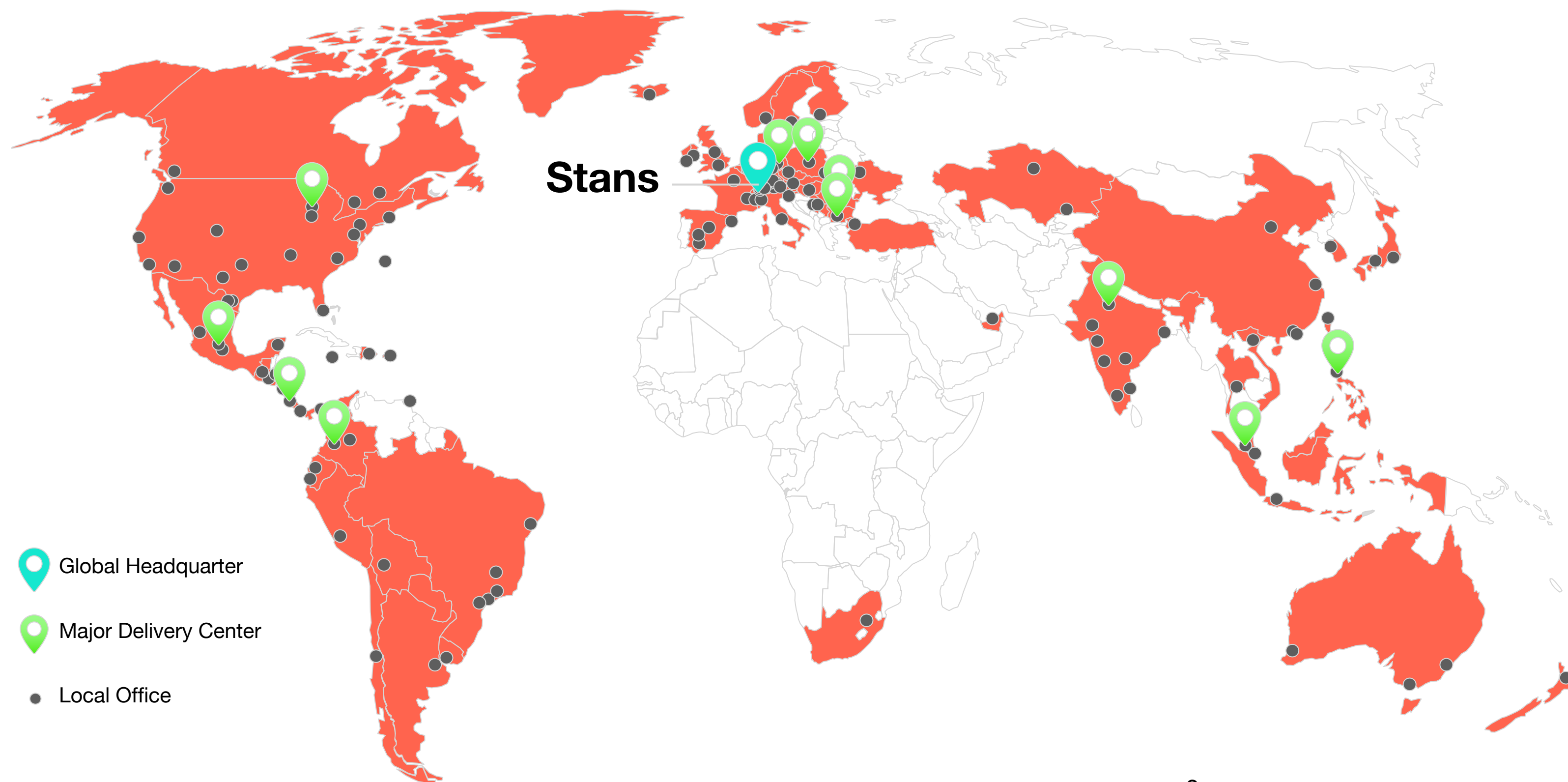


**Unlocking the value of technology**

**software** **ma**

# Accelerated digital transformation drives opportunity for the prepared

SoftwareOne professionals in [90](#) countries, backed by [30+](#) years of unmatched software and cloud solutions experience make us the ideal partner to help our clients.



- Deliver planning, migration, deployment and managed services for your applications and workloads.
- Procure and optimise software and the clouds it runs on.



---

1997

2015

2019

...

- Established in **1997** as a software solution provider
- In one of **Top 10** Turkey Software Integrators
- **Comparex** acquisition in 2015 & **SoftwareOne** acquisition in 2019
- **~75 Employees, ~50** Technical Resources, offices in **2 cities**

**all in **one** software.**



# AI-Ready Infrastructure Strategy & OpenShift Bare Metal Optimization

# Agenda

- Differences Between VM Server and Bare Metal
- Why Deploy OpenShift on Bare Metal
- Vakıf Katılım Solution Architecture
- NVIDIA MIG on OpenShift
- What We Learned & The Road Ahead

# Differences Between VM Server and Bare Metal



## Architectural Difference

In a VM-based structure, there is a layered architecture consisting of  
Hardware → Hypervisor → VM → OpenShift Node.

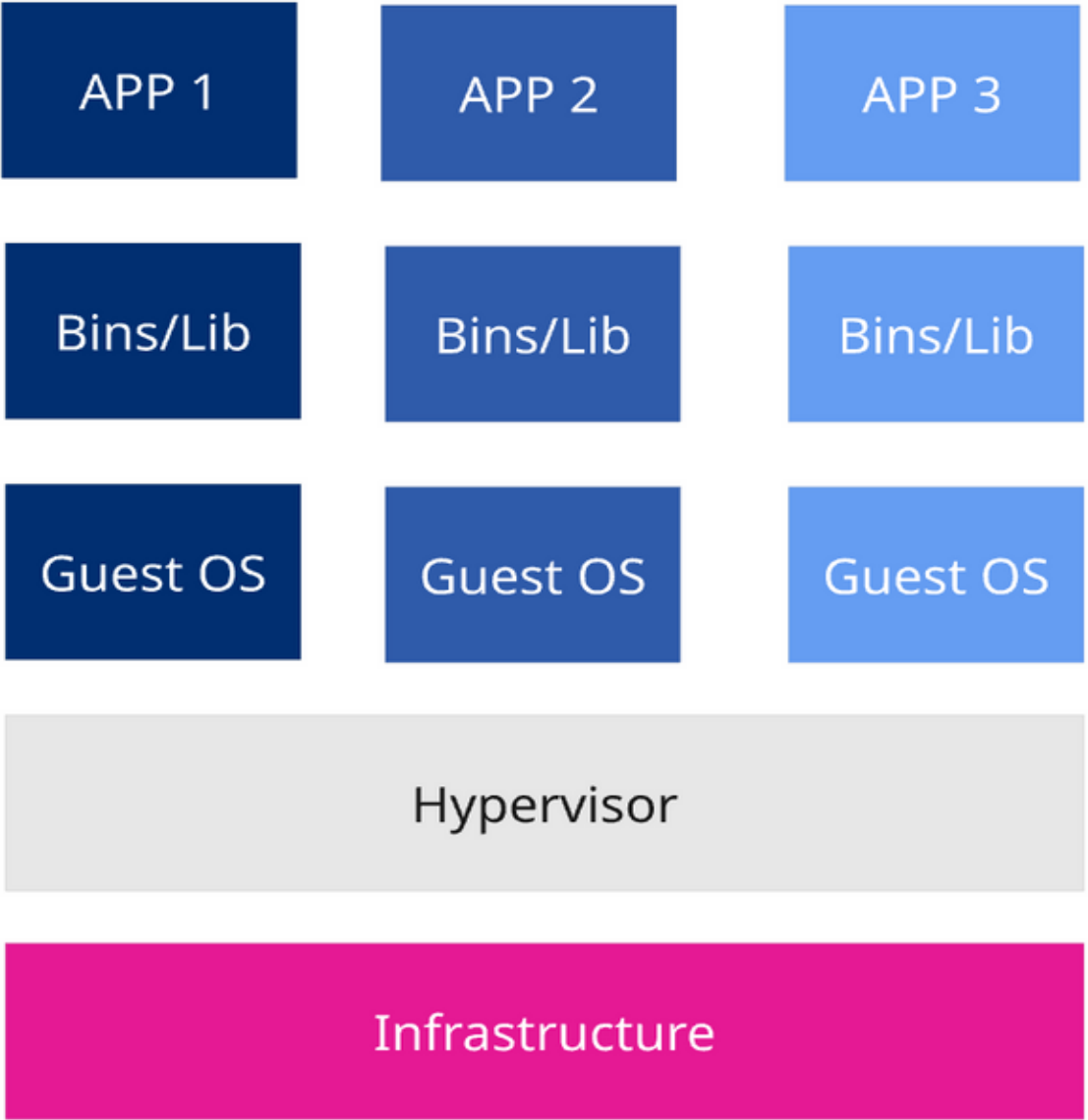
In bare metal, there is direct access from  
Hardware → OpenShift Node.

## Cost Difference

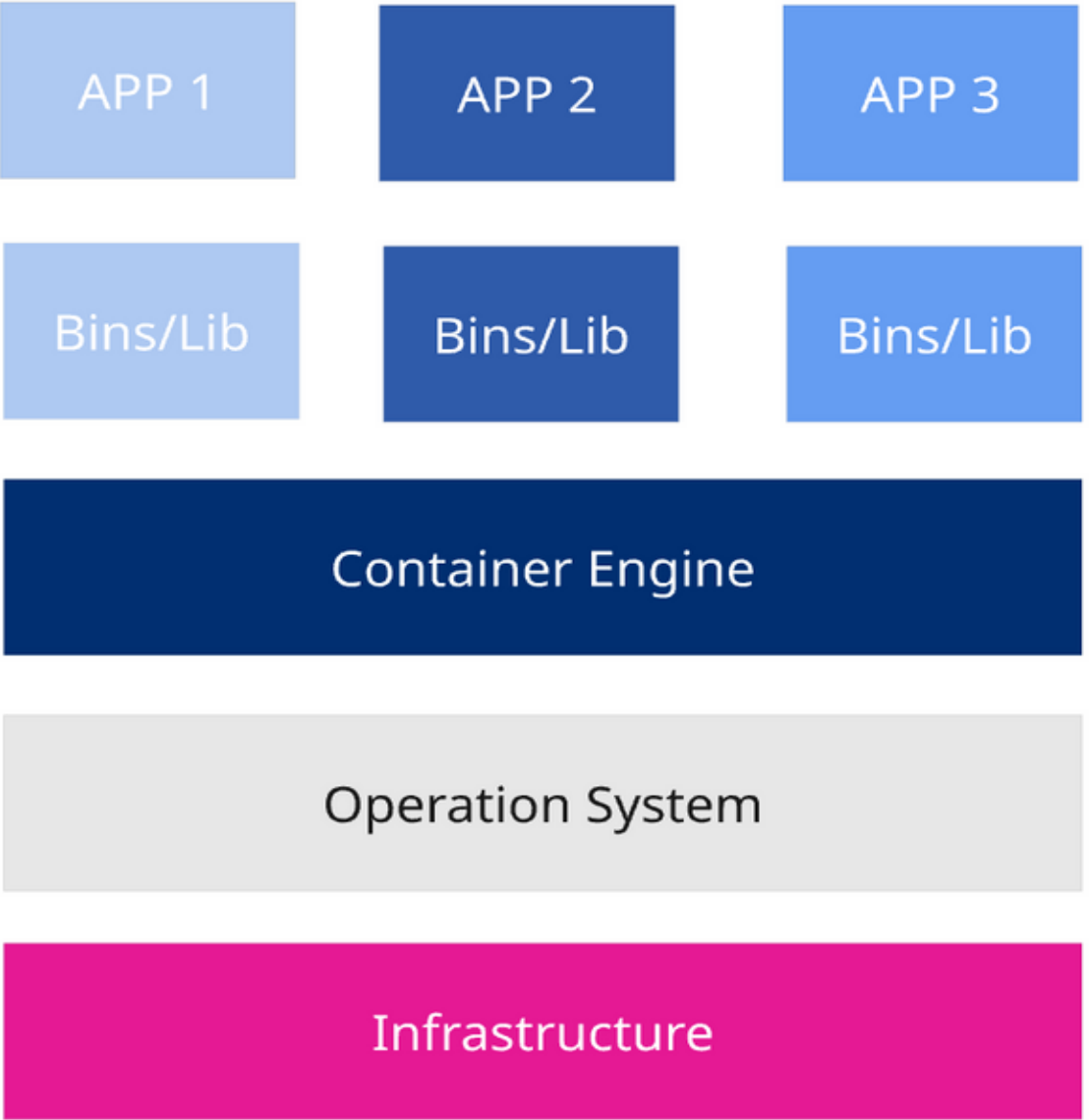
In VM-based installations, costs increase due to hypervisor licenses, additional management tools, and resource overhead.

In bare metal, since there is no hypervisor layer, license costs decrease, hardware is used more efficiently, and overall TCO (Total Cost of Ownership) is reduced.

### VIRTUAL MACHINES SERVER



### BAREMETAL SERVER



# Why Deploy OpenShift on Bare Metal



## High Performance

---

- Predictable latency
- Direct hardware Access

## Cost Efficiency

---

- Reduced TCO
- No hypervisor licensing
- Troubleshooting effort

## Operational Simplicity

---

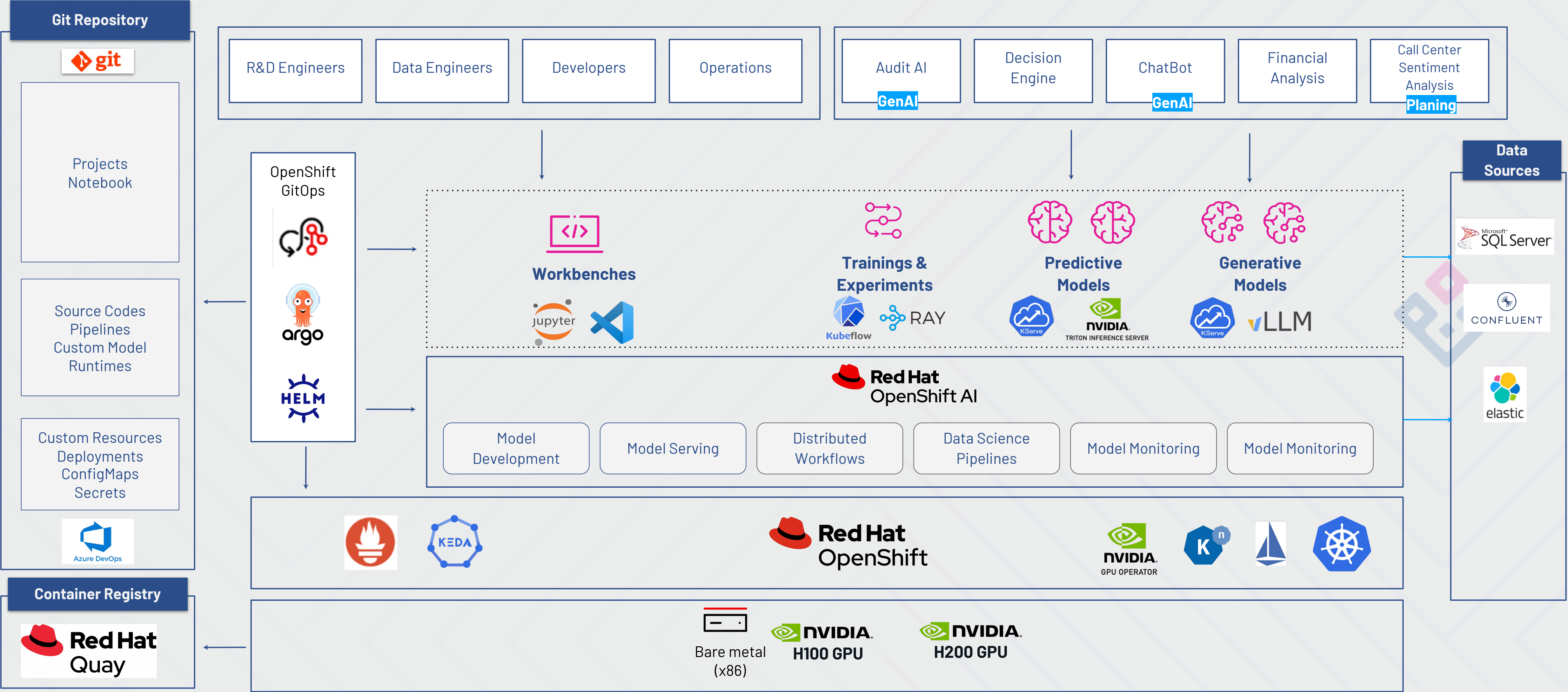
- Easier troubleshooting
- Fewer layers

## AI/ML Ready

---

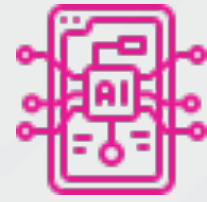
- GPUs direct access for AI/ML use cases
- Heavy IO
- High throughput network devices

# Vakıf Katılım Solution Architecture

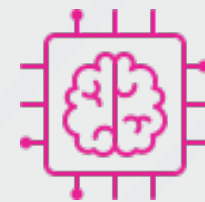


# AI Capabilities at Vakıf Katılım

---



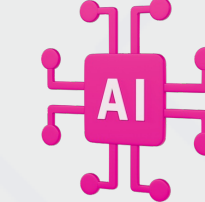
Training  
(RAY)



Predictive AI  
(NVIDIA)



Generative AI  
(VLLM)



RAG Pipeline  
(Enterprise  
Knowledge)



## RAG - Enterprise Knowledge Access

---

- QDMS, Jira and internal documents unified
- Instant and accurate Information retrieval
- Institutional knowledge becomes AI-accessible

## Agent - Financial Analysis Automation

---

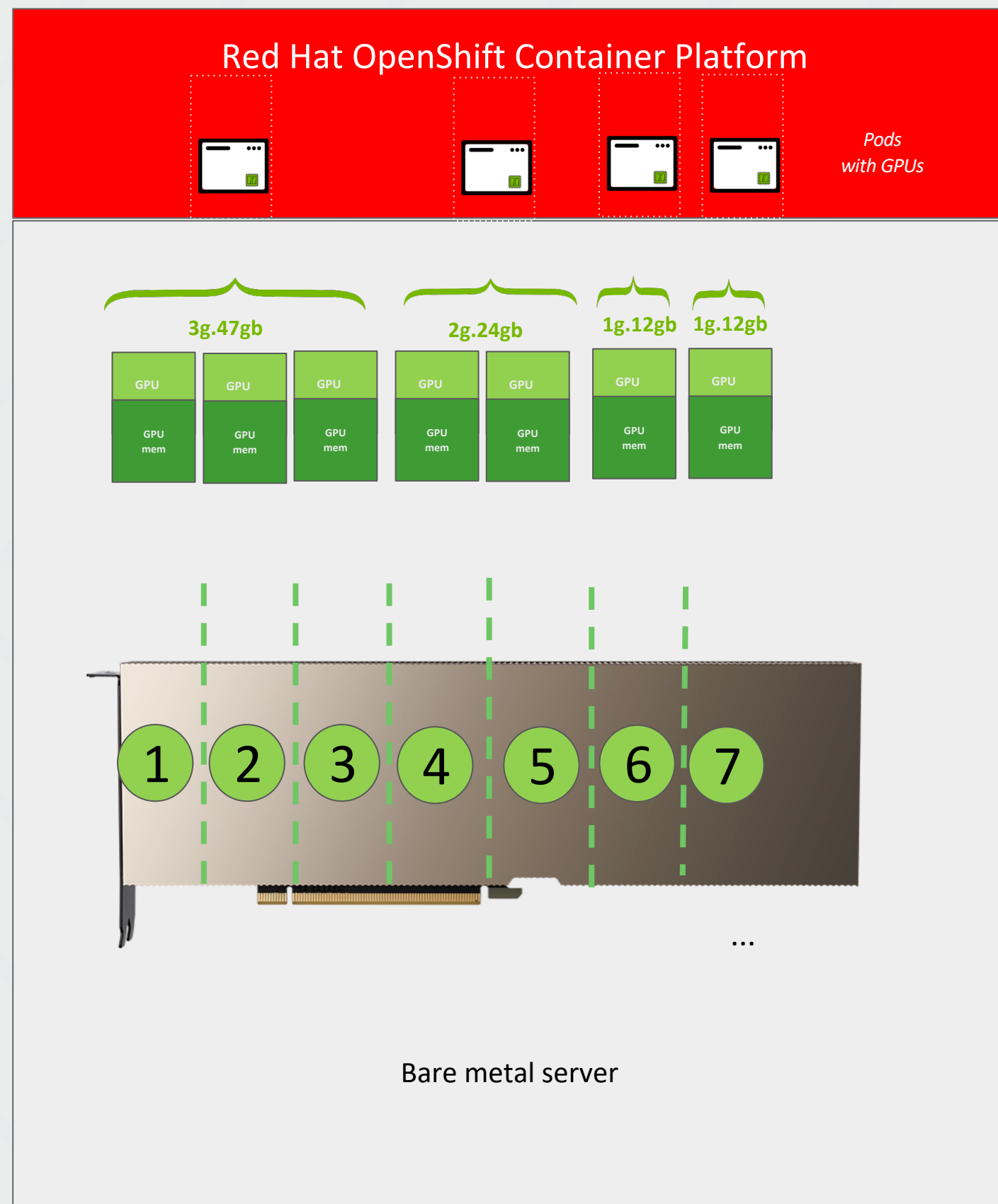
- Analysis time reduced from hours to minutes
- Agent executes its own analysis pipeline
- Human oversight for exceptional or risky cases

## Human in the Loop

---

- Human approval at critical decision points
- Risky or inconsistent outputs automatically blocked
- Safe and controlled AI adoption across business units

# NVIDIA MIG on OpenShift



- Launching multiple pods per physical GPU
- Optimize GPU utilization and cost
- Configured on the NVIDIA H100 & H200 GPU accelerators.
- MIG partitions a single NVIDIA GPU into up to seven independent GPU instances with guaranteed Quality of Service.
- MIG speeds up both development and deployment of AI models
- Small GPU instances are good for Notebooks and biggest instances for training
- Advertisement strategy: mixed (heterogeneous, diagram example)



# What We Learned & The Road Ahead

---

- Standardized AI-ready infrastructure
  - Faster model development & deployment
  - GPU efficiency increased by 30-45%
  - Enterprise knowledge accessible via RAG
  - Automated financial analysis (Agent)
  - Safe & controlled AI adoption
- Expand enterprise RAG capabilities
  - Multi-model serving (1 GPU -> multi-models)
  - Wider AI agent adoption across business units
  - Strengthen governance & responsible AI
  - Continuous optimization of GPU resources

software **one**

Thank You!