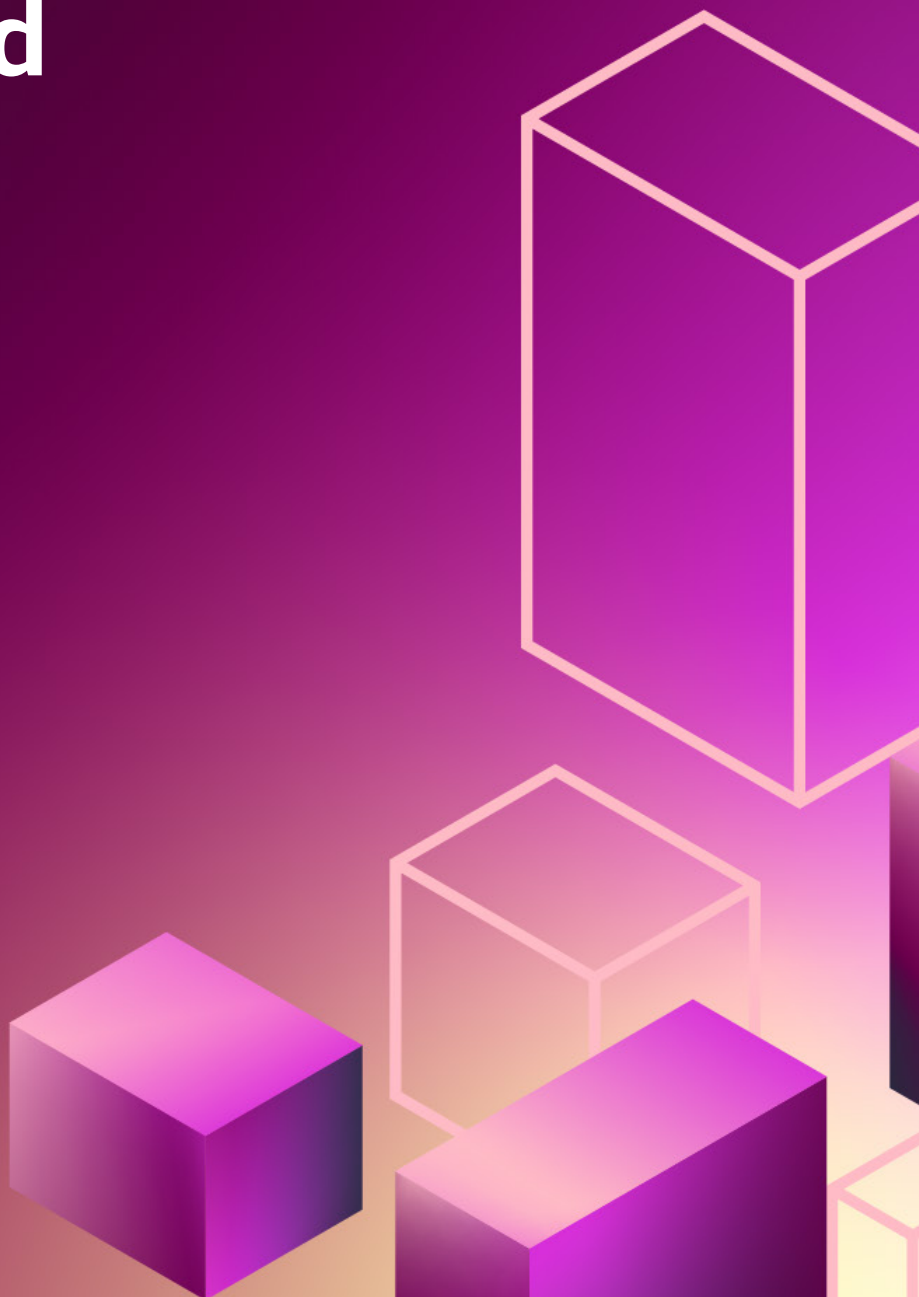
Seamless Kubernetes on premises and in the cloud



Executive summary

This whitepaper addresses current challenges with modernizing on-premises systems with Kubernetes. Amazon Web Services (AWS) public cloud services offer several advantages; however, some customers may need to run select workloads on premises for a number of reasons, including low latency and data locality, both of which are covered in depth in this whitepaper. To help address these challenges, we present a set of solutions powered by **Amazon Elastic Kubernetes Service** (Amazon EKS), specifically **Amazon EKS Anywhere**, which is a solution to run Kubernetes on premises that allows for perpetual on-premises container and infrastructure management—with AWS-supported tooling and opinionated packaging—as a vehicle for migrating to the cloud.

This whitepaper is intended to help accelerate organizations' decision makers' evaluation processes for on-premises solutions that best fit business requirements.

Preface

In his book *The Paradox of Choice: Why More Is Less*, psychologist Barry Schwartz argues that the freedom of choice has not given us freedom but rather has paralyzed us. A glance at more than 2,000 open-source projects in the Cloud Native Computing Foundation (CNCF) can seem daunting. And the explosion of hosting options and management tools means that cloud architects need to make more decisions in their infrastructure management than ever before. As a result, we have seen customers paralyzed with choice at nearly every architectural decision point, such as deciding on logging providers, storage interfaces, container runtimes, and more.

To reduce decision fatigue with running containers on premises, AWS launched Amazon EKS Anywhere, a deployment option for Amazon EKS that allows customers to create and operate Kubernetes clusters running on their infrastructure virtually anywhere. Amazon EKS Anywhere is bundled with opinionated architectural patterns, tooling, and open-source projects to accelerate your Kubernetes adoption with AWS-recommended approaches and to alleviate the paralysis of choice.

Introduction

AWS has made cloud computing a broadly accessible and reliable form of IT infrastructure. You no longer need to commit to a significant upfront investment to build and operate data centers before developing applications. Lowering the risk of experimentation with minimal upfront costs means ideas can go from developing prototypes to running applications sooner.

On-premises environments can benefit from recent innovations in infrastructure management, application packaging, and open-source tooling, minimizing risk and maximizing existing infrastructure investments.

Data center modernization

One of the primary benefits of the cloud is that someone else manages the "undifferentiated heavy lifting," a term coined by Jeff Bezos in 2006 in reference to the value that **Amazon Elastic Compute Cloud** (Amazon EC2) provides to its customers. Undifferentiated heavy lifting refers to work that is not unique to your business—it's simply the price of admission. For example, managing application infrastructure is critical to your company's success, but it is by no means unique or differentiated work.

Infrastructure has changed a lot since 2006—from physical servers and virtual machines (VMs) to containers and serverless. Amazon has been at the forefront of many of these innovations, helping customers adopt the appropriate technologies and operate them at scale. AWS will partner with your organization to accelerate innovation efforts and support your application needs.

Kubernetes has standardized container and infrastructure management through a common API interface and collection of open-source software. Amazon EKS Anywhere brings the combined Kubernetes and AWS partnership to your data center. Modernization isn't simply deploying specific software; it's also the potential to leverage AWS Partners' expertise to accelerate your innovations.

Despite the amount of undifferentiated work associated with managing data centers, customers still have several reasons to run workloads in on-premises or edge environments. These include:

- **Capitalizing on existing investments ›**
- **Data locality requirements ›**
- **Latency-sensitive compute ›**

Let's cover these in depth.

Capitalizing on existing investments

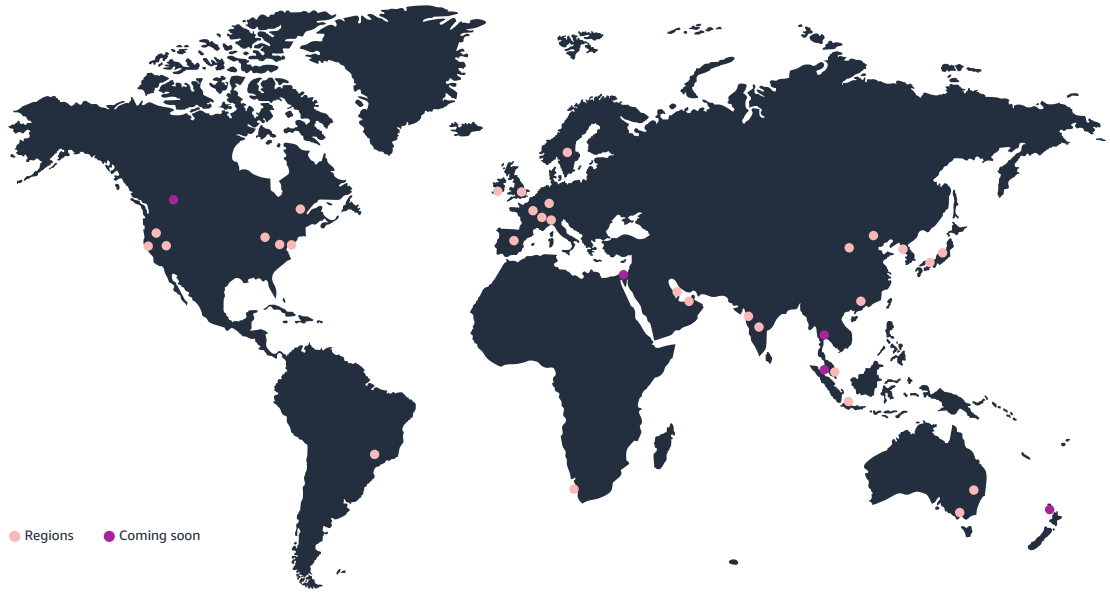
Your data center is a long-term investment. There is significant capacity planning, training, and process creation that require specialization for your company and industry. You have invested in the hardware and space required to build a data center, and you've equally invested the time to train people to operate the hardware you've purchased.

This is not something that can be easily migrated away from before seeing your return on investment (ROI). The amount of time to see returns varies for every company. ROI is often calculated using your company's hardware refresh cycles, project funding cycles, or general industry trends.

Once you have realized your ROI, you can choose to keep the existing processes and expertise to reinvest in your data center or partner with an external services provider with new investments. No matter what decision you make, you can still modernize with Kubernetes on premises to gain efficiencies, deployment speed, and stability. Additionally, if you choose to migrate your on-premises workloads to the cloud later, running Kubernetes on premises can greatly streamline the migration process.

Figure 1: AWS Global Infrastructure Map

The AWS Cloud spans 99 Availability Zones within 31 geographic AWS Regions around the world, with announced plans for 15 more Availability Zones and five more AWS Regions in Canada, Israel, Malaysia, New Zealand, and Thailand.



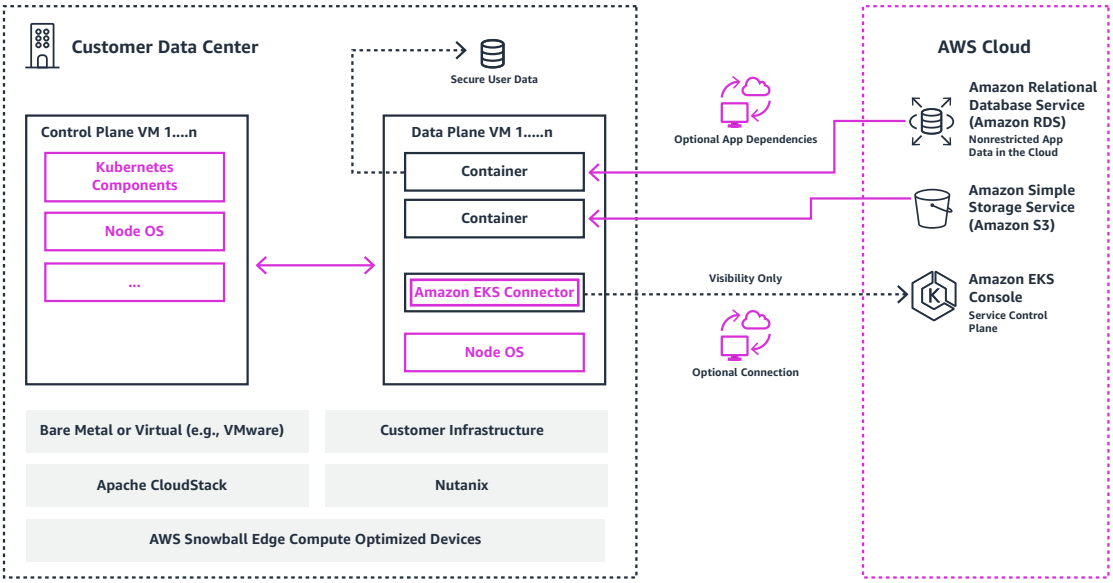
Data locality requirements

Different countries and industries have varying levels of risk tolerance. We believe that the world will continue to gain trust in managed infrastructure services. However, the curve of adoption takes longer for some countries and industries.

If your business operates in a country with data residency requirements, you must comply with local regulatory rules to do business. Some countries have determined that the best way to keep data safe is by physically storing the magnetic fields or electric charges—collectively "data"—in data centers within geographic borders.

If you're unable to leverage technology outside of these regulatory requirements, running infrastructure on premises may be the best option; for example, gaming customers that need to keep data and operations within national borders without dependencies on resources in other countries or manufacturing customers that want to run their data in factories because of the sensitivity of the workload and to avoid being subject to wide-area network connections.

Figure 2: With Amazon EKS Anywhere, customers can store secure data on premises while leveraging the cloud for unrestricted data, reducing their overall operational overhead.



Infrastructure that is required to maintain compliance with standards such as the Health Insurance Portability and Accountability Act (HIPAA) and the PCI Security Standards Council (PCI SSC) should be managed carefully. Modernizing the data and compute infrastructure in verified locations can reduce migration overhead while increasing operability.

Many applications and organizations run workloads that forbid internet access. They neither publish publicly available endpoints nor rely on internet-hosted infrastructure to provide value. With this architectural pattern, managing a data center could be the best option. This architectural pattern is supported in AWS-managed infrastructure, but if there is no local **AWS Availability Zone** in your location, then managing a data center is often the best choice.

Latency-sensitive compute

Some environments don't have regulatory restrictions, but they may have physical limitations. Every mile of wire between your CPU and your data adds 8.2 microseconds of latency. It doesn't matter if that mile is from another data center, cloud provider, or wires within the data center. If latency matters, location matters.

For applications that need real-time processing like augmented reality (AR), autonomous vehicles, game streaming, and real-time trading, **AWS Outposts**, a family of fully managed solutions delivering AWS infrastructure and services to virtually any on-premises or edge location for a truly consistent hybrid experience, and **AWS Local Zones**, infrastructure deployment that places compute, storage, database, and other select AWS services close to large population and industry centers, are available for this need.

This is especially true for applications that serve functions like routing for telecommunications. You may need the business logic to be near the routing layer to make fast decisions for customer calls or to serve data from the closest Point of Presence (PoP).

Industries that process large amounts of data, such as the film and entertainment industries, develop petabytes of information requiring rendering or processing. The data is often created in a studio where artists create digital assets or filmmakers record directly to network storage. Moving this data takes time. Adding compute close to the storage is architecturally the simplest and most performant solution.

Amazon EKS for on premises, in the cloud, and at the edge

Today, 80 percent of all containerized applications running in the cloud run on AWS.[1] Customers choose Amazon EKS to run containers because of its security, reliability, scalability, and ease of management. Through Amazon EKS, AWS enables customers to run Kubernetes wherever their applications require. Customers can run the same AWS-managed Amazon EKS offering in **AWS Regions**, **AWS Outposts**, AWS Local Zones, and **AWS Wavelength Zones** for a consistent Kubernetes experience across environments. And customers that want to leverage existing investments have the flexibility to run Amazon EKS Anywhere on their hardware.

With this breadth of deployment options, Amazon EKS meets customers where they are so they can modernize at their own pace with a consistent set of solutions, whether they are migrating to the cloud or modernizing their on-premises and edge environments.

Figure 3

At the root of all Amazon EKS deployment options is **Amazon EKS Distro**, an open-source distribution of Kubernetes services built and maintained by AWS. This is the same Kubernetes distribution used by Amazon EKS in the cloud and on premises with **Amazon EKS Anywhere**. AWS offers a number of choices and locations for deploying Amazon EKS clusters:

- **AWS Regions** – Deploy in a single AWS Region across one or more Availability Zones powered by AWS-managed data centers.
- **AWS Local Zones** – In industry centers and close to large populations, this type of infrastructure deployment includes compute, storage, database, and other select AWS services. With AWS Local Zones, you can run latency-sensitive applications closer to end users.
- **AWS Wavelength Zones** – Within 5G networks, AWS Wavelength can embed AWS compute and storage services, providing mobile edge computing infrastructure for developing, deploying, and scaling ultra-low-latency applications. AWS Wavelength Zones allow for ultra-low-latency applications for 5G devices.
- **AWS Outposts** – In virtually any on-premises data center or colocation space, Outposts creates a consistent hybrid experience. This fully managed service offers the same AWS infrastructure, AWS services, APIs, and tools and runs on AWS-provided hardware in your data center. Outposts lets you run AWS infrastructure and services on premises for a consistent hybrid experience.
- **Amazon EKS Anywhere** – On your own infrastructure with the support of AWS, customers can create and operate Kubernetes clusters. Run Amazon EKS Anywhere on **bare-metal** servers, use a virtualization layer such as VMware vSphere, Apache CloudStack, or Nutanix, or on AWS Snowball Edge.

[1] Containers at AWS

Benefits of Amazon EKS Anywhere

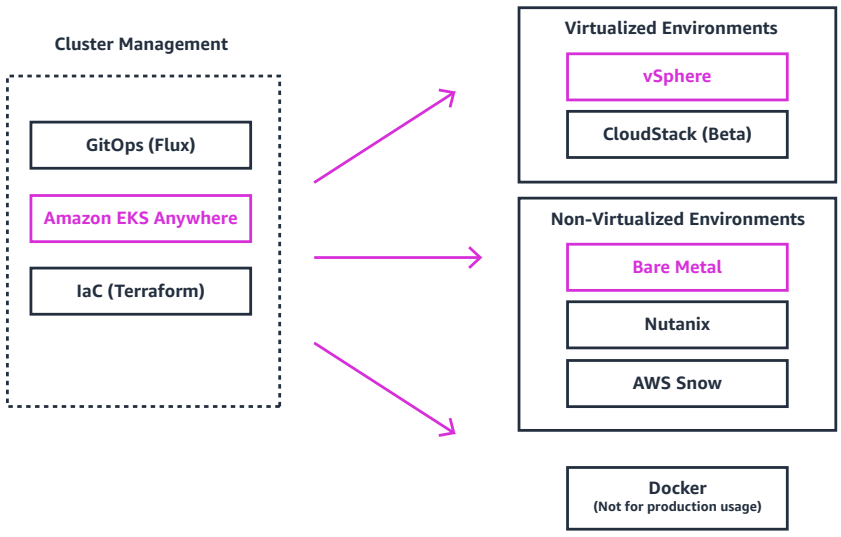
Amazon EKS Anywhere is open source. You don't need to be an AWS customer to use it or have existing contracts to get started. Instead, AWS offers enterprise support through a subscription. When compared to the other deployment options on AWS, Amazon EKS Anywhere is distinct, as it solves a combination of the three problems outlined in this whitepaper:

- **Capitalizing on existing investments**
- **Data locality requirements**
- **Latency-sensitive compute**

Deployment options – Amazon EKS Anywhere gives you the freedom to bring your own compute.

There are a number of options for deploying Amazon EKS Anywhere: directly on bare-metal servers or by using a virtualization layer such as vSphere, CloudStack, or Nutanix, or on Snowball Edge. You can even run Amazon EKS Anywhere on your local machine for non-production usage, giving teams an easy way to have a consistent experience when developing and testing applications.

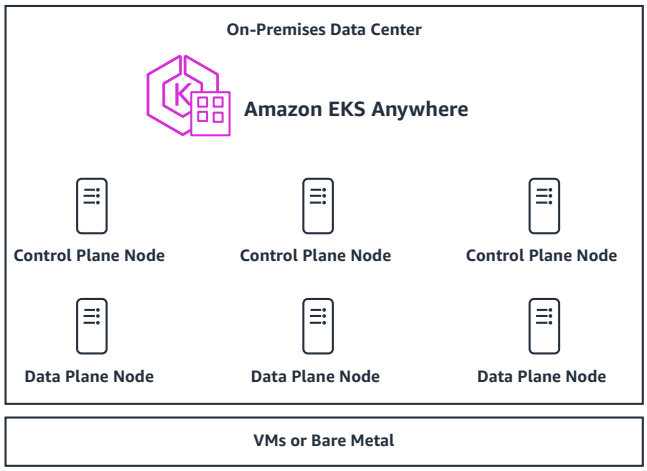
Figure 4



Flexible connectivity – Amazon EKS Anywhere allows you to connect back to Amazon EKS on AWS or run in disconnected environments.

Notably, Amazon EKS Anywhere builds on Amazon EKS Distro, allowing organizations to leverage their own infrastructure with the ability to administer control plane nodes where workloads are running. A notable advantage is the ability to run workloads in a "disconnected" mode without needing to communicate back to AWS data centers.

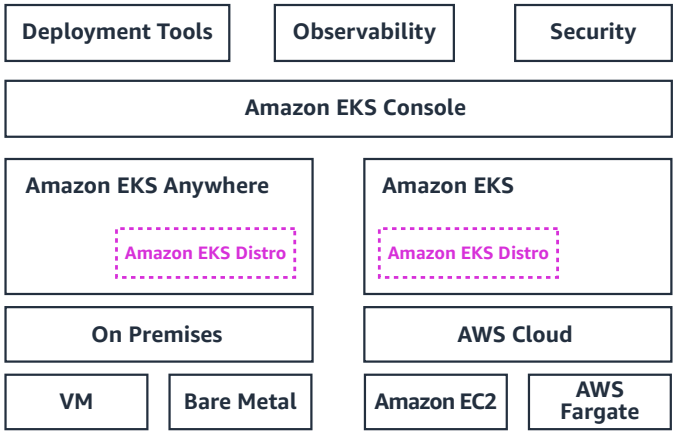
Figure 5



Consistent tooling and opinionated packaging – Avoid the paradox of choice with AWS-supported tooling.

With consistent tooling and opinionated packaging, tooling across on-premises and cloud environments will be similar. This means you can use the same command line interface (CLI) tool, eksctl, to work with both Amazon EKS and Amazon EKS Anywhere clusters or choose other open-source solutions like Terraform and **cdk8s**. In addition, you can use the **Amazon EKS Connector** to see clusters visually represented in the Amazon EKS console.

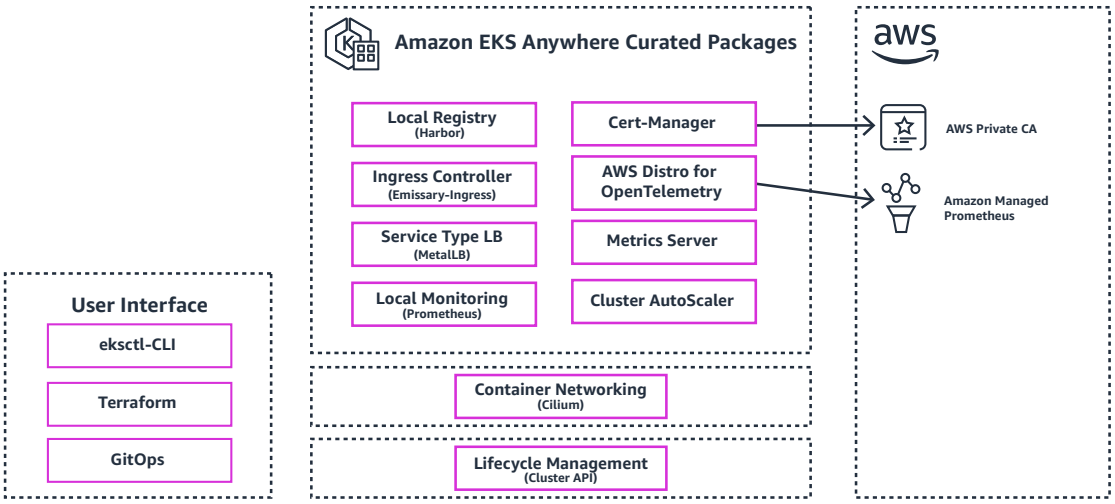
Figure 6



Curated Packages – Extend Kubernetes with capabilities from third-party vendors and open-source projects.

To help customers overcome decision fatigue when deciding on tooling (including open-source projects), Amazon EKS Anywhere introduced Curated Packages. These are Amazon-curated software packages that extend the core functionalities of Kubernetes on your Amazon EKS Anywhere clusters. Like Amazon EKS add-ons, Amazon EKS Anywhere Curated Packages are versioned and upgraded with your clusters for compatibility. And they are supported directly by AWS.

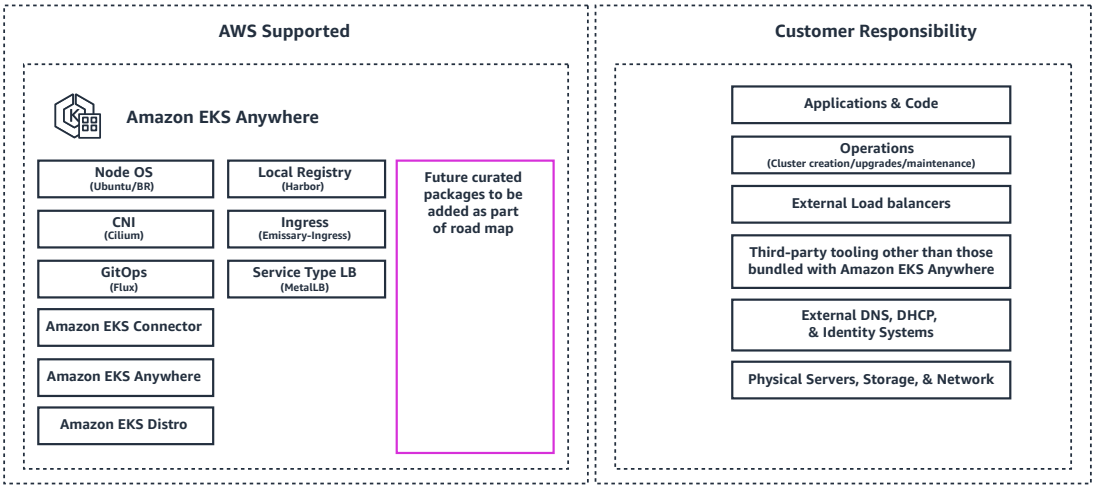
Figure 7



Shared support – One contact for all support needs.

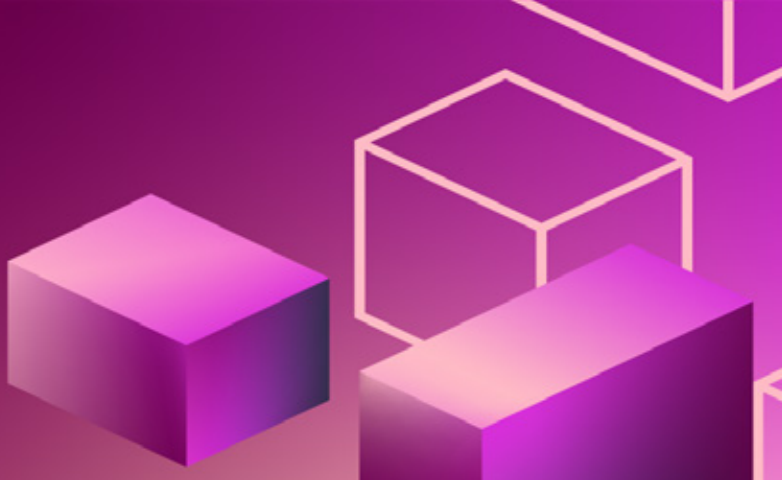
For customers, a single point of contact for support is critical. With AWS supporting the complete stack, from Amazon EKS Distro to tooling, such as the CLI, to a number of curated packages, you have one centralized place for getting the support you need.

Figure 8



CONCLUSION

Ready to get started?



As Kubernetes adoption continues to grow, more applications will be built with containers and cloud-native technologies. Wherever you are in your cloud modernization journey, AWS can provide the appropriate services and support to help your organization make the right decision for your business. Amazon EKS Anywhere is available as open-source software that you can download, install on your existing hardware, and run in your own data centers.

To get support and additional paid features for Amazon EKS Anywhere clusters, you can purchase an **Amazon EKS Anywhere Enterprise Subscription**. An AWS Enterprise Support Plan or AWS Enterprise On-Ramp Support Plan is a prerequisite for purchasing an Amazon EKS Anywhere Enterprise Subscription.

Learn more about AWS 24/7 Enterprise Support ›

Get in touch with an EKS Anywhere specialist ›

Visit our documentation and get started for free ›

Interested in learning more about AWS container services? Watch Amazon EKS, Amazon EKS Anywhere, and other containers-focused content on **Containers from the Couch**, or visit the **Amazon EKS Anywhere** service page for more information.