

# Improving Resource Efficiency in Kubernetes Clusters



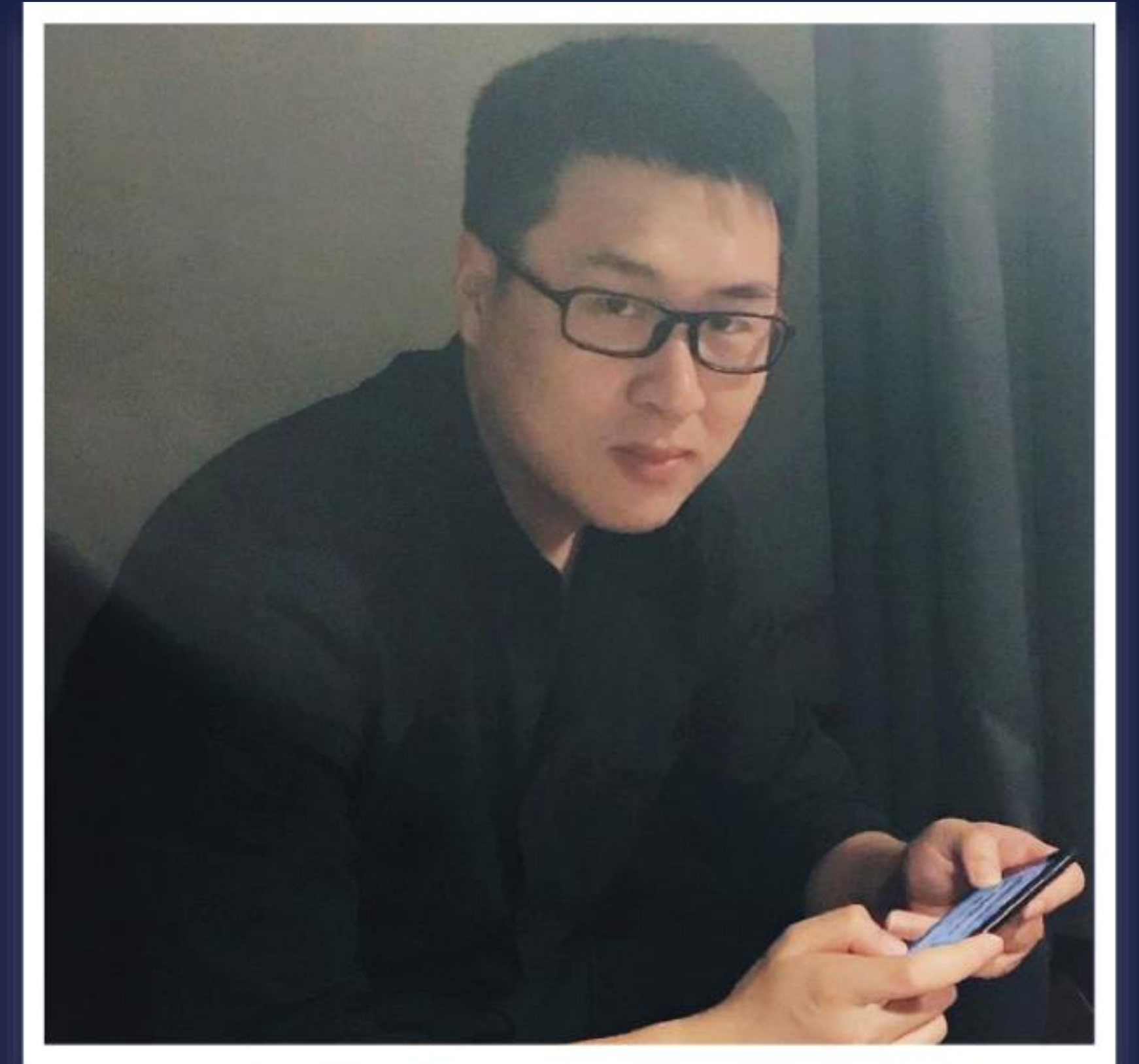
Tongkai Yang  
[tongkai.ytk@antfin.com](mailto:tongkai.ytk@antfin.com)



扫码关注公众号,获取 316 北京站 PPT

# About Me

- ❖ 杨统凯
- ❖ 蚂蚁金服 - 系统部 - 资源调度
- ❖ Scheduling & Resource Optimization



# Should I Care About Utilization?

- ❖ Cluster Scale

- > 1k Cpu Nodes
- > 200 Gpu Nodes

- ❖ Cluster Cost

- > 100 million

# Considerations?

- ❖ Cost

- Research & Development
  - Operations

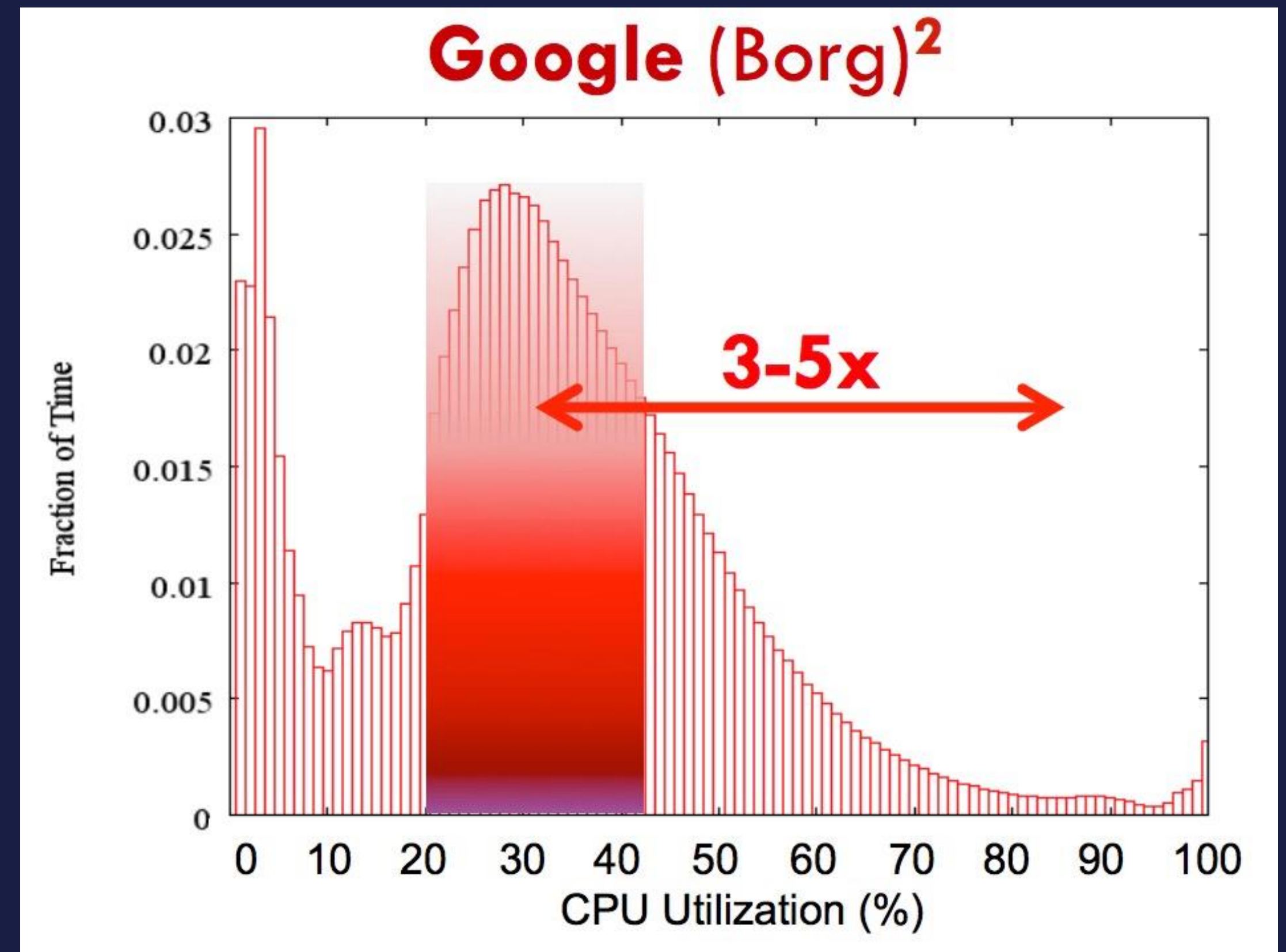
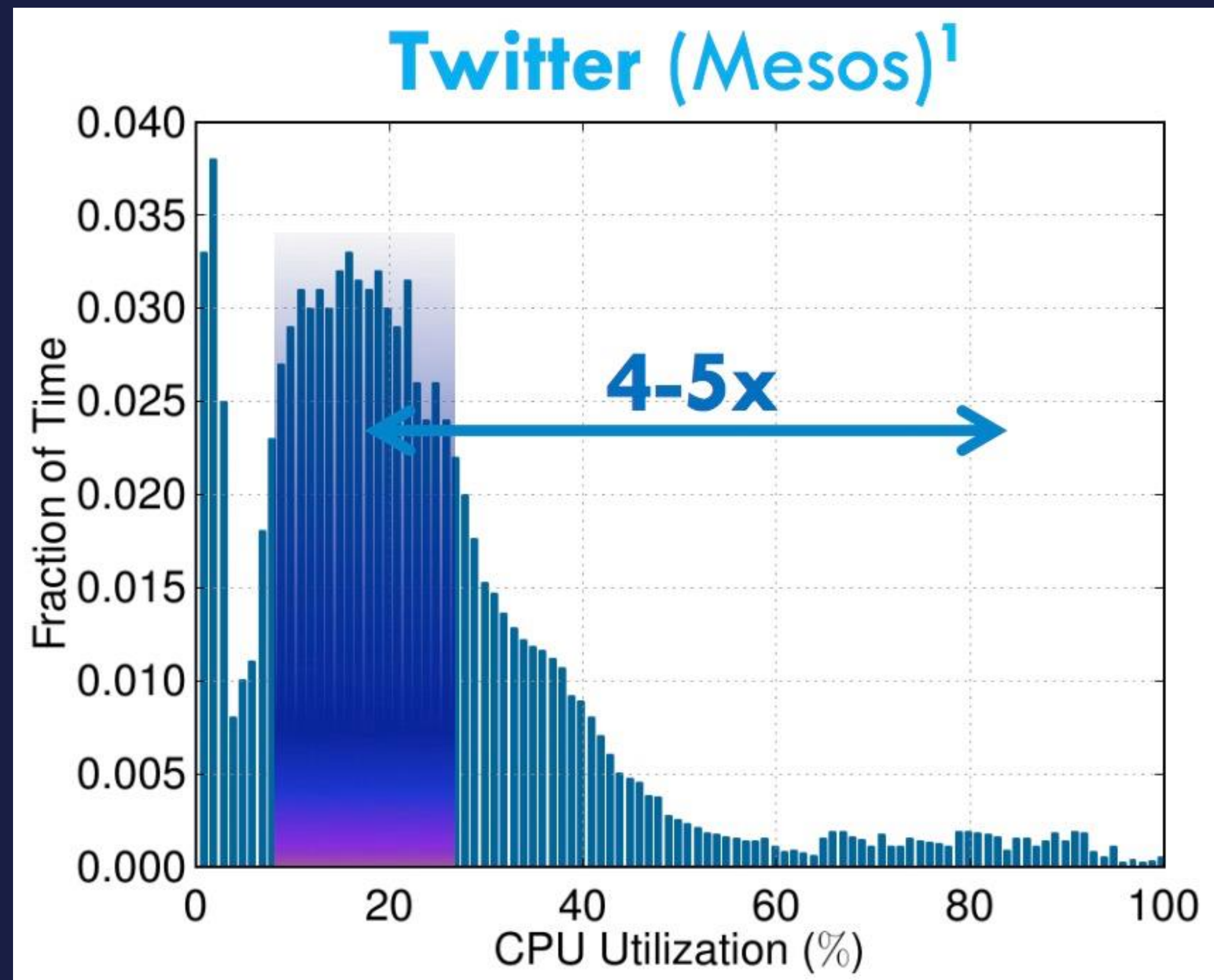
- ❖ Stability

- Hire or Fire

- ❖ Architecture



# Current



Ref: [http://cs1.stanford.edu/~christos/publications/2015.christina\\_delimitrou.phd\\_thesis.slides.pdf](http://cs1.stanford.edu/~christos/publications/2015.christina_delimitrou.phd_thesis.slides.pdf)

# Current

- ❖ High Cpu Allocation vs. Low Cpu Utilization
- ❖ Cpu Limited vs. Memory Unallocated
- ❖ Resource Pools vs. Lack of Colocation
- ❖ Imbalance Nodes Cpu Utilization
- ❖ Mostly Guaranteed Pods

# Why?

- ❖ User`s Fault?
- ❖ Cluster Management Fault?

# How is the Cluster Workload?

- ❖ Business Types

- Long-running Services
- Streaming Jobs
- AI/batch Jobs
- Testing/CI/CD Tasks



# How is the Cluster Workload?

## ❖ Resource Feature

- Long-running Services
  - – Regular Usage
- Streaming Jobs
  - – Regular & Busy Cpu Usage
- AI/batch Jobs
  - – Irregular & Busy Cpu/Memory Usage
- Testing/CI/CD Tasks
  - – Irregular & Busy IO/Cpu

# What's Kubernetes Offer?

- ❖ User Definition Resource
  - ❖ Static
  - ❖ Always Overcommitted
- ❖ Pod QoS Class
  - ❖ No Explicit Expression
  - ❖ No Resource SL0
- ❖ Scheduling
  - ❖ Static Quota Scheduling
- ❖ Pod Priority & Preemption

# What We Do? – Resource

- ❖ Resource SLO Definition
  - PRODUCTION / PREEMPTIBLE
- ❖ Resource Quota
  - Match Business Organization
  - Sharing & Transfer
- ❖ Resource Isolation & Priority
  - ❖ Cpu / Cache / Memory / Net / BlkIO
- ❖ Resource Auto-Profiling
  - ❖ Cpu / Memory
  - ❖ VPA & HPA

PRODUCTION

SLO Guaranteed

NOT Preemptible

High Priority

Support Auto-Profiling

PREEMPTIBLE

No SLO

Preemptible

Low Priority

# What We Do? – Scheduling

- ❖ Scheduling Strategy

- ❖ Running Cpu Quota

- ❖ Priority & Preemption

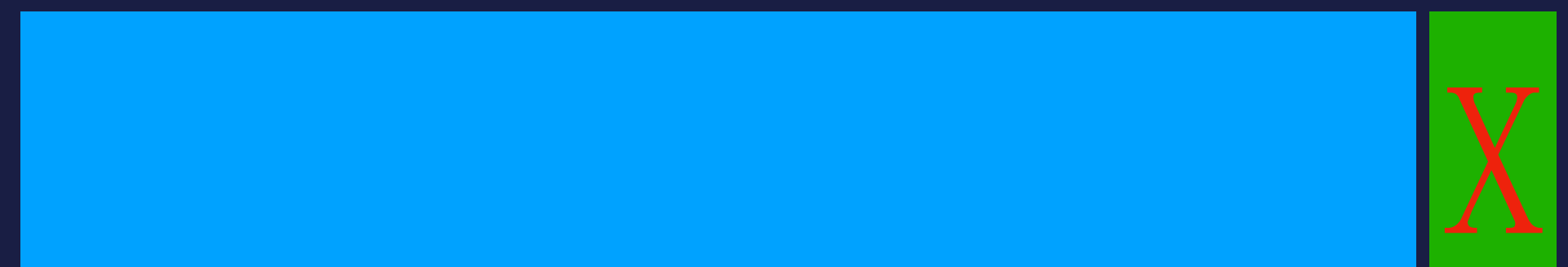
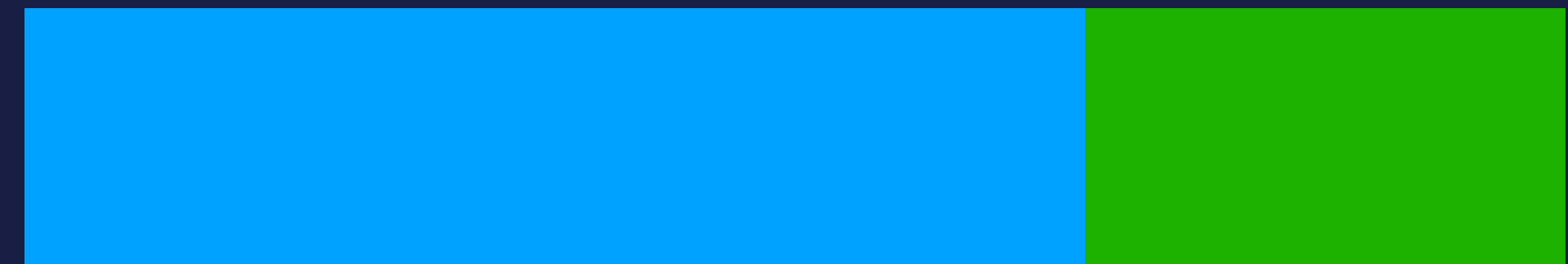
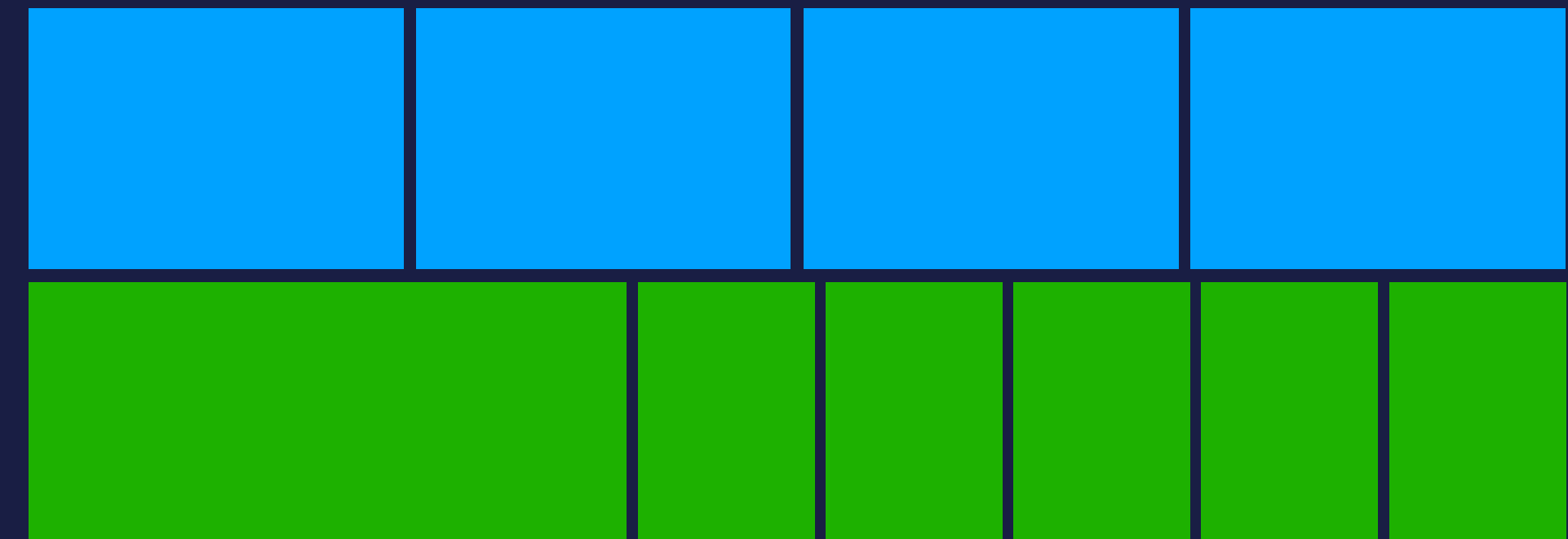
- ❖ Balance Algorithm

- ❖ Colocation Strategy

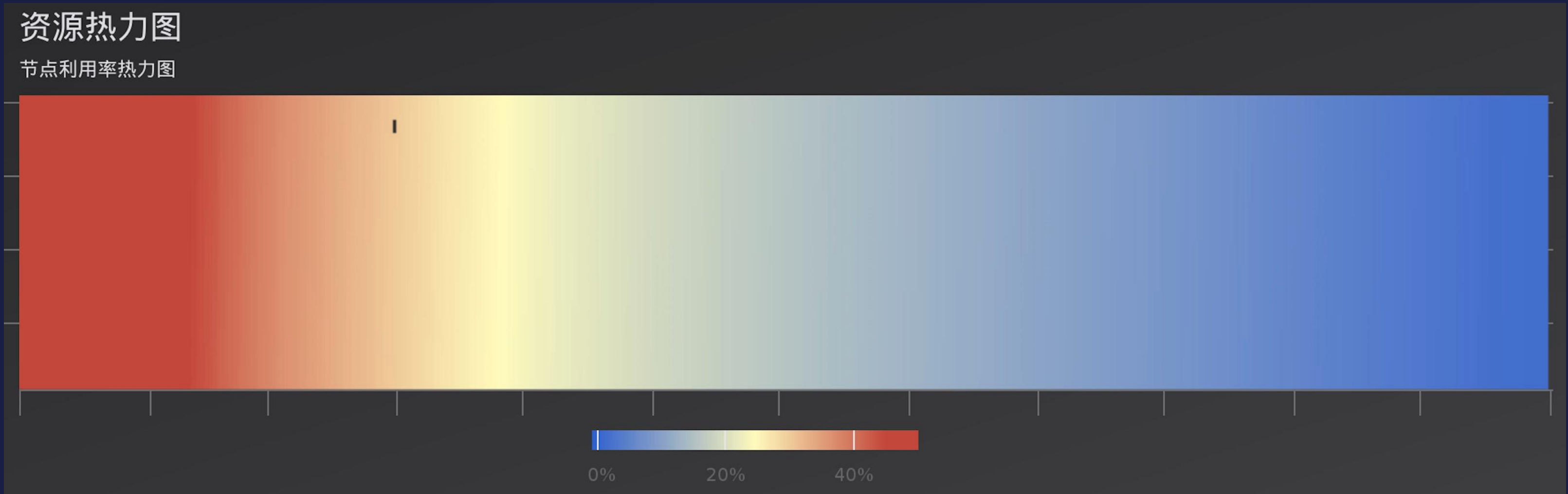
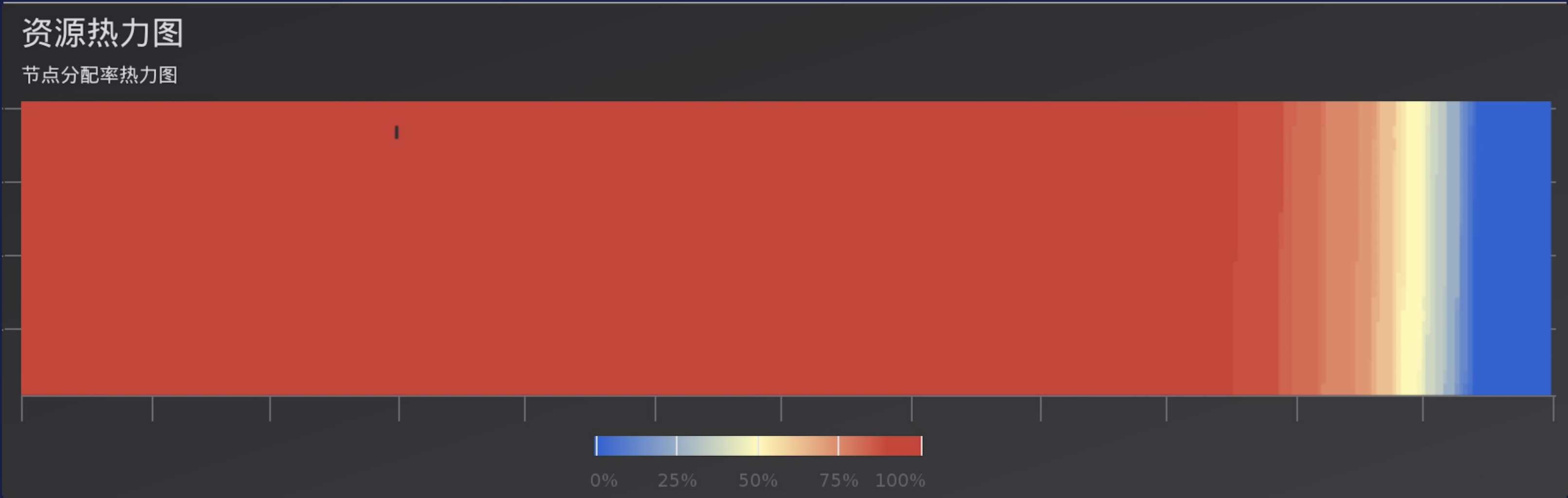
- ❖ Co-locate Service & Jobs

- ❖ Reschedule

- ❖ Eviction



# What We Got?





# Thank You !



扫码关注公众号,获取 316 北京站 PPT

## We Are Hiring !

[tongkai.ytk@antfin.com](mailto:tongkai.ytk@antfin.com)

