

Challenges and Opportunities in Alibaba Scheduling and Resource Management

Ding Yu (叔同)

Ding Yu (叔同)

- 2010–Now Alibaba
 - Sr. Staff Software Engineer at Alibaba Group
 - Responsible for stability in 11/11. Was leading High Availability Architecture Team
 - Director of the Cluster Management / Scheduling Team



Agenda

- What is Sigma
- Typical Scenarios and Architecture
- Problems & Evolution
- Challenges and Opportunities

Overview – Alibaba Infrastructure

Database / Storage / Middleware /
Compute Platform

Resource Scheduling / Cluster Management /
Container

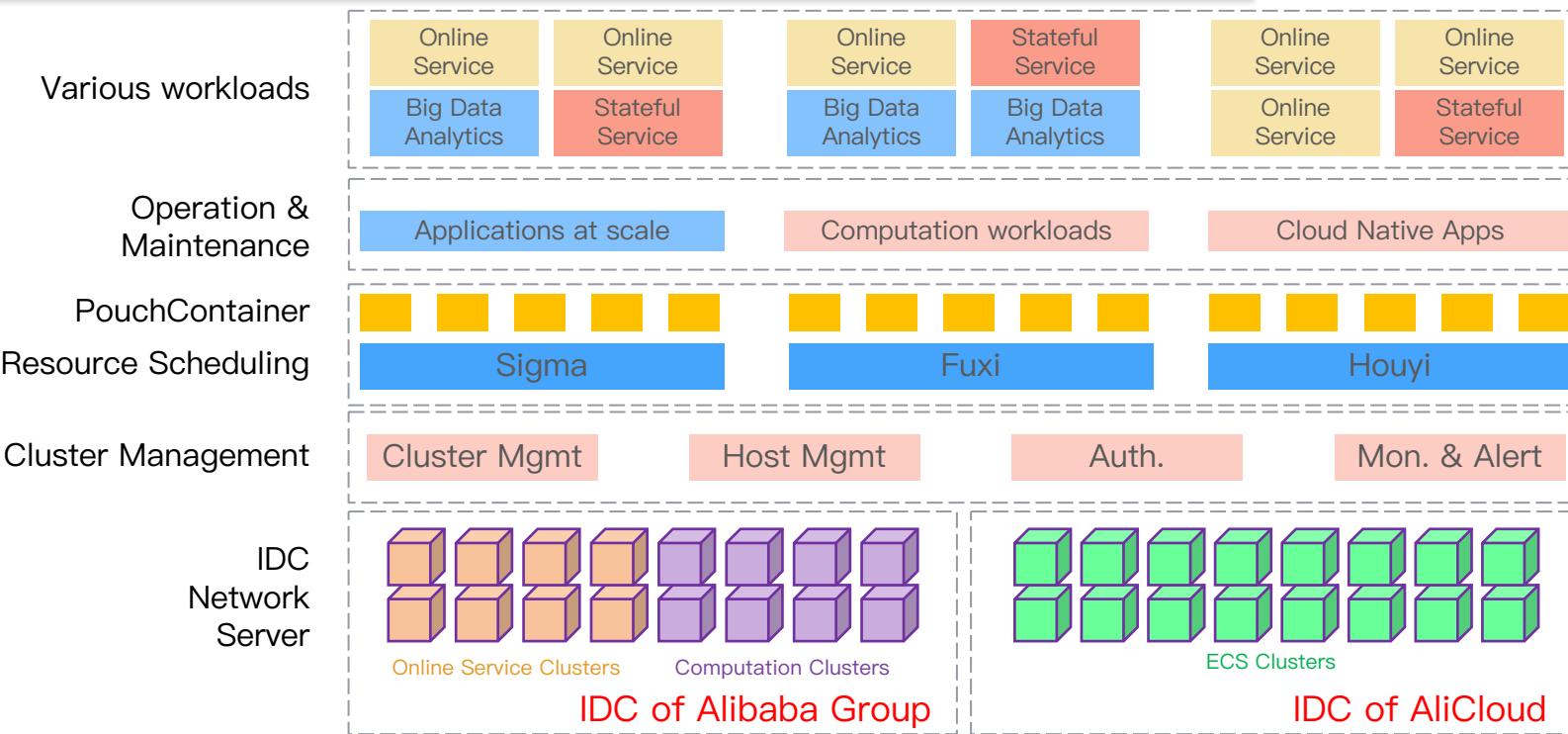
Foundation Software (OS / Virtualization)

Data Center (IDC / Networking / Servers)

Sigma Overview

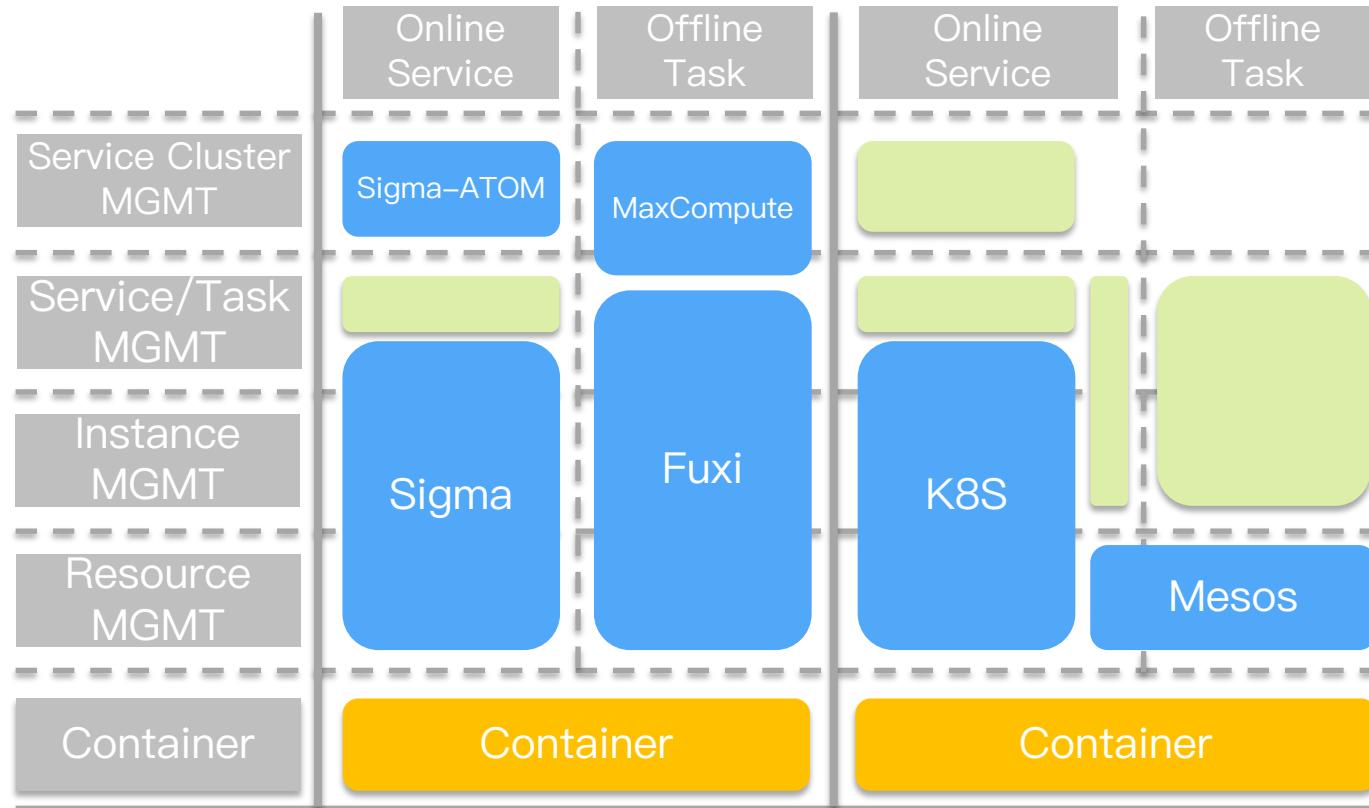
- Developed since 2012
 - to improve the datacenter utilization significantly, with best cloud native solution
- For all online services in Alibaba Group, managing the life–cycle of every container from all Business Units
- Support various resource requirements on different infrastructures
- Support various jobs with different priorities in the same cluster
- Easily integrated with other non–prod job schedulers (e.g. Fuxi, Yarn)
- Currently we have a cluster (about 10K servers) to colocate of mixed workloads
 - The cpu avg util is about 40%. Also used as temporary buffer in Singles Day (11/11)
- Use AliCloud resources as temporary buffer to support sales promotion, e.g. Singles Day (11/11)

Tech stacks in our infrastructure



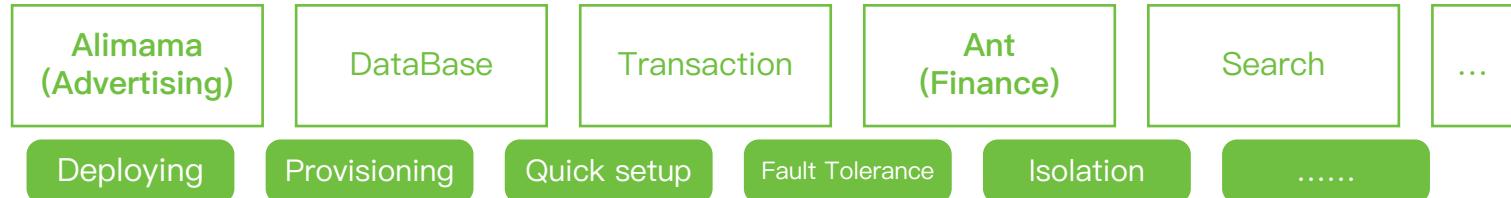
To make datacenter as a computer, unified scheduling for lower cost

Sigma's position as in open source solution

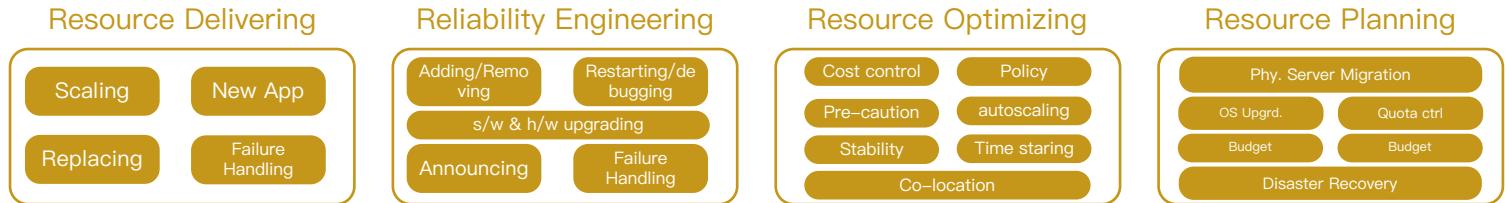


Scenarios and Responsibilities

Business Fields



Scenarios



Capabilities

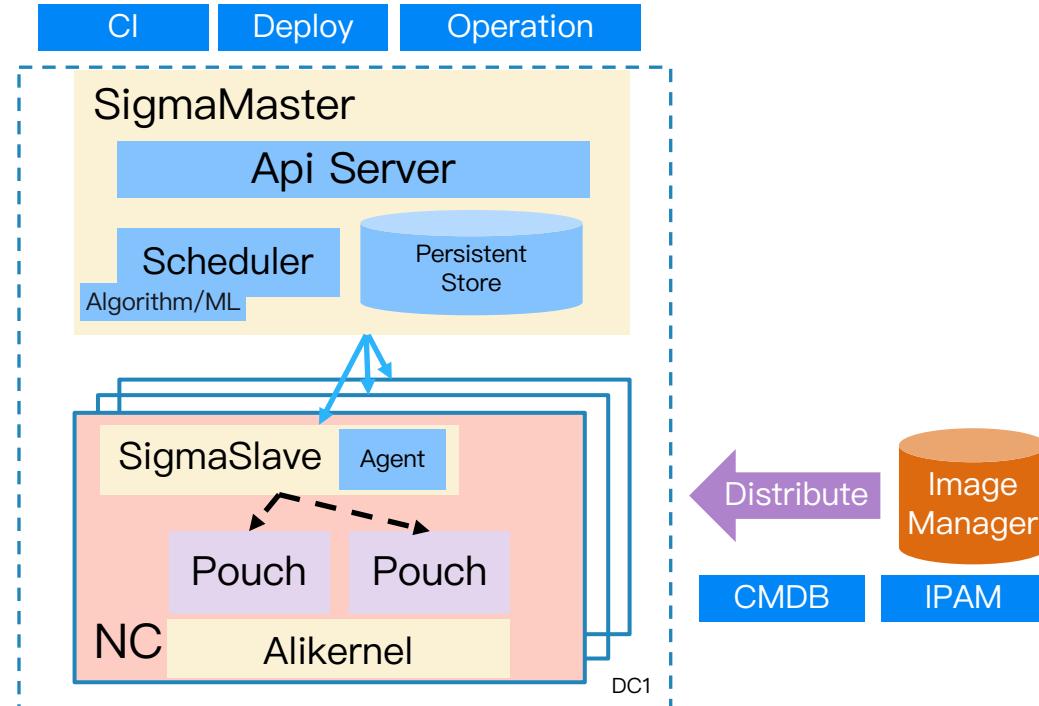
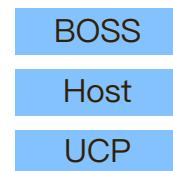


Infrastructure

IDC (Networking / Phy. Server / IP /)

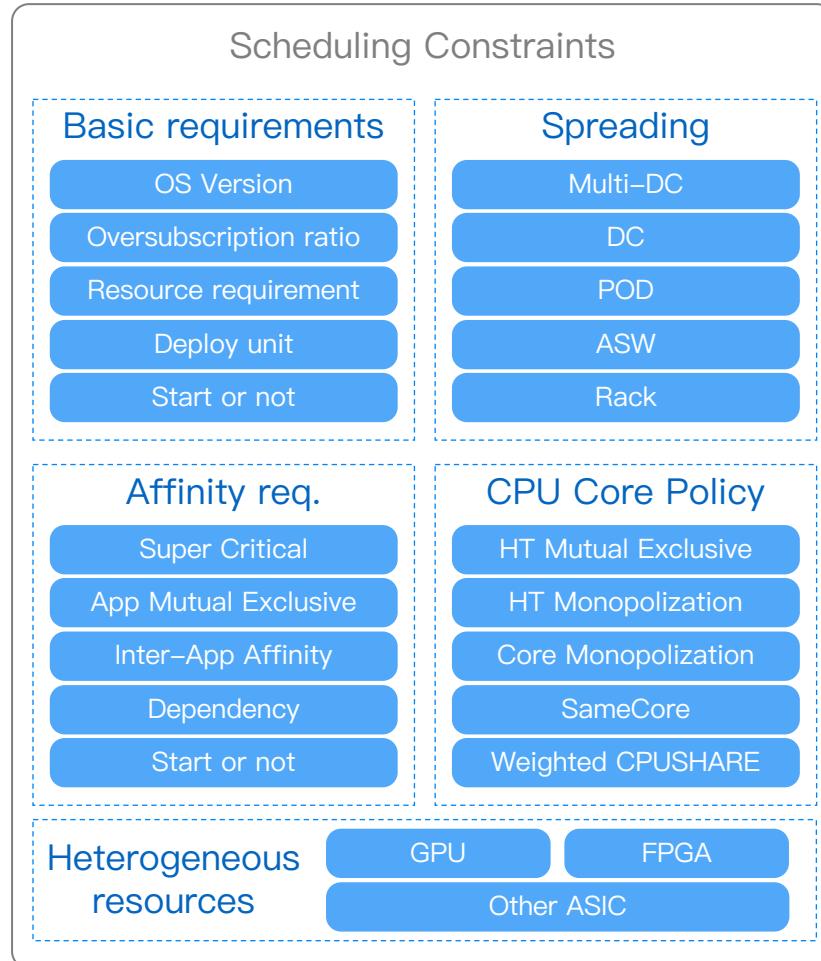
Sigma for Online Services

- 3-layer Brain (Cluster /Host / Kernel)
- Hundreds of thousands of machines
- Horizontal scaling in minutes
- Vertical scaling in seconds

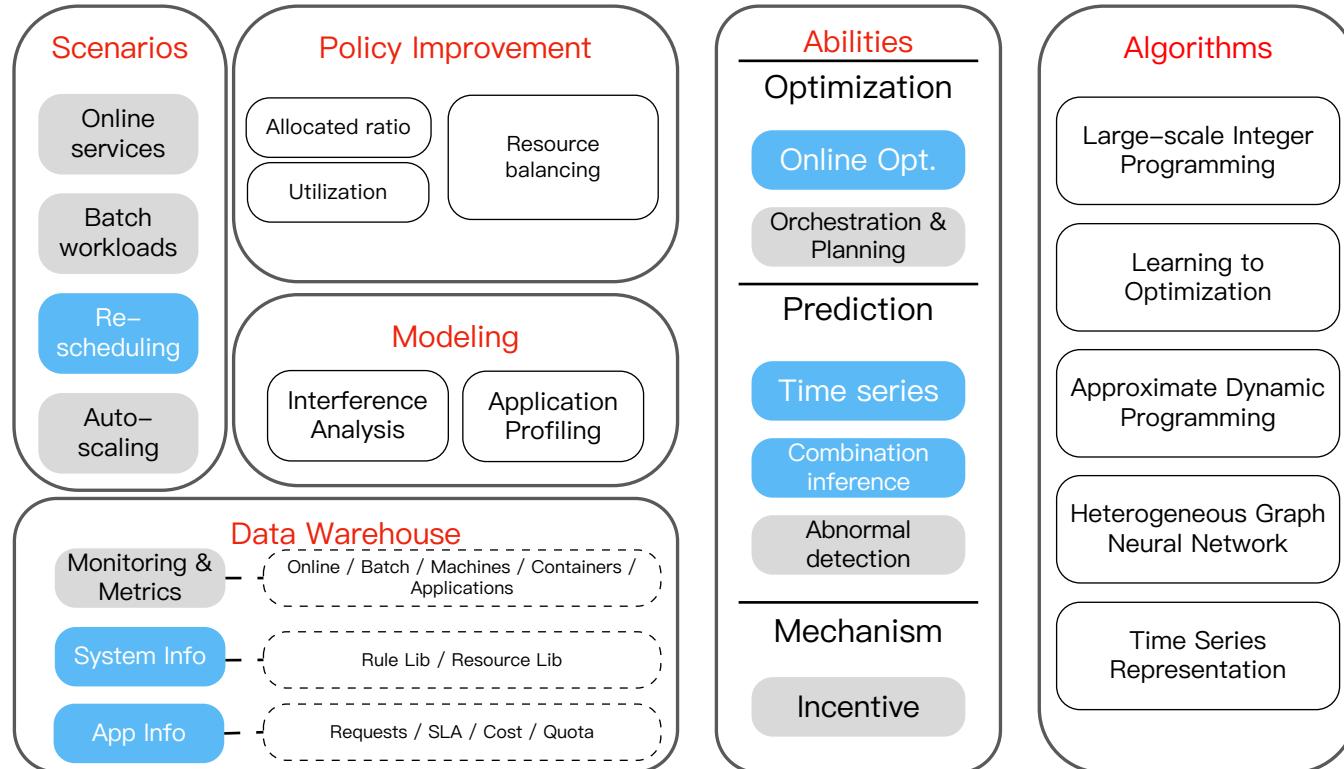


Sigma scheduling policies

- Resource Constraints (CPU/Mem/Disk/Network IO/etc.)
- Priority based on the importance of applications
 - Transaction Applications/Core Services first
- Multi-level Disaster Tolerance: Region/DC/Rack/etc.
- Optimizing Allocation at run time
 - Monitoring/Analyzing/Restarting/Migrating/etc.
- Many affinity/anti-affinity/dependencies to consider

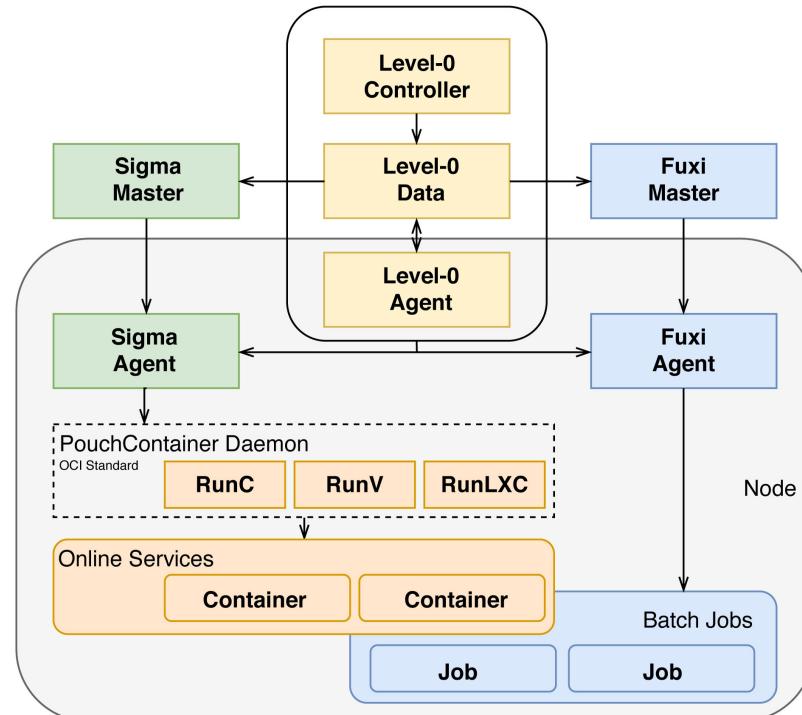


Algorithms boosting scheduling and managing



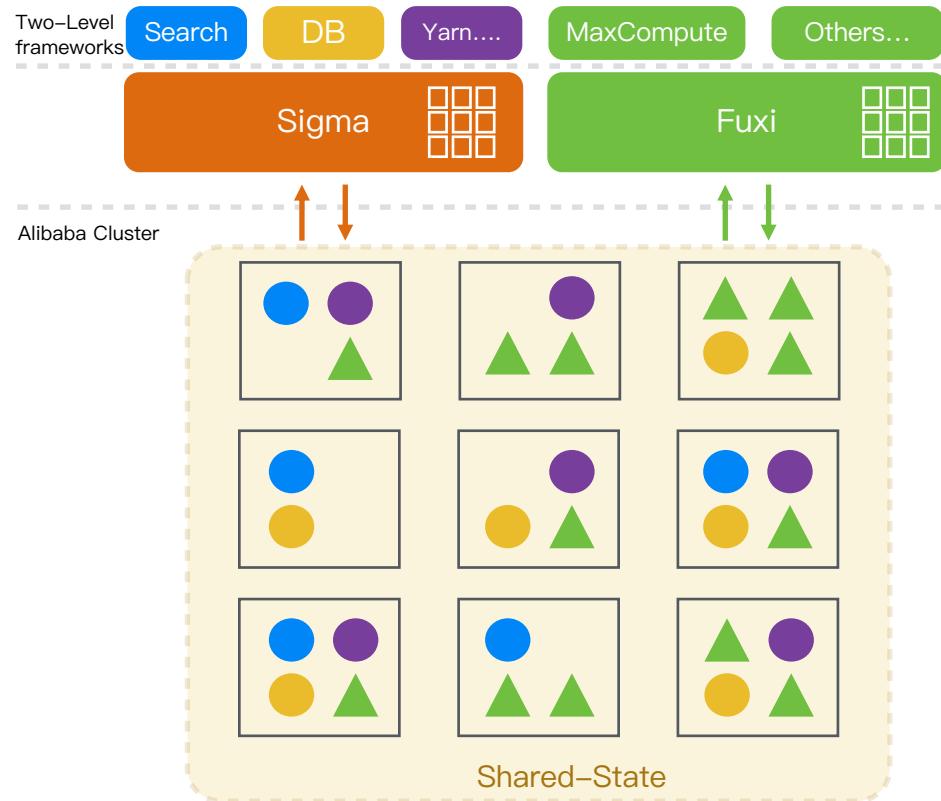
Colocation of Sigma and Fuxi

- Tens of thousands machines
- 40%+ avg. Utilization
- 30%+ cost saved



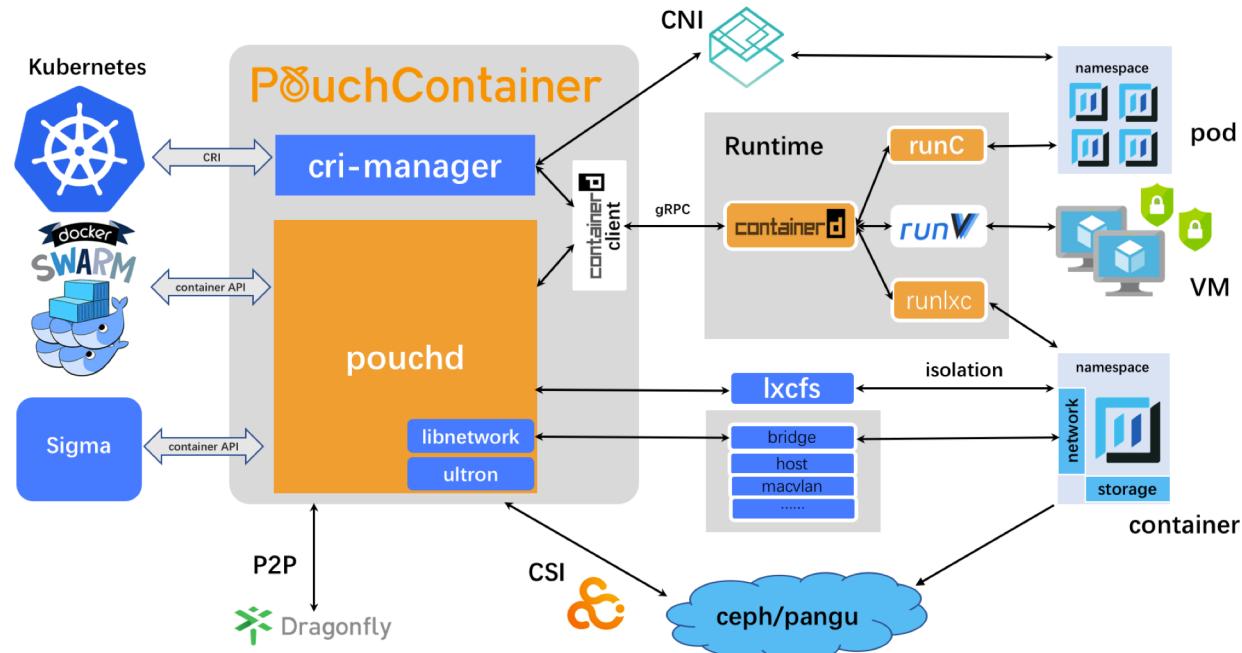
Colocation overview

- Various workload types
- Priority based resource sharing



PouchContainer – An Efficient Enterprise-class Rich Container Engine

- Since 2011
- Rich container
- Strong isolation
- Stability
- Kernel compatibility
- Enterprise level



Tech Evolution

- Computation & Storage Separation
- Dynamic Scheduling Test Bed & Intelligent Scheduling
- Supporting workloads of multiple priorities
- Application resource profiling, runtime analysis
- Interference detection & analysis, host level resource optimization, vertical scaling
- CPU share & NUMA

Tech Evolution of the industry

- Data Center
 - Network: 25G → 100G
 - RDMA, NVMe
 - Heterogeneous computing, FPGA, GPU, AI ASICs
- System Software
 - Unikernel, User Mode IO, Serverless
- PaaS/Middleware
 - FaaS
 - Service Mesh
 - Cloud Native
 - AI/ML

Challenges ahead

- Very large scale: **1%** improvement means a lot
- Competition is fierce; Need to iterate rapidly
- How to evolve ourselves while industry is evolving?
- Build the world class, high efficient next generation cluster management OS
- Build best cloud native solutions with enterprise level container & management platform

Challenges ahead

- More aggressive resource sharing for various workload
- How to select non-prod jobs to efficiently exploit the reusable resource in prod jobs clusters, without affecting SLO?
- How to balance the prod jobs SLO, non-prod jobs SLO (task-eviction rates), DC limitation (power, bandwidth etc.) and resource utilization?
- How to improve all Alibaba datacenters resource utilization?
- How to continuously reduce the cost in supporting 11/11? Down to 0 additional resource/cost

The challenges ahead

- The Advantages in Alibaba
 - Many large scale & high fault-tolerance scenarios
 - e.g. e-Commerce, Entertainment, LBS
 - AliCloud is the No.1 cloud vendor in China, we can get the user feedback soon
 - It means the innovation of software infrastructure can iterate faster

THANKS

shutong.dy@alibaba-inc.com

WeChat:18657182390