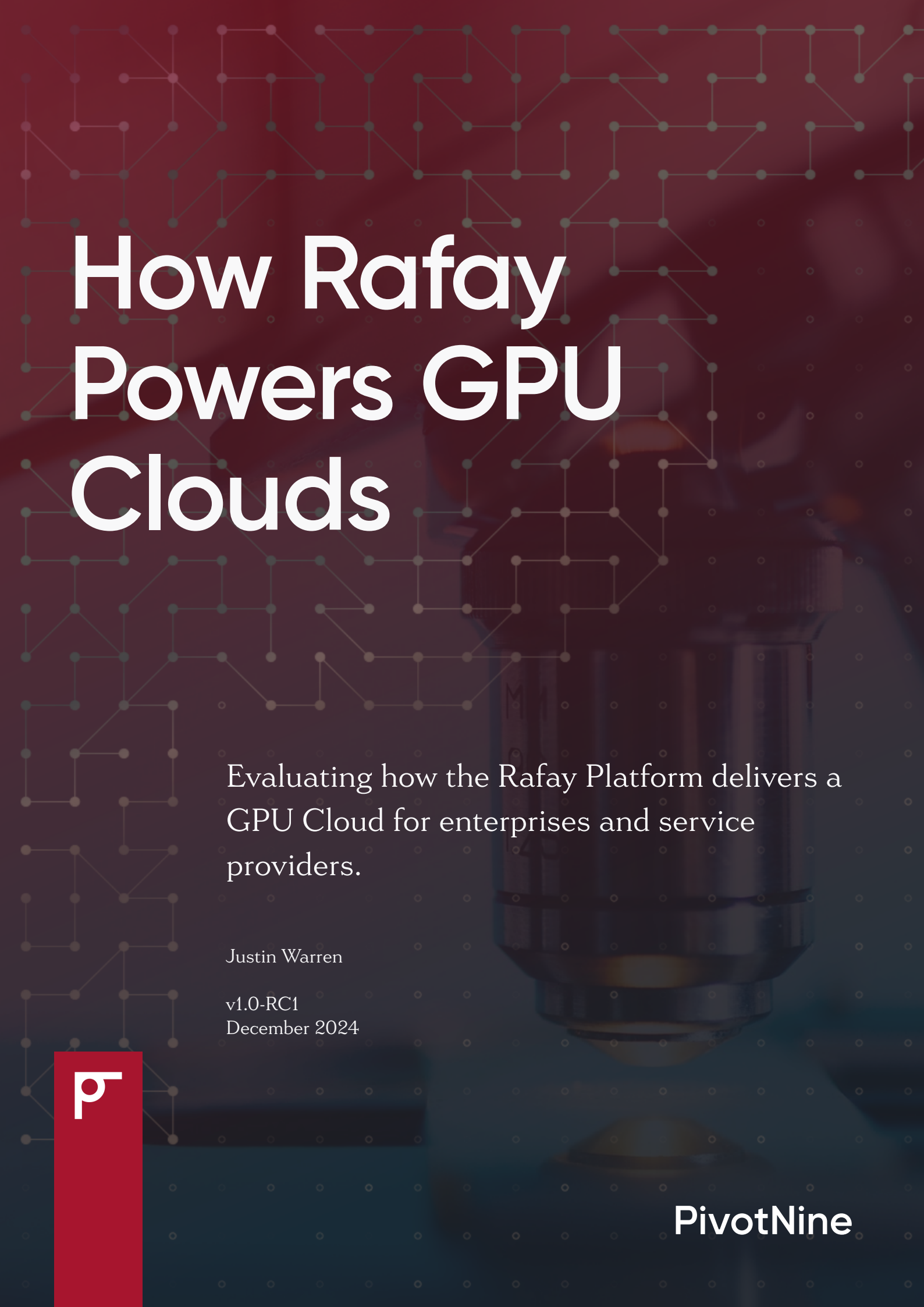

How Rafay Powers GPU Clouds

Evaluating how the Rafay Platform delivers a GPU Cloud for enterprises and service providers.

Justin Warren

v1.0-RC1
December 2024

PivotNine

Contents

Executive Summary

The Rafay Platform is a superior choice for enterprise, MSP, and large commercial customers looking to deliver self-service GPU consumption, complete with AI applications, to their internal or external users. With Rafay, customers can speed up the return on their investments in GPU infrastructure by quickly delivering a fully functional GPU Cloud. Rafay's approach enables high return on investment by simultaneously reducing costs and increasing the value delivered by GPU-based workloads.

Increase Enterprise Value

The purpose of AI projects is to deliver new value to the organization. Yet lack of access to scarce GPU resources means data science teams and AI developers are often blocked from completing their work. Projects are delayed, and teams become frustrated as they are forced to wait. No amount of cost-cutting will help if the value never arrives.

A well-managed GPU Cloud removes barriers and constraints that prevent teams from delivering the value they want to create. Internal developers and data scientists can make good use of enterprise investments in GPU infrastructure, rapidly converting ideas into shipped products. Service providers can quickly deliver AI workspaces to their customers who are eager to build with AI.

With Rafay providing highly automated, self-service access to GPU resources, teams can work autonomously without becoming blocked. Turnkey AI workbenches reduce batch setup times and ensure teams work securely, in line with organizational policy and governance, without wasting time reinventing common approaches.

A GPU Cloud can become a strategic resource that the organization uses to leverage its own specialized datasets to deliver unique, differentiated value. While others struggle with commodity systems aimed at the lowest common-denominator, leading organizations can focus on what sets them apart.

Lower Costs

Rafay helps organizations to pool their scarce and costly GPU resources to achieve high utilization with efficient, automated workload allocation. Silos of GPU infrastructure dedicated to specific teams tend to have low utilization. Lack of access to resources can mean teams purchase temporary access to transient GPU resources at high spot prices, which increases velocity but at substantial cost.

By pooling resources into a shared GPU Cloud, the large up-front costs of GPU infrastructure can be amortized over a larger number of projects and teams, reducing overall costs and increasing the utilization of this precious resource. The effective cost per project decreases by sharing common administrative overhead across a unified platform.

More importantly, costs are reduced by minimizing the waste of idle resources. For GPU resources obtained via subscription, idle resources are still billed. Every minute GPUs go underutilized is an immediate cash cost. For organizations with capital investments in GPU infrastructure, depreciation is rapid and economic costs will be recognised in expenses even if the GPUs sit idle. Organizations can avoid paying too much for resources that go unused by improving resource utilization with a GPU Cloud. The savings can be reinvested elsewhere.

Overview

Background on GPU Clouds

A GPU Cloud is a specialized form of compute infrastructure that uses cloud-style operations to provide access to a shared pool of Graphical Processing Units, or GPUs. Mature GPU Clouds also provide templated AI workbenches designed to address specific AI use cases such as fine-tuning models and inferencing.

Where CPUs are general-purpose computing devices, GPUs are optimized for certain kinds of large-scale mathematical manipulations, such as vector and matrix operations, originally used in graphical applications. Modern approaches to machine learning and AI workloads make great use of these mathematical operations. Greater performance and efficiency is achieved when using GPUs[1] rather than CPUs to perform these calculations.

Readers will be familiar with public cloud infrastructure such as AWS, Microsoft Azure, and Google Cloud which provide multi-tenant access to a shared pool of infrastructure over a network. The distinctive as-a-service operational model of the cloud—self-service, on-demand access to infrastructure and services, charged on a consumption basis—is now being applied to pools of GPU resources in a similar way to CPU resources. We use the term GPU Cloud to refer to systems optimized for delivery of GPU resources in a cloud-like fashion in order to distinguish them from the more generic *cloud* terminology.

Why Use a GPU Cloud?

The rapid rise in popularity of AI, particularly since the public release of OpenAI's ChatGPT, has meant demand for GPU resources dramatically outstrips the available supply. The challenge manifests itself in two main areas.

Firstly, GPU resources are both expensive and scarce. When organizations purchase GPUs, the large capital outlay (often many millions of dollars) demands that the GPUs should be utilized as much, and as quickly, as possible. This is no different to any other capital-intensive purchase.

Secondly, making use of GPUs also requires highly trained, specialist staff. Data scientists and AI/ML developers are relatively scarce and command high rates of pay. Like the GPUs themselves, organizations wish to maximize the utilization of these expensive resources.

With both GPUs and data scientists in high demand and short supply, it is vital for organizations to maximize their investment in both. Data scientists that cannot get access to the GPUs they need are under-utilized. They often become frustrated that they are unable to access the tools they need to do their job. Meanwhile, GPUs that sit idle depreciate rapidly without delivering value.

[1]There are other forms of specialized compute devices, such as Tensor Processing Units (TPUs) that perform a very similar function to GPUs. We include these devices in the more general term GPU unless distinguishing between them aids understanding.

A well-architected GPU Cloud provides a good way to match these resources with each other; data scientists with GPUs and AI workbenches. By pooling each resource and using the GPU Cloud to match demand with available supply, organizations can make better use of both GPUs *and* data science staff. Beyond the obvious monetary savings, staff happiness is also a benefit as highly skilled staff do not enjoy having their time wasted unnecessarily.

Cloud service providers can use a GPU Cloud to provide multi-tenant access to GPU resources for customers that want to access these resources but with minimal administrative overhead. Customers are familiar with using cloud platforms in this way, and a GPU Cloud is a natural extension of the concept.

Enterprises are also familiar with cloud platforms, but many wish to maintain their own private GPU infrastructure for a range of entirely valid reasons. They can also benefit from a GPU Cloud that is able to pool private resources and still provide the same cloud-like operational experience to internal teams. While they will accept some administrative overhead, they would still prefer to purchase and use a GPU Cloud platform, not build and maintain one themselves.

A Platform Experience

Data scientists and developers don't want to waste time building and configuring infrastructure. They need the infrastructure to do their job, but it isn't what's really important to them. A platform is merely the ground on which they stand.

Data scientists want to create AI workbenches on-demand to work on an interesting data problem. They want to use models that have already been vetted by the organization, choosing the model that best suits their needs. A GPU Cloud provides that experience by shifting the burden away from the data scientists onto the platform itself. Models can be vetted once, and then every data scientist of the GPU Cloud can access the resources they need quickly, without fuss.

Similarly, developers are focused on writing AI-enabled applications. They want a platform that supports their preferred workflow, but without forcing them to know all the infrastructure details. They are application developers, not infrastructure developers. The platform should support them, not weigh them down.

How data scientists and developers experience the platform is vital to its success. The best experience is one they almost never notice. They want a GPU Cloud that gives them what they need, when they need it, and otherwise stays out of their way.

Solution Evaluation

Perceptual Fingerprints

PivotNine's Perceptual Fingerprint™ visually summarizes how well an offering aligns with customer needs. We have looked at the two key markets Rafay is targeting to assess how well it meets customer needs in these markets.

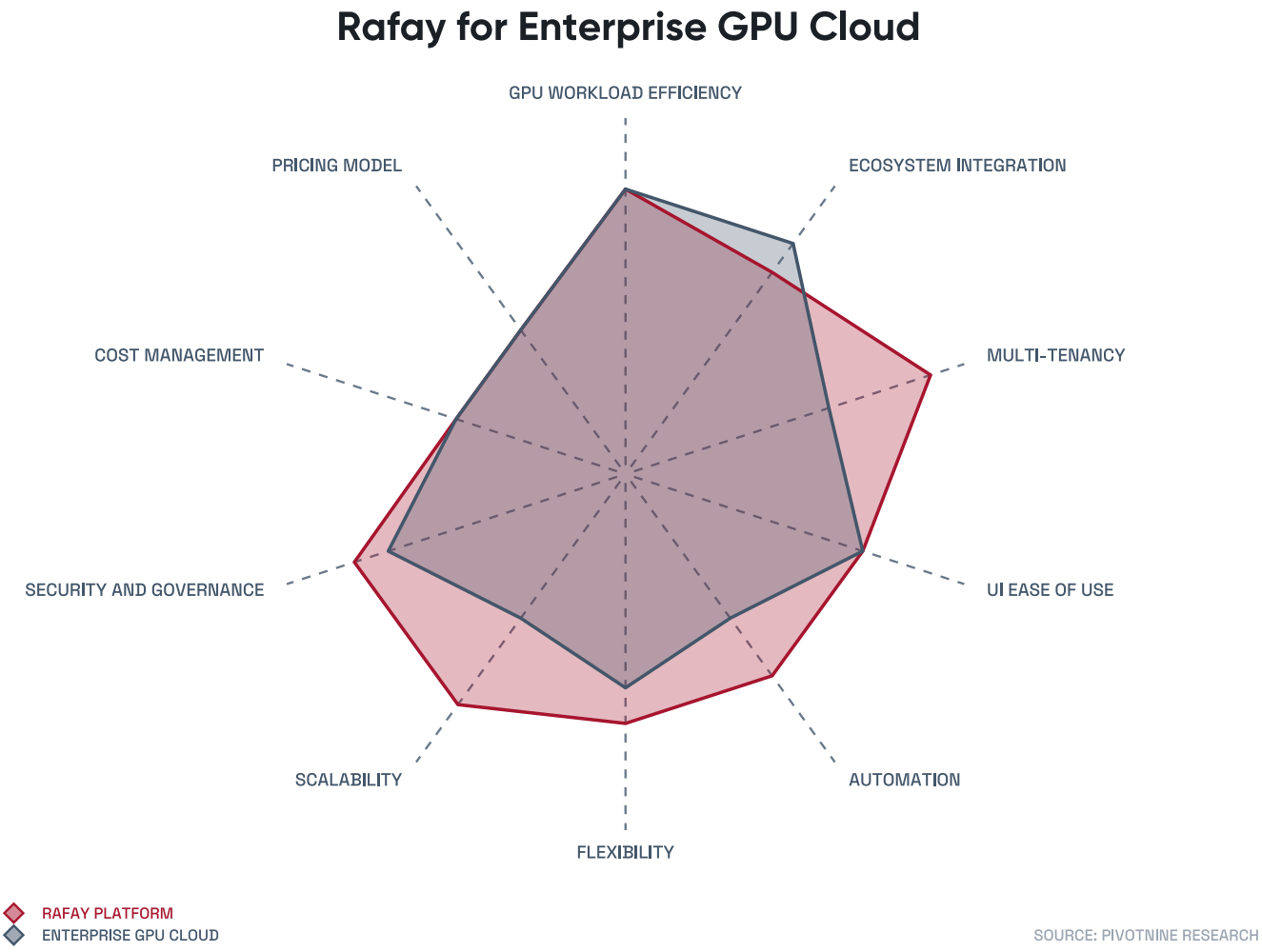

Figure 3.1: The PivotNine Perceptual Fingerprint™ of the Rafay Platform's capability set for enterprise GPU Clouds.

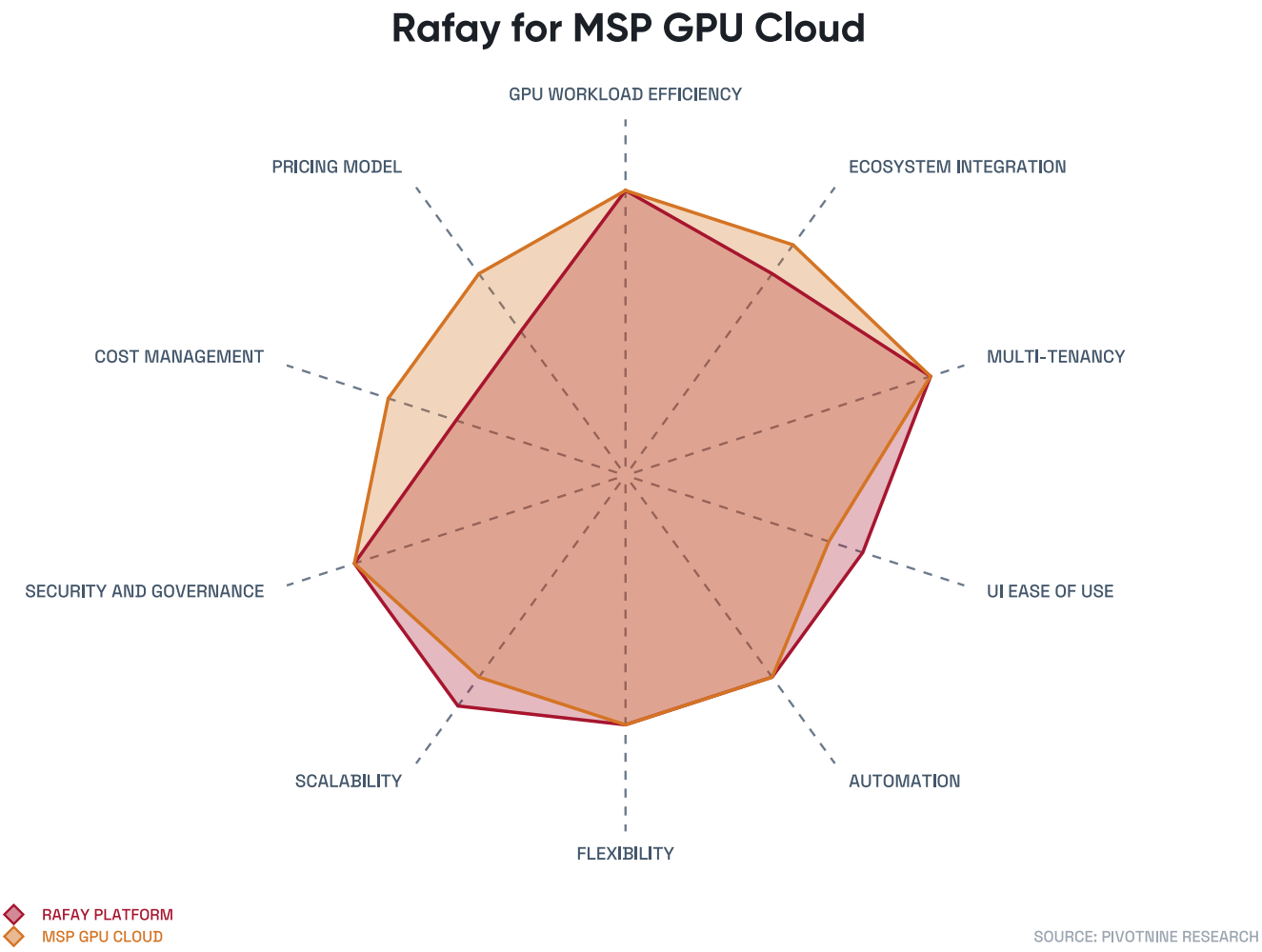

Figure 3.2: The PivotNine Perceptual Fingerprint™ of the Rafay Platform's capability set for service provider GPU Clouds.

Evaluation Detail

Here we provide commentary on the major criteria we have used to assess how the Rafay Platform delivers a Platform-as-a-Service (PaaS) experience to enterprises and GPU Cloud providers. We briefly describe why each criterion is relevant and how it aligns with customer needs. We highlight particularly noteworthy aspects of the solution relating to each criterion and how it benefits customers.

For more information on the evaluation criteria, contact PivotNine's research team via email at research@pivotnine.com.

GPU Workload Efficiency

GPUs are an expensive and scarce resource. Data scientists and AI developers are also relatively expensive and scarce. To maximize an organization's return on investment, it is important that the utilization of GPU cluster resources is maximized while also maximizing the amount of productive work done by data scientists and AI developers.

The Rafay Platform collects GPU resources into a centrally managed pool that can be efficiently allocated based on diverse needs. The Platform supports "spatial partitioning" of GPUs with techniques such as NVIDIA's multi-instance GPU (MIG).

By combining silos of underutilized resources into a shared pool that can be addressed through configurable SKUs (e.g. small, medium, and large), GPU consumers are freed from waiting for busy resources while otherwise idle resources are put to work. This simultaneously increases the value delivered by data scientists and developers, while also reducing cost and waste.

Helping developers and data scientists to work on what they want to work on *when* they want to work on it can also contribute to improved staff satisfaction. There is nothing quite as frustrating as being unable to do your job because of arbitrary friction or bureaucracy.

Deployment Options

GPU scarcity means that accessing GPUs when and where they are available is critical to improving utilization and reducing unwanted idle time.

The Rafay Platform supports self-hosted, cloud, and hybrid- and multi-cloud deployments. Customers can access GPUs wherever they deem most appropriate. This is particularly important, given GPU availability is highly fluid. Customers may wish to change how they make use of resources as their projects evolve over time, as they purchase new GPU capacity, or when they retire surplus resources. Rafay provides significant flexibility in this regard.

The Rafay Platform is available as both SaaS and self-hosted software, providing customers the choice to align operations with their privacy and security requirements. SaaS provides benefits to those who prefer lower operational overhead, but SaaS is not always appropriate, particularly in heavily regulated industries.

Sovereign Security

Some AI workloads need to operate on highly sensitive or private data. Access to this data must necessarily be limited to secured locations, adding a constraint on possible utilization efficiency. By allowing for secure resource management of sovereign and self-hosted GPU clusters, the Rafay Platform is able to provide the benefits of a publicly accessible GPU Cloud, while ensuring appropriate security and privacy controls are in place.

Ecosystem Integration

Developers and data scientists have a preferred set of tools. The Rafay Platform provides first-class support for key tools in this ecosystem, from low-level tools such as Python and Jupyter notebooks, through to full MLOps driven environments.

For teams that have standardized on the Ray compute engine, the Rafay Platform provides an integrated Ray-as-a-Service offering for MLOps. Platform engineering teams can provision Ray-based environments to teams or individuals as isolated “virtual clusters” that share physical GPU cluster resources. The Rafay Platform also supports Kubeflow-based environments, providing a cross-platform experience similar to AWS Sagemaker or VertexAI without locking the organization to a specific vendor’s tooling or infrastructure.

The Rafay Platform supports Run:AI, either as an out-of-the-box experience to deploy and consume Run:AI on a Rafay GPU Cloud, or to complement Run:AI's offerings. Teams can choose the mode that best suits their preferred workflow. The Platform supports NVIDIA Inference Microservices (NIMs) as well as the open source kServe. By supporting multiple inferencing options, Rafay provides customers with flexibility and choice while maintaining consistency for common components. MSPs can also leverage the Platform's turnkey support for NVIDIA Cloud Functions (NVCF) to offer serverless access to their GPU hardware to NVIDIA's customer base via the NVCF global control plane.

The Rafay Platform also provides a low-code framework for rapidly adding more specialized AI components such as verticalized agents, co-pilots, and document translation services. For complex environments that use a combination of standardized tools and bespoke or emerging methods, Rafay provides a unified approach that integrates them all.

Platform Operations

Operating a GPU Cloud requires tools for two key audiences: infrastructure platform teams that operate the GPU Cloud itself, and data scientists and developers using the GPU Cloud. Rafay has a comprehensive set of tools for both.

Platform teams can operate a fleet of clusters from a single-console, helping to maintain organizational standards and consistency with great efficiency. A small team of specialist platform engineers can use the Rafay Platform to support a large number of data science teams working independently without compromising on security or governance.

For data scientists and developers, the Rafay Platform provides self-service access to GPU-enabled workspaces with a curated set of standard tools already installed and configured. Customers can develop their own "golden image" environments that align with common working patterns unique to their organization. Data scientists can also adapt these environments to their own needs, subject to their organizations' policy and governance arrangements.

The Rafay Platform is designed to help the platform teams collaborate with data scientists and developers on environment standards. Feedback on environment configurations can be provided via developer-preferred processes such as GitOps or via the Platform's robust APIs. By providing solid support for these processes within the Platform itself, it reduces unnecessary friction that can hamper collaboration. This eases the burden on both teams as they maintain the platform, allowing them to more quickly adapt to changing business needs.

Multi-Tenancy

Multi-tenancy is a vital part of robust shared service design. The Rafay Platform has superior support for complex multi-tenancy arrangements, including multiple levels of delegation and self-service.

Some customers prefer highly centralized architectures where a core team implements standardized resources that are then consumed by downstream teams as a commodity with a low-touch, self-service retail experience. Other customers prefer to define high-level policy and governance centrally, while delegating more detailed decisions to downstream teams, sometimes through several layers of granularity. Rafay supports both approaches without loss of functionality.

The Rafay Platform has multiple features to provide fine-grained resource partitioning and access control. Virtual clusters (vClusters) provide full cluster capabilities to tenants while sharing the resources of the host cluster infrastructure. Kata containers can be used to wrap workloads in microVMs to protect against container escapes without losing performance or agility. Additional controls —including container, network, and cluster policies—also help to isolate tenants from one another to maintain privacy and security.

Resource quotas help ensure resources are shared fairly between tenants, based on organizational policy. Multi-tenant cost allocation and visibility is supported, providing views into resource consumption partitioned by tenant or sub-tenant. MSP customers will particularly appreciate the granularity on offer, allowing them to tailor their pricing and packaging to their target customers.

Network segmentation and secure remote access help ensure communication channels are limited to authorized users that can only access the resources they are supposed to. Tenants can be provided with limited administrative access while broader, more sensitive access is limited to platform administrators. All activity is audited, providing tenants and platform owners with the assurance that what is supposed to happen—and only that—actually does happen.

The Rafay Platform's flexibility is particularly well-suited to MSPs looking to provide a GPU Cloud service to their customers. Customizable dashboards and automated reporting mechanisms, including predictive analytics, help customers feel in control of their environments without having to assume full responsibility for every aspect of the platform.

Cost Management

The Rafay Platform has an integrated cost management system that provides visibility into the GPU Cloud using cost management profiles. Cost visibility is controlled via RBAC ensuring cost information is subject to the same security controls as other resources. The Platform provides a preconfigured FinOps role for rapid cost management setup, but customers with mature FinOps practices can customize the system to align with their established practices.

The Platform provides built-in cost reporting, and supports charge back groups for cost allocation. Costs can be allocated based on resource allocation or shared equally across tenants within a cluster. Reports automatically take charge back groupings into account, and RBAC integration limits access to cost reports to appropriate personnel.

Cost information can also be exported to external FinOps systems via the API. This feature will be of particular interest to MSPs looking to integrate GPU Cloud cost data into their existing billing systems.

API-Driven Automation

Scalable efficiency requires automation, and API-driven automation is the industry standard approach. Rafay's background in large-scale Kubernetes cluster management is evident in its mature support for highly automated operations at scale.

The Rafay Platform supports a HashiCorp validated Terraform provider for those customers with established Infrastructure-as-Code approaches to infrastructure management. The Platform also supports GitOps workflows by combining YAML format configuration with customer CI/CD tools and the Rafay CLI tool.

For more advanced customers, the Platform also has a comprehensive API. Developers building customized automation tooling can explore the API via the industry standard Swagger tool. Access to the API is controlled via the same RBAC mechanisms used for the rest of the platform.

For customers with large environments, the Platform provides fleet plans that can be applied to collections of clusters as a group. Fleet configuration can be aligned with the organizational approach to multi-tenancy discussed above to balance the tension between delegated responsibility

Flexibility

The world of AI is changing extremely rapidly. Customers need to be able to respond quickly to changes, which requires a flexible, adaptable system. However, customers also want a system that is easy to use and does not require months of complex configuration before it starts to deliver value. The Rafay Platform manages this tension well.

The Rafay Platform is a highly flexible platform that provides customers with plenty of options across multiple dimensions. The Platform supports multiple public cloud Kubernetes services, including AWS EKS, Azure AKS, and GCP GKE, as well as self-hosted Kubernetes clusters. Customers can configure the Platform as a highly centralized managed service where data scientists consume commodity services, or as a delegated multi-tenant platform with a mix of commodity options and highly customizable self-service environments.

Customers can pool GPU resources across environments, or isolate specific sets of resources with specialized needs. Workloads can be relocated throughout the development lifecycle to ensure they are making the best use of available resources at all times. By providing a consistent experience regardless of deployment location, the Platform helps organizations to onboard new teams rapidly while maintaining operational efficiency.

Whichever approach best suits the situation, the Platform can overlay organizational policy and governance to provide consistency where it is needed, and flexibility where it is desirable. As customers evolve their use of the Platform, the system can evolve with them, ensuring it stays relevant at all times.

Scalability

Quickly scaling both up and down is a key feature of cloud-style operations, and is well supported by the Rafay Platform. New clusters can be created quickly and easily, in a variety of locations, helping to provide scale where it is required. GPUs can be added or removed dynamically, and workloads allocated as the required GPUs become available. Workloads can be closely matched to the available resources in a dynamic, automated way.

As experiments complete, their resources can be rapidly redeployed to waiting jobs, minimizing waiting time. For cloud resources that are unused, they can be deprovisioned to reduce costs, helping organizations to avoid paying for unused, expensive GPUs.

By linking scalability to cost management, resource consumption in the GPU Cloud can be closely aligned with organizational budgets and objectives. Cross-platform compatibility helps to reduce the risk of stranded assets, while automation reduces the switching costs of decisions. The Rafay Platform helps organizations to achieve more than mere scale but economies of scale.

Security and Audit

AI workloads are of particular concern to cyber security teams as they require access to large quantities of often sensitive data, particularly for Retrieval Augmented Generation (RAG) approaches.

The Rafay Platform provides robust and comprehensive identity and access management (IDM) and role-based access control (RBAC) mechanisms. They are fundamental to the platform and how it is able to deliver a scalable and flexible multi-tenant solution. Rafay integrates with a large number of external single-sign-on (SSO) services including AWS SSO, Azure Entra ID, ADFS, Okta, Duo, and Keycloak.

The Platform supports policy-driven network isolation of workloads, minimizing the attack surface and reducing the potential for lateral movement. Similarly, cluster level policies can be used to govern workload privileges and permissions. Resource quotas are also supported, helping to ensure that GPU resources are shared appropriately.

A centralized, immutable audit trail is automatically maintained for the GPU Cloud, capturing all user activities, both via GUI and API. Audit logs also capture policy enforcement events, and cluster and node activity events. The Platform can stream the audit logs to external enterprise SIEM systems such as Splunk, Datadog, AWS Cloudwatch and SumoLogic.

The Rafay Platform is Service Organization Control (SOC) 2, Type II compliant and audits are performed annually by independent auditors.

PivotNine's Opinion

The Rafay Platform is a superior choice for enterprise, MSP, and large commercial customers looking to increase the return on their investments in GPU infrastructure by operating a GPU Cloud.

Strengths

The Rafay Platform's support for open standards allows customers to customize and extend the platform to better meet their unique requirements with lower risk of vendor lock-in. For a fast-moving field like AI, this helps customers to standardize on known-good approaches as they become established while maintaining the flexibility to rapidly explore the frontier of the field at lower risk of stranded investment.

The Platform's support for complex, federated multi-tenancy with robust security and delegated self-service mechanisms makes it ideal for MSPs looking to provide a commercial GPU-as-a-Service offering. Similarly, enterprises with a shared-services IT approach can adopt Rafay to achieve high utilization of scarce GPU resources while simultaneously increasing the value delivered by data science and AI developer teams.

Future Opportunities

No vendor product is perfect. In a fast-moving field like AI, we expect vendors to continually enhance their offerings to meet evolving customer needs. We see the following areas as emerging areas of customer need that Rafay would do well to invest in.

Security Enhancements

The Rafay Platform's RBAC controls are strong, and we would like to see them extended to provide even greater protection against malicious insiders. Quorum-based authorization (two to sign, three-of-five to sign, etc.) of sensitive administrative actions such as environment removals would provide additional protection against accidental and malicious damage.

Rafay has already made good progress integrating with external tools, such as its support for Aqua Security's tfsec. We would like to see more pre-built integrations with common enterprise security tools, particularly SIEM and SOAR tools, so that security teams can more readily add Rafay into their existing workflows.

We would also like to see more governance assistance tooling, such as assurance reporting and software bill of materials reports to assist with auditing the platform. Having good security is necessary; being able to clearly demonstrate that the system remains secure is even better.

Network Device Integration

Rafay does not currently integrate closely with network hardware devices in large GPU clusters used for training models; however, the company is actively working to address this gap. GPU clusters are more sensitive to network issues than other workloads, and network telemetry information is invaluable for management and troubleshooting of expensive training runs.

However, there are few organizations involved in training base models. Most Rafay customers will build upon foundational models to fine-tune or adapt them to their own needs, or perform inferencing using small portions of a larger pool of GPU resources.

Customer-Defined Templates

Pre-defined environments are at the core of as-a-Service offerings, particularly for customers using platform engineering or DevOps management structures. Rapidly sharing and adopting these templated approaches would help customers to avoid needlessly reinventing the same standard solutions, particularly as known-good approaches become standardized. This mirrors the experience of data scientists using models obtained from repositories such as HuggingFace Hub.

We see a particular benefit to MSPs or systems integrators. They are well-placed to observe common patterns across their customer base and can identify when a standardized approach would benefit many customers. A marketplace would allow them to promote these standardized templates without redeveloping the same solutions for each customer. Customer self-service from an ever-growing catalog of desirable standards has intuitive appeal.

Cost Management Enhancements

Rafay has developed its own cost visibility tools that are built into the platform. These provide a baseline set of tools for organizations to build upon, especially where other cost management tooling is absent. However, more sophisticated customers that have already invested in FinOps tools will prefer to bypass Rafay's interfaces and work with cost data more directly.

These customers would benefit from pre-built integrations into commonly-used FinOps tools, reducing the effort of adopting Rafay. However, we acknowledge that this integration work is challenging, as customers often have their own bespoke approaches to cost allocation and charging. A pre-built integration is necessarily a starting point for further customization rather than an end in itself. Rafay could show leadership here by establishing a baseline of sensible defaults based on its experience with existing customers, helping new customers to avoid early missteps in adopting FinOps.

About

The Author

Justin is the founder and principal analyst at PivotNine, a boutique analyst and consulting firm based in Melbourne, Australia. He covers enterprise data center, cloud, and information security technologies. An IT industry veteran with extensive global experience, Justin has worked with enterprise organisations including ANZ bank, IBM, NetApp, Pure Storage, Telstra, and VMware as well as a variety of startups including Isovalent, Illumio, Pulumi, and Solo.io.

His writing has appeared in a variety of well-known outlets including including *Forbes.com*, *The Register*, *iTnews.com.au*, and *CRN Australia*. He has appeared as a guest host on *theCUBE* and as a contributing lead analyst at GigaOm.

Justin holds an MBA from Melbourne Business School, and is a graduate member of the Australian Institute of Company Directors.

PivotNine

PivotNine Pty Ltd is an analysis and consulting firm specialising in enterprise and developer technology.

We assist clients with the evaluation and selection of enterprise and developer technologies for datacenters, cloud, and information security. We conduct custom research on behalf of our clients to better inform their decision-making, and to ensure they are making an informed choice both now, and into the future.

PivotNine assists vendors with product marketing, marketing strategy, and communicating their product benefits to enterprise buyers.

Become a Client

PivotNine welcomes new clients. To learn more about how we can help you to buy, sell, and use enterprise technology, contact us today.

sales@pivotnine.com · pivotnine.com/contact

PivotNine