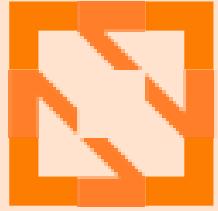




KubeCon



CloudNativeCon

Europe 2022

WELCOME TO VALENCIA



Service Mesh at Scale: How Xbox Cloud Gaming Secures 22k Pods with Linkerd

Chris Voss
Senior Software Development Engineer @ Microsoft





Chris Voss
Microsoft (Xbox Game Streaming)

- Senior Software Development Engineer at Microsoft
- Joined Xbox Cloud Gaming in 2018
- Our team manages and operates the infrastructure for Xbox Cloud Gaming services
- Contacts:
 - <https://www.linkedin.com/in/christopher-voss-9916111/>
 - <https://twitter.com/hatfarmer>





Xbox Cloud Gaming (Beta) with Xbox Game Pass

Play 300+ high-quality console games on the devices you already have
with Xbox Game Pass Ultimate.





PLAY THE GAMES YOU WANT, WITH THE
PEOPLE YOU WANT, ON THE DEVICES
YOU ALREADY OWN





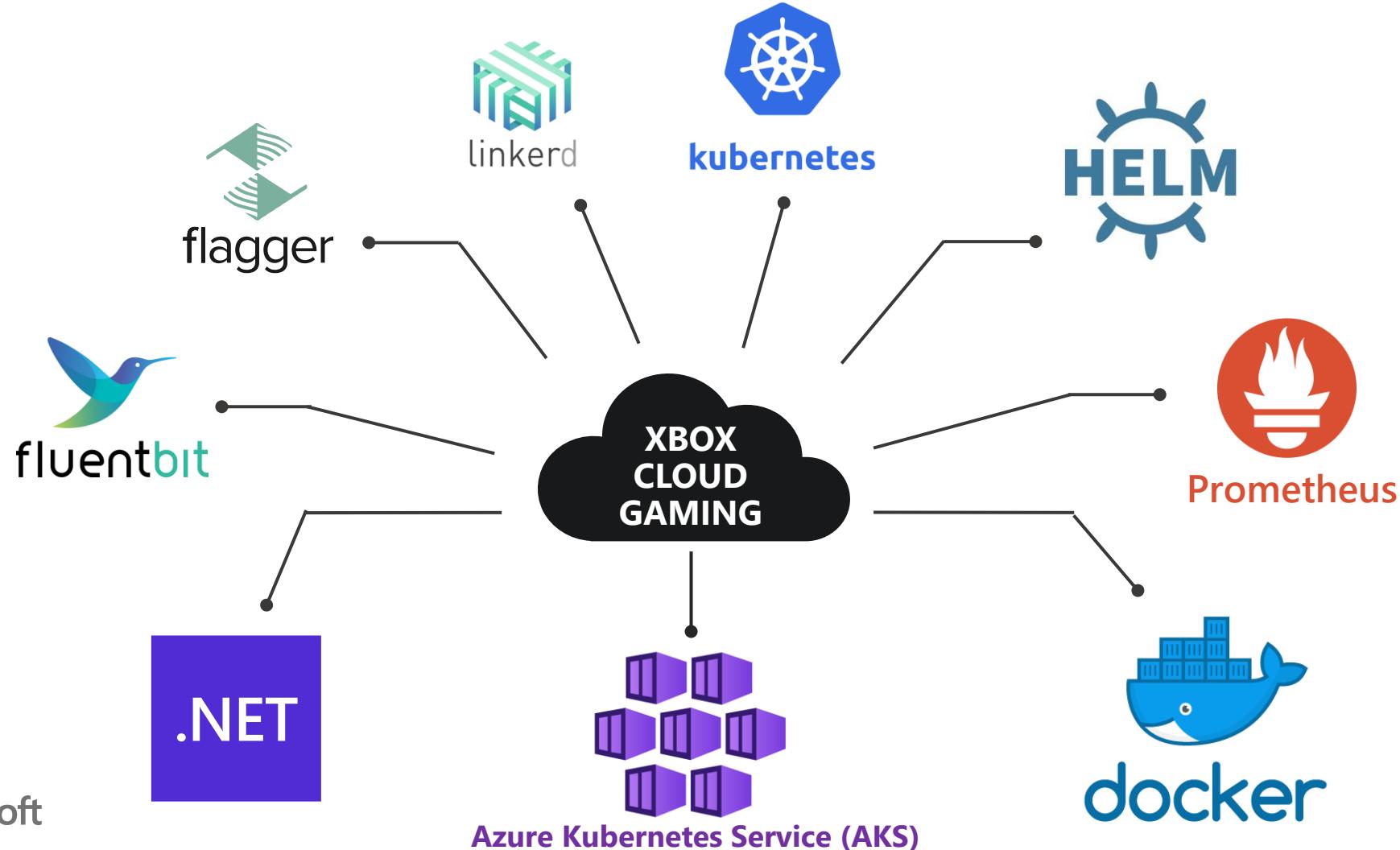
What Our Footprint Looks Like

- 26+ AKS clusters across several Azure regions
- 50+ Microservices
- 700 – 1000 pods per cluster
- More than 22k pods around the world
- Servers based on custom Xbox Series X hardware deployed in data centers around the globe





Our Tech Stack





What is a Service Mesh?

- Provide controls around traffic inside of your clusters
 - Sometimes out of
- The Service Mesh Interface defines specs around:
 - Traffic Access Control
 - Traffic Metrics
 - Traffic Specs
 - Traffic Split
 - <https://smi-spec.io/>





Why a Service Mesh?

- Simplify our service-to-service TLS (mTLS) story
 - Previously, we managed our own solution.
- Code flighting and Progressive Deployments
- Observability
- Code-free instrumentation





The Service Mesh Search Begins

Important to understand what you are wanting to get out of the service mesh.

Started in 2020

We explored a few different options

Our Requirements

- Implements the Service Mesh Interface (SMI)
 - Enables Canary deployment
- Efficient resource utilization (especially CPU)
- Low latency impact
- Observability
- Setup/maintenance ease

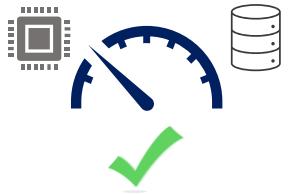




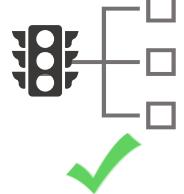
Selecting and Prototyping Linkerd

Requirements for the proof of concept

Efficient Resource Utilization



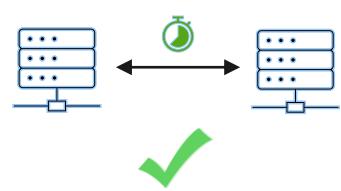
Traffic Split



Observability



Low latency



SMI compliant





Scaling our business with Linkerd

- mTLS (with cert rotation)
- High Availability Mode
- Prometheus metrics
- Service Mesh Interface Extension
 - Canary Releases
- Work closely with Linkerd community
 - <https://github.com/linkerd/linkerd2/issues?q=is%3Aissue+author%3Ababirishges+is%3Aclosed>





mTLS

- Zero-config mTLS = simple
- Secured 50+ Microservices
- 22k pods across 26+ clusters running across the globe





Observability

- Code-free visibility
- HTTP response codes
- Latency monitoring
- Canary deployment status
- Request volume
- Comes with its own Prometheus

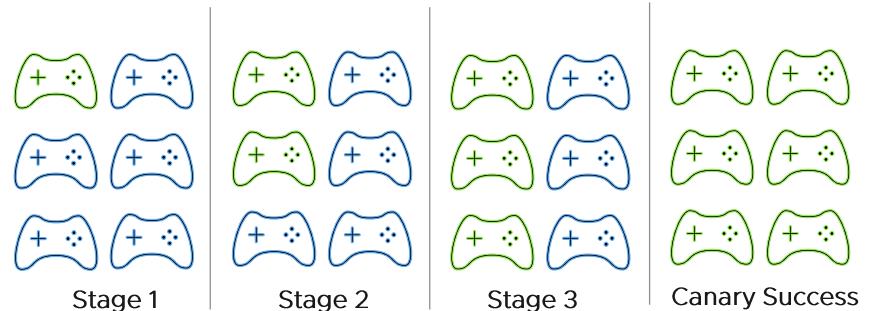




Canary Deployments

- Leveraging Flagger to progressively migrate traffic
- Integrated into Azure DevOps release pipelines for auto-rollback
- Flexible canary evaluation

Canary Deployment Stages





Key Learnings from Flagger

- How to handle canary deployments for different types of workloads
 - Non-http
 - Asynchronous
 - CronJob type workloads
- Not all services have constant or high traffic volumes
- Traffic to liveness / readiness endpoints skew canary evaluation





Engineering and Cost Savings

- Engineers freed from supporting in-house mTLS
- Reduced spend on alternative observability solutions

The above two things alone resulted in saving thousands of dollars per month.





Next Up in Our Linkerd Journey

- Service-to-service auth
- Multi-cluster communication and failover
- Fault injection, chaos testing





Abereham Wodajie
Microsoft (Xbox Game Streaming)

- Software Development Engineer at Microsoft
- Joined Xbox Cloud Gaming in 2018
- He was unable to make it to Valencia, but I wanted to make sure you all know who he is





Thanks for Joining!

Any questions?

Feel free to try Xbox Cloud Gaming right now!

xbox.com/play

