master servers in the cluster
- **nodes_count:** Modifies the total number of servers in the cluster

After the successful integration, users can select from the database services that are supported in Portworx Data Services.

Benefits of Automating Portworx Data Services on Equinix Metal

When you automate the deployment of Portworx Data Services directly on Equinix Metal servers using the Terraform module, it provides the following benefits:

1. **User experience—Zero touch deployment:** The zero-touch deployment and complete end-to-end automation improves the user experience of deploying Portworx Data Services on Equinix Metal. This allows for faster onboarding of data services on Equinix Metal clusters.
2. **Manageability and protection for databases:** With ad-hoc and policy-based backup of data services, companies can protect their data from ransomware attacks and other threats. This level of protection ensures that businesses can continue to operate without fear of losing critical information. Additionally, the rapid deployment of databases at different Metal edge locations allows businesses to scale their operations quickly and efficiently. Finally, the ability to scale up and down analytics and database clusters seamlessly gives companies the flexibility they need to adapt to changing business requirements.
3. **Flexibility—On-premises and cloud:** With the ability to deploy data services to a target Kubernetes cluster anywhere and everywhere, businesses can move their applications and data to any public cloud provider with ease. Moreover, the absence of vendor lock-in for moving databases to on-premises provides flexibility and freedom of choice to the customers.



Validating Apache Kafka on Equinix Metal and AWS Bare Metal

To demonstrate the compatibility of the automated solution with industry standard databases, Kafka was deployed using Portworx Data Services, and performance was tested.

Apache Kafka is an open-source distributed event platform designed for streaming real-time analytics, data integration, and data pipelines for use in various applications such as internet of things (IoT) and artificial intelligence/machine learning. Since Apache Kafka requires seamless scalability, persistent storage with data protection, and capacity management, Kafka is increasingly being moved into Kubernetes.

Kafka was deployed on the Portworx Data Services/Equinix Metal environment, and its performance was evaluated. The performance was also compared by deploying Kafka on Rook/Ceph/Equinix Metal and on Portworx Data Services/AWS bare metal.

Deploying Apache Kafka on Kubernetes with Portworx Data Services enables you to:

- Enhance the management of a scalable Kafka cluster by reducing broker failover time and enabling fully automated storage capacity expansion
- Improve security with encryption-at-rest for Kafka broker data
- Facilitate seamless migration of Kafka applications and their associated data across various environments, as well as across different cloud providers or data centers

Test Environment

The following table lists the bare metal configurations. Equinix Metal was used for Kafka deployment on both Portworx Data Services and Rook/Ceph. AWS bare metal was used for Kafka deployment on Portworx Data Services.

Component	Equinix Metal—n2.xlarge.x86 (Portworx Data Services/Kafka deployed on Rook/Ceph storage)	AWS Bare Metal—z1d.metal (Portworx Data Services/Kafka deployed on AWS Bare Metal)
vCPU	64	48
CPU Type	2× 2.20GHz Intel Xeon Gold 5120 processor	2× 4.0GHz Intel Xeon Scalable processor (Skylake 8151)
Memory	384GB	384GB
Storage	1× 3.8TB (NVMe SSD) 1× 240GB (for metadata)	2× 900GB (NVMe SSD) 1× 150GB (EBS allocation for metadata)
Network	40Gbps	25Gbps

TABLE 1 Bare Metal Server Configuration

During validation, a replication factor of 2 was applied in both the Portworx and Kafka environments. All the tests were performed with different types of brokers (2, 4, and 8) and producers (4, 8, and 16).



Kafka Performance Comparison: Test Results

The test results demonstrate the value provided by deploying Kafka on Portworx Data Services/ Equinix Metal. Deploying Kafka on Portworx Data Services/Equinix Metal server has several advantages over deploying Kafka on AWS bare metal or Rook/Ceph. Table 2 lists these advantages.

Kafka Deployment Environment	Performance (Throughput and Run Time)	Cost Effective	Data Manageability	Data Protection	Reliability and Resiliency	Cloud Lock-in	User Experience
Portworx Data Services on Equinix Metal	Better than Rook/Ceph and comparable to AWS bare metal	✓	✓	✓	✓		Automated
Rook/Ceph on Equinix Metal	Not as good as Portworx Data Services/ Equinix Metal						Manual
Portworx Data Services on AWS bare metal	Slightly better than Portworx Data Services/ Equinix Metal		✓	✓	✓	✓	Manual

TABLE 2 Advantages of deploying Kafka on Portworx Data Services/Equinix Metal

Performance of Kafka on Portworx Data Services and Rook/Ceph on Equinix Metal

As seen in Figures 4 and 5, the throughput and run time performance of Kafka on Portworx Data Services is much better (44% and 36% faster respectively) than Kafka deployed on Rook/Ceph.

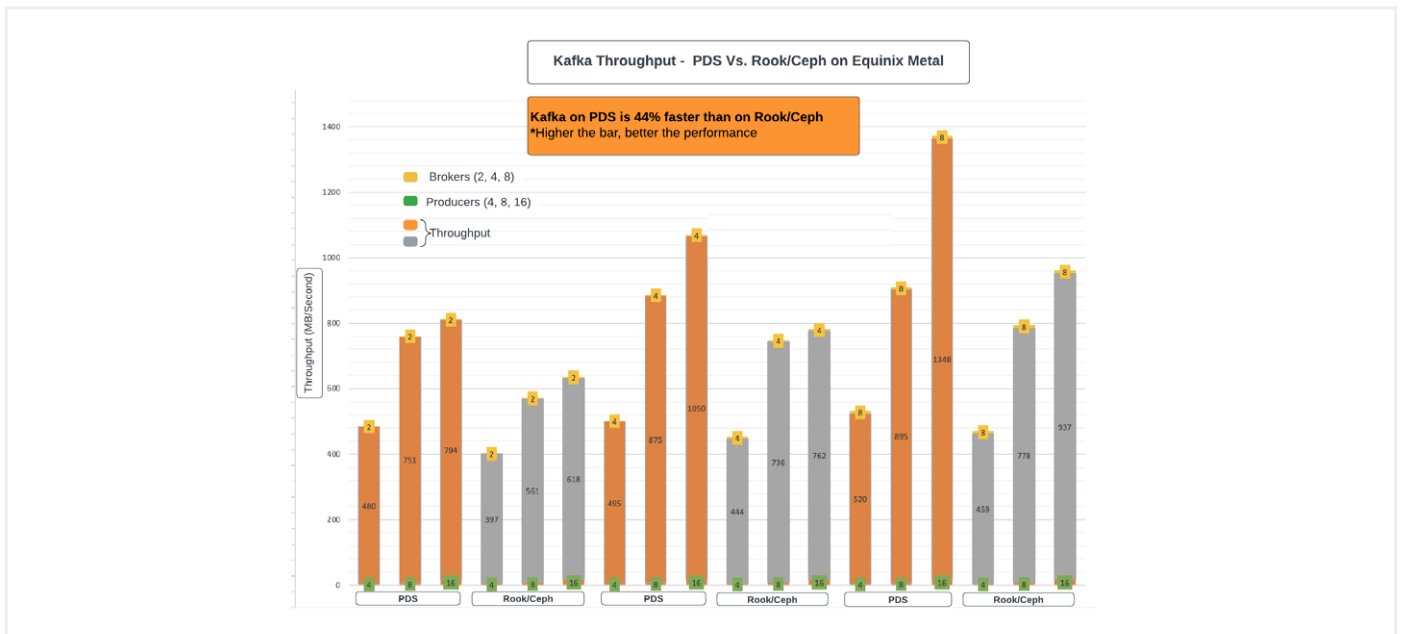


FIGURE 4 Throughput comparison of Kafka deployed on Portworx Data Services and Rook/Ceph



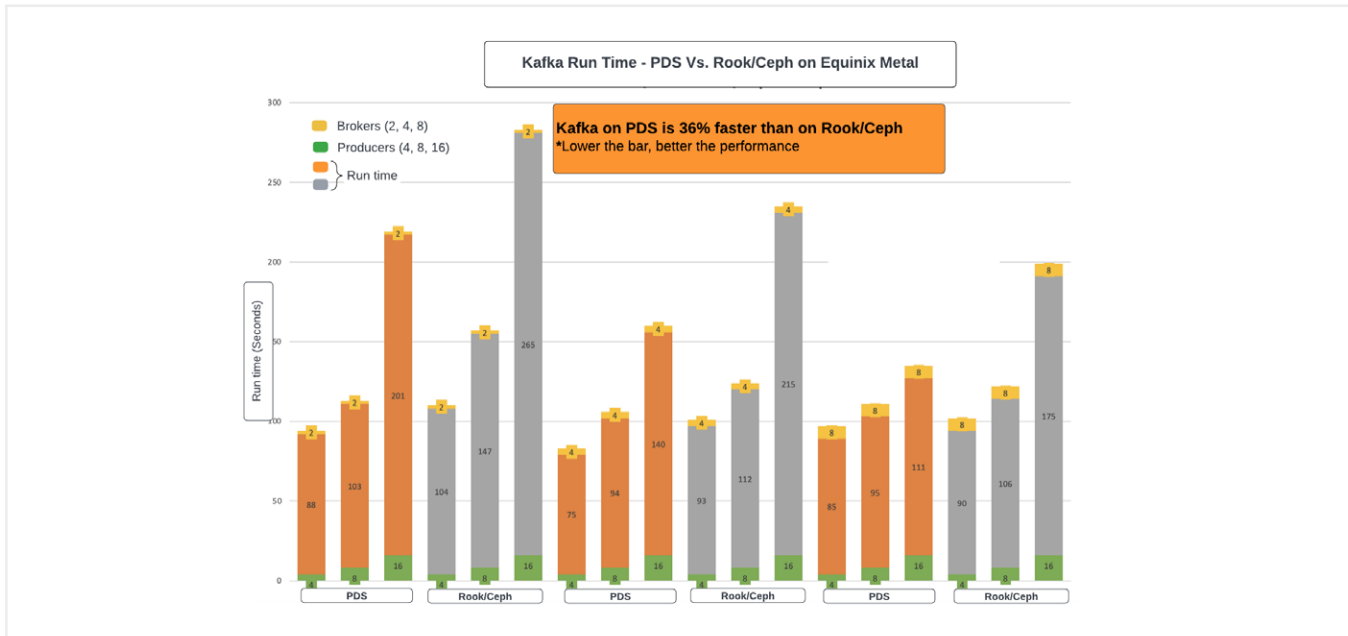


FIGURE 5 Run time comparison of Kafka deployed on Portworx Data Services and Rook/Ceph

Performance of Kafka on Portworx Data Services/Equinix Metal and Portworx Data Services/AWS Bare Metal

Figures 6 and 7 demonstrate that deploying Kafka on AWS bare metal results in slightly better performance compared to Portworx Data Services/Equinix Metal deployment. However, to deploy Kafka on Portworx Data Services/AWS Bare Metal, organizations would need to pay approximately 36% more per day.

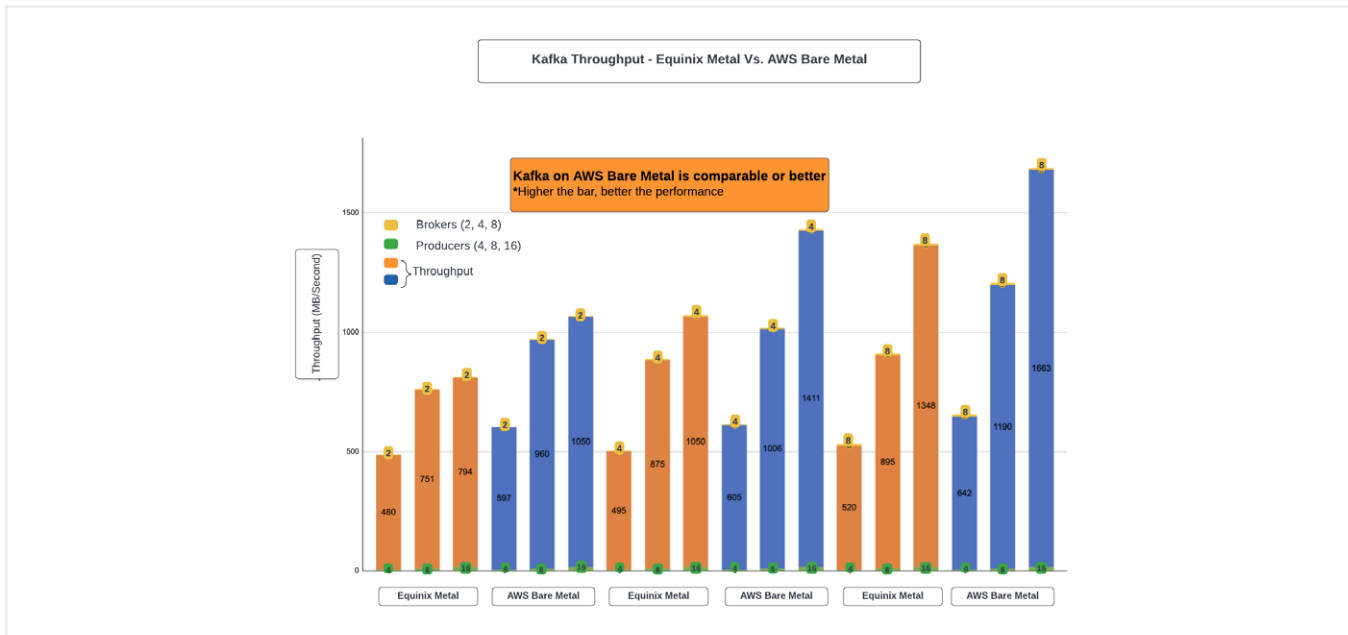


FIGURE 6 Throughput comparison of Kafka/Portworx Data Services deployed on Equinix Metal and AWS bare metal

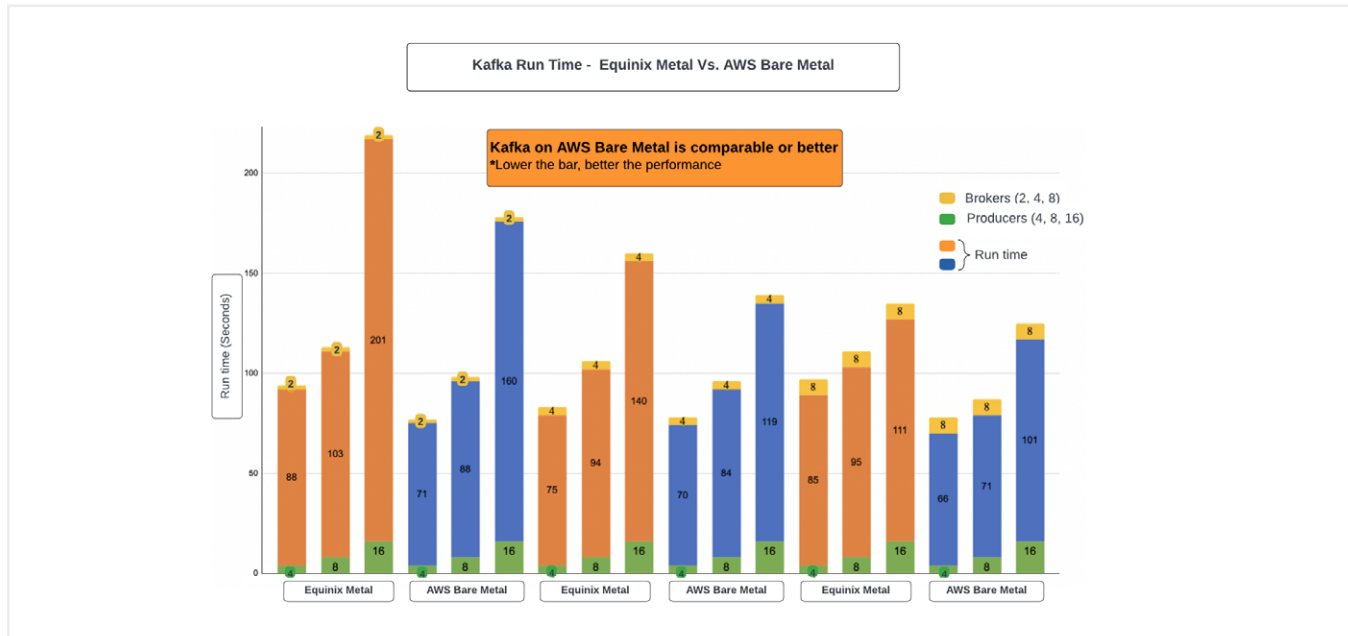


FIGURE 7 Run time comparison of Kafka/Portworx Data Services deployed on Equinix Metal and AWS bare metal

Benefits of Deploying Kafka on Portworx Data Services/Equinix Metal Over Rook/Ceph/Equinix Metal

The test results show that deploying Kafka on Portworx Data Services/Equinix Metal has several advantages over Rook/Ceph/Equinix Metal:

- **Faster throughput:** The throughput for Kafka deployed on Portworx Data Services is 44% faster than Kafka on Rook/Ceph (see Figure 4).
- **Faster run time performance:** The run time performance of Kafka deployed on Portworx Data Services is 36% faster than Kafka deployed on Rook/Ceph (see Figure 5).
- **Additional advantages:** In addition to having better throughput and run time performance, Kafka on Portworx Data Services provides superior data protection and high reliability features when compared to Rook/Ceph.

Benefits of Deploying Kafka on Portworx Data Services/Equinix Over Portworx Data Services/Amazon Bare Metal

While the test results show that deploying Kafka on Portworx Data Services on AWS bare metal is slightly better than Equinix Metal, the Equinix Metal deployment has several advantages:

- **Cost-effectiveness:** While Portworx Data Services/Kafka on AWS Bare Metal is comparable or better (see Figure 6 and Figure 7), Amazon bare metal deployment comes with the disadvantage of being expensive. To achieve the same level of performance as Portworx Data Services/Equinix Metal with AWS bare metal deployment, organizations would need to pay approximately 36% additionally per day (see [Appendix A](#)).
- **No cloud lock-in:** The AWS bare metal platform locks users, making it tough to switch to another platform. Portworx Data Services on Equinix Metal server, however, eliminates the risk of cloud lock-in and provides full control over your infrastructure.



Conclusion

Portworx Data Services on Equinix Metal server provides a simplified, flexible, and scalable solution for running data services on containerized applications. The Terraform module provided by Pure Storage streamlines the deployment of Portworx Data Services on the Equinix Metal server, eliminating the need for manual installation and configuration of individual components. This automated approach facilitates the swift deployment of database services. Portworx Data Services deploys data services using curated images that employ industry-standard best practices, thus ensuring optimal performance, reliability, and uptime for your applications.

As validated in this white paper, deploying Kafka on Portworx Data Services over Rook/Ceph provides better throughput and run time performance (44% and 36% faster respectively), superior data protection and manageability, high reliability and resiliency.

When compared to Kafka on Portworx Data Services/AWS bare metal, Kafka on Portworx Data Services/Equinix environment offers a cost effective alternative (36% less expensive than AWS bare metal per day), without the hassles of Cloud lock-in. Also, Portworx Data Services deployments provide complete data protection, reliability, and resiliency features.

As businesses expand their operations to multiple global locations as part of their digital transformation strategy, Portworx Data Services on Equinix Metal offers a balance between the customizable control that is available in co-location and the flexibility and agility that customers get in public cloud.

Appendix

The following table provides a comparison of the deployment costs for Kafka on Equinix Metal and AWS Bare Metal.

Component	Equinix Metal—n2.xlarge.x86 (For Kafka deployed on Portworx Data Services and Rook/Ceph)	AWS Bare Metal—z1d.metal (For Kafka deployed on Portworx Data Services and AWS Bare Metal)
vCPU	64	48
CPU Type	2× 2.20GHz Intel Xeon Gold 5120 processor	2× 4.0GHz Intel Xeon Scalable processor (Skylake 8151)
Memory	384GB	384GB
Storage	1× 3.8TB (NVMe SSD) 1× 240GB (for metadata)	2× 900GB (NVMe SSD) 1× 150GB (EBS allocation for metadata)
Network	40Gbps	25Gbps
Number of nodes	7	7
Cost per hour per node	\$3.25	\$4.46
Cost per day	\$546	\$750
Portworx Data Services license cost	\$0.35 per hour	\$0.35 per hour

TABLE 3 Cost calculations: Equinix Metal Vs AWS Bare Metal



The following table lists the Equinix Labs and Pure Storage GitHub repositories for downloading the Terraform modules.

Repository Name	Site	URL
Portworx Enterprise	Equinix Labs GitHub	Portworx Enterprise Module
Portworx Data Services	Equinix Labs GitHub	Portworx Data Services Module
Portworx Data Services (end-to-end)	Pure Storage GitHub	End-to-End Deployment of Portworx Data Services
Documentation for automatically deploying Portworx Data Services on Equinix Metal	Portworx Documentation Site	Automatically Deploying Portworx Data Services on Equinix Metal

TABLE 4 GitHub Repositories for downloading the Terraform modules

About the Authors



Bikash Roy Choudhury, As a Director for Solutions for Pure Storage, Bikash Roy Choudhury is responsible for designing and architecting solutions for DevOps workflows relevant across industry verticals including high tech, financial services, gaming, social media and web-based organizations. He has also worked on validating solutions with Rancher/Kubernetes, GitLab, Jenkins, JFrog Artifactory, IBM Cloud Private and Perforce using RESTful APIs and integrating them with data platforms in private, hybrid, and public clouds. In his current role, Bikash drives integrations with strategic DevOps partners, including Rancher, Mesosphere, Perforce, GitLab, and JFrog.



Keerthivasan Suresh is a solutions architect at Pure Storage, where he defines and executes solutions related to DevOps and open source integrations. He has 13 years of experience in managing IT infrastructure, developing REST APIs using Python, and implementing Infrastructure as Code (IaC) using Ansible, Terraform, and Puppet. His areas of expertise also include private cloud and Kubernetes. He is passionate about building DevOps solutions and formulating data management strategies that effectively address customer issues.

