Volcano: Bring Batch Capability Into Kubernetes

Da Ma (@k82cn)

Huawei Expert

极客邦科技
会议推荐2019

5月

QCon 北京

全球软件开发大会

大会：5月6-8日
培训：5月9-10日

QCon 广州

全球软件开发大会

培训：5月25-26日
大会：5月27-28日

6月

GTLC GLOBAL TECH LEADERSHIP CONFERENCE 上海

技术领导力峰会

时间：6月14-15日

GMTC 北京

全球大前端技术大会

大会：6月20-21日
培训：6月22-23日

7月

ArchSummit 深圳

全球架构师峰会

大会：7月12-13日
培训：7月14-15日

10月

QCon 上海

全球软件开发大会

大会：10月17-19日
培训：10月20-21日

11月

GMTC 深圳

全球大前端技术大会

大会：11月8-9日
培训：11月10-11日

AiCon 北京

全球人工智能与机器学习大会

大会：11月21-22日
培训：11月23-24日

12月

ArchSummit 北京

全球架构师峰会

大会：12月6-7日
培训：12月8-9日

Geekbang. 极客邦科技 | TGO 鲲鹏會

TGO鲲鹏会

汇聚全球科技领导者的高端社群

全球12大城市

850+高端科技领导者

使命 Mission

为社会输送更多优秀的科技领导者

愿景 Vision

构建全球领先的有技术背景优秀人才的学习成长平台

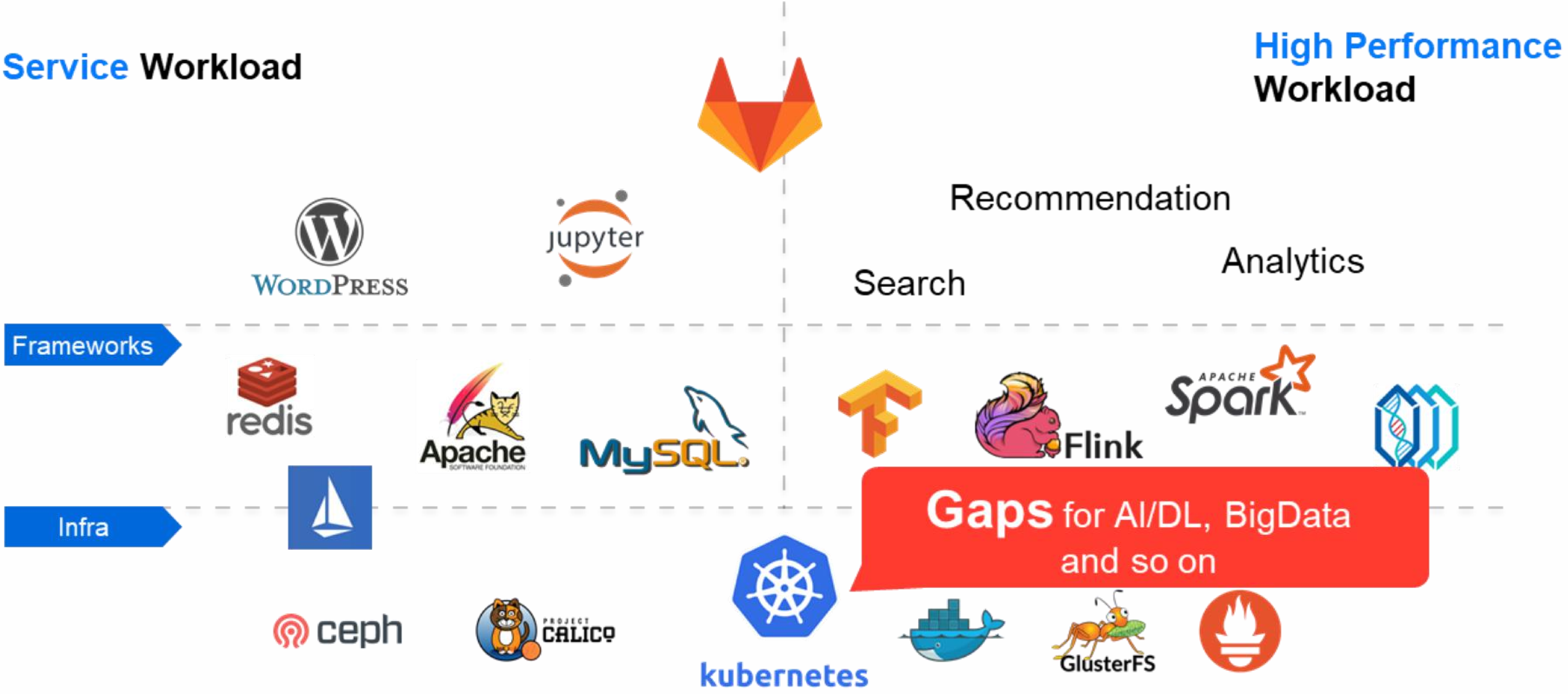
扫描二维码，了解更多内容

About The SPEAKER

Da Ma

Software Architect

- Kubernetes SIG-Scheduling co-Leader
- Volcano & kube-batch creator
- Expert at Huawei (now)
- Ex-IBM Spectrum CE/L3 Team/Tech Lead
- Jilin University master’s degree, majoring in grid computing and distributed system



Service Workload

High Performance Workload

Recommendation

Search

Analytics

Frameworks

Infra

Gaps for AI/DL, BigData and so on

Gaps

- Job/Queue Management
  - Queue status/configuration
  - Hierarchical queue
  - Job with multiple pod template
  - Lifecycle management of Job, e.g. restart, suspend/resume
  - Error Handling, e.g. restart job if pod failed (MPI, TFJob)
  - Indexed Job
  - Task dependency, e.g. Spark (executor/driver)
  - Delay Pod Creation
  - …
- Runtime
  - Singularity
  - …
- Scheduler
  - Coscheduling
  - Faire-share
  - Queue
  - Preemption/Reclaim
  - Reserve/Backfill
  - Topology (network, accelerator)
  - …
- Others
  - Throughput
  - Round-trip
  - Data locality (Data Aware Scheduling)
  - …

Volcano: A Kubernetes Native Batch System

Website: https://volcano.sh

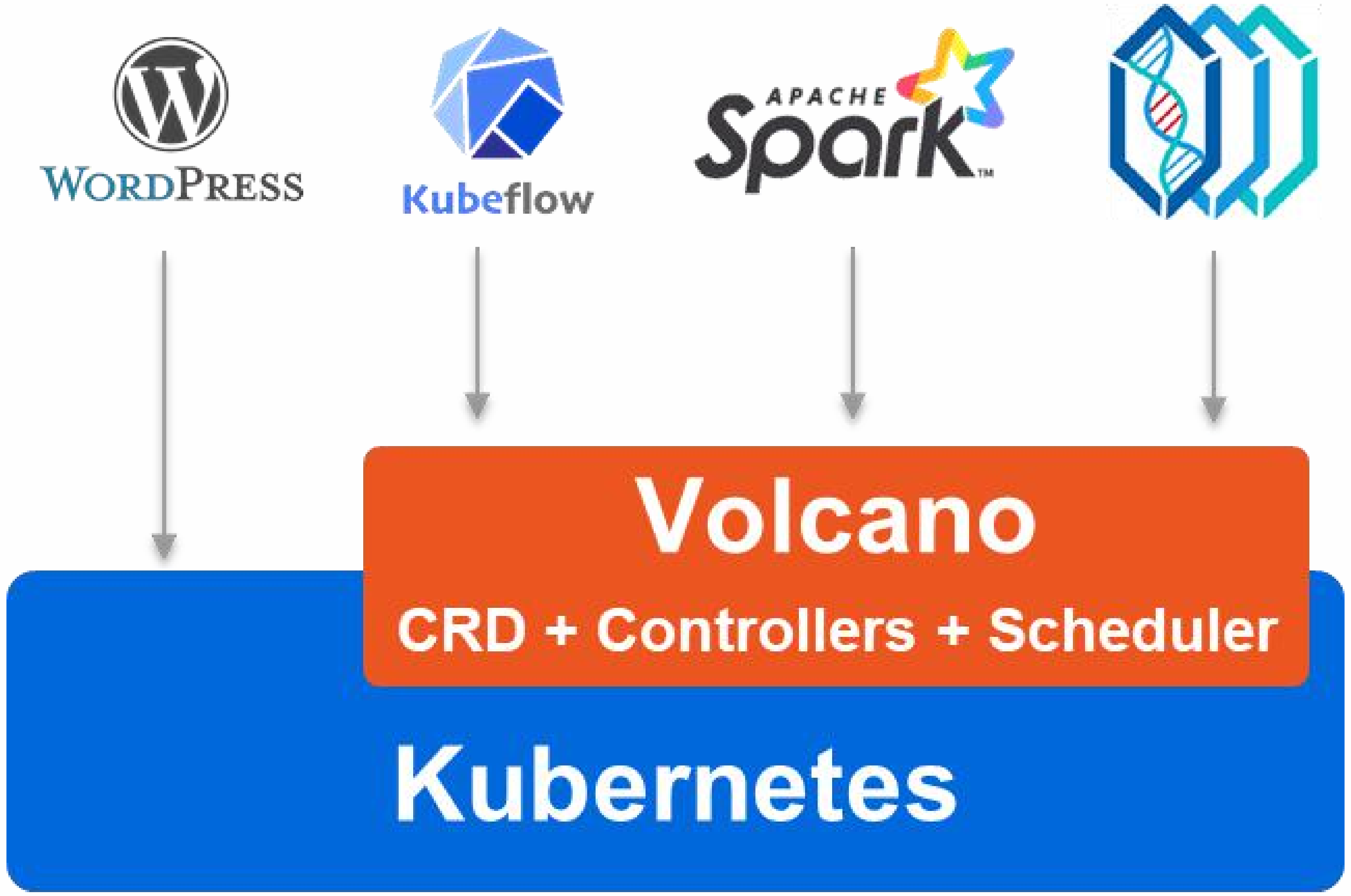
Github: http://github.com/volcano-sh/volcano

Twitter: https://twitter.com/volcano_sh

Slack: http://volcano-sh.slack.com

Email: volcano-sh@googlegroups.com

Overall Architecture



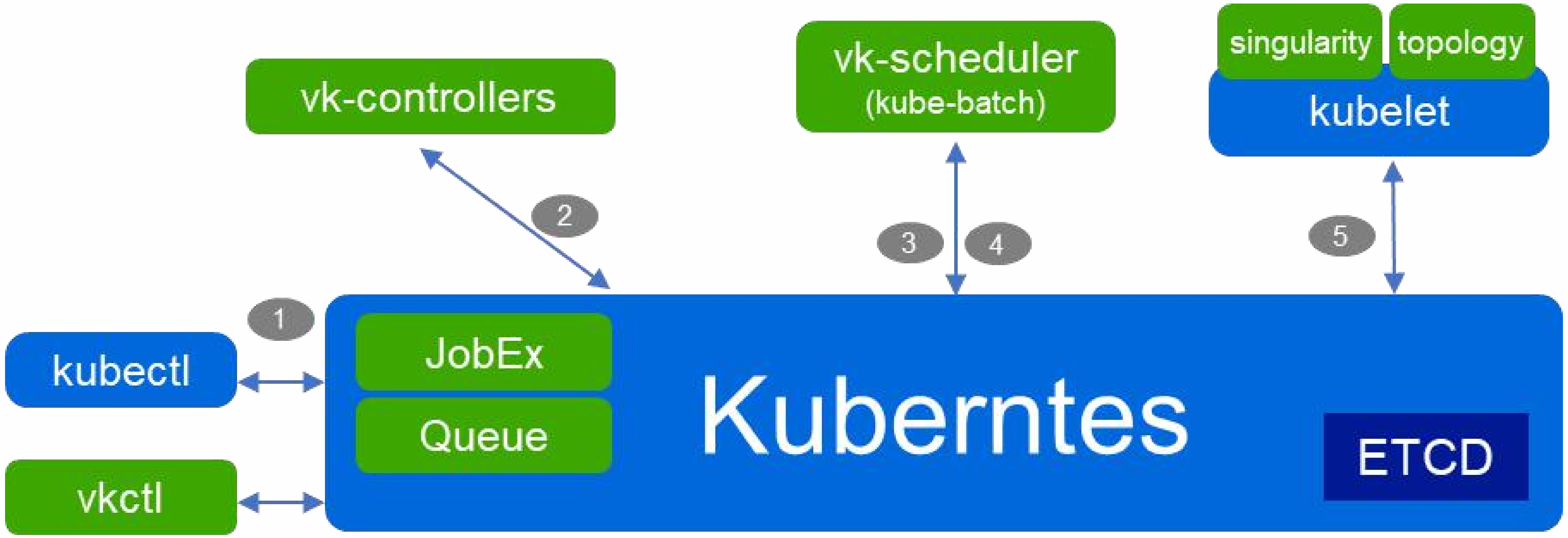
Domain frameworks:

- Deployment/Installation of framework in k8s
- Map framework's terms/concepts into common concept, e.g. Job, Queue
- Enable related features for frameworks, e.g. gang-scheduling for TensorFlow training

Common Service for high performance workload:

- Batch scheduling, e.g. fair-share, gang-scheduling
- Enhanced job management, e.g. multiple pod template, error handling
- Accelerator, e.g. GPU, FPGA
- kubectl plugins, e.g. show Job/Queue information

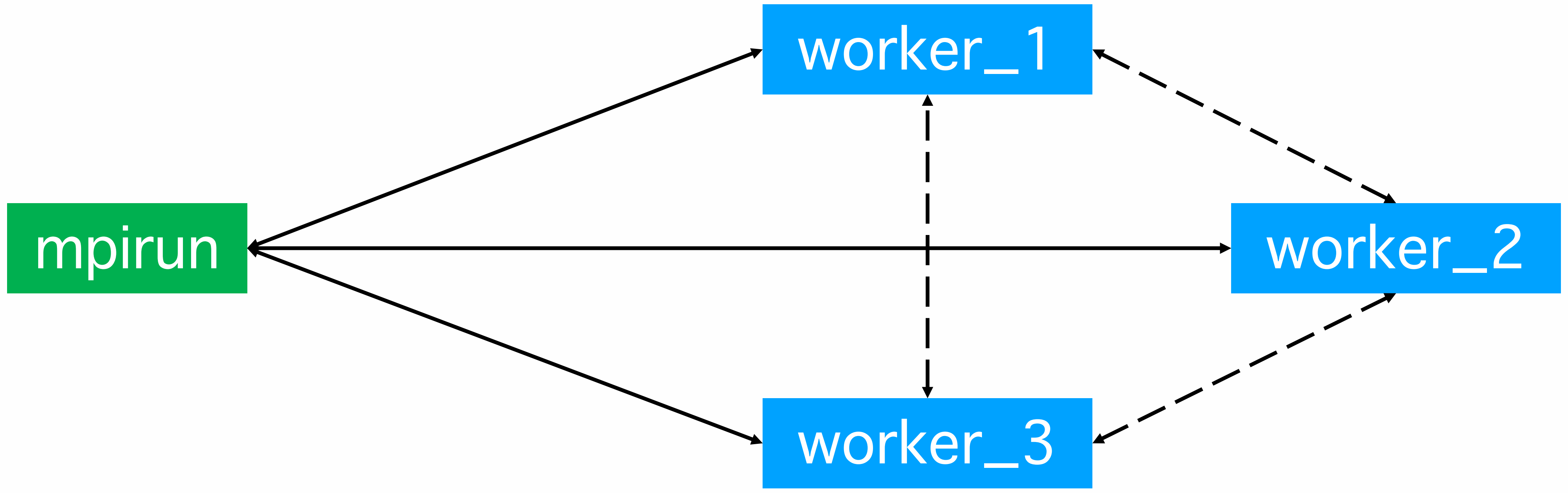
Overall Architecture

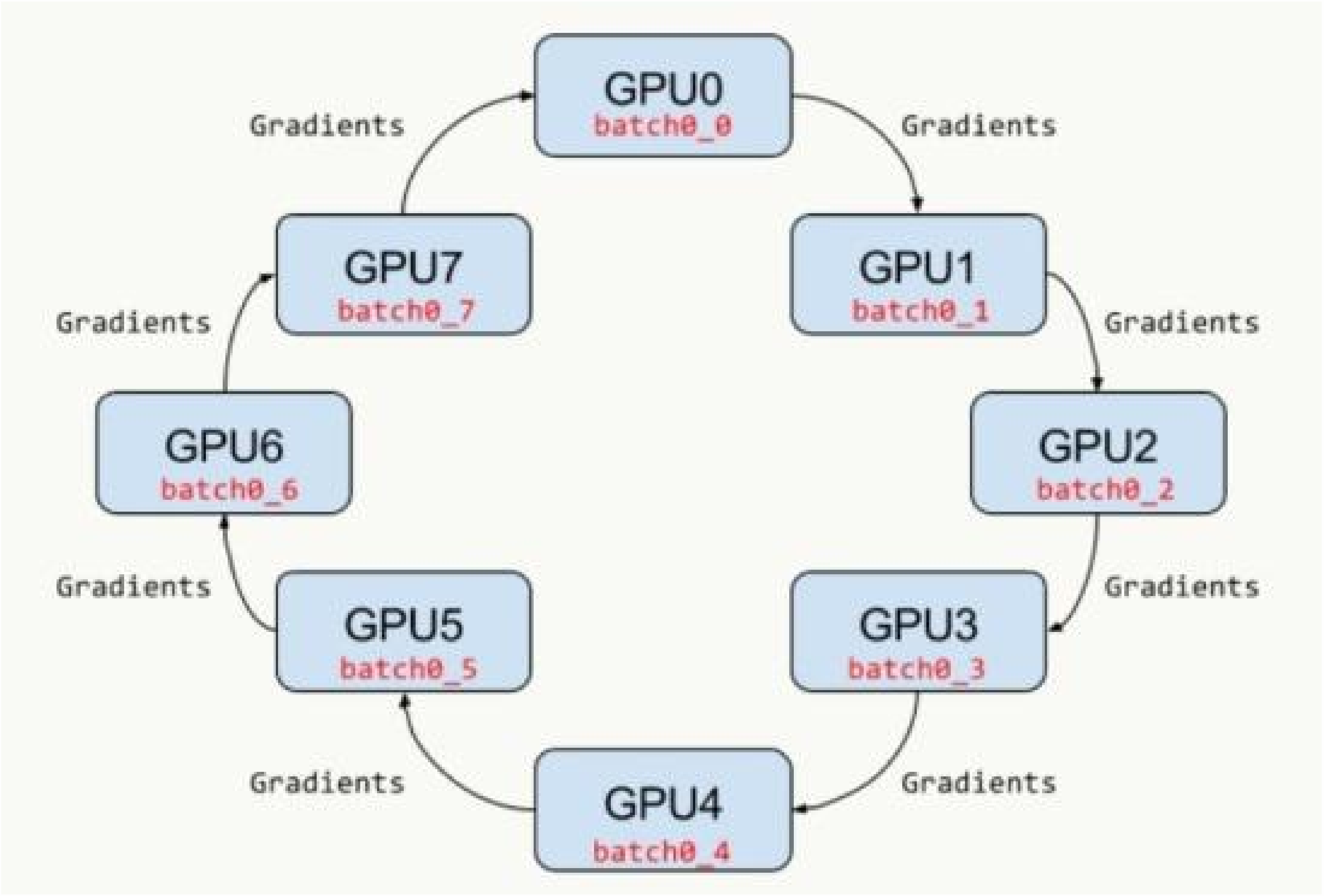


- The policy in vk-scheduler is pluggable, e.g. DRF, Priority, Gang
- vk-controllers includes JobExController, QueueController
- Volcano handles high performance workload

- Kubectl creates a JobEx object in apiserver if all admission passed
- JobExController create Pods based on its replicas and templates
- vk-scheduler get the notification of Pod from apiserver
- vk-scheduler chooses one host for the Pod of JobEx based on its policy
- kubelet gets the notification of Pod from apiserver; and then start the container

Scenarios: MPI

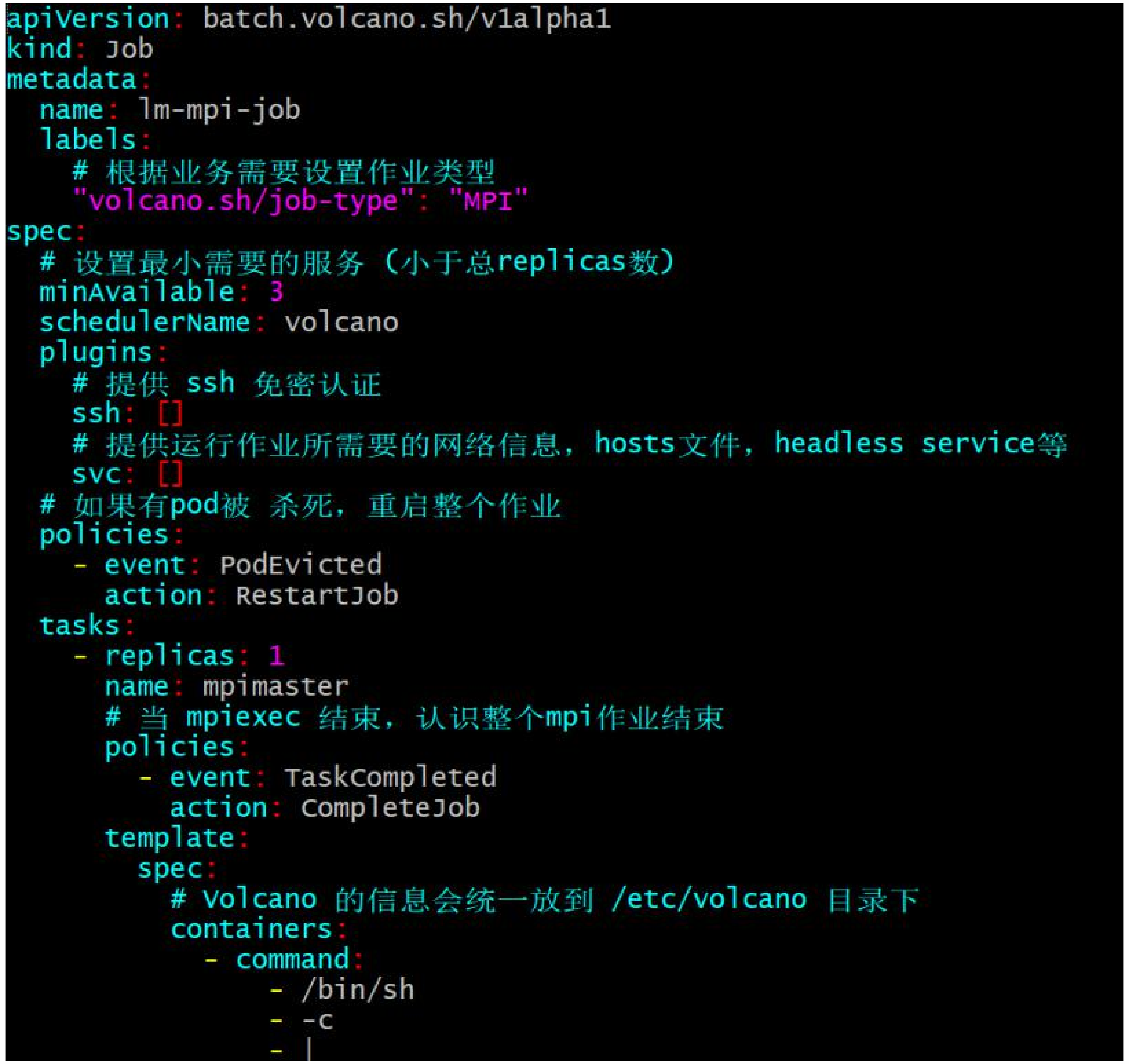




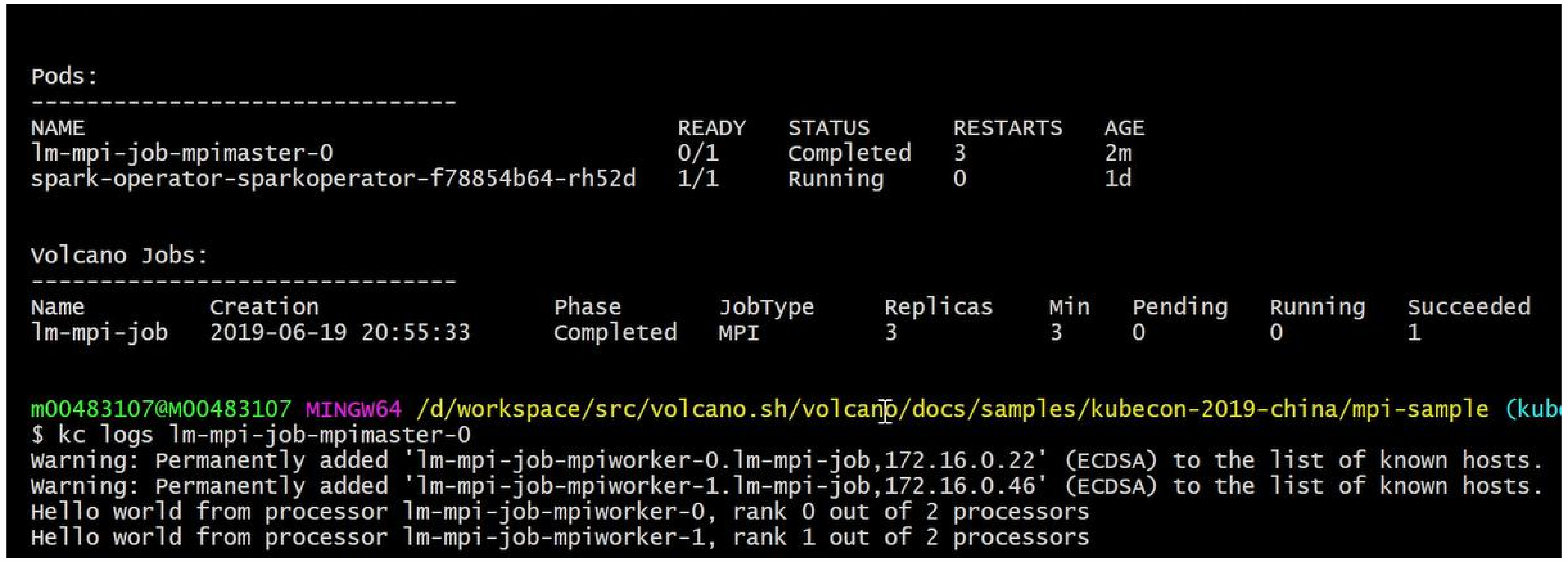
- Multiple Pod Template
- Lifecycle Policy
- Gang-scheduling

- ssh or kubectl
- Complete job when mpirun completed
- Headless service

Scenarios: MPI



```
apiVersion: batch.volcano.sh/v1alpha1
kind: Job
metadata:
  name: lm-mpi-job
  labels:
    # 根据业务需要设置作业类型
    "volcano.sh/job-type": "MPI"
spec:
  # 设置最小需要的服务 (小于总replicas数)
  minAvailable: 3
  schedulerName: volcano
  plugins:
    # 提供 ssh 免密认证
    ssh: []
    # 提供运行作业所需要的网络信息, hosts文件, headless service等
    svc: []
  # 如果有pod被 杀死, 重启整个作业
  policies:
    - event: PodEvicted
      action: RestartJob
  tasks:
    - replicas: 1
      name: mpimaster
      # 当 mpiexec 结束, 认识整个mpi作业结束
      policies:
        - event: TaskCompleted
          action: CompleteJob
      template:
        spec:
          # Volcano 的信息会统一放到 /etc/volcano 目录下
          containers:
            - command:
                - /bin/sh
                - -c
                - |
```

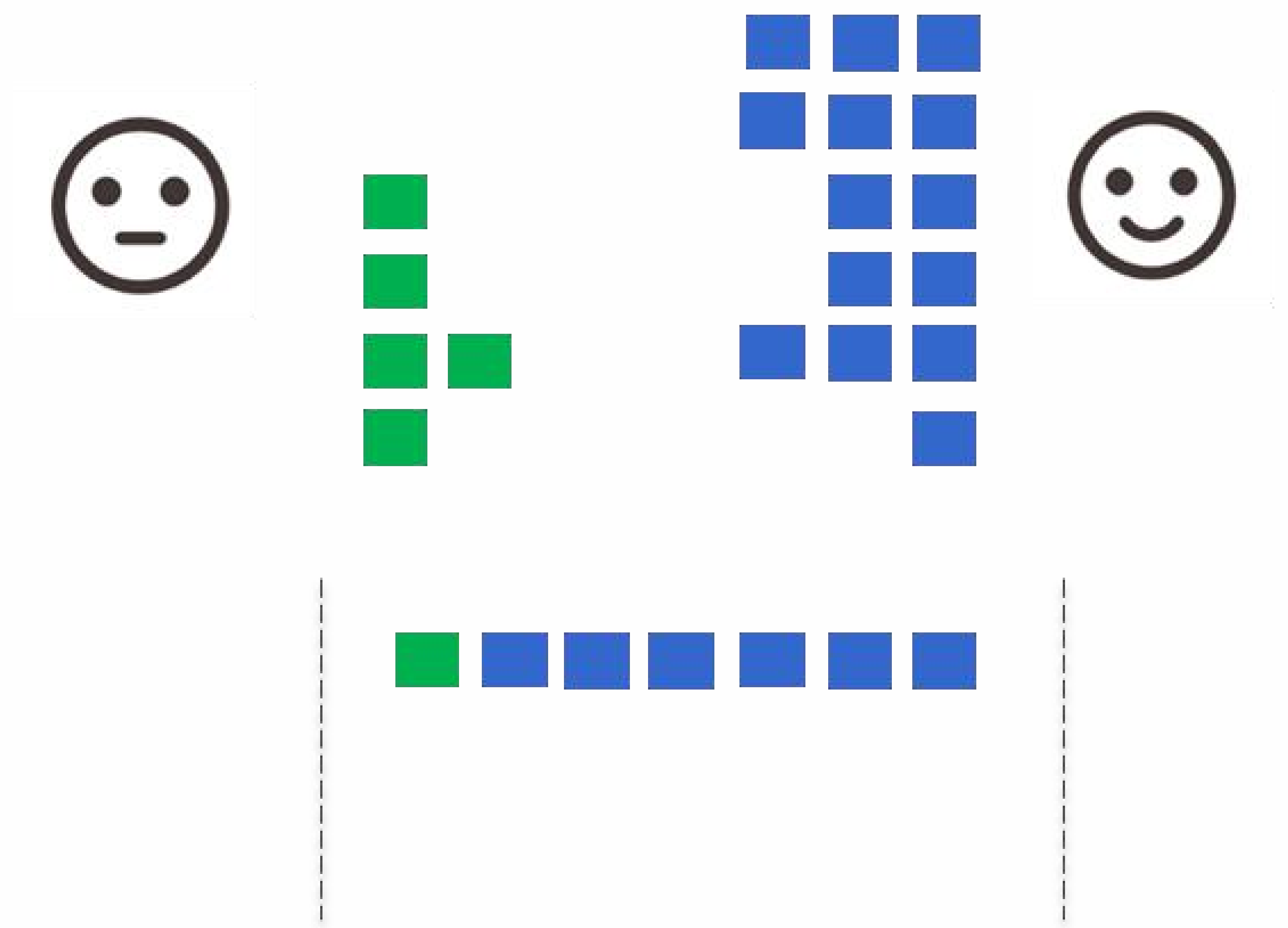


```
Pods:
--------------------------------
NAME                                  READY   STATUS      RESTARTS   AGE
lm-mpi-job-mpimaster-0                0/1     Completed   3          2m
spark-operator-sparkoperator-f78854b64-rh52d   1/1     Running     0          1d


Volcano Jobs:
--------------------------------
Name         Creation              Phase       JobType     Replicas    Min    Pending    Running    Succeeded
lm-mpi-job   2019-06-19 20:55:33   Completed   MPI         3           3      0          0          1

m00483107@M00483107 MINGW64 /d/workspace/src/volcano.sh/volcano/docs/samples/kubecon-2019-china/mpi-sample (kub
$ kc logs lm-mpi-job-mpimaster-0
Warning: Permanently added 'lm-mpi-job-mpiworker-0.lm-mpi-job,172.16.0.22' (ECDSA) to the list of known hosts.
Warning: Permanently added 'lm-mpi-job-mpiworker-1.lm-mpi-job,172.16.0.46' (ECDSA) to the list of known hosts.
Hello world from processor lm-mpi-job-mpiworker-0, rank 0 out of 2 processors
Hello world from processor lm-mpi-job-mpiworker-1, rank 1 out of 2 processors
```

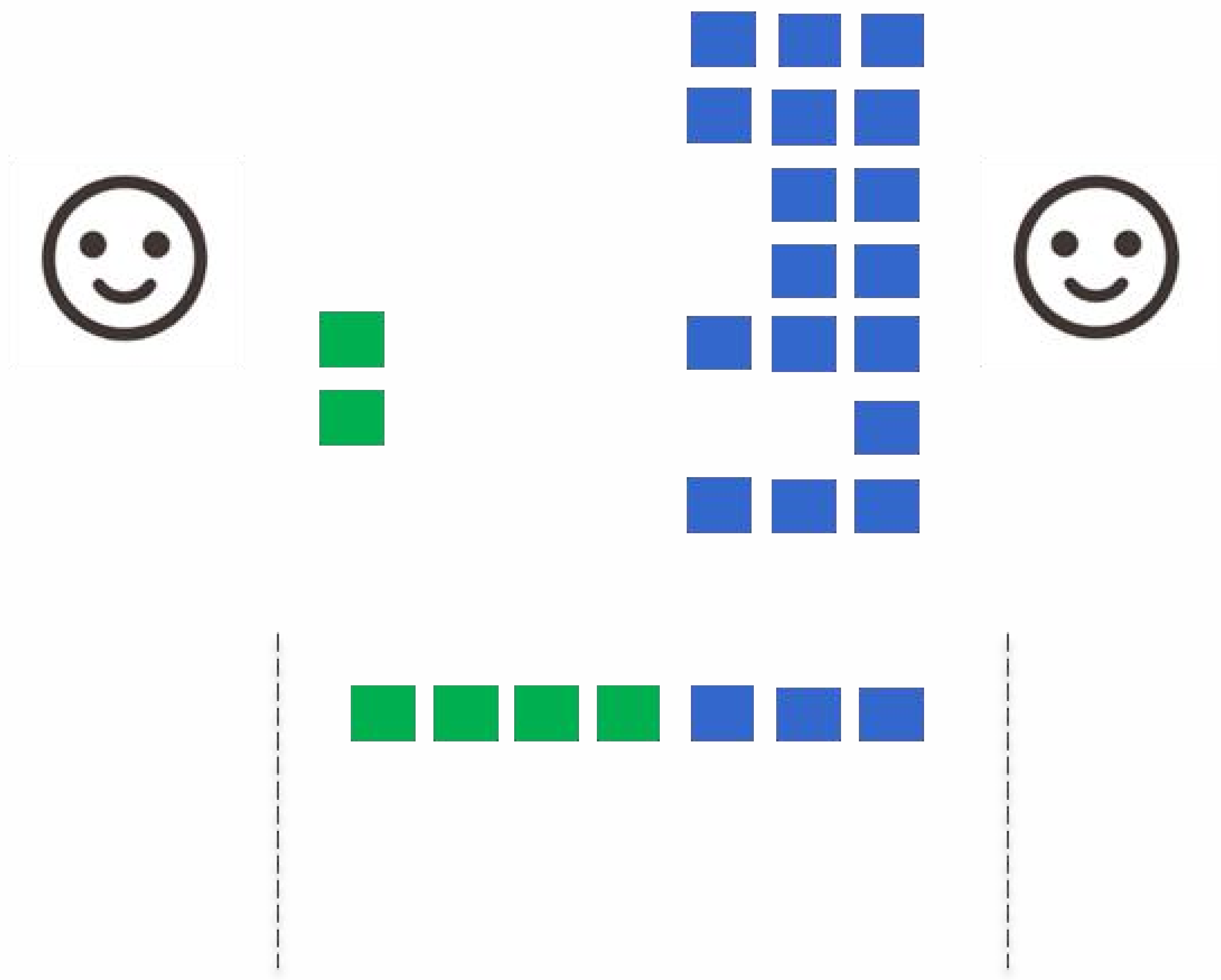
- The workers are deleted by job controller because of sshd
- The pod of mpiexec/mpirun will not be deleted for output
- The pod of mpiexec/mpirun may restart few times because of network setup delay

```
m00483107@M00483107 MINGW64 /d/workspace/src/volcano.sh/volcano/docs/samples/kubecon-2019-china/mpi-sample (kubecon-2019-china-demo)
$ watch.sh
Nodes:


---------------------------------------------------------------------------------------------------------------
                | 192.168.23.24          | 192.168.51.67          | 192.168.57.163         | 192.168.63.50          |
---------------------------------------------------------------------------------------------------------------
 Allocatable    | cpu: 3920m             | cpu: 3920m             | cpu: 3920m             | cpu: 3920m             |
                | mem: 6265300Ki         | mem: 6265300Ki         | mem: 6265300Ki         | mem: 6265300Ki         |
---------------------------------------------------------------------------------------------------------------
 Idle           | cpu: 2420m             | cpu: 1920m             | cpu: 3420m             | cpu: 2920m             |
                | mem: 3643860Ki         | mem: 4168148Ki         | mem: 5741012Ki         | mem: 4168148Ki         |
---------------------------------------------------------------------------------------------------------------



Pods:
-------------------------------
NAME                                           READY   STATUS    RESTARTS   AGE
spark-operator-sparkoperator-f78854b64-rh52d   1/1     Running   0          1d


Volcano Jobs:
-------------------------------
No resources found

m00483107@M00483107 MINGW64 /d/workspace/src/volcano.sh/volcano/docs/samples/kubecon-2019-china/mpi-sample (kubecon-2019-china-demo)
$
```

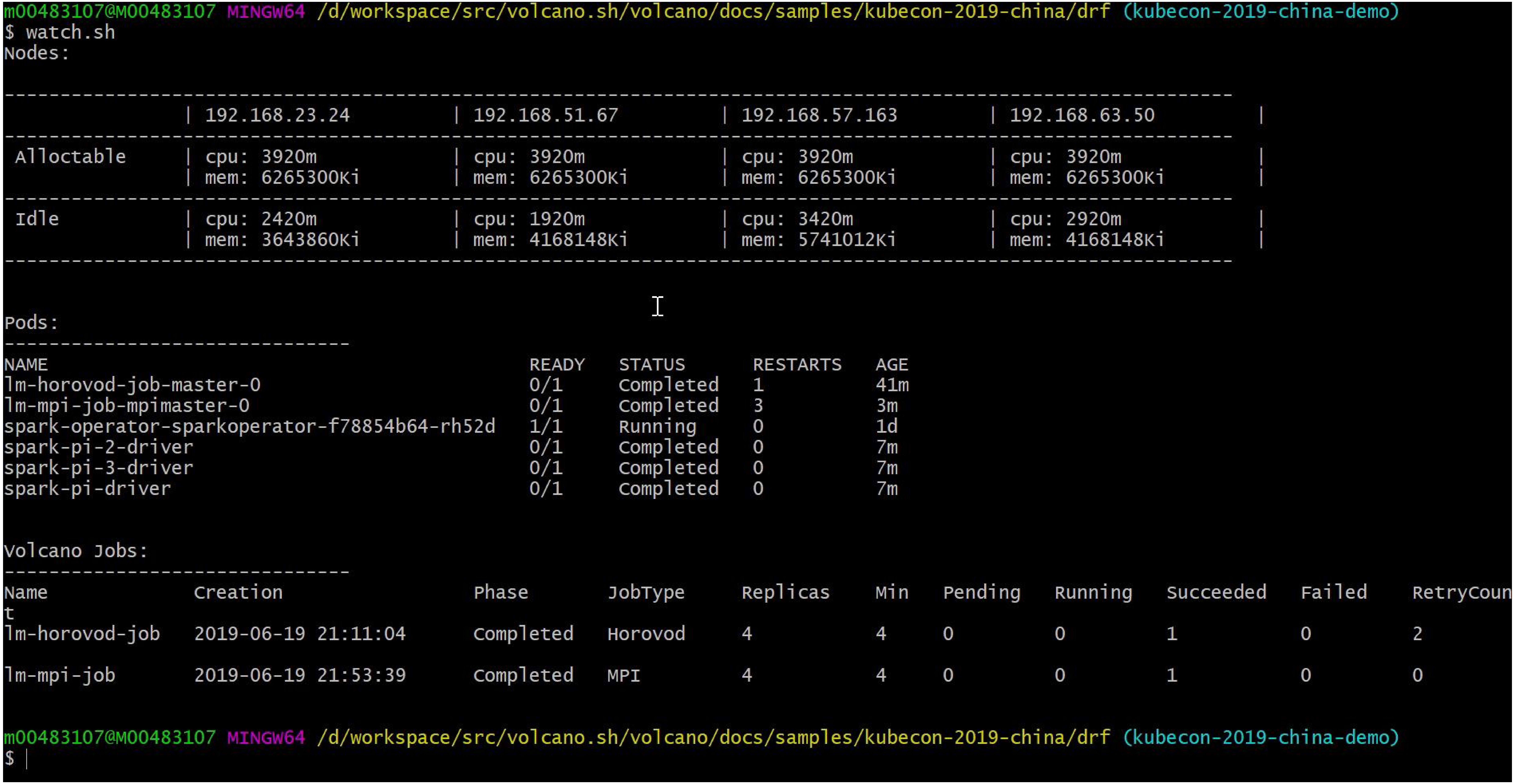
Scenarios: Faire Share



The more workload, the more resources???

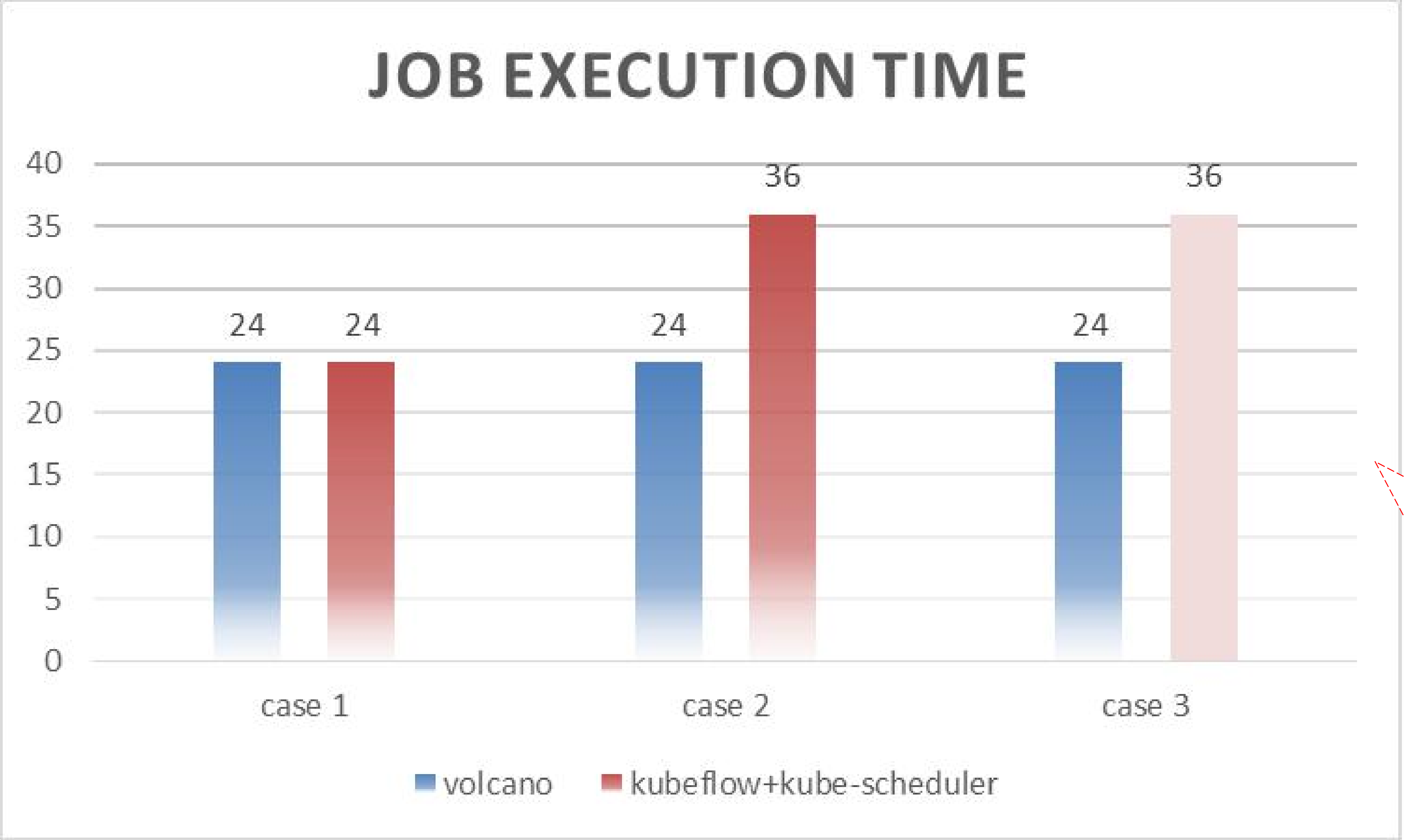


Share resources by weight !!!



```
m00483107@M00483107 MINGW64 /d/workspace/src/volcano.sh/volcano/docs/samples/kubecon-2019-china/drf (kubecon-2019-china-demo)
$ watch.sh
Nodes:

-------------------------------------------------------------------------------------------------------------
                | 192.168.23.24         | 192.168.51.67         | 192.168.57.163        | 192.168.63.50         |
-------------------------------------------------------------------------------------------------------------
 Allocatable    | cpu: 3920m            | cpu: 3920m            | cpu: 3920m            | cpu: 3920m            |
                | mem: 6265300Ki        | mem: 6265300Ki        | mem: 6265300Ki        | mem: 6265300Ki        |
-------------------------------------------------------------------------------------------------------------
 Idle           | cpu: 2420m            | cpu: 1920m            | cpu: 3420m            | cpu: 2920m            |
                | mem: 3643860Ki        | mem: 4168148Ki        | mem: 5741012Ki        | mem: 4168148Ki        |
-------------------------------------------------------------------------------------------------------------


Pods:
-------------------------------
NAME                                            READY   STATUS      RESTARTS   AGE
lm-horovod-job-master-0                         0/1     Completed   1          41m
lm-mpi-job-mpimaster-0                          0/1     Completed   3          3m
spark-operator-sparkoperator-f78854b64-rh52d    1/1     Running     0          1d
spark-pi-2-driver                               0/1     Completed   0          7m
spark-pi-3-driver                               0/1     Completed   0          7m
spark-pi-driver                                 0/1     Completed   0          7m


Volcano Jobs:
-------------------------------
Name             Creation                  Phase       JobType     Replicas    Min   Pending   Running   Succeeded   Failed    RetryCoun
t
lm-horovod-job   2019-06-19 21:11:04       Completed   Horovod     4           4     0         0         1           0         2

lm-mpi-job       2019-06-19 21:53:39       Completed   MPI         4           4     0         0         1           0         0


m00483107@M00483107 MINGW64 /d/workspace/src/volcano.sh/volcano/docs/samples/kubecon-2019-china/drf (kubecon-2019-china-demo)
$
```

Gang-scheduling: Job Execution Time

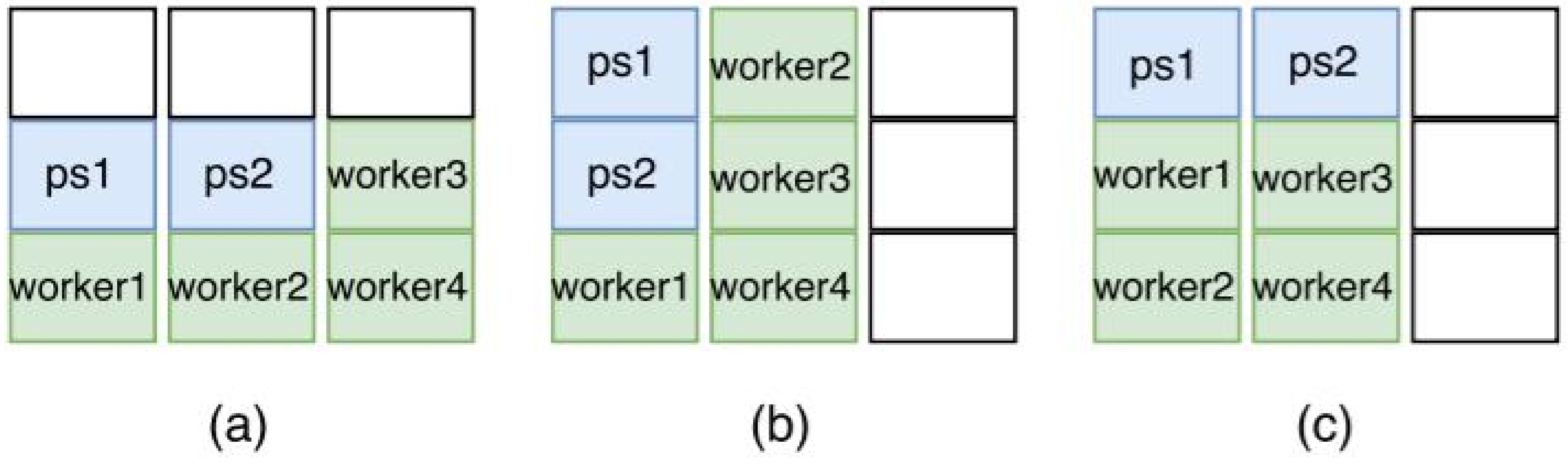


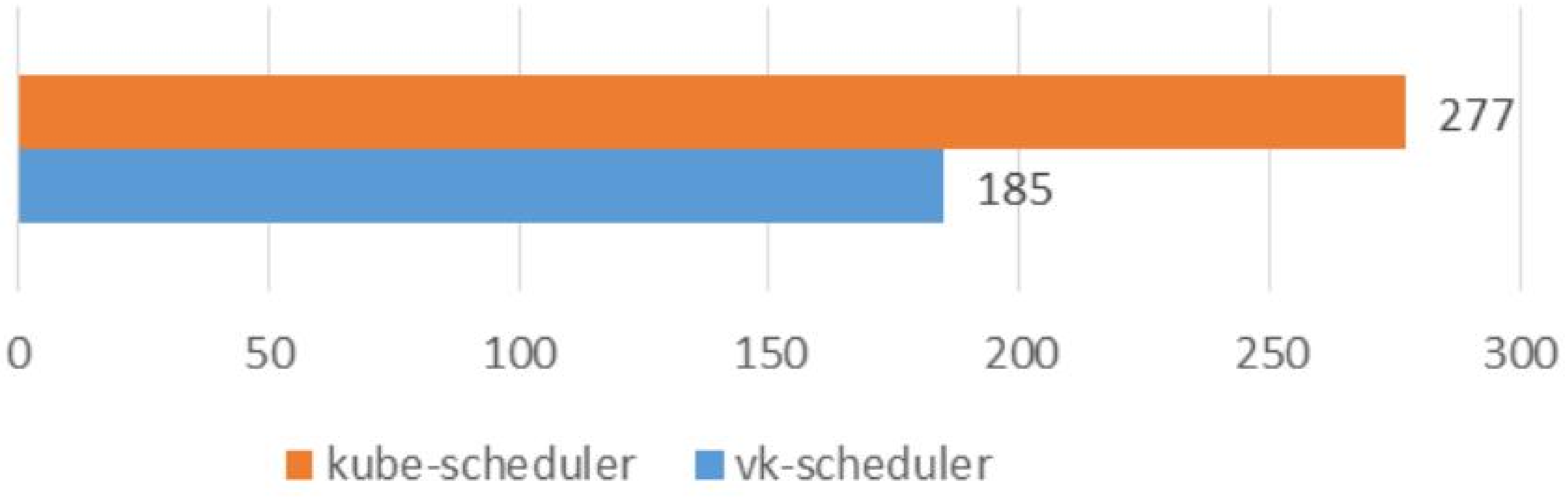
- Case 1: 1 job with 2ps + 4workers
- **Case 2: 2 jobs with 2ps + 4workers**
- **Case 3: 5 jobs with 2ps + 4workers**

- No enough resource for 2 Jobs to run concurrently; one of them **wasting** resources without Gang-Scheduling !
- 2 of 5 jobs was finished because of deadlock (+20 hours)

http://status.openlabtesting.org/builds?project=theopenlab%2Fvolcano

Task-Topology + Binpack





- The execution time of 3 jobs in total; 2ps + 4workers for each job
- The execution time is unstable when tested by default scheduler
- The improvement dependent on data exchanges between pods
- Task-topology within a Job also improved scheduler’s performance
- Open Source at volcano-sh/volcano#272

Reference: “Optimus: An Efficient Dynamic Resource Scheduler for Deep Learning Clusters”

Integrations

Framework	Status	API
MPI	Done	Volcano Job
Horovod	Done	Volcano Job
Kubeflow/tf-operator	Done	PodGroup
Kubeflow/arena	Done	Volcano Job
Spark-Operator	On-going	PodGroup
Cromwell	On-going	Volcano Job
PaddlePaddle	On-going	Volcano Job
...	On-going	Volcano Job / PodGroup

Pipeline

- GPU Share/Topology
- Job Management
- Queue Management
- Hierarchical Queue
- Preemption/Reclaim
- ......

Call for Contribution

Geekbang> 极客邦科技 | InfoQ 极客传媒

InfoQ官网 全新改版上线

促进软件开发领域知识与创新的传播

精选内容

华为面试改革重编程：黑客攻击 Git 私有库索要赎金 | Q 新闻

华为面试改革，增加 90 分钟网上编程和 30 分钟当面编程；德国摧毁全球第二大暗网交易平台；黑客攻击 Git 开发者私...

作者：赵钰莹

文化 & 方法　开源　华为

世界第三移动操作系统：KaiOS 的前世今生

作者：Nick Summers | 译者：无明

移动　硬件　前端

技术债务就像俄罗斯方块：我如何通过修改代码来搞定它？

作者：Eric Higgins | 译者：无明

文化 & 方法　语言 & 开发

Facebook：全球最大社交网络，向未知转型

作者：德川

文化 & 方法　企业动态

推荐内容

热点

1 华为面试改革：增加 90 分钟网上编程 30 分钟当面编程环节，充分考察实际动...

2 快速上手几个 Linux 命令：每家公司都有自己的黑话

3 京东刘海锋：过去十年架构领域最重要的三个变化

4 这一天还是来了：亚马逊利用 AI 监控员工工作效率，并生成解雇指令

5 8 年上市、3 年微服务改造、2 年中台战，回顾这家公司 12 年技术架构转型路

6 埃森哲被诉：两年花两亿，连个可用的网站都不能交付？

订阅 InfoQ 每周精要
资深编辑为你推荐一周好内容

查看往刊 >

QCon 全球软件开发大会 | 广州站
全领域实战中的防坑指南

华为面试改革重编程；黑客攻击 Git 私有库索要赎金 | Q 新闻

作者：赵钰莹

华为面试改革，增加 90 分钟网上编程和 30 分钟当面编程；德国摧毁全球第二大暗网交易平台；黑客攻击Git 开发者私有库，索要赎金。

Apache Spark 实现可扩展日志分析，挖... 大潜力

作者：Dipanjan Sarkar | 译者：...

关注InfoQ网站
第一时间浏览原创IT新闻资讯

免费下载迷你书
阅读一线开发者的技术干货

THANKS

Geekbang 极客邦科技 InfoQ