Data Sheet

F5 Distributed Cloud Customer Edge (CE) Deployable Software

Key Benefits

Integrate enterprise-grade networking and security
Unified L3 routing, network segmentation, TLS/SSL security, L4-7 distributed load balancing, advanced network firewall, service policy, distributed denial-of-service (DDoS) protection, and web application and API protection (WAAP) in a single software package.

Reduce network complexity
Simplify network connectivity across environments, without needing to manually provision networking services like VPN, NAT, or SD-WAN. Solve connectivity issues in case of IP address overlap.

Enable consistent security in any environment
Provide consistent network and application security policy definition, deployment, and management on any cloud, on-premises, or edge location.

Deploy across multiple platforms
Support for on-premises virtual machine (VM) and bare metal platforms, with automated deployment for all major public cloud providers.

Simplify operations across environments
SaaS console centrally manages connectivity, security and upgrades along with dashboard providing rich observability and insights

F5 Distributed Cloud Services at the Customer Edge

Hybrid and multicloud application deployment can be an operational and architectural challenge for many enterprises. Differing toolsets for each provider make maintaining connectivity and consistent application security difficult to manage, which can increase the time it takes to deploy a new application.

To solve this, enterprises need a solution that can seamlessly connect applications and provide consistent security policy enforcement across disparate environments and can be controlled from a central console.

This can be done by deploying Customer Edge (CE) software to every environment where apps are hosted. When deployed as an F5® Distributed Cloud Mesh site, CEs automatically create a secure network to connect different environments over the internet or a private network, with centralized management.

For enterprises deploying applications at the edge, the deployment experience also differs across environments. Each location will differ in areas like APIs, image formats, and access policies, which slows application deployments and upgrades. This is amplified as apps are deployed in more environments.

When CEs are deployed as a Distributed Cloud App Stack site, enterprises get an F5-managed Kubernetes platform to host applications. It exposes a Kubernetes API endpoint, allowing management of apps on each site or simultaneously across multiple sites.

Features

Cloud orchestration
Automated creation of virtual private clouds (VPCs), virtual networks, and other cloud-native networking, routing, and security objects while deploying a CE site on a public cloud.

Automated, encrypted network transport
Connect multiple CEs over the public Internet or via the F5 Global Network using IPSec tunnels and native TLS encryption across workloads.

Network segmentation
Isolate traffic from different VLANs and VPCs into multiple segments across environments and manage access within and across segments using advanced network firewall policies.

Web application firewall (WAF)
Detect signature-based attacks, security violations, and false positive events, run threat campaigns, and apply preventive actions.

Distributed load balancing
Discover applications at one CE site and advertise them to clients on other CE sites.

Automated service discovery
Automated discovery of services and apps using Kubernetes API, Consul, or DNS.

Distributed API protection
Get insights into APIs being used across your apps, classify endpoints with Personal Identifiable Information (PII) and secure access using policies

What is a CE?

F5 Distributed Cloud Services Customer Edge (CE) is a Kubernetes-based, integrated software stack which is managed centrally via the Distributed Cloud Console. CEs can be deployed as a virtual machine (VM) or as a standalone containerized service in any environment. It orchestrates the local control plane and data plane components to route, encrypt, and secure traffic. A CE operates as a highly available edge gateway that can securely extend networks across sites, without the need to establish physical network connectivity.

CEs are generally deployed in clusters of three control nodes to ensure high availability, with support for multiple worker nodes. Single-node CE sites are also supported, but F5 recommends three-node clusters for redundancy. Worker nodes are not available in single-node deployments. CEs can be deployed as either **Distributed Cloud Mesh sites** or **Distributed Cloud App Stack sites**.

- **Distributed Cloud Mesh sites:** Each CE can automatically connect to other CEs, as well as regional edges (REs) running on points of presence (PoPs) on the F5 Global Network, using IPSec tunnels. They can be used as L3 gateways to reach other networks with CEs deployed, or as distributed load balancers that can discover applications on one CE site and advertise them to any other CE site, or the public Internet through REs. They serve as points of policy enforcement as traffic passes through each site, and can deliver WAAP services such as Distributed Cloud Bot Defense and Distributed Cloud API Security locally without routing traffic externally. Security and networking policies can be centrally managed and pushed to every site with a CE to reduce operational overhead.
- **Distributed Cloud App Stack sites:** CEs can also act as a managed Kubernetes cluster for hosting applications with Distributed Cloud App Stack, in addition to providing connectivity as a mesh site. An App Stack site can host applications in both containers and VMs, and apps can be managed either individually on a single site, or simultaneously across multiple sites, by exposing a Kubernetes API endpoint.

CEs are managed centrally through the Distributed Cloud Console. In this console, CEs can be created, deleted, and upgraded, connectivity can be established between sites, and security policies can be configured and deployed.

Features (cont.)

SaaS-based lifecycle management

Receive continuous updates to security, networking, and app hosting services, managed through the F5 Distributed Cloud Console.

Application hosting with App Stack

Distributed Cloud App Stack is integrated within each CE to provide a consistent managed Kubernetes platform for hosting applications on the cloud, on-premises, or at the edge.

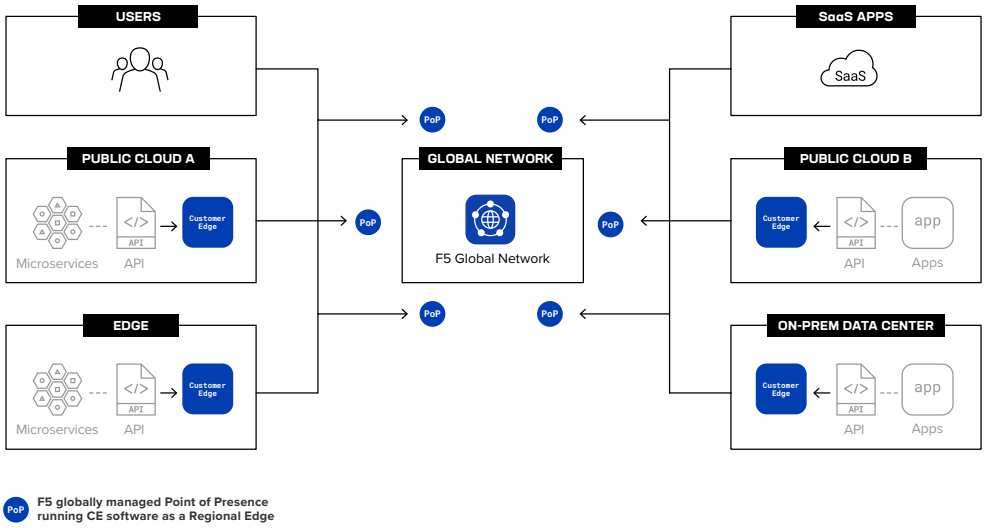

Figure 1: A CE can be deployed in any customer environment to deliver security and networking services locally.

Understanding CE Clustering

CEs are generally deployed in clusters of three control nodes to ensure high availability. In the event that one control node fails, high availability is guaranteed through the other two control nodes. Additional worker nodes can be deployed as part of the same cluster to provide additional throughput if needed.

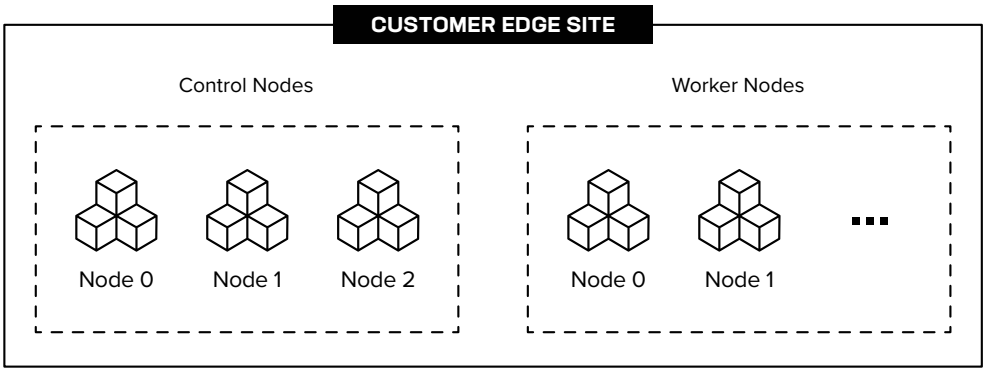

Figure 2: F5 recommends deploying CEs in three-node clusters in each environment, with support for worker nodes to assist with throughput.

In a multi-node configuration, two of the three CE control nodes create one tunnel each to REs running on the two nearest PoPs. If one control node with a tunnel fails, that tunnel is reassigned to the third CE control node. For single-node configurations, the single CE control node creates tunnels to the two nearest REs.

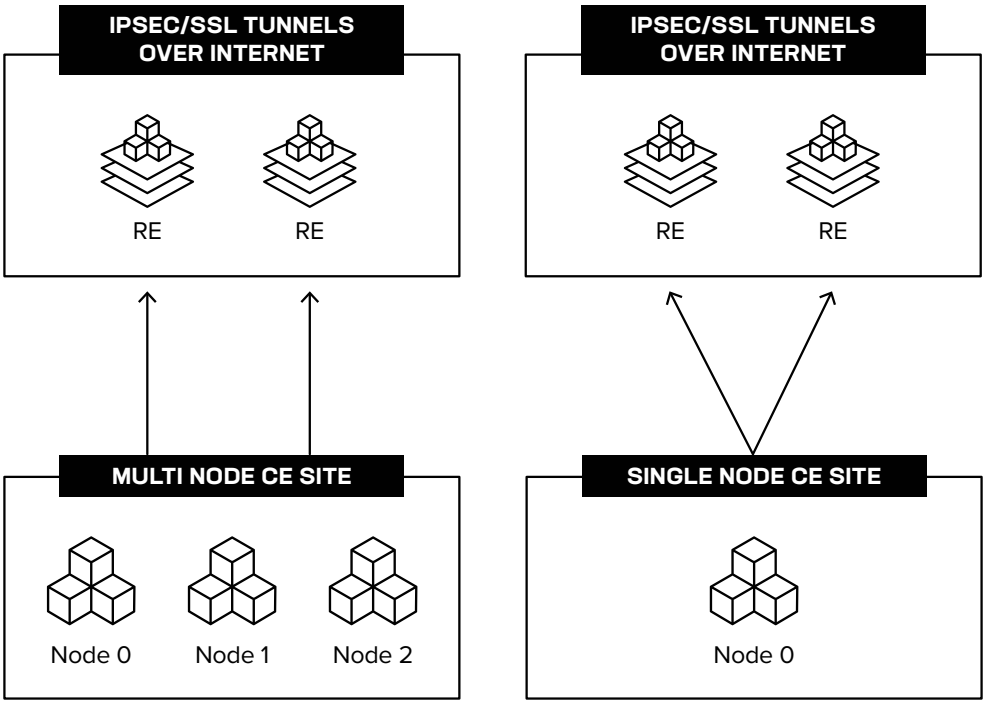

Figure 3: A multi-node CE deployment will automatically connect two CEs to the two closest REs, with the third CE available to take over connections in case of a CE failure.

A CE in a single-node deployment will connect to the two closest REs, with no redundancy.

CE Components

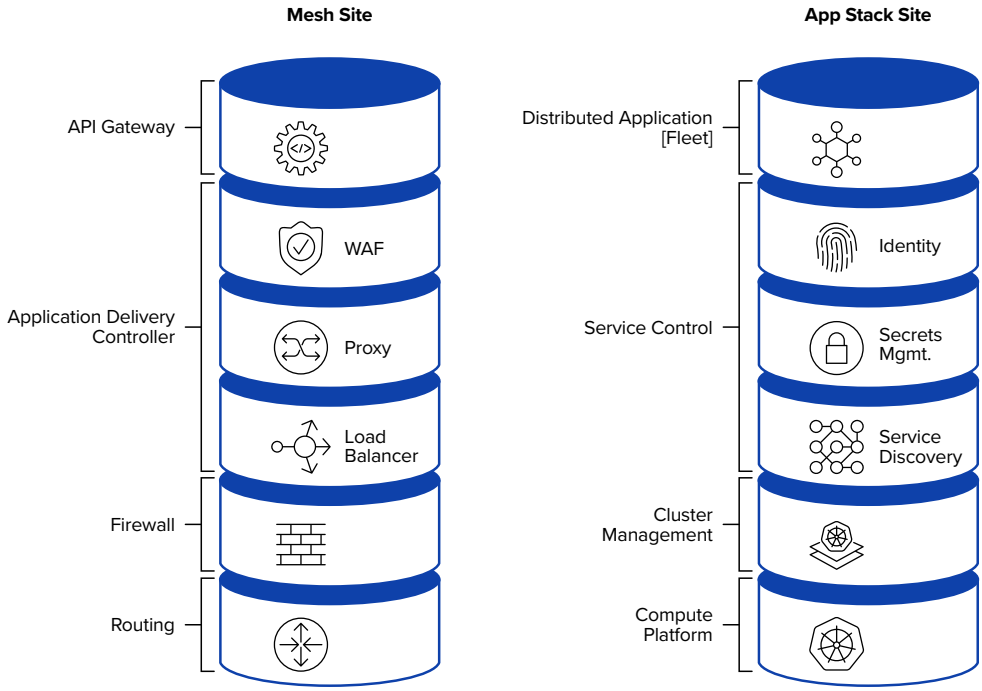

Figure 4: Each Distributed Cloud Mesh and Distributed Cloud App Stack site are comprised of several services to enable local security, networking, and app hosting.

CEs are comprised of several different components.

Distributed Cloud Mesh site components:

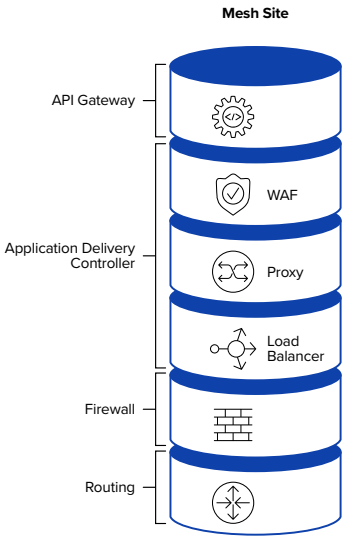

Figure 4a: Mesh Sites provide API gateway, application delivery controller, firewall, and routing services.

- **API Security**
  Mesh sites can discover API endpoints for application services and perform behavioral analysis on endpoint logs via advanced machine learning. Distributed Cloud API Security provides endpoint learning, schemas, shadow API and sensitive data detection, and OpenAPI Specification (OAS) generation.
- **Web Application Firewall (WAF)**
  Mesh sites include a WAF that integrates behavioral analysis and signatures to assess threats. It detects signature-based attacks, distinguishes good bots from malicious bots, runs threat campaigns, detects security violations and false positive events, and applies associative preventive actions. It can also mask sensitive data in request logs, deliver customized blocking response pages, and custom-define a list of allowed response codes.
- **Proxy**
  Dynamic Reverse Proxy on Mesh sites enables connectivity to remote SaaS apps privately, without needing to create any complex routing relationships or advertising public-IP space inside existing corporate networks, by implementing forward proxy and network address translation (NAT) capabilities.
- **Load Balancer**
  Mesh sites can discover apps on any CE site and publish it to other sites or to the public internet. The load balancer integrates services like L7 routing, URL filtering, service policy, TLS offload, fault injection, caching, and TCP optimization.
- **Network Firewall**
  Mesh sites have an integrated network firewall which can implement forward proxy policies, enhanced firewall policies, and access control lists (ACLs), and use tags to dynamically manage access within and across the network segments. Service insertion of Palo Alto Networks Firewall is also supported.
- **Routing**
  Mesh sites enable networks to be routable across sites using IPSec tunnels, supports network segmentation, implements NAT to solve IP address overlap between connected networks, enables segment routing over IPv6 (SRv6) and border gateway protocol (BGP) for route advertisements.

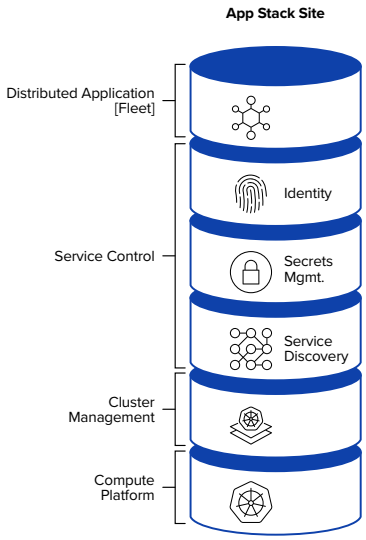

Figure 4b: App Stack Sites, in addition to the services delivered by Mesh Sites, provide distributed application management, service control, cluster management, and compute platform services.

Distributed Cloud App Stack site components:

- **Distributed Application Fleet Management**
  App Stack sites can be centrally managed across environments, enabling users to manage a large number of App Stack sites through a single console.
- **Identity Control**
  App Stack sites allow each app instance to receive a unique identity via a public key infrastructure (PKI) certificate, which is issued and maintained throughout the applications lifecycle. This identity is used for mutual transport layer security (mTLS) between apps, retrieving application secrets and accessing keys stored in the key management service (KMS).
- **Secrets Management**
  App Stack sites offer cryptographically secure, double-blinding system storage for customer secrets. Secrets are not stored in the clear, so even in the event of a breach, that information will not be lost. App Stack sites can also be integrated with enterprise secrets management products like HashiCorp Vault or CyberArk.
- **Service Discovery**
  App Stack sites can automatically discover service endpoints and publish virtual IP addresses (VIPs). With apps and services able to migrate to different environments, automated service discovery removes the need for manual proxy configuration. Discovery can be done via the DNS method, using Kubernetes native discovery, or HashiCorp Consul.
- **Cluster Management**
  App Stack sites support clustering multiple compute and storage resources into a single site. Application workloads and cloud services can autoscale as soon as a new node is added to the cluster. Using underlying autoscaling capabilities provided by the cloud provider, scale the number of nodes within a site depending on demand and configuration constraints through Distributed Cloud Services from F5.
- **Compute Platform**
  App Stack sites provide a managed Kubernetes cluster for hosting applications, in addition to providing connectivity as a mesh site. An App Stack site can host applications in both containers and VMs.

Ecosystem Support

CEs can be deployed on most major public cloud providers and on-premises virtualization platforms:

Provider Type	Provider	Automated Deployment	Manual Deployment
Public Cloud	Amazon Web Services (AWS)	Yes	Yes
	Microsoft Azure	Yes	Yes
	Google Cloud Platform (GCP)	Yes	Yes
	Oracle Cloud Infrastructure (OCI)	No	Yes
On-Premises	VMware	No	Yes
	Kernel-based virtual machine (KVM)	No	Yes
	Bare metal servers	No	Yes
	F5 rSeries	No	Yes
	Red Hat OpenShift	No	Yes

While some providers only support manual deployment, all CE services and upgrades are managed from the Distributed Cloud Console once sites are registered.

CE Deployment Requirements

A CE can be deployed as a VM instance on public cloud, as a VM on VMware and KVM, directly on a bare metal server, and in any regular or cloud managed Kubernetes cluster as a pod.

F5 recommends deploying a minimum of three CEs for high availability in production environments. Additional worker nodes can be added to improve L7 performance for any deployment or extra hosting capacity for Distributed Cloud App Stack deployments.

Minimum resources required per node:

- 4 vCPUs
- 14 GB RAM
- 80GB of disk space for Distributed Cloud Mesh nodes OR 100GB for Distributed Cloud App Stack nodes

F5 offers CEs in three different sizes, depending on performance requirements. Recommended instance sizes for different providers are:

Provider	Small Node	Medium Node	Large Node
AWS	t3.xlarge	t3.2xlarge	m5.4xlarge
Microsoft Azure	Standard_D3_v2	Standard_D4_v2	Standard_D5_v2
GCP	n1-standard-4 t2d-standard-4 a2-highgpu-1g	n1-standard-8 t2d-standard-8 a2-highgpu-2g	n1-standard-16 t2d-standard-16 a2-highgpu-4g
OCI, VMware, KVM, Red Hat OpenShift, F5 rSeries	4 vCPUs 16 GB RAM	8vCPUs 32 GB RAM	16vCPUs 64 GB RAM

Note: NVIDIA A2 GPUs are supported for bare metal deployments on HP DL360 and Dell R650 servers.

Deployment Workflow

Automated Deployment

On AWS, Microsoft Azure, and Google Cloud Platform, automated deployment of CEs can be managed via the Distributed Cloud Console. This method also automates virtual private cloud (VPC) and virtual network (VNet) creation, as well as networking configuration, security groups creation, and other required objects for site operation. In this mode, CEs can also be deployed on existing VPCs/VNets.

Manual Deployment

CEs can also be manually deployed on any supported provider. This enables greater flexibility of deployment topologies, which can be useful for brownfield deployments. On AWS and Microsoft Azure, CE images are published on the respective marketplaces. For other providers, CE images are available for download, and can be manually installed in supported environments. Once installed, users can create tokens in the Distributed Cloud Console while creating CE nodes. These tokens allow CEs to authenticate and register with a user's tenant.

F5 has compiled deployment instructions for cloud, on-premises data centers, Kubernetes clusters, and bare metal servers in the Site Management documentation.

Upgrades

Upgrades to CEs are managed via the Distributed Cloud Console. CEs are upgraded separately from the instance operating system (OS). The site dashboard shows the current and the latest available software and OS version. An upgrade link is shown if a new version is available, which starts the upgrade.

Once the upgrade is started, CE clusters are upgraded one node at a time. The OS upgrade will automatically rollback to the previous version if the upgrade is unsuccessful.

CE versions are either based on a standard release crt-<version-date> or an LTS release lts-<version-date>. The standard releases (crt-xxxx) are regular software releases with a six-month support window offering bug fixes as well as new features better suited for customers who want to consume the latest features and functionality.

LTS releases (lts-xxxx) are intended to offer a stable software that is supported for an extended duration for a total of 12 months with only critical bug fixes and security vulnerabilities delivered in maintenance releases.

Full Feature List

Feature	Description
Site Mesh Groups (SMG)	Direct site-to-site IPSec tunnels over the Internet or private network, with sites connected in full-mesh or hub-spoke configuration
Data Center Cluster Groups (DCG)	Direct site-to-site, IP-in-IP tunnels over private network in full-mesh connection configuration
Network Segmentation	Isolate traffic from different VLANs and VPCs into multiple segments across environments
Enhanced Firewall	Network policies, ACLs, forward proxy policies
Cloud Orchestration	Automated creation of cloud VPC/VNet and other cloud-native networking, routing, and security objects while deploying a CE site on a public cloud
Cloud Connect	Onboard VPCs across different accounts onto F5 Distributed Cloud Services
Network Address Translation (NAT)	Connect networks with overlapping CIDRs together with automatic NAT
Private Connectivity to Public Cloud	Provides support for AWS Direct Connect and Microsoft Azure ExpressRoute
Cloud Link	Orchestrate pre-provisioned direct connections, create a network, and then connect, deliver, secure, and operate networks and apps across hybrid environments

Accelerated Networking	Azure Accelerated Networking support for increased network performance
DPDK	Intel DPDK support on bare metal servers with supported NICs
L3 Enhanced Performance mode	Supports jumbo frames and reserves more CPU/memory resources for L3 processing
L3/L4 DDoS Mitigation	Detect and mitigate large-scale, volumetric, network-targeted attacks in real-time
BGP Peering	BGP peering support to advertise load balancing VIPs and routes
F5 BIG-IP Advanced WAF	Traffic steering policy to forward traffic to BIG-IP Advanced WAF for inspection and security enforcement
Palo Alto Networks Firewall Support	Traffic steering policy to forward traffic to Palo Alto Networks Firewall for inspection and security enforcement
Distributed Load Balancer	Discovers services at a CE site and advertise them to the same or other sites or the internet over PoPs
Dynamic Reverse Proxy	Connect to SaaS providers privately without the need to create complex routing relationships and without the need to advertise public IP space inside organizations' corporate networks
IP Reputation	Blocks traffic from F5-published threat sources
Web Application Firewall	Detects signature-based attacks, distinguish good bots from malicious bots, runs threat campaigns, detects various security violations, detects false positive events, and applies preventive actions
API Discovery	Provides endpoint learning, including request and response schemas and sensitive data detection. It also provides show inventory and shadow sets and OpenAPI Specification (OAS) generation
Service Policies	Allow/deny requests based on L7 parameters
Malicious User Detection	Analyze and block users with suspicious activities
Managed Kubernetes	Provides a Kubernetes API endpoint, application runtime, and scheduling layer for deploying applications on App Stack sites
Virtual Kubernetes	Provides a Kubernetes API endpoint for deploying applications on an associated namespace across multiple App Stack sites that are part of a virtual site
Service Discovery	Automated discovery of services/apps running on external Kubernetes clusters
Offline Survivability	Allows CE to function in case of upstream connectivity outage

GPU support	GPU support on App Stack sites for running AI/ML workloads
Persistent Volumes	PersistentVolume and PersistentVolumeClaim support on App Stack sites
Virtual Machine	Deploy apps as VMs on App Stack sites

Offline Survivability

CEs require control plane connectivity to REs and the global controller (GC) to exchange routes, renew certificates, and decrypt blindfolded secrets. To enable business continuity during an upstream outage, Offline Survivability can be enabled on all sites in a full-mesh site mesh group (SMG) or data center cluster group (DCG). The feature is not supported for hub-spoke SMG.

Offline Survivability mode enables CEs to continue normal operations for seven days without connecting to REs or the GC. With the offline survivability feature enabled on a CE site, the local control plane becomes the certificate authority in case of connectivity loss. The decrypted secrets and certificates are cached locally on the CE. Because of this, Offline Survivability is not turned on by default, as it may not align to existing security policies, so users can opt in.

In Offline Survivability mode, there are three components that ensure the CE remains active:

- **Routing:** Routes are exchanged via BGP within a site, and across sites in a mesh group or a DCG. In Offline Survivability mode, the local control plane allows local traffic load balancing for the site to continue. If two or more sites in a mesh group have Offline Survivability enabled, and the SMG is a full-mesh type group with control plane enabled, load balancing across local and remote endpoints in those sites continues to function, even when connectivity with the RE is lost. The same is also true for sites within a DCG.
- **Identity Management:** Certificates for services are issued via a certificate authority local to the site when they start or restart without connectivity to the GC. If services restart, they get new certificates and continue functioning.
- **Secret Management:** Secrets that are decrypted from the platform when connectivity to GC was intact are cached locally on the site. This enables services to obtain decrypted secrets even when the connectivity is lost.

When offline in Offline Survivability mode, logs older than five minutes and metrics older than two hours are lost.

Load Balancer Features

A CE can operate as an L4 or L7 load balancer. With multiple CEs deployed across different providers, they act as a distributed load balancer which can discover applications at one CE site and advertise them to clients on other CE sites.

LB Type	Features
HTTP or HTTPS	Automatic certificate
	Custom certificate
	Auto HTTP to HTTPs redirect
	Custom listen port and port range
	Configurable TLS version
	Custom cipher suits
	Mutual TLS
	CRL verification
	HTTP v1 and v2 support
	L7 Routing (match HTTP headers, method, path, port, and redirect, send a direct response, or select an origin pool)
	Origin server subsets (load balance request to a subset of origin servers based on route match)
	Header manipulation (add/remove header in request/response. Add request and certificate parameters to headers)
	SPDY support
	WebSocket support
	Retry policy
	Request mirroring (mirroring traffic to another CE site)
	Cross-origin resource sharing (CORS)
	Custom error responses
	Request buffering
	Response compression
	Idle timeout

TCP	Custom listen port and port range
	SNI
	Automatic certificate
	Custom certificate
	Configurable TLS version
	Custom cipher suits
	Mutual TLS
	CRL verification
	Proxy protocol (supported only for traffic from LB to the origin pool)
General	Advertise to Internet (advertise services over PoPs using anycast IP or user-specified public IP)
	Advertise to Internet (advertise services over PoPs using anycast IP or user-specified public IP)
	Custom private VIP (user-defined IPv4 and IPv6 VIP)
	Load balancing algorithms • Round Robin • Least Active Request • Source IP Stickiness • Cookie-Based Stickiness (HTTP only) • Ring Hash Policy (HTTP only) • Random
	Trusted client IP headers
	Health checks (HTTP, TCP, custom)
	Re-encryption to origin
	Circuit breaker (temporarily disables traffic to the origin server if it returns error responses above a given threshold)
	Outlier detection (temporarily removes origin servers that consistently perform badly (passive health check)

More Information

Learn more about deploying CEs to any environment.

Resources

CE High Availability for App Delivery