Palantir

Kubecon 2019

Palantir
Technologies

Kubernetes on Ephemeral Infrastructure

DATE /

23.05.19

SPEAKERS /

01 – Greg
DeArment
Head of Production
Infrastructure

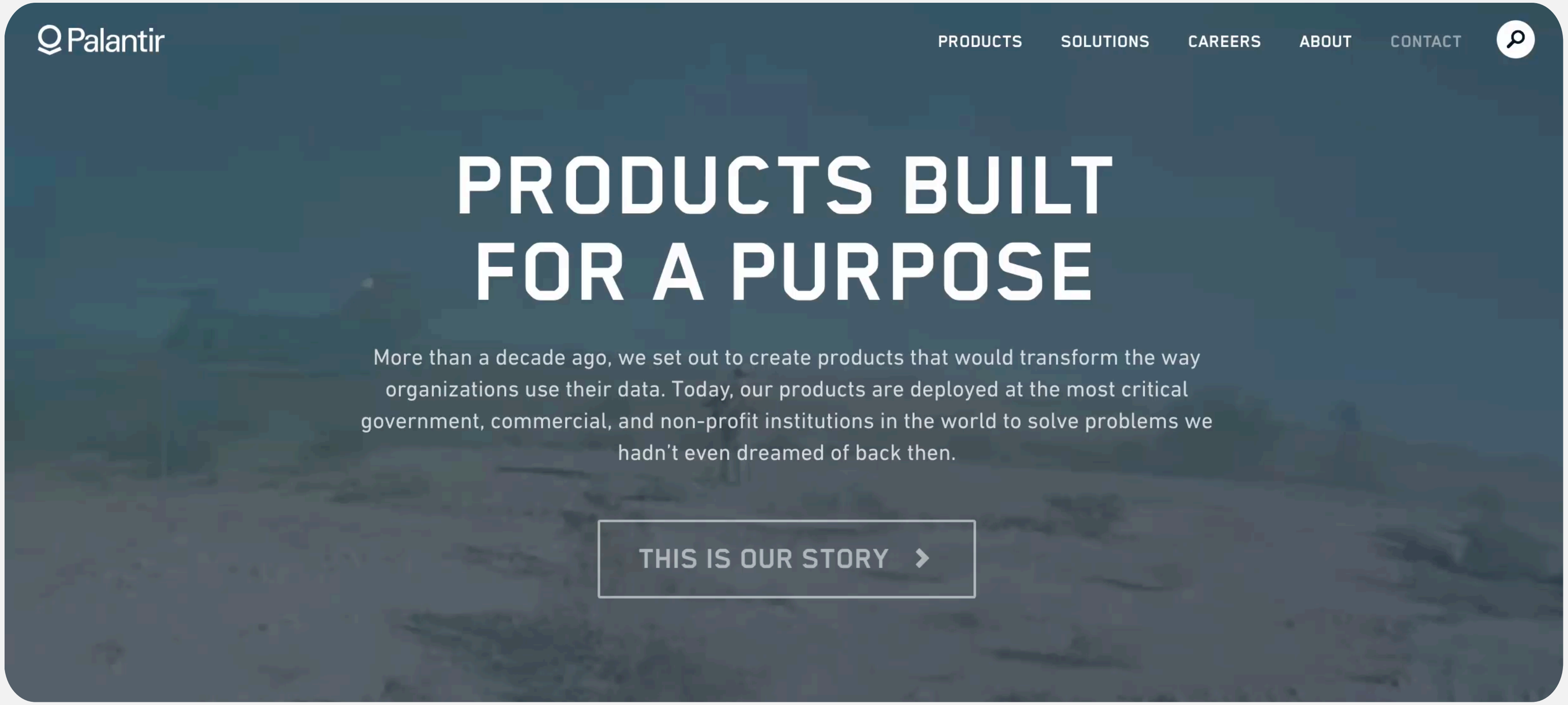
02 – Vlad
Ungureanu
Tech Lead, Cloud
infrastructure

Palantir

KubeCon

Problem Solving Software

Our products are deployed at the most critical government, commercial, and non-profit institutions in the world



PRODUCTS BUILT FOR A PURPOSE

More than a decade ago, we set out to create products that would transform the way organizations use their data. Today, our products are deployed at the most critical government, commercial, and non-profit institutions in the world to solve problems we hadn't even dreamed of back then.

Cyber Defense Software

Quietly powering the world’s most important organizations.

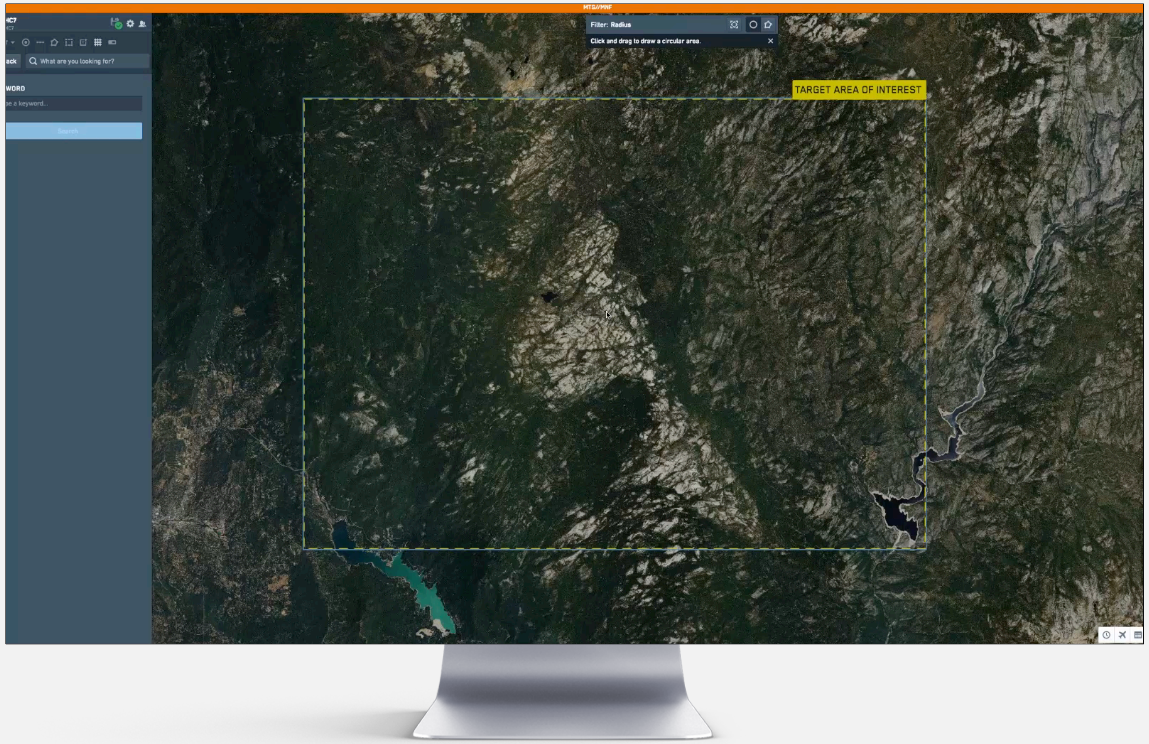
We secure entire states and integrate and analyze petabytes of data from every single IT system in every department in the country.

Supporting Anti–terrorism Missions

KubeCon

We help with everything from intelligence operations to mission planning in the field.

Palantir Gotham



Filter: Radius

Click and drag to draw a circular area.

Palantir

Bringing Technology to the Paddock

Ferrari uses Palantir Foundry to increase performance + reliability.

Palantir Foundry

Fast + Smart

Bringing Technology to the Paddock

KubeCon

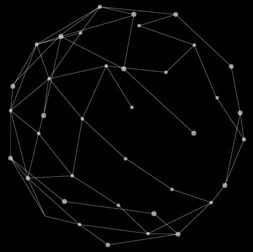
Ferrari uses Palantir Foundry to increase performance + reliability.

Palantir Foundry

Fast + Smart

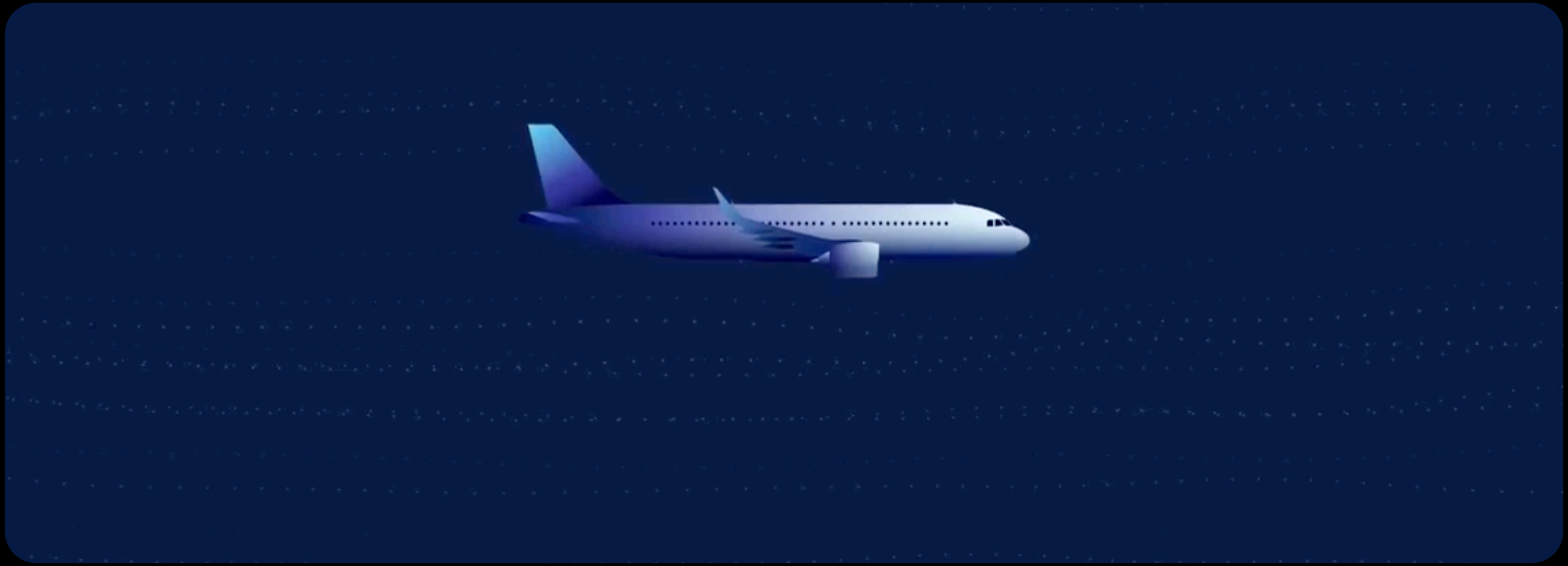
Disrupt the Aviation Industry

Palantir and Airbus founded Skywise to help make air travel safer and more economical.



skywise.

Airbus engineers use this data to make operations more efficient.



Cancer Research

Syntropy brings together the greatest minds and institutions to advance research toward the common goal of improving human lives.

Syntropy

Products Built for a Purpose

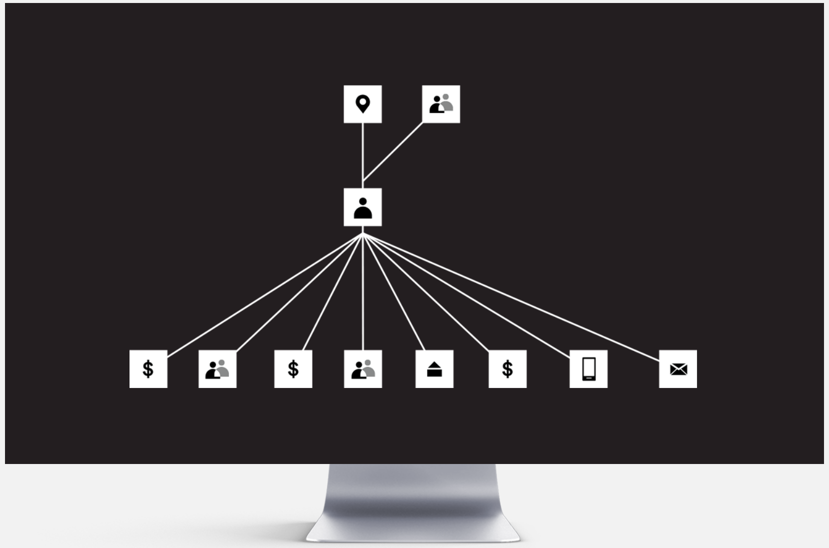
Solving the hardest data problems of the world's most important organizations.

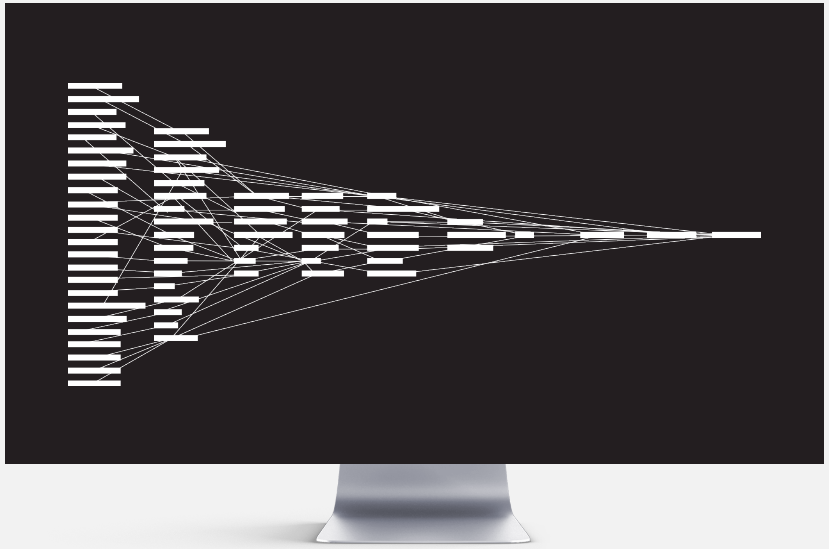
Palantir Gotham

Integrate, manage, secure, and analyze all of your enterprise data.

Palantir Foundry

Amplify and extend the power of data integration.





Products Built for a Purpose

Solving the hardest data problems of the world's most important organizations.

Palantir Gotham

Integrate, manage, secure, and analyze all of your enterprise data.

Palantir Foundry

Amplify and extend the power of data integration.

PALANTIR GOTHAM

Integrate, manage, secure, and analyze all of your enterprise data.

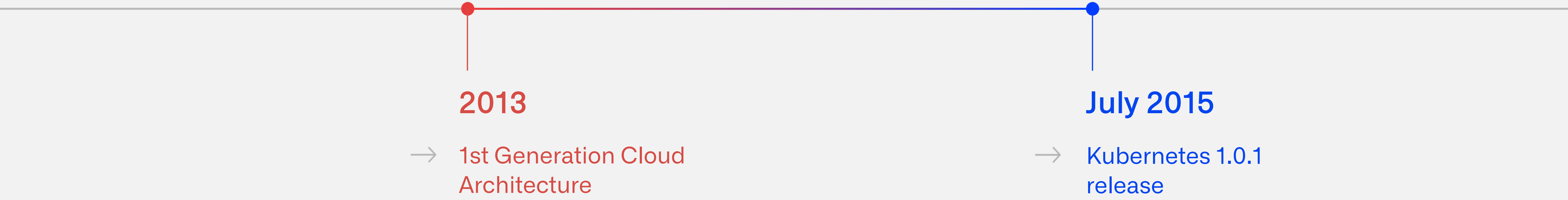
Join us. Visit palantir.com/careers

We build game-changing software.

We're solving real-world problems.

Palantir in the Cloud

Designing our cloud architecture for the future.



2013

→ 1st Generation Cloud Architecture

July 2015

→ Kubernetes 1.0.1 release

Palantir in the Cloud

Why Rubix?

- Scale challenges with patching infrastructure in **first gen architecture**
- Forced reboots across fleet due to **CVEs in AWS**
- Chaos engineering influences from **Netflix et al**

Ephemeral Infrastructure?

Infrastructure had to be both immutable and ephemeral.

What if we destroyed and rebuilt every host across every environment, every 48 hours?

December 2016

→ Kubernetes 1.5 release

2017

→ Rubix is Introduced

Rubix Today

Kubernetes is used as the scheduling and execution engine for all of our distributed compute frameworks.

→ Operating Environments /

9 AWS Regions &
4 Continents

→ K8s Nodes /

2500+

→ Largest Clusters /

800+ nodes

→ Destroy and Rebuild /

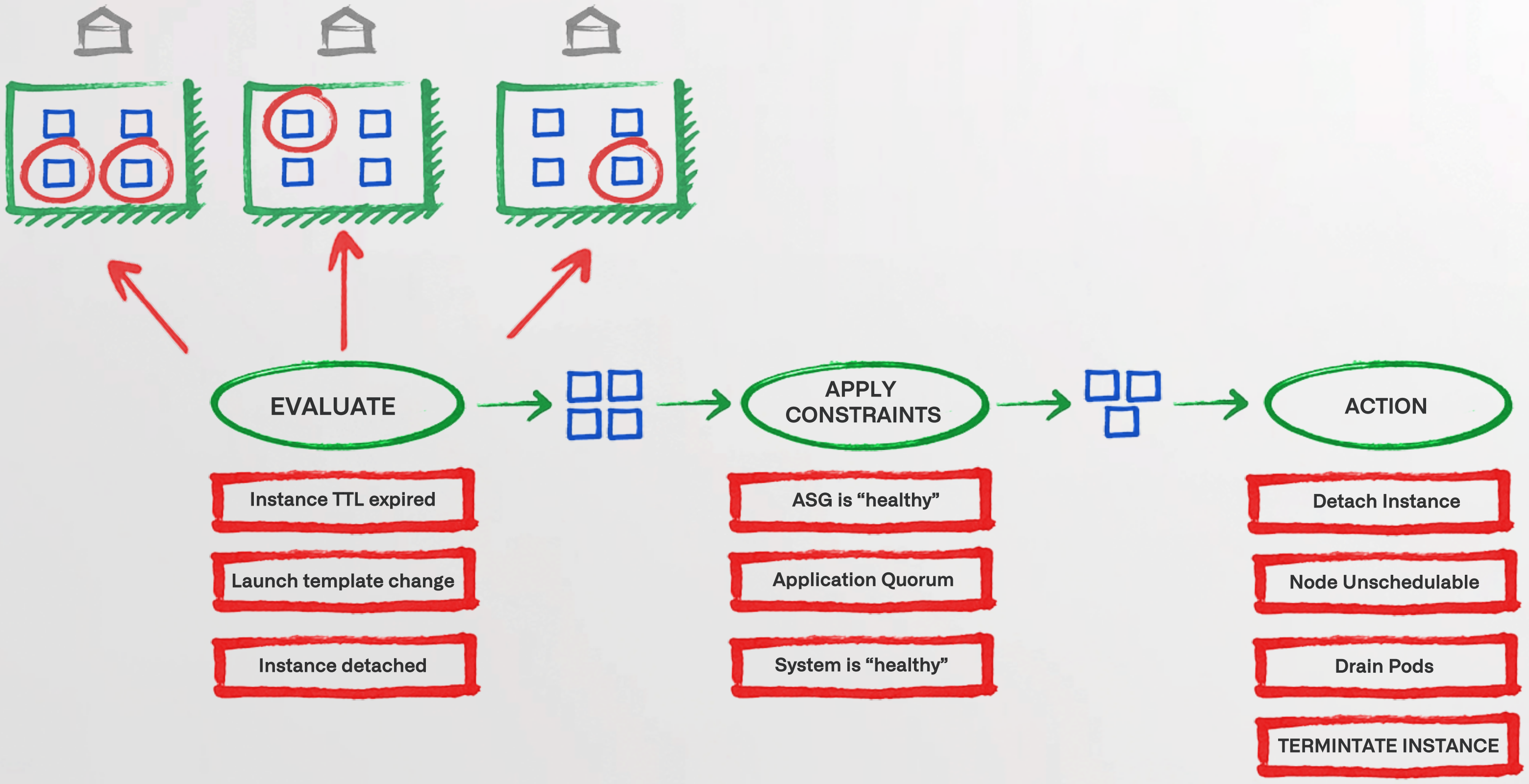
10,000+ Ec2 instances per day

→ Pods Run /

1,000,000 per day

Termination Automation

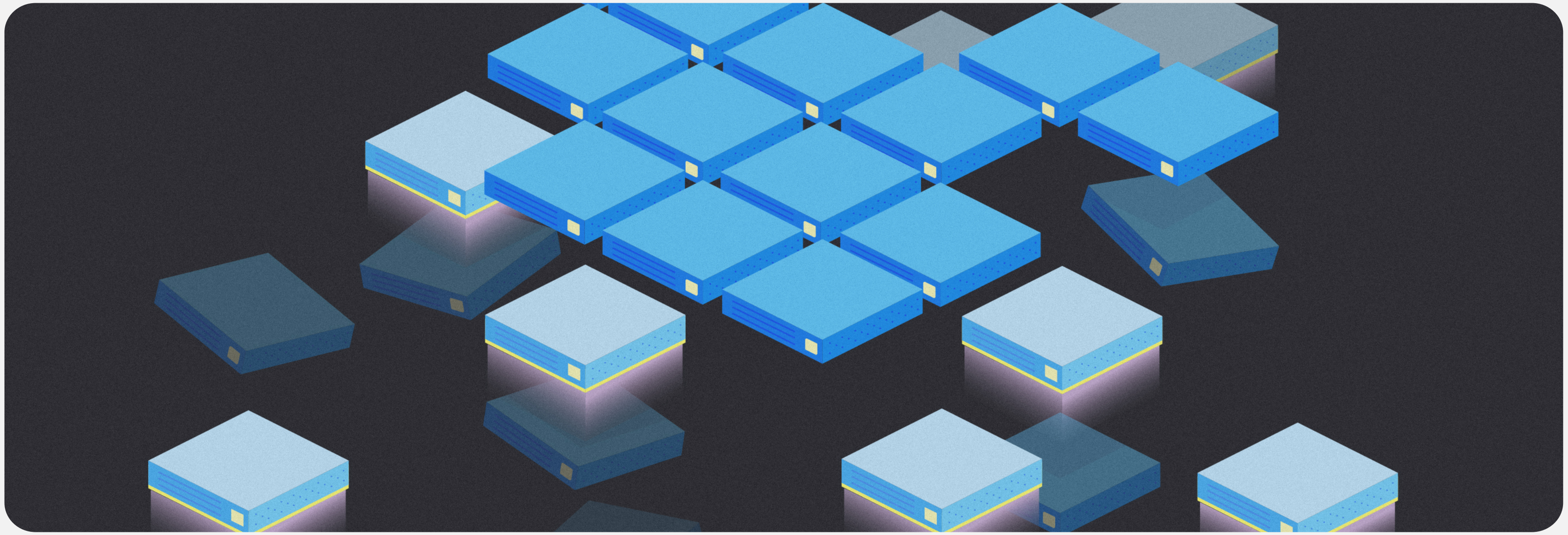
Welcome to–
Termination 101



Benefits

Ephemeral Infrastructure
has its advantages.

- **CVE mitigation**, upgrades “free”
- **Autoscaling easier** / required solving all the same problems
- Security posture substantially improved

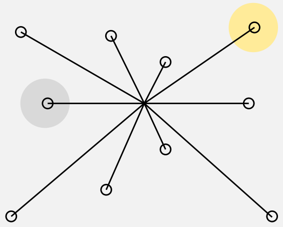


Vlad Ungureanu speaks on the challenges associated with this architecture.

Challenges

Consensus / Leader-based Services

Handling services that require consensus or leadership election

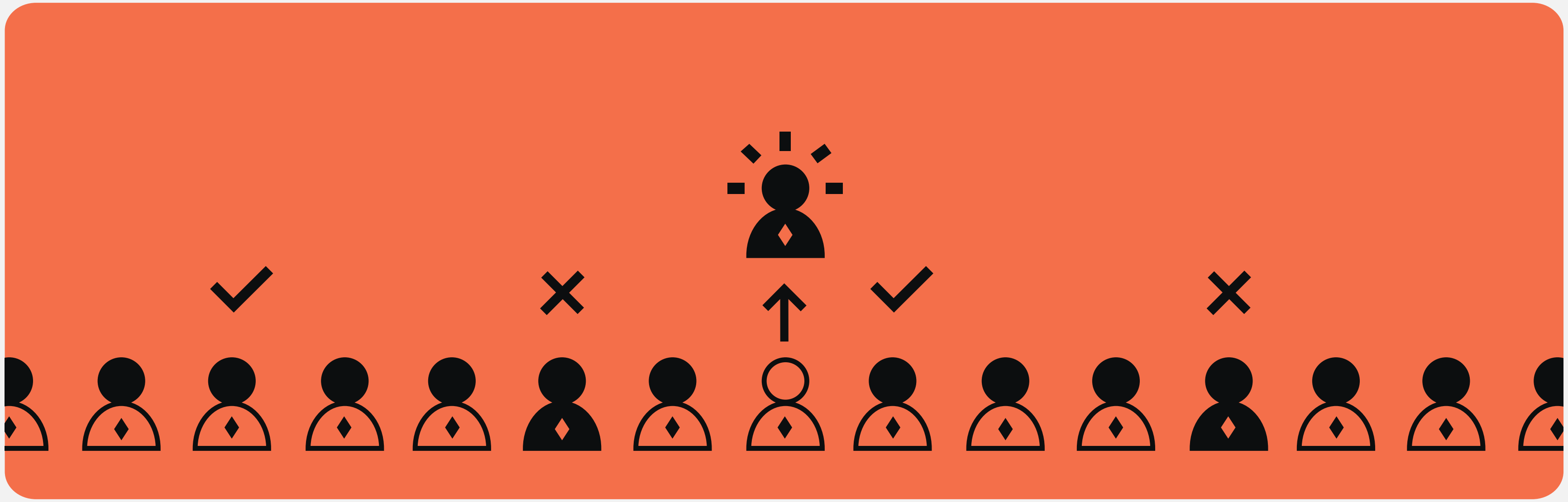


→ etcd

→ Vault

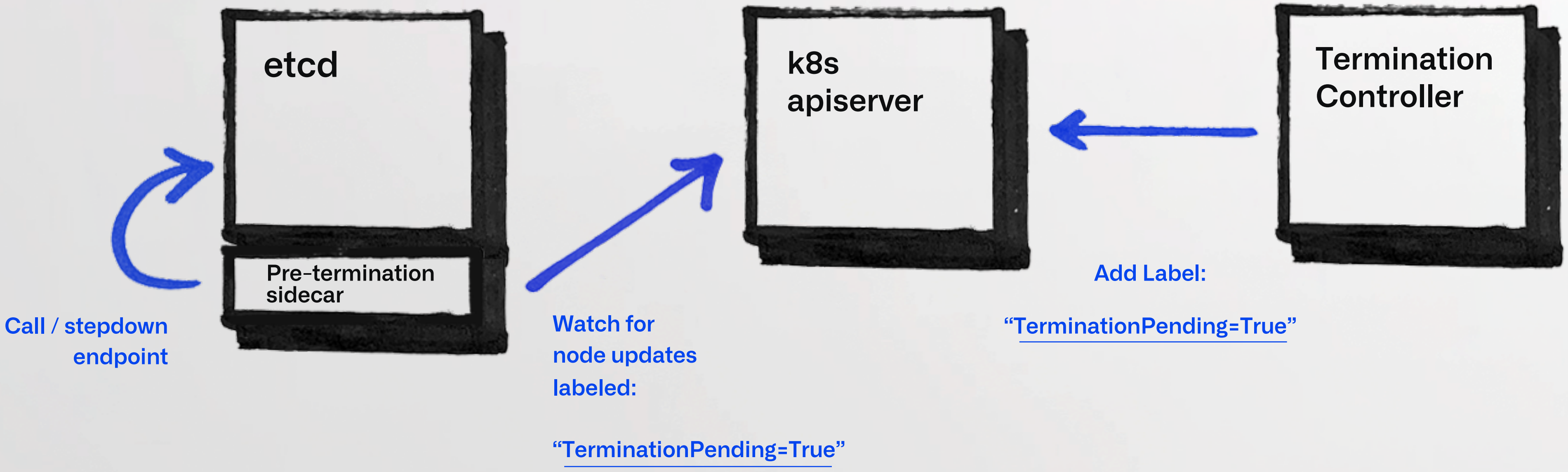
→ k8s-controller–manager

→ k8s-scheduler



Pre–Termination Actions Framework

We developed a framework that allows an action to be executed before the host is terminated.



Triggering Leadership Election

KubeCon

We call the move leader endpoint before terminating the host

etcd

→ move leader API endpoint

HashiCorp Vault

→ step down API endpoint

THINGS THAT JUST WORKED:

Controller-manager ✓

Scheduler ✓

The Cloud: Not Actually Infinite

We found the edge
of the cloud.

→ Sometimes AWS doesn't have enough capacity

→ Multi-instance type ASGs allow you to hedge

The Cloud: Hardware Failures are Real

Everyone knows hardware failures happen in the cloud

→ Touching **10k hosts per day,**
they happen a lot more frequently.

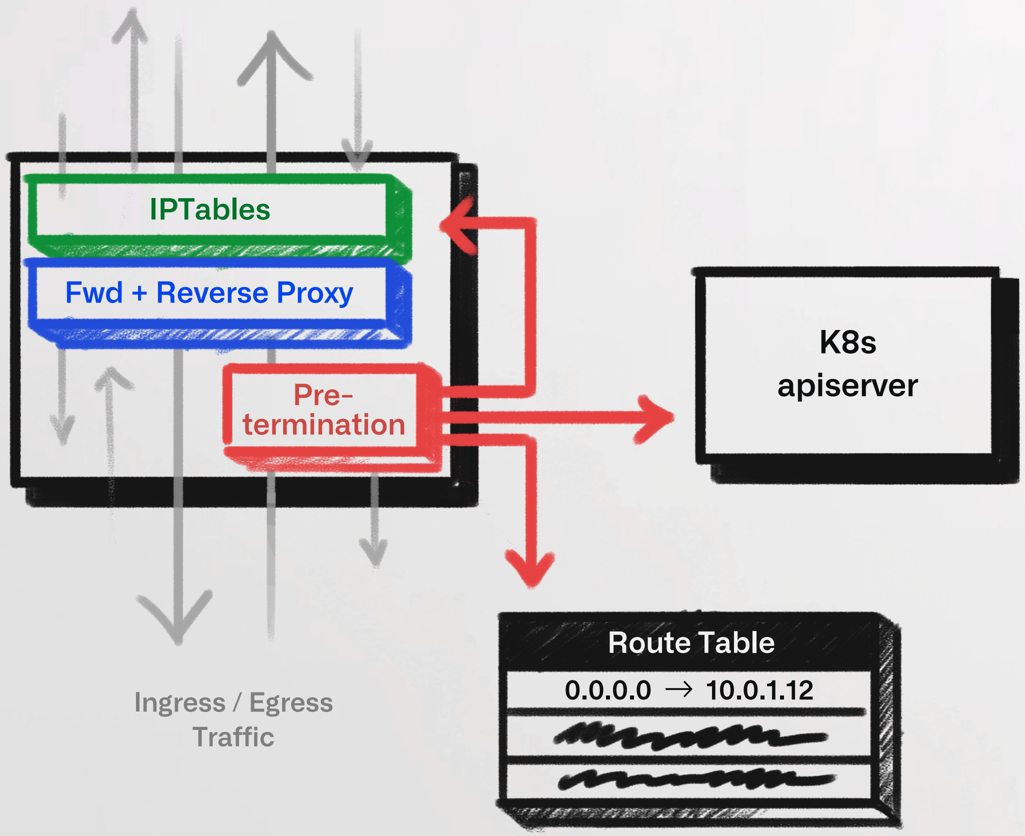
Host Initialization Matters

We use Vault to generate a PKI keypair at machine boot.

- The time from **Scale Up** ⟶ **Running Containers** matters
- **Images Pulls** are expensive in both time and bandwidth
- **Identity Management** causes load spikes

Ingress / Egress Networking

- — Active TCP connections need reset
- — RouteTable entries need updated (Egress)
- — New incoming connections need to be rejected



Container Networking

Solving challenges in the k8s network area

Calico → Lyft | cni-ipvlan-vpc-k8s → Cilium

Calico (Past)

→ Lack of Pod IP garbage collection in ~v2.5

→ BGP route syncing latency increased with cluster size

Lyft | cni-ipvlan-vpc-k8s (Present)

→ Large scale ups cause AWS API rate limiting

→ Pre-allocate IPs on ENIs

Cilium (Future)

Questions? – please stop
by our booth ☺

Thank You.