TECHNICAL WHITE PAPER: **March 2026**

F5 BIG-IP Container Ingress Services on vSphere Kubernetes Service on VMware Cloud Foundation

Reference Architecture

vmware by Broadcom

Table of contents

Executive Summary

F5 BIG-IP with Container Ingress Services (CIS) integrates with VMware vSphere Kubernetes Service (VKS) running on VMware Cloud Foundation (VCF) by watching Kubernetes resources (such as Ingress, Services) and automatically programming BIG-IP to deliver enterprise-grade application connectivity for workloads deployed on the VKS platform. In this joint solution, platform teams standardize on VCF for consistent compute, storage, and networking operations while CIS provides a Kubernetes-native control plane that translates application intent into BIG-IP virtual servers, pools, health monitors, and security policies—enabling scalable north–south traffic management, centralized policy enforcement, and consistent L4-L7 services across clusters and environments. Customers choose this combined approach to accelerate app delivery without sacrificing proven BIG-IP capabilities such as advanced load balancing, TLS offload, granular traffic control, high availability, and integration with existing operational processes, while also improving security posture and governance through repeatable, automated configuration aligned to Kubernetes deployments. This technical paper illustrates the deployment architecture for our jointly validated solution.

Benefits of using VKS with modern applications:

Lower TCO: With VKS organizations can reduce silos, leverage existing tools and skill sets without having to retrain staff and/or change existing processes. Utilizing unified lifecycle management across infrastructure components to stay up-to-date with the most recent patches and minimizing security risks.

Operational Simplicity: VKS is engineered for unparalleled operational simplicity, leveraging the familiarity of existing vSphere tools, skills, and workflows. This design philosophy significantly reduces the learning curve for IT teams and streamlines management processes. With VKS, organizations benefit from automated cluster provisioning, which accelerates deployment times and minimizes manual configuration errors. Furthermore, its robust capabilities extend to automated upgrades and comprehensive lifecycle management. This integrated approach ensures consistent operations, reduces overhead, and frees up valuable resources to focus on innovation rather than infrastructure maintenance.

Run and Manage Kubernetes at Scale: Effortlessly deploy and manage Kubernetes clusters at scale, leveraging a built-in, Cloud Native Computing Foundation (CNCF) certified Kubernetes distribution. VKS provides fully automated lifecycle management, streamlining operations from initial setup to ongoing maintenance and upgrades. This comprehensive approach ensures that organizations can harness the power of Kubernetes for their containerized applications with unparalleled efficiency and reliability, without the complexities typically associated with large-scale Kubernetes deployments.

vSphere Kubernetes Service

vSphere Kubernetes Service (VKS) is the Kubernetes runtime built directly into VMware Cloud Foundation (VCF). With CNCF certified Kubernetes, VKS enables platform engineers to deploy and manage Kubernetes clusters while leveraging a comprehensive set of cloud services in VCF. Cloud admins benefit from the support for N-2 Kubernetes versions, enterprise grade security, and simplified lifecycle management for modern apps adoption.

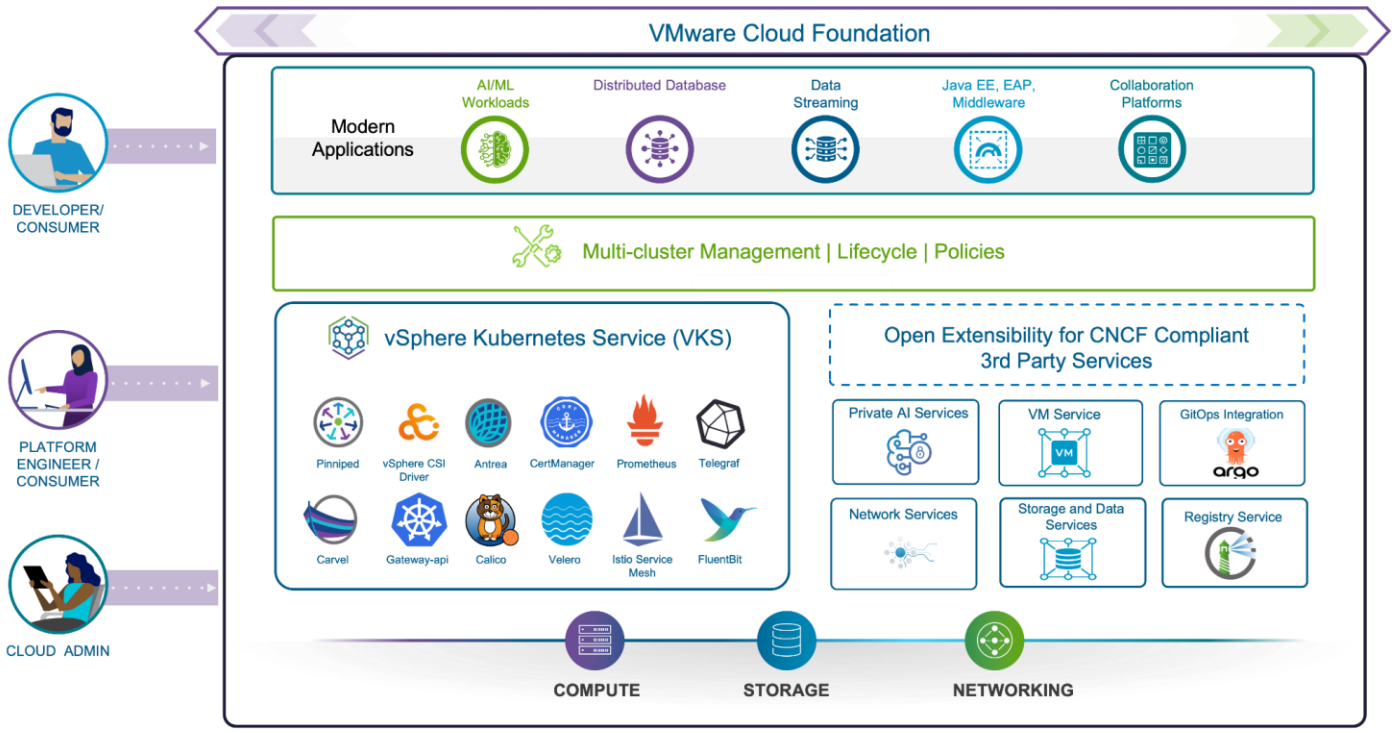

Figure 1: VKS on VCF Ecosystem

F5 BIG-IP Container Ingress Services

F5 BIG-IP Container Ingress Services (CIS) lets you publish and protect your Kubernetes applications in F5 BIG-IP device from Kubernetes. The F5 BIG-IP becomes an external L4/L7 load balancer and ingress controller using the Kubernetes native CLI/API. On top of exposing externally the Kubernetes services it adds the expected functionalities from BIG-IP:

- Advanced TLS encryption including safe key storage with Hardware Security Module (HSM) or Network & Cloud HSM support.
- Advanced WAF, L7 bot and API protection.
- L3-L4 High performance firewall with IPS for protocol conformance.
- Behavioural DDoS protection with cloud scrubbing support.
- Visibility into TLS traffic for inspection with 3rd party solutions.
- Identity-aware ingress with Federated SSO and integration with leading MFAs.
- AI inference and agentic support thanks to JSON and MCP protocol support.

Moreover, F5 BIG-IP CIS allows to expose applications hosted in multiple Kubernetes clusters into a single VIP in the BIG-IP, thus bringing application-aware multi-cluster for Kubernetes, an industry first.

F5 BIG-IP also provides integration with third-party IPAM solutions such as Infoblox for VIP address management.

The installation is performed using standard Helm charts and the publishing and configuration of applications, including its protections, is performed using either standard Ingress resources or F5 CRDs.

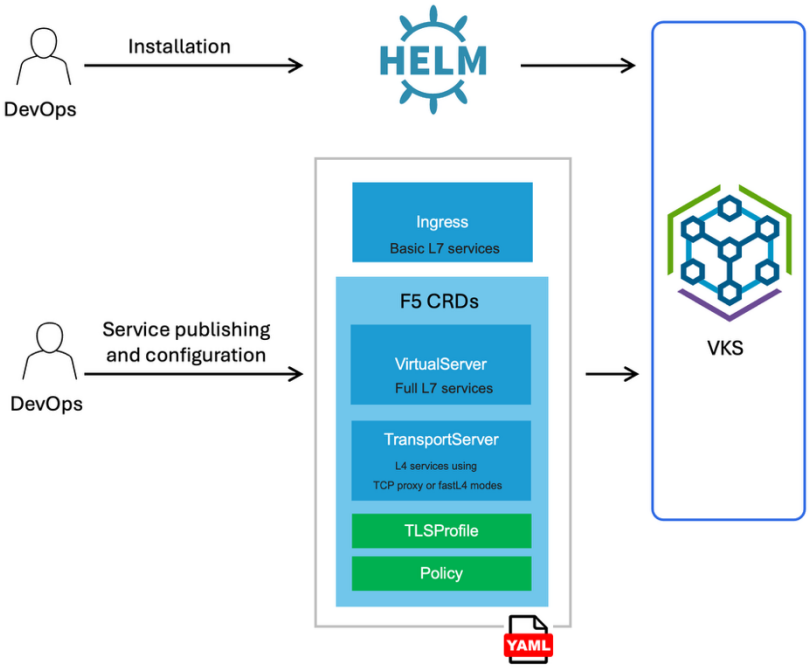

Figure 2: Kubernetes workflows for Installation, service publishing and configuration with F5 BIG-IP and F5 CIS

When using F5 CRDs all functionality available in F5 BIG-IP is exposed to Kubernetes.

The VirtualServer and TransportServer CRDs are basically L7 and L4 VIPs respectively and simple service publishing typically does not require the use of additional resources.

The TLSProfile and Policy CRDs allow to setup advanced services which require customizations or need to expose any of advanced functionalities of the BIG-IP. These CRDs can be referenced and shared from multiple VirtualServer and TransportServer resources. The following code block shows an example of a Policy CRD.

```
apiVersion: cis.f5.com/v1
kind: Policy
spec:
  l7Policies:
    waf: /Common/WAF_Policy
  l3Policies:
    firewallPolicy: /Common/AFM_Policy
    dos: /Common/dos
    botDefense: /Common/bot-defense
    allowSourceRange:
      - 1.1.1.0/24
      - 2.2.2.0/24
  logProfiles:
      - /Common/Log all requests
```

```
    - /Common/local-dos
  htmlProfile: /Common/html
iRules:
  secure: /Common/irule1
  insecure: /Common/irule1
```

Solution Architecture

From a compute point of view, the F5 BIG-IP sits outside of the VKS cluster as either a Virtual Machine, an Appliance or a Scale-Out Chassis. The configuration is the same regardless of the form-factor.

From a network point of view and independently of the form-factor, the F5 BIG-IP is both an External Load Balancer and Ingress controller thanks to supporting Direct-to-Pod communication with both VKS-supported CNIs (Antrea and Calico). This eliminates the need of an in-cluster ingress controller, thus simplifying the traffic flow, reducing latency and providing better Pod monitoring and session affinity (a.k.a. stickiness). A second tier Ingress Controller can optionally be deployed for additional NetOps-DevOps teams separation. F5 CIS can also operate with the more traditional NodePort mode.

With respect to the control-plane (dotted lines in the next figure), F5 CIS acts as a controller which continuously reconciles the configuration set by DevOps in Kubernetes (the source of truth) with the configuration in the F5 BIG-IP. The interaction between F5 CIS and F5 BIG-IP is done using F5 AS3 (Application Services 3 Extension) declarative API.

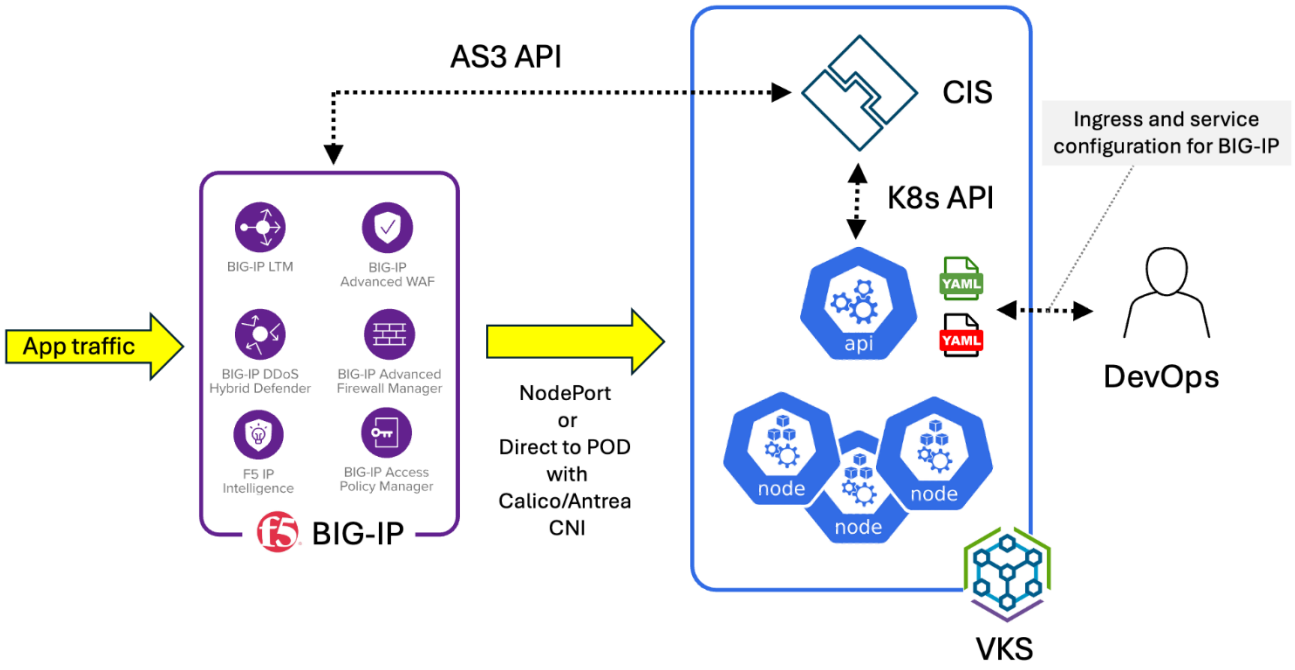

Figure 3: Detail of F5 BIG-IP and CIS interaction with VKS for App and DevOps flows

A fully deployed architecture of VMware VKS with F5 BIG-IP and CIS is a highly redundant solution at all levels:

- F5 BIG-IP is deployed in high availability mode with F5 ScaleN horizontal scaling. Moreover, F5 BIG-IP provides end-to-end monitoring for Pods when using direct-to-Pod networking.
- VMware VKS, being a Kubernetes platform, provides high reliability and redundancy for applications.
- Optionally, F5 CIS can discover the availability of the Pods of a given application in multiple VKS clusters and publish it as a single customer-facing application. Hence the application availability does not depend on the availability of a single cluster.

Please note in the next figure that, as indicated previously, the use of an in-cluster Ingress Controller is optional. This is the case regardless of using a single or multiple VKS clusters.

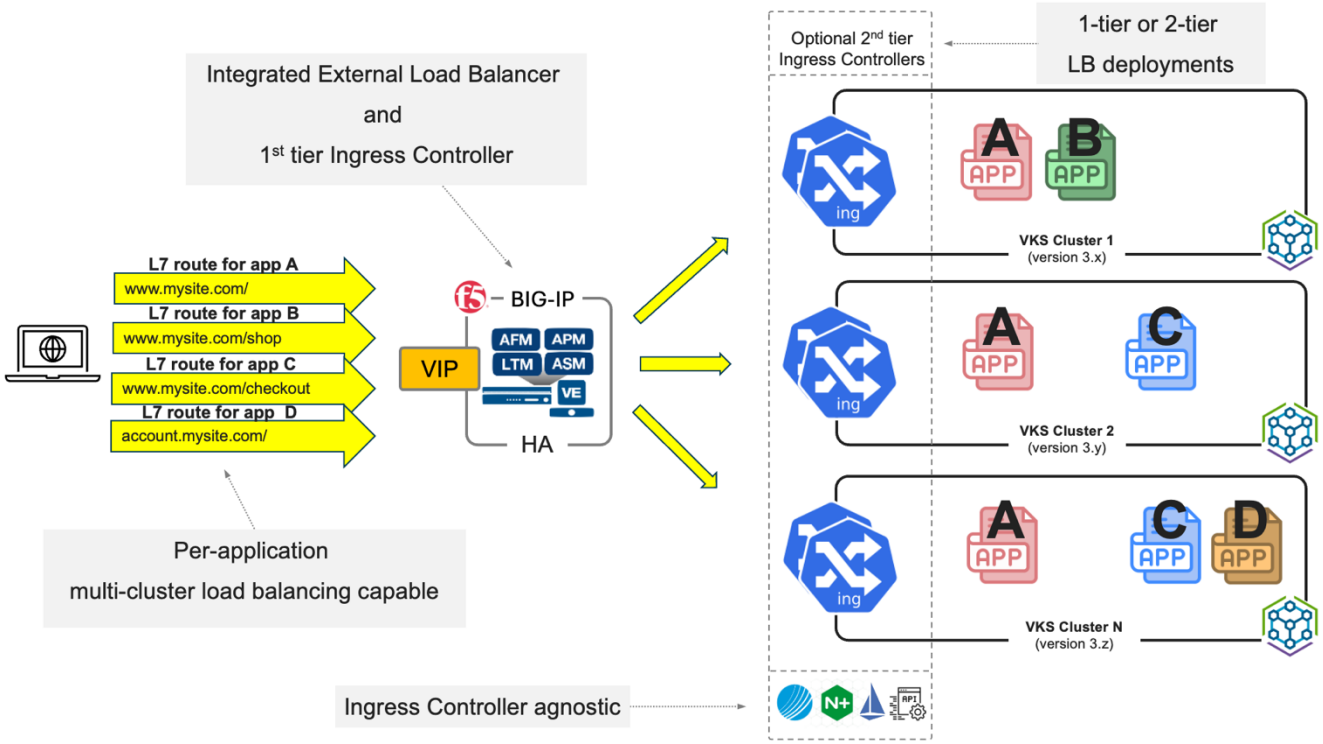

Figure 4: Overall deployment of F5 BIG-IP with VMware VKS showcasing its ability to publish an application from multiple clusters

The use cases for multi-cluster deployments are as follows:

- Facilitate VKS migrations by allowing applications to be served by multiple VKS clusters.
- Increase capacity.
- Increase high availability, for example during maintenance windows.

- Blue-green (A/B testing) deployment of applications across clusters.
- Sharding of applications across clusters by placing specific applications in specific clusters.
- Disaster Recovery sites.
- Local geo-redundancy by having VKS clusters in stretched data centers or in different Availability Zones of the same region.
- Split large VKS clusters into smaller ones.
- L4 multi-cluster load balancing is also possible, typically used for database applications.

Solution Validation

Below is the list of software versions used during this validation.

Component	Version
VMware Cloud Foundation	9.0
vSphere Kubernetes Service (VKS)	3.6
vSphere Kubernetes release (VKr)	1.35
F5 BIG-IP	v17 and v21
F5 Container Ingress Service (CIS)	2.20.3
F5 AS3	3.56.0

The following combination of CIS features and CNIs have been tested with VMware VKS.

	Ingress	CRD	Multi-Cluster
Antrea – direct to Pod using NodePortLocal	✓	✓*	N/A
Calico – direct to Pod using ClusterIP	✓	✓	✓
Antrea and Calico – access to Node using NodePort	✓	✓	✓

Both Antrea in NodePortLocal mode and Calico in ClusterIP mode allow to have direct-to-Pod traffic.

For CIS multi-cluster or TransportServer support with Antrea NodePortLocal, please consult your F5 representative.

* CIS can operate in either Ingress mode or CRD mode but not both concurrently. If this is needed, additional instances of CIS can be run for the same BIG-IP.

The solution has been validated when using either NSX or vSphere networking.

When deploying BIG-IP in a NSX network, the following must be considered:

- The BIG-IP cannot have a leg in the same VPC segment where the VMware VKS cluster is because VPC is system-managed by NSX.
- Specifically for Calico, ClusterIP mode cannot be used in an NSX network because this would require the BIG-IP to have one leg in the same VPC segment as VMware VKS which is not possible at present. Use vSphere networking if Calico is to be used in ClusterIP mode.

Install F5 BIG-IP CIS on vSphere Kubernetes Service

Please follow the instructions in https://clouddocs.f5.com/containers/latest/userguide/vmware-vks/.

Conclusion

Customers benefit from the joint F5 BIG-IP CIS and VMware VKS solution by combining Kubernetes-native automation with proven, enterprise application delivery—so teams can deploy modern apps faster while maintaining consistent security, availability, and performance across clusters. By standardizing on VMware VKS for streamlined Kubernetes operations and using CIS to translate in-cluster intent into centrally managed BIG-IP traffic and policy controls, organizations reduce configuration drift, simplify day-2 operations, and extend existing BIG-IP investments into cloud-native environments. The result is a more secure, scalable, and operationally consistent application platform that helps customers modernize with confidence.