

# Kubernetes at the Edge

How Roche is on the way to manage 1000+ clusters in Hospitals & Labs

Alexander Hungenberg  
Tech Lead Edge Infrastructure, Roche

13.06.2024

# I'm Alex - nice to meet you!



32 years old

Excited to be back in one of my favourite cities!

Lived here for a long time, but now moved to the also beautiful canton of Aargau a month ago

Working in Kaiseraugst (next to Basel) as tech lead of the edge infrastructure team in Roche

# Table of contents

1. Who is Roche?
2. What is the Edge & why is it important?
3. Why is it hard?
4. How are we doing it?
5. (Random Deep Dive: Secure Boot)

# Roche at a glance

Who we are and what we do



**128 years**

founded in Basel in 1896



**A leader in  
healthcare R&D**

with CHF 13.2 billion invested  
in 2023



**3 Nobel prizes** and  
**44 Prix Galien**

since 1974



**CHF 58.7 billion\***

in Roche Group sales in 2023



**45 Roche medicines  
& 90 diagnostics\*\***

on the WHO List of  
Essential Medicines & Tests



**>100,000**

dedicated employees  
worldwide



**>22 million people**

treated with our medicines  
in 2023



**29 billion tests**

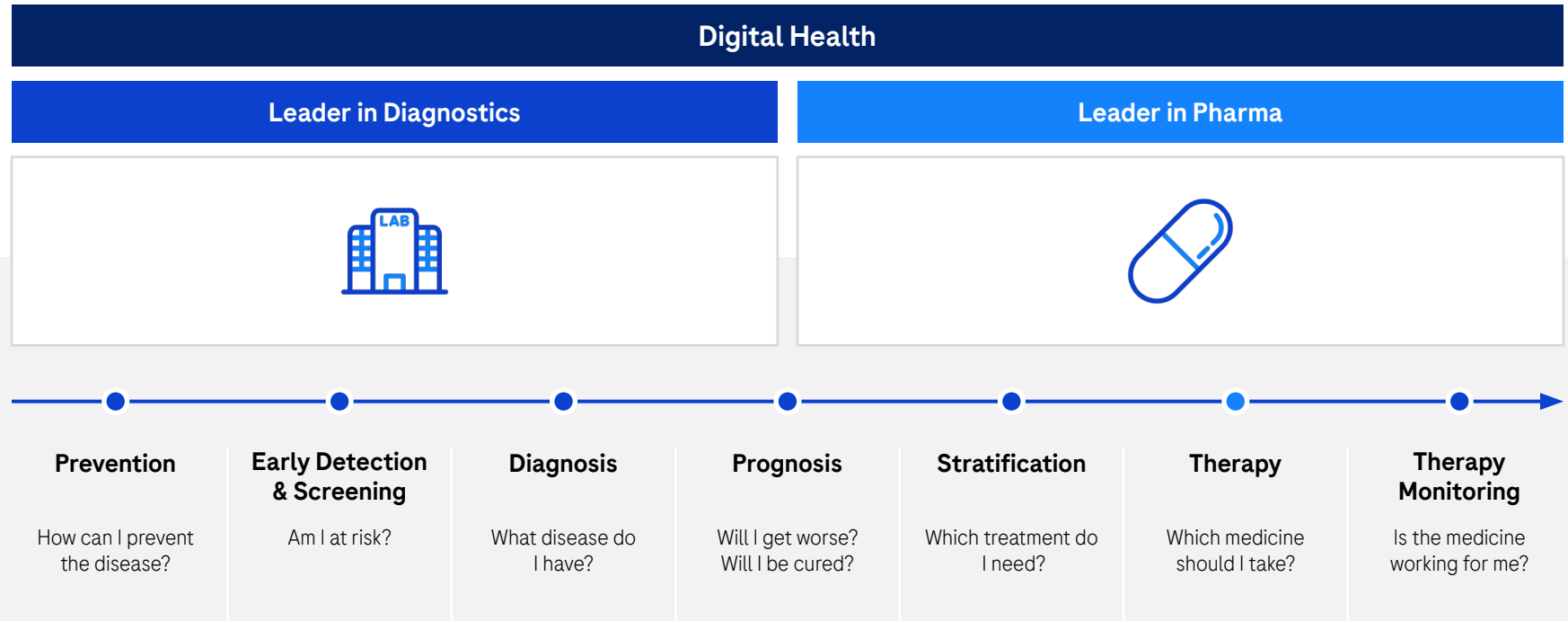
conducted with our  
Diagnostics products in 2023

\*Unless otherwise stated, all growth rates and comparisons to the previous year are at constant exchange rates (CER; average rates 2022) and all total figures quoted are reported in CHF.

\*\* Medicines and tests that have either been developed or acquired by Roche

# The Major Divisions

Helping patients through the continuum of healthcare

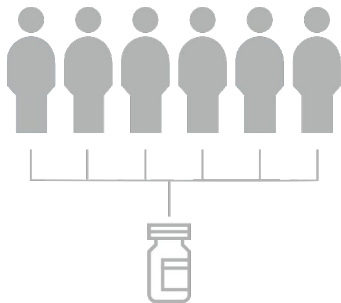


# Why Software Matters in Healthcare?

Personalized Healthcare

Past

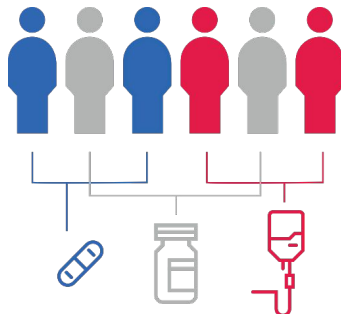
“One size fits all” drugs



Present

Companion  
diagnostics

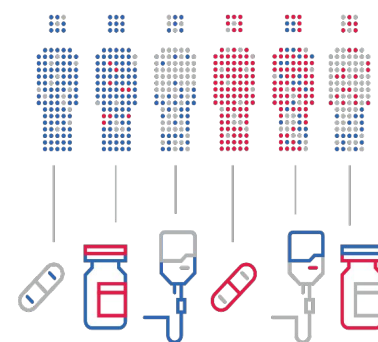
Targeted medicines



Future

Meaningful data  
and advanced  
analytics

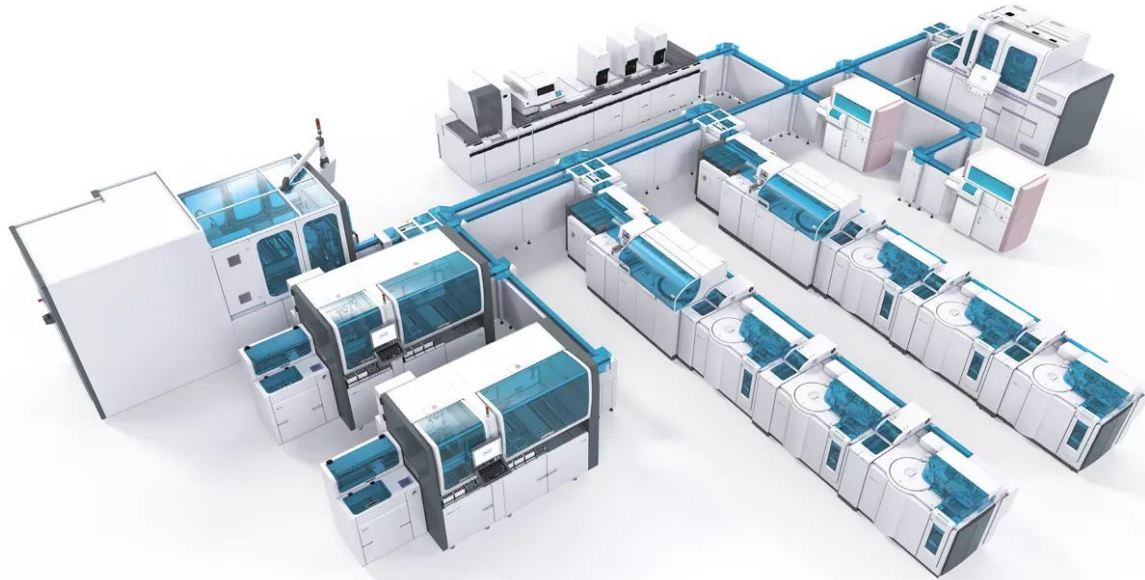
Individualised treatments



Smarter, more efficient R&D ; Improved access  
and truly personalised care

**What is the Edge?**

# What is a Lab?





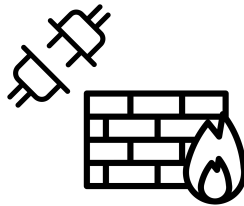
# What is a Lab?



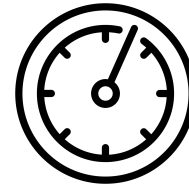
*Edge computing is a [...] model that **brings computation and data storage closer to the sources of data***



Privacy



Connectivity



Bandwidth & Latency

<sup>1</sup> source: [https://en.wikipedia.org/wiki/Edge\\_computing](https://en.wikipedia.org/wiki/Edge_computing)

**Why is it hard?**

# How Edge Computing Works Today?

- you want to deploy a software?
  - build your own hardware and sell it as a device. Try to find a place to physically put the machine!
- you want to do updates?
  - either, by sending new hardware
  - or: Load it on a USB stick, get someone in a car to drive there and install it

This costs 

... and makes many interesting use cases for software simply unfeasible.

# Changing That is Hard, Because...

... a selection of reasons

- not our networks
- deployed worldwide and on-site. Staff needs to go on-site for installation, but they are not developers
- connectivity constraints:
  - Check out a great technical talk from two of my colleagues at the KubeCon Paris collocated event: [YouTube: Meshing it up Securely](#)
- cost constraints - small labs can't afford to pay tens of thousands of \$ for compute infrastructure
  
- we want to have remote operations, but still comply with all local laws & regulations. Which often require data to stay within a certain geographical / political region.

# Last But Not Least

## Deploying SaaS in the Cloud



dev

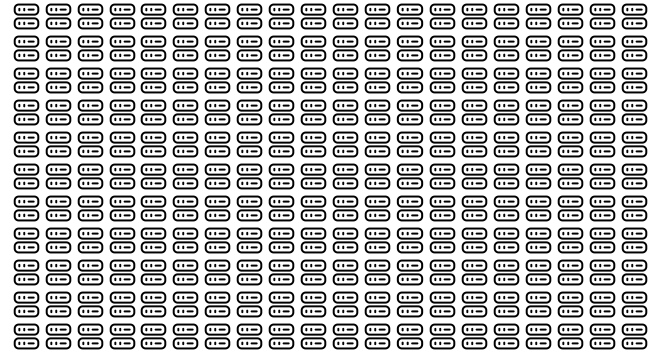


staging



prod

## Deploying to Edges



If you're curious how we took GitOps to the next level for that, check out our talk from GitOpsCon North America:

[YouTube: When GitOps Is Not Enough](#)

**How are we doing it?**

# How Does Roche Run Modern Software at the Edge?

## TLDR; we put Kubernetes clusters in labs!

we settled on Kubernetes as our software framework for running workloads on the edge

We treat it as the "new OS" which developers have to build applications for. It's no longer for Windows or for Linux. It's for Kubernetes!

... and as such it comes with major advantages:

- horizontal scalability on cheap hardware
- high-availability deployments possible with multi-nodes
- tremendous ecosystem of software around it
- widely used -> great documentation and resources available

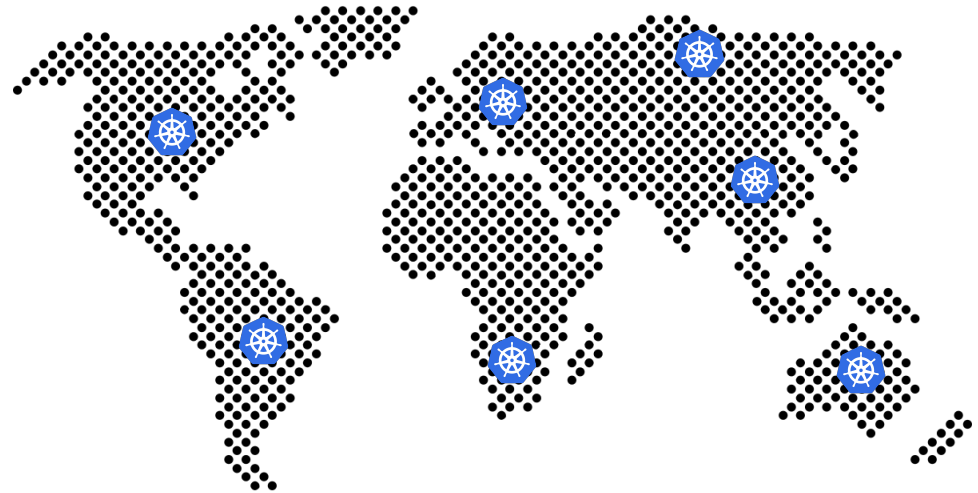




# How Are We Solving It?

... from an organizational perspective?

- We built an **internal platform** team to provide a compute platform that can run applications
- This team, in collaboration with others, is responsible to roll out infrastructure:  
... to **labs, hospitals, pharmacies, and doctor offices.**
- So that app developers don't have to worry about it - and they can just build Docker images



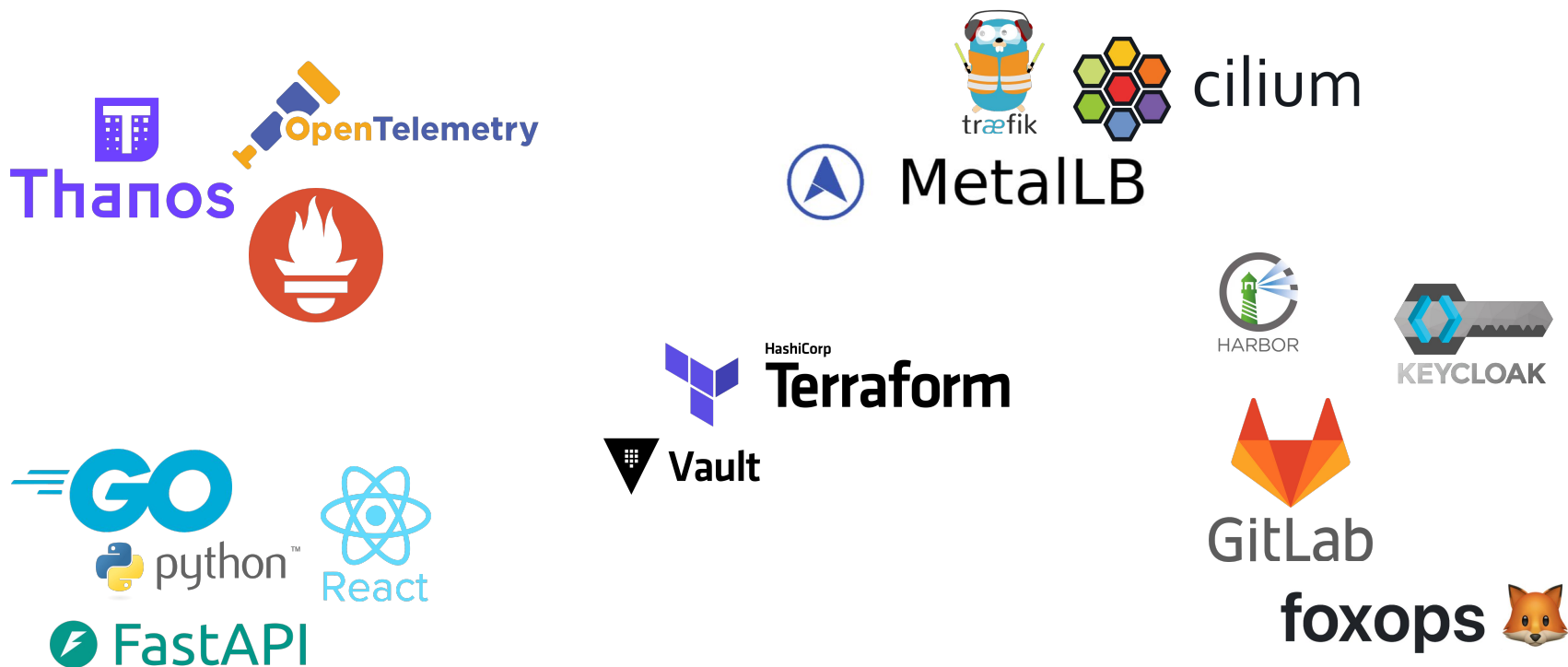
# Kubernetes Is Not Enough!

while everything revolves around K8s, there are a bunch of challenges that need to be solved around it:

- Which **hardware** to run on?
- Which **operating system** to run on?
- How to **keep** that **OS up-to-date** and prevent configuration drift over time?
- How to **bootstrap** a new deployment?
- Which **Kubernetes distribution** to use?
- Which **CNI** (network implementation) should we use?
- How can we **monitor** the cluster and workloads?
- How can we **connect remotely** for debugging problems?
- How can we **maintain an overview** of all clusters out there?
- ...

# Tech Stack

Built on Open Source Technologies, Giving back to the Community!



The image displays a collection of open-source technology logos arranged in a grid-like fashion. The logos include:

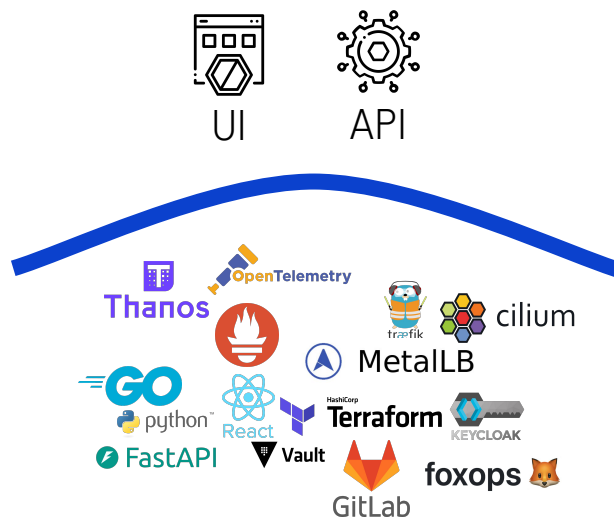
- Thanos**: A purple icon of a building with a grid pattern.
- OpenTelemetry**: A blue and yellow icon of a microscope.
- Prometheus**: A red circular icon with a white flame.
- Grafana**: A red circular icon with a white flame.
- Traefik**: A blue and white icon of a dog wearing a life vest.
- MetalLB**: A blue circular icon with a white triangle.
- Cilium**: A colorful hexagonal icon.
- Harbor**: A green and white icon of a lighthouse.
- Keycloak**: A grey icon of a key with a blue and white hexagonal head.
- Terraform**: A blue icon of three stacked blocks.
- HashiCorp**: A blue icon of three stacked blocks.
- Vault**: A black icon of a downward-pointing triangle with a white grid.
- Go**: A blue icon of the word "GO" with three horizontal lines to the left.
- Python**: A blue and yellow icon of the Python logo.
- React**: A blue icon of an atom.
- GitLab**: An orange and red icon of a fox head.
- Foxops**: A brown and orange icon of a fox head.

# The Art of Hiding Complexity

And while our platform stands on the shoulders of giants

... our main intention is to **fully hide** all (most of) those tools from end users - all while making it as easy as possible for development teams to build applications for it.

By building custom APIs and UIs which set up these clusters with an opinionated configuration



# The Operating System

We picked Talos as our operating system.

It's purpose-built for Kubernetes and comes with a few advantages:

- extremely hardened from security perspective (no shell, only a small API is exposed)
- immutability prevents configuration drift over time



## **Bonus Topic: Talos & Secure Boot**

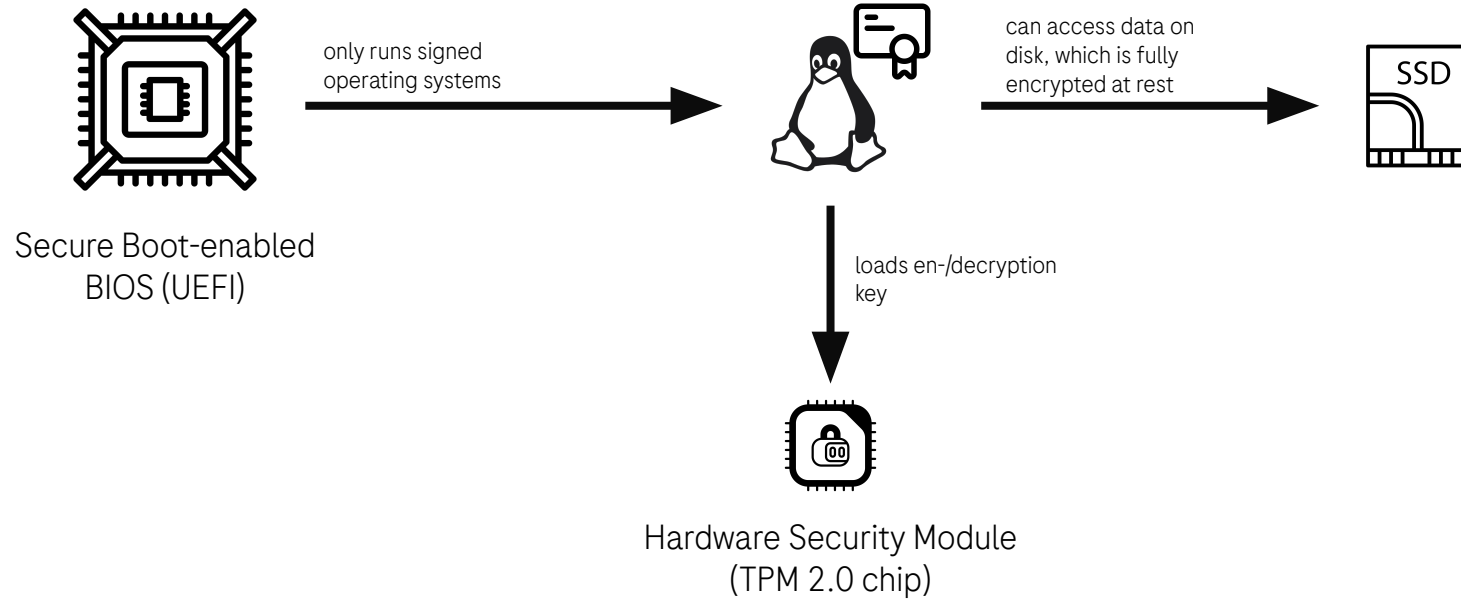
# Talos & Secure Boot

One major security issue we are facing is "physical access security".

Labs are not datacenters or airport security zones. Our edge clusters and applications host sensitive patient data. How can we protect that in the (unlikely) scenario where someone unplugs the machine and takes it home?

With the help of SideroLabs (the company maintaining Talos OSS), we contributed a feature to enable **Secure Boot & Full-Disk Encryption**

# What Is Secure Boot?







**The End**

# We are hiring!

near Barcelona, Spain!



Our team is located in Switzerland, Canada, Spain and other sites!

Currently we're looking for great **senior/principal software engineers** in St Cugat - close to Barcelona, Spain!\*

If you know someone over there who should work on one of the most interesting and impactful deep tech software projects in healthcare, please recommend them and reach out!

\* Posting not yet online, but will be soon 🕶️

**Doing now what patients need next**

**Q&A**