Dave Cuza

Ingeniero de Software de Infraestructura | Sistemas Distribuidos, Kubernetes, Python
Toronto, Ontario, Canadá | Ciudadano español, con permiso de trabajo en la UE y en Canadá
dave@cuza.dev | cuza.dev | linkedin.com/in/dave-cuza

Perfil profesional

Ingeniero de software de infraestructura con unos 8 años administrando Kubernetes en producción, desde 2017. En el equipo de Compute Infrastructure de Yelp desarrollo y mantengo los sistemas distribuidos en Python que despliegan y ejecutan todos los servicios de Yelp: PaaSTA, la plataforma como servicio open source de Yelp, y Tron, su scheduler de trabajos batch, también open source. Administro la flota donde corren, unos 40 000 pods de producción en miles de nodos sobre Amazon EKS. Lideré la migración que distribuyó las cargas de trabajo de PaaSTA en varios namespaces, para lo cual implementé rolling updates entre namespaces, algo que Kubernetes no ofrece de forma nativa. Trabajé en el service mesh interno de Yelp, basado en SmartStack y Envoy, que se encarga del enrutamiento de peticiones y del balanceo de carga. También trabajé en la migración de Kubernetes self-managed a EKS y en el cambio del autoescalado de nodos de Clusterman a Karpenter. Hago guardias on-call para la flota y actualmente trabajo en integrar PaaSTA con herramientas de IA.

Habilidades técnicas

Lenguajes y desarrollo: Python, Bash, Java, C/C++, C#/.NET, PHP, JavaScript, código de reconciliación para Kubernetes, herramientas de plataforma y de línea de comandos.
Sistemas distribuidos y orquestación: Kubernetes, Amazon EKS, Karpenter, service mesh (Envoy, SmartStack), descubrimiento de servicios, enrutamiento de peticiones, balanceo de carga, scheduling, autoescalado, instancias spot y on-demand, consolidación y bin packing, aprovisionamiento y ciclo de vida de nodos, aislamiento de cargas de trabajo, rolling deployments, scheduling de trabajos batch, OPA Gatekeeper, CoreDNS, VPC CNI, Helm, Docker.
Nube e infraestructura como código: AWS (EKS, EC2, VPC, IAM), GCP, Azure, Terraform, Flux, GitOps, centros de datos on-premise.
Confiabilidad y observabilidad: SLOs, alertas, guardias on-call, respuesta a incidentes, depuración en producción, gestión de capacidad, conocimientos prácticos de Prometheus, Grafana, Alertmanager y PromQL.
Pipelines de despliegue: Jenkins, GitLab CI, GitHub, pipelines de CI/CD, automatización de despliegues.

Experiencia profesional

Yelp Inc. | Site Reliability Engineer, Compute Infrastructure
Toronto, Ontario, Canadá | Mayo 2021 a la actualidad

- Mantengo y mejoro PaaSTA, la plataforma como servicio open source de Yelp, escrita en Python, que despliega y ejecuta todos los servicios de la empresa. Esto incluye el código de reconciliación que convierte la configuración de cada servicio en recursos de Kubernetes y los mantiene sincronizados, de modo que los equipos despliegan sin escribir YAML de Kubernetes.
- Trabajé en el service mesh interno de Yelp, tanto en su versión con SmartStack como en la actual con Envoy. El mesh se encarga del descubrimiento de servicios, el enrutamiento de peticiones y el balanceo de carga entre los servicios de Yelp. Mi trabajo incluyó el código de PaaSTA que registra los servicios en el mesh y la resolución de incidentes de enrutamiento, descubrimiento y balanceo de carga durante las guardias.
- Trabajé en la migración de las cargas de trabajo de Yelp desde clústeres de Kubernetes self-managed, con su propio control plane y etcd, hacia Amazon EKS. PaaSTA movió los servicios stateless a EKS con un cambio de configuración.
- Trabajé en el cambio del autoescalado de nodos de Clusterman, el autoescalador interno de Yelp, a Karpenter, y ajusté la consolidación, el bin packing y la selección de instancias para las cargas de Yelp. Karpenter reacciona en segundos a los pods pendientes, mientras que Clusterman revisaba el clúster una vez por minuto.
- Lideré la migración que sacó las cargas de trabajo de PaaSTA de un único namespace compartido y las distribuyó en varios namespaces de Kubernetes, con lo que quedaron aisladas entre sí. Implementé rolling updates entre namespaces, algo que Kubernetes no soporta de forma nativa, para mover servicios en ejecución de un namespace a otro.
- Mantengo Tron, el scheduler de trabajos batch open source que Yelp usa en producción. Agrego funcionalidades y corrijo problemas de confiabilidad y de operación.
- Administro todos los clústeres de Kubernetes de Yelp sobre Amazon EKS, unos 40 000 pods de producción en miles de nodos, además de los clústeres de desarrollo y staging. En la flota corren los servicios stateless de Yelp, trabajos batch de Spark y Tron, pipelines de Flink y cargas de machine learning con GPU, sobre instancias spot y on-demand. También corren Kafka y unos 90 clústeres de Cassandra de producción, con casi un petabyte de datos.
- Creo y elimino clústeres, actualizo las versiones de EKS en toda la flota de producción y administro add-ons compartidos como CoreDNS, VPC CNI, kube-proxy y Metrics Server.

- Defino los estándares de confiabilidad que siguen los equipos de servicio para PodDisruptionBudgets, probes, y requests y límites de recursos, y los hago cumplir con políticas de OPA Gatekeeper.
- Defino los SLOs y configuro las alertas de la plataforma de cómputo.
- Hago guardias on-call para la flota de EKS, Karpenter, los add-ons compartidos, PaaSTA y Tron. Resuelvo fallos de Kubernetes, del ciclo de vida de los nodos, de scheduling, de capacidad, de red y de trabajos batch. Estandaricé los runbooks de los incidentes recurrentes.
- Administro la infraestructura de EKS con Terraform, despliego la configuración de los clústeres con Flux y trabajo con las redes VPC e IAM de AWS.
- Trabajo en integrar PaaSTA con herramientas de IA.
- Creé un entorno local de PaaSTA donde los ingenieros pueden ejecutar, probar y depurar PaaSTA en su propia máquina.
- Soy mentor de otros ingenieros del equipo de Compute Infrastructure, y escribo y reviso propuestas de diseño.

Inteloom Incorporated | Ingeniero DevOps
Gatineau / Ottawa, Canadá | Agosto 2019 a mayo 2021

- Lideré la migración de GOOMTool, de Docker Compose en un solo servidor a Kubernetes administrado en DigitalOcean.
- Monté y administré los clústeres de Kubernetes de alta disponibilidad, los servidores VMware ESXi y los sistemas Linux donde corría producción en el centro de datos on-premise, en un equipo de infraestructura de dos personas.
- Creé y mantuve los pipelines de CI/CD y la automatización de la infraestructura.
- Diseñé e implementé la solución de respaldo del centro de datos.

The Furniture Guys | Ingeniero Full Stack
Toronto, Ontario, Canadá | Septiembre 2018 a agosto 2019

- Fui el único ingeniero de la empresa, a cargo del desarrollo backend y frontend, QA, operaciones y soporte técnico.
- Desarrollé una API en Python y Django para la optimización de materiales.
- Administré la infraestructura en AWS, con servicios en Docker sobre EC2 y pipelines de pruebas y despliegue en AWS CodePipeline y AWS CodeDeploy.

Universidad de Matanzas | Líder Técnico
Matanzas, Cuba | Septiembre 2016 a julio 2018

- Lideré el producto desde la planificación hasta la entrega.
- Administré Kubernetes self-managed en producción durante cerca de un año, con pipelines de CI/CD, contenedores Docker y un GitLab instalado en nuestros propios servidores.

Universidad de Matanzas | Desarrollador Web y Administrador de Sistemas
Matanzas, Cuba | Septiembre 2013 a junio 2016

- Desarrollé aplicaciones web con Symfony y ASP.NET MVC, y administré servidores VMware ESXi y sistemas Linux.

Educación

Universidad de Matanzas Camilo Cienfuegos
Ingeniería Informática | Matanzas, Cuba | 2011 a 2016

Información adicional

Idiomas: español e inglés, ambos a nivel nativo.
ACM-ICPC: participante en la Final Regional del Caribe en representación de la Universidad de Matanzas, lugar 20 en 2011.