



Cloud to Edge: *Architecting for the Next Generation*



WESLEY REISZ

Cloud Native Architect /
Engineering Leader
Chairperson QCon SF &
Co-host of The InfoQ Podcast



Toyota ~2 million vehicles sold in North America alone

Toyota predicts that the car-to-cloud data stream will reach 10 exabytes per month by 2025.

WESLEY REISZ

Cloud Native Architect /
Engineering Leader
Chairperson QCon SF &
Co-host of The InfoQ Podcast

wes@wesleyreisz.com

@wesreisz



AGENDA

DEFINE EDGE

Edge overloaded term

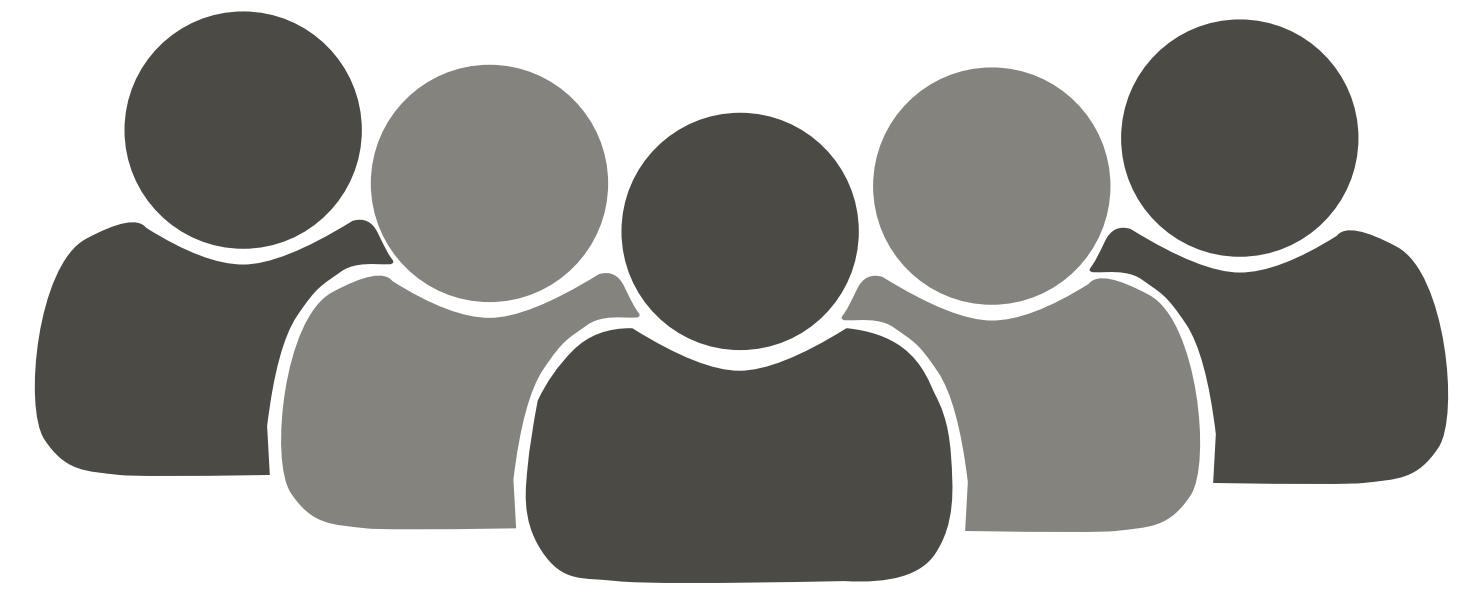
EDGE-NATIVE APPLICATION ARCHITECTURE

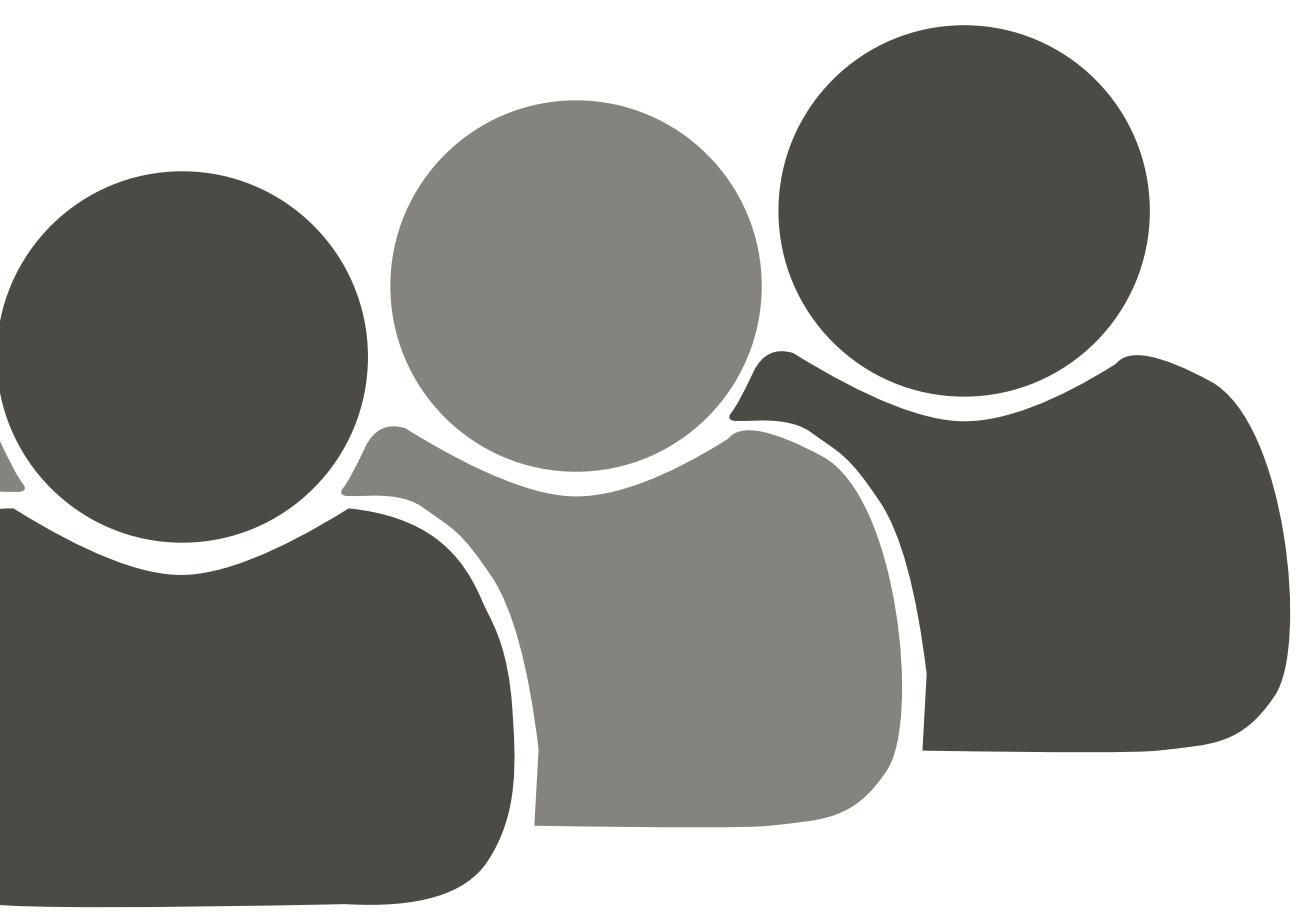
Production Lessons from the Edge

TL;DR

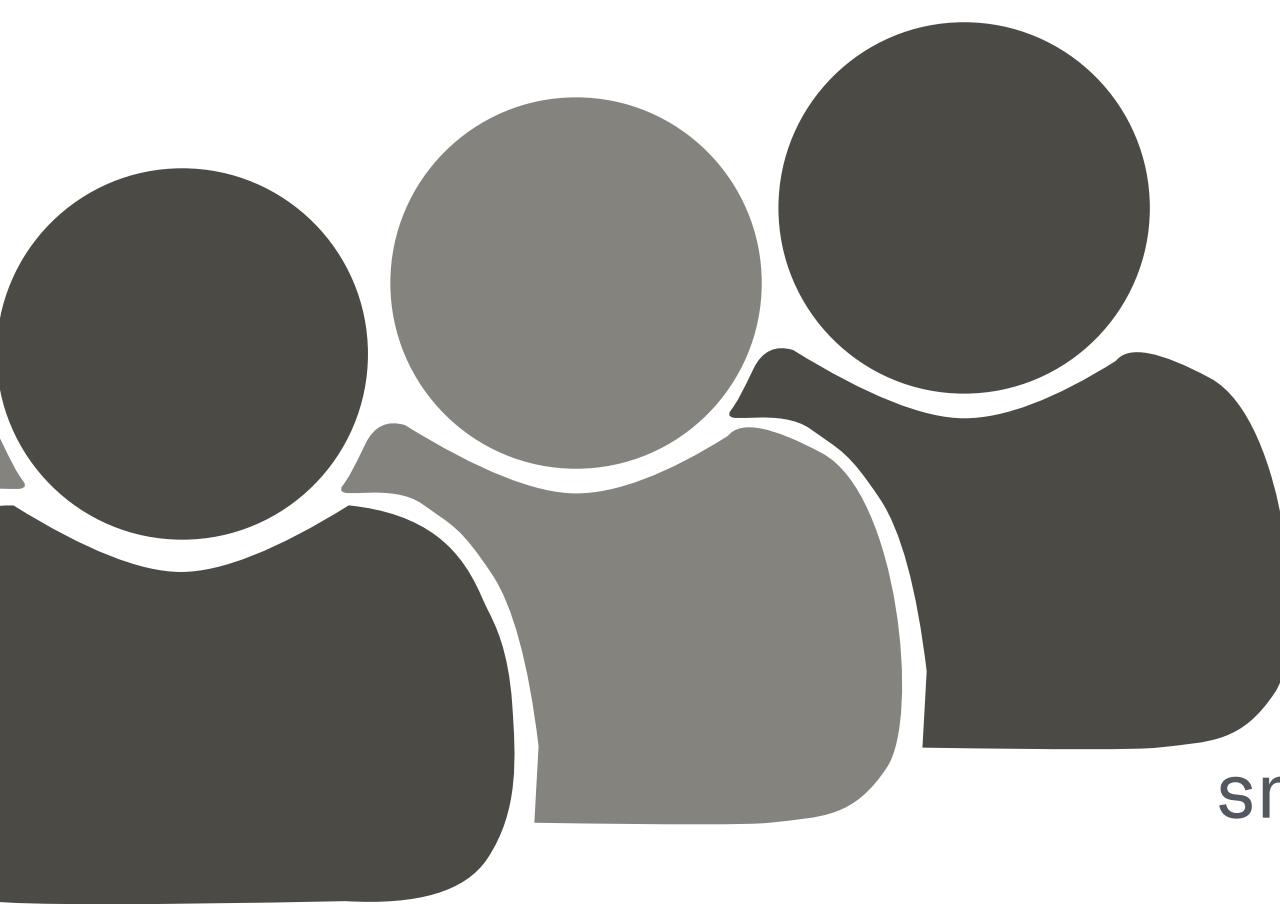
Modern applications leverage Cloud Native, next generation architectures will open completely new use cases leveraging *Cloud Native* and *Edge Native Architectures* (across the infrastructure and device edge).

DEFINING EDGE





DEVICE EDGE



smart-phones, wearables and automobiles.

Gateway devices such as IoT aggregators, switching and routing devices.

On-premise server platforms.

LAST MILE

INFRASTRUCTURE EDGE

Regional Data Centers / CDNs

Network access equipment, such as cellular radio base stations and xDSL access sites.

Switching equipment and other service gateway functionality



Edge is running workloads between the user and the cloud. It can be divided into two parts (using the the *last mile* as a boundary). **Device edge** is closest to the user and **Infrastructure edge** closest the cloud.

EDGE-NATIVE APPLICATION ARCHITECTURE

WHY

Cost: Intelligent routing and optimization of compute

Performance Cost Optimization: Reduce latency / Improve performance

Legislation: Regulations & Compliance (GDPR)

New Use Cases: Such as the ability to influence the energy market through the aggregation of consumer devices.

Examples: Inferencing at the Edge, Reduce backhaul, Federated Learning, Enhanced CDN, Better Routing

What Can I Do at the Edge

Cloudflare:

Service Worker API implementation for the Cloudflare platform. Brings a server less style approach to running JavaScript workloads on their Points of Presence.

Source: <https://www.infoq.com/presentations/cloudflare-workers/>

Facebook Live:

Stream is sent via RTMP (Real-Time Messaging Protocol) to a geographically local PoP. The connection is forwarded over an internal Facebook network to a Facebook data-centre. When you see a live stream in your feed and you click on it the player requests the manifest. If it isn't already on your local PoP the request goes to the data centre to get the manifest, and then fetches the media files in 1 sec clips. As they get sent back they are cached on the PoP if they aren't there already.

Source: <https://www.infoq.com/podcasts/sachin-kulkarni-facebook-live/>

Chick fil-A:

We think of our Edge Computing environment as a “micro private cloud”. By this, we mean that we provide developers with a series of helpful services and a place to deploy their applications on our infrastructure.

Source: <https://medium.com/@cfatechblog/edge-computing-at-chick-fil-a-7d67242675e2>

Tesla:

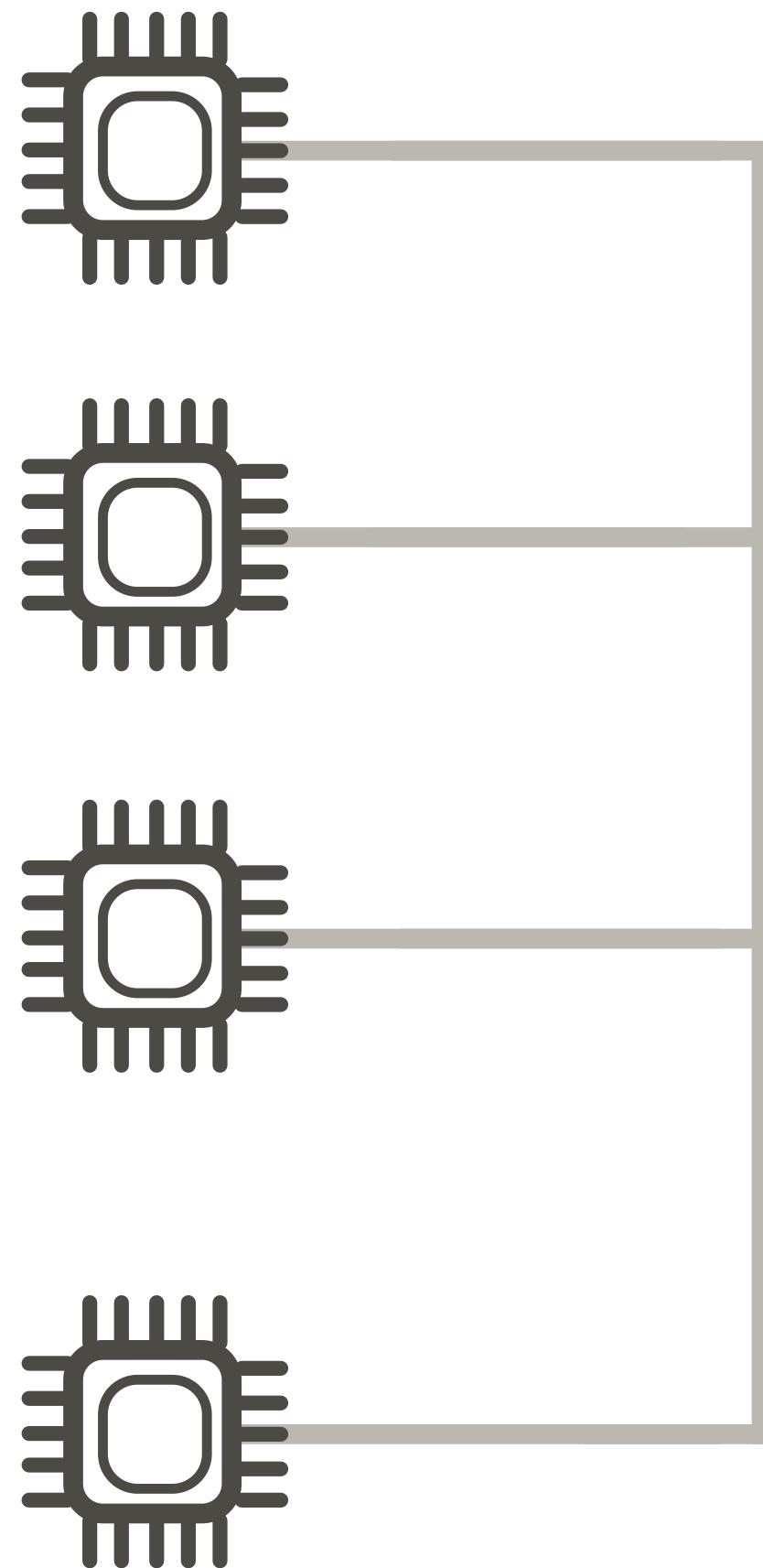
A Virtual Power Plant (VPP) is a network of distributed energy-resources (often solar, wind, and batteries) that are aggregated to provide smarter and more flexible power generation, distribution, and availability. Tesla's VPP consists of vertically integrated hardware and software, including both cloud and edge computing.

Source: <https://www.infoq.com/news/2020/03/tesla-vpp>

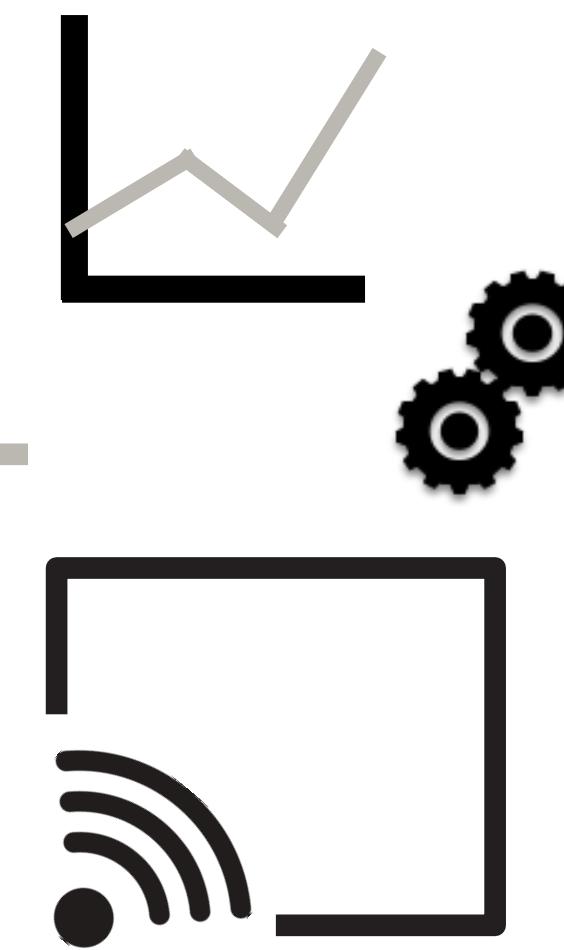


Source: <https://www.teslarati.com/tesla-energy-sun-home-wheels/>

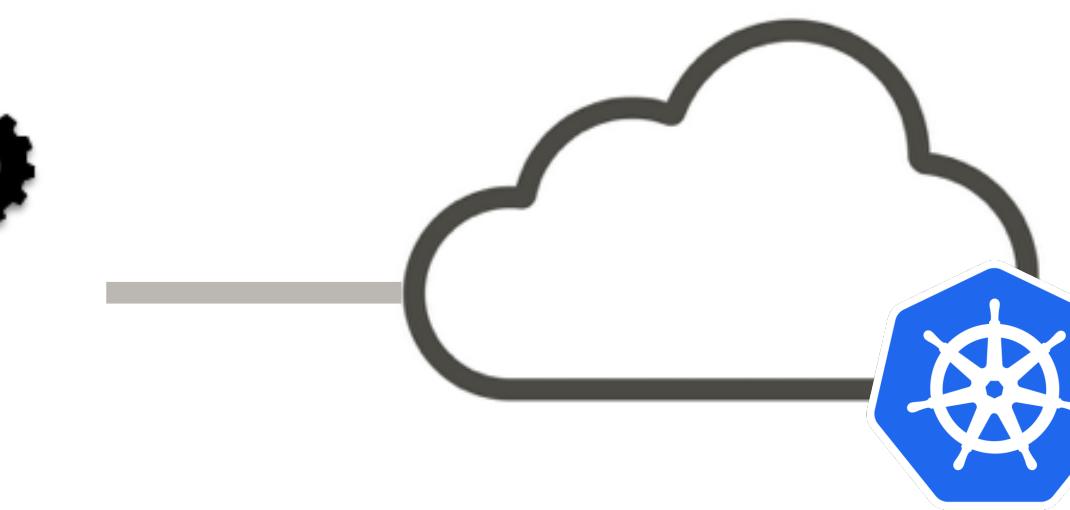
IOT/ SENSORS



EDGE



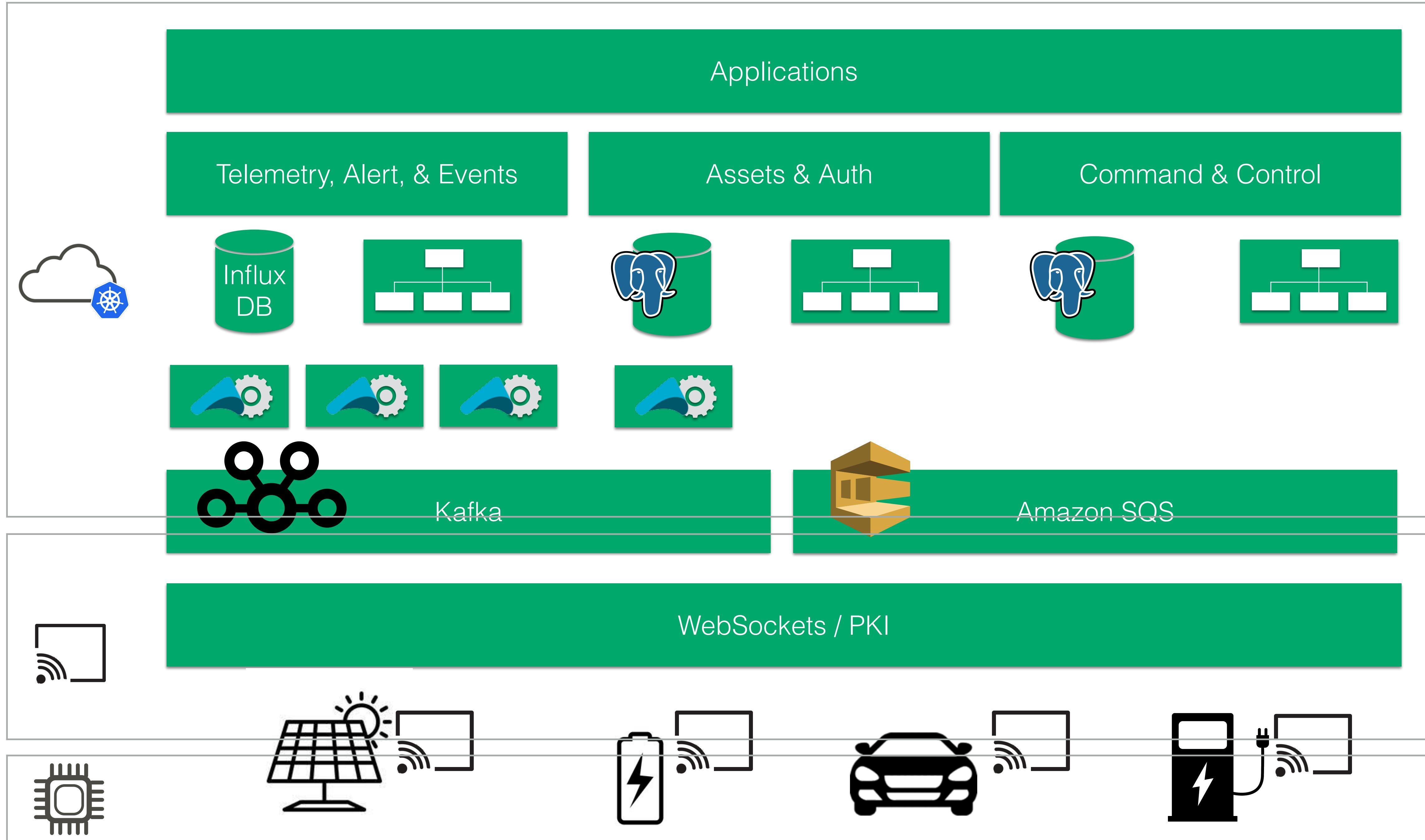
CENTRALIZED CLOUD



Lower Latency / Less Compute

High Latency / More Compute





Tesla is able to aggregate and make bids into the energy market (on behalf of consumers) every 5 seconds.

KEY TAKEAWAYS

- **Edge** is running workloads between the user and the cloud. It can be divided into two parts (using the the **last mile** as a boundary). **Device edge** is closest to the user and **Infrastructure edge** closest the cloud.
- Edge works **reduce network costs**, **improve speed/latency**, and **enables new use cases** around *machine learning*, *performance*, *compliance*, & *cost optimization*.
- It's time to **rethink** how we deploy workloads. No longer is it just a defacto Cloud deployment.



AGENDA

DEFINE EDGE

Edge overloaded term

EDGE-NATIVE APPLICATION ARCHITECTURE

Actual Stories from the Edge

QUESTIONS

WESLEY REISZ

Cloud Native Architect /
Engineering Leader
Chairperson QCon SF &
Co-host of The InfoQ Podcast

wes@wesleyreisz.com

@wesreisz



Executive Team Presentation

Technical Outcomes – CxO Speak



15 Minutes

Presentation Guidelines

- ❖ Candidate selects relevant topic – hiring manager + candidate agrees beforehand
- ❖ Candidate presents why?, what?, how? in non-technical terms to executive team
- ❖ Candidate can use any combination of slides, whiteboard, demo
- ❖ First 5 minutes: audience will not interrupt the session. Last 10 minutes: interviewer(s) ask why should I care? questions