

高效AI基础设施：业务灵活弹性 + 云端 GPU即时供应

莫源 / 阿里云 & 姜伟 / CloudPilot AI

Content 目录

- 01** 业界挑战
- 02** 解决方案
- 03** 架构&实现
- 04** 总结
- 05** Q&A

AI

Part 01

业界挑战

AI



业界挑战



Stable Diffusion



deepseek



blender®

成本高

云上GPU费用较高，
如，AWS us-east-1 T4的花费高达
\$383.98/月。

资源确定性差

传统集群伸缩关键时刻可能无法弹性足够的计算资源，
导致无法处理业务突发峰值。

业务/节点弹性不足

无法及时根据业务需求扩缩容导致资源的浪费/不足。

启动速度受限

GPU业务通常镜像或所需运行文件巨大，导致启动速度较慢。

Part 02

解决方案

AI



解决方案

业务弹性伸缩

KServe配置简单，根据业务的请求并发量弹性业务，确保高效响应。

节点弹性伸缩

Karpenter 提供节点资源弹性，最大化资源利用率，多实例类型选择提升资源供给确定性，同时支持Spot实例以降低成本。

数据读取加速

利用Fluid的分布式缓存，高性能读取业务初始化文件。



镜像加速

利用云厂商/开源的延迟加载技术（如 AWS SOCI, nydus等），极速启动大镜像业务。

-如AWS T4, 相比按量, Spot实例能节省70%, 其他云厂商类似。
-GPU 业务通常涉及特定的容器镜像和模型参数, 例如运行大模型时需要使用vLLM 镜像和 DeepSeek 模型参数文件。

-根据统计显示, 在镜像启动的时候, 平均只有 6.4%的数据被真正使用到。
-Spot 实例是未被使用的资源, 价格低至 On-Demand 实例的 20-30%, 但是存在随时被中断的风险。
-Spot 价格和中断率查询工具: Spot Insights <https://spot.cloudpilot.ai/>

解决方案



ecs.r6.2xlarge

1934.5/月
按量

1272.0/月
包月

406.6/月
竞价(spot)

ecs.gn7i-c8g1.2xlarge/NVIDIA A10

6958.7/月
按量

4575.6/月
包月

2436.0/月
竞价(spot)

数据来源：

- <https://spot.cloudpilot.ai/alibabacloud?instance=ecs.r6.2xlarge#region=cn-beijing>

- <https://www.aliyun.com/price/product?spm=a2c4g.11186623.0.0.2a4e1f9a2Za2KO#/ecs/detail/vm>

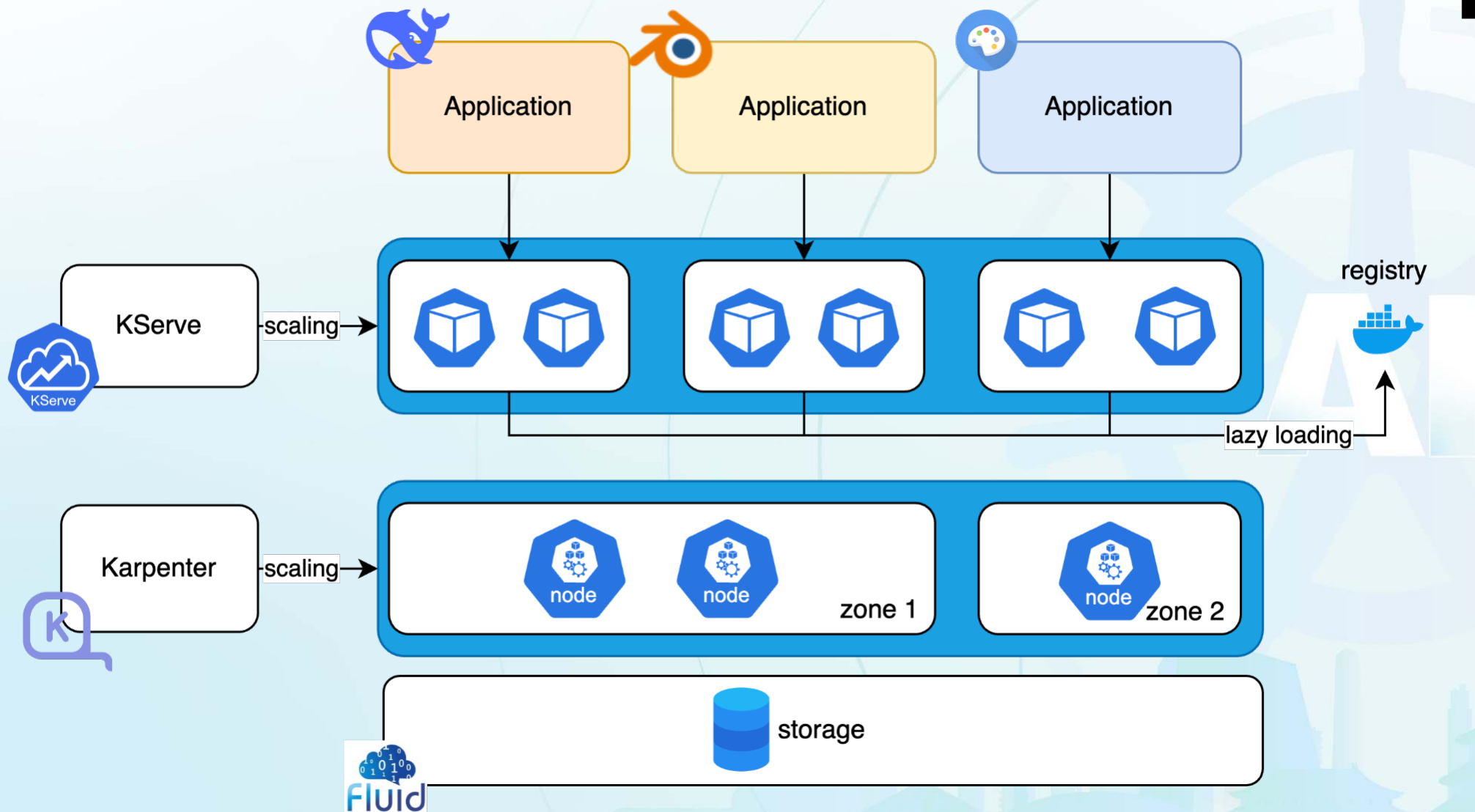
Part 03

架构&实现

AI



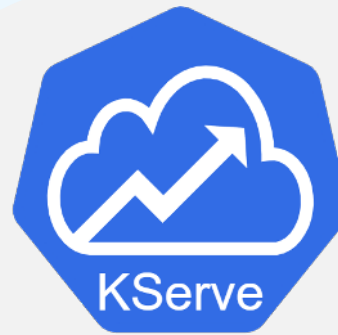
总体架构



KServe介绍

支持多种推理框架如
Pytorch、TF、
XGBoost、HF

灰度发布



KServe是一个开源的云原生
模型服务平台，旨在简化在
K8s上部署和运行机器学习模
型的过程。

弹性扩容能力

统一数据面API

Github 地址：
<https://github.com/kserve/kserve>

KServe介绍



Transformer Service

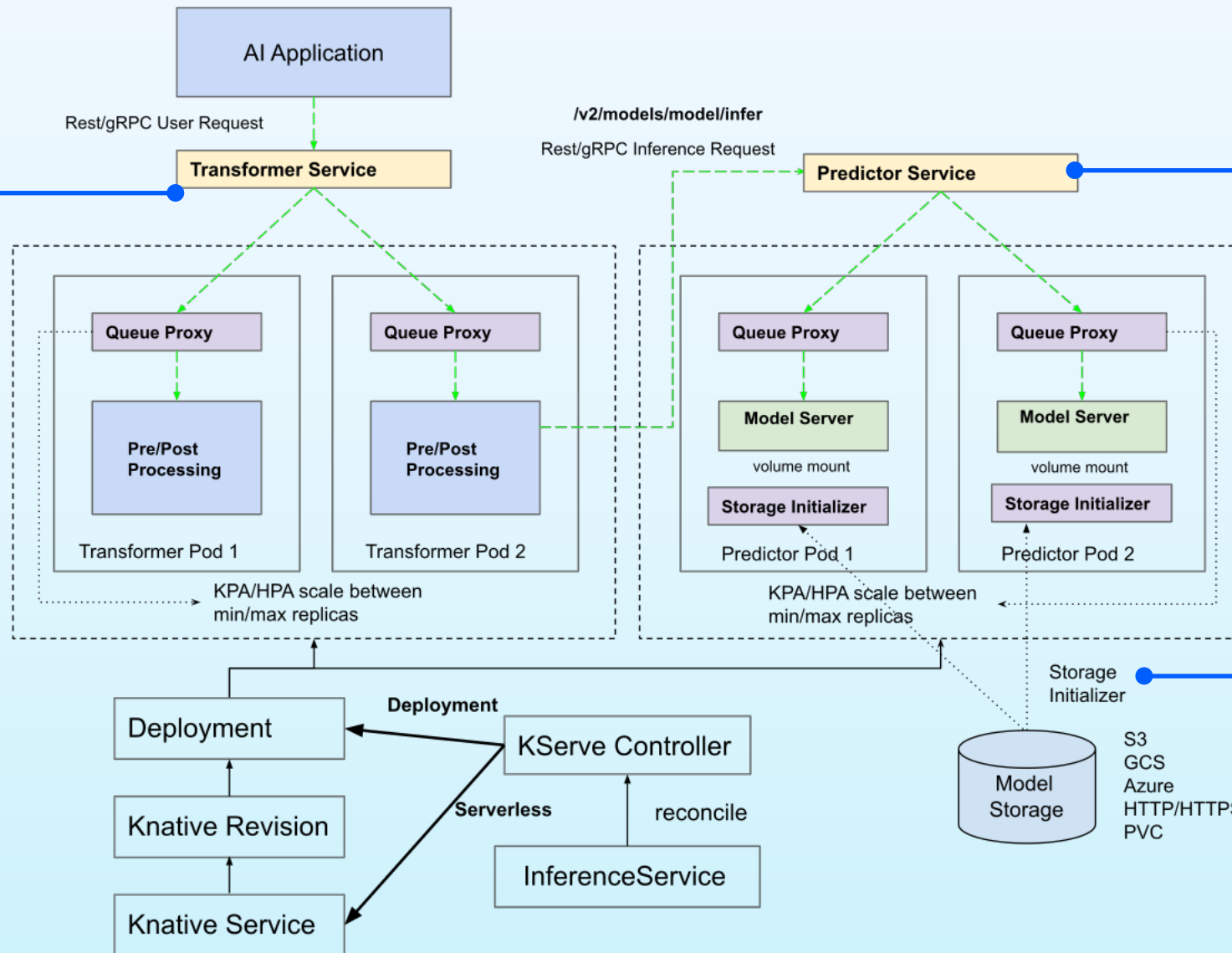
负责数据的Pre/Post处理

Predictor Service

负责提供推理服务。本次分享主要使用此Service

Storage Initializer

负责加载应用所需的数据/模型。目前支持S3、GCS、Azure云盘、HTTP/HTTPS网页链接及PVC



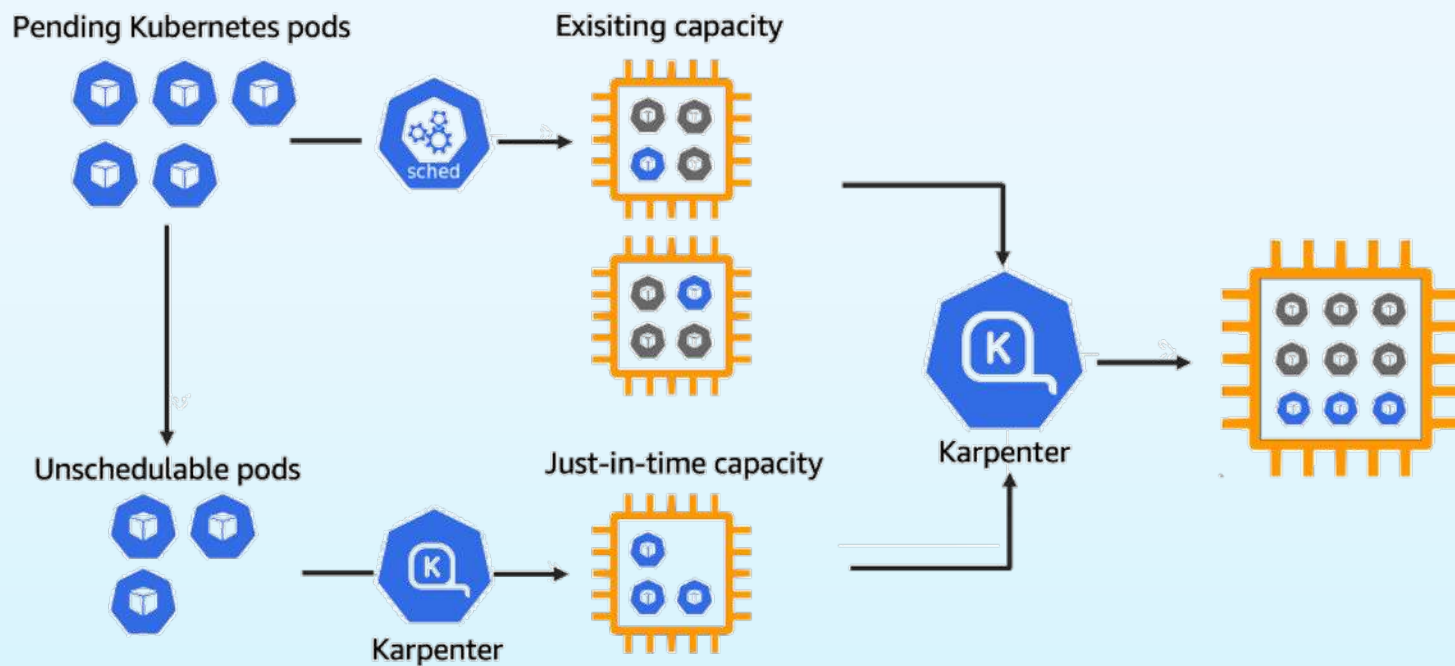
KServe介绍



```
apiVersion: serving.kserve.io/v1beta1
kind: InferenceService
metadata:
  name: deepseek-r1
spec:
  predictor:
    containerConcurrency: 10
    model:
      modelFormat:
        name: huggingface
      args:
        - --model_name=deepseek-r1
        - --model_id=deepseek-ai/DeepSeek-R1-Distill-
Qwen-32B
      env:
        - name: HF_TOKEN
          valueFrom:
            secretKeyRef:
              name: hf-secret
              key: HF_TOKEN
              optional: false
    resources:
      limits:
        nvidia.com/gpu: "1"
      requests:
        nvidia.com/gpu: "1"
```

**KServe 通过
InferenceService CR 即可快速配置服务，拉起服务和弹性服务。**

Karpenter介绍



Karpenter 架构图

专为 Kubernetes 设计
自动管理集群的节点扩缩容。

无需节点组
直接根据集群需求创建节点。

灵活选型
可自动选择最适合的实例类型和规格。

智能资源匹配
如 Pod 需要 GPU 资源，
Karpenter 自动弹性 GPU 实例。

Karpenter介绍



```
apiVersion: karpenter.sh/v1
kind: NodePool
metadata:
  name: ecsnodepool
spec:
  disruption:
    budgets:
      - nodes: 95%
    consolidationPolicy: WhenEmptyOrUnderutilized
    consolidateAfter: 1m
  template:
    spec:
      requirements:
        - key: karpenter.k8s.alibabacloud/instance-
category
          operator: In
          values: ["ecs.g5", "ecs.g6"]
        - key: kubernetes.io/arch
          operator: In
          values: [ "amd64" ]
        - key: kubernetes.io/os
          operator: In
          values: ["linux"]
        - key: karpenter.sh/capacity-type
          operator: In
          values: ["spot"]
      nodeClassRef:
        group: "karpenter.k8s.alibabacloud"
        kind: ECSNodeClass
        name: defaultnodeclass
```

NodePool 关键特性

- 用于管理节点 (Node) 配置的一种 Karpenter 资源。
- 声明式定义：设定节点的硬件规格、容量等参数。
- 精细化控制：支持标签、污点等配置。
- 自动化管理：优化节点创建与生命周期管理。

Karpenter介绍

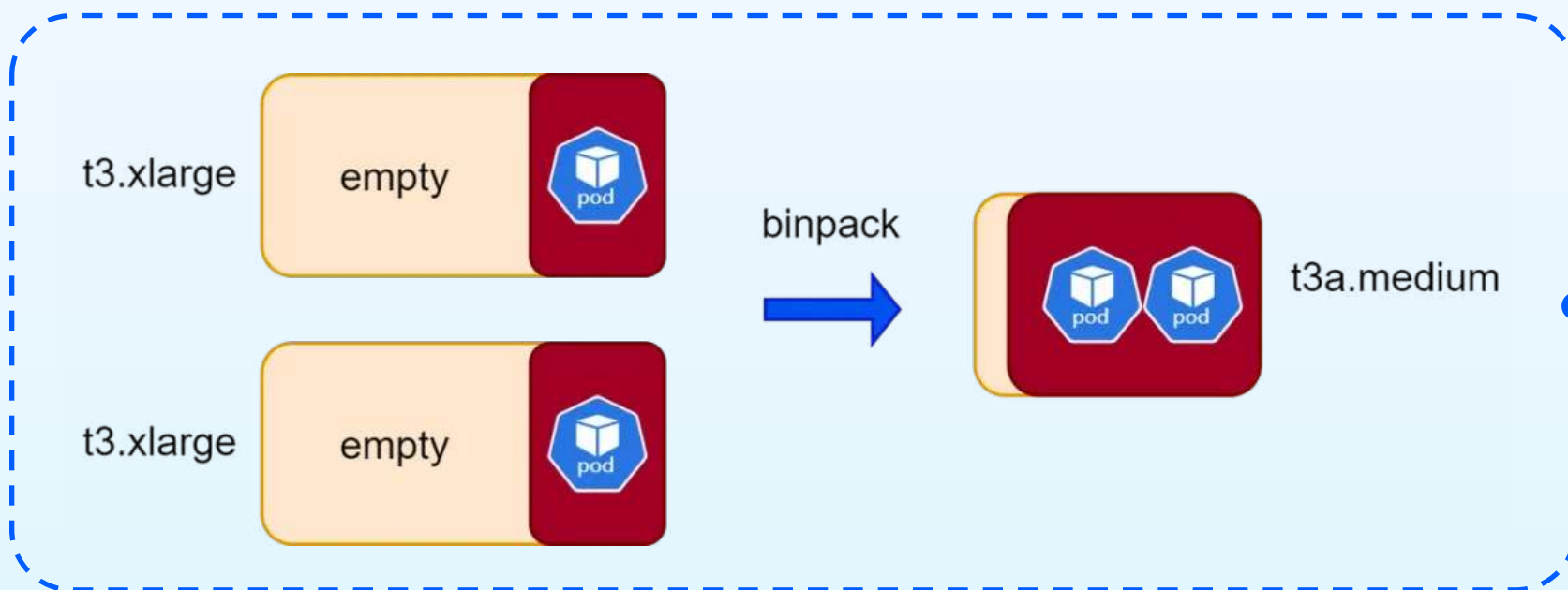


NodeClass 关键特性

- 用于定义节点 (Node) 启动配置的一种资源类型。
- 声明式管理：设定 AMI 镜像、磁盘挂载、标签等属性。
- 灵活配置：支持启动参数等多维度配置。

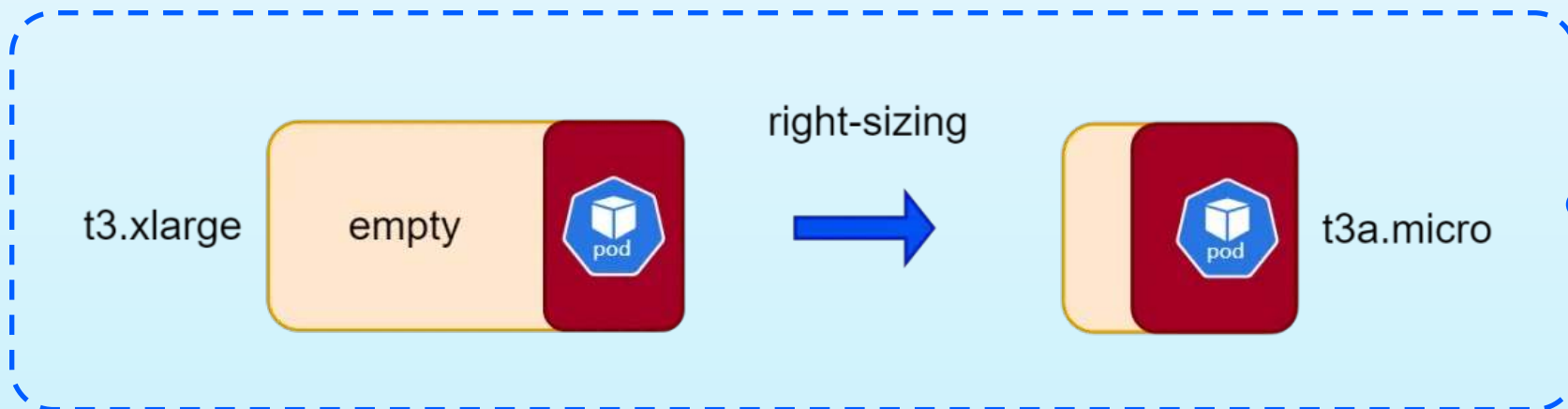
```
apiVersion: karpenter.k8s.alibabacloud/v1alpha1
kind: ECSNodeClass
metadata:
  name: defaultnodeclass
spec:
  vSwitchSelectorTerms:
    - tags:
        karpenter.sh/discovery: "${CLUSTER_NAME}"
  securityGroupSelectorTerms:
    - tags:
        karpenter.sh/discovery: "${CLUSTER_NAME}"
  imageSelectorTerms:
    - alias: ContainerOS
```

Karpenter介绍



Binpack

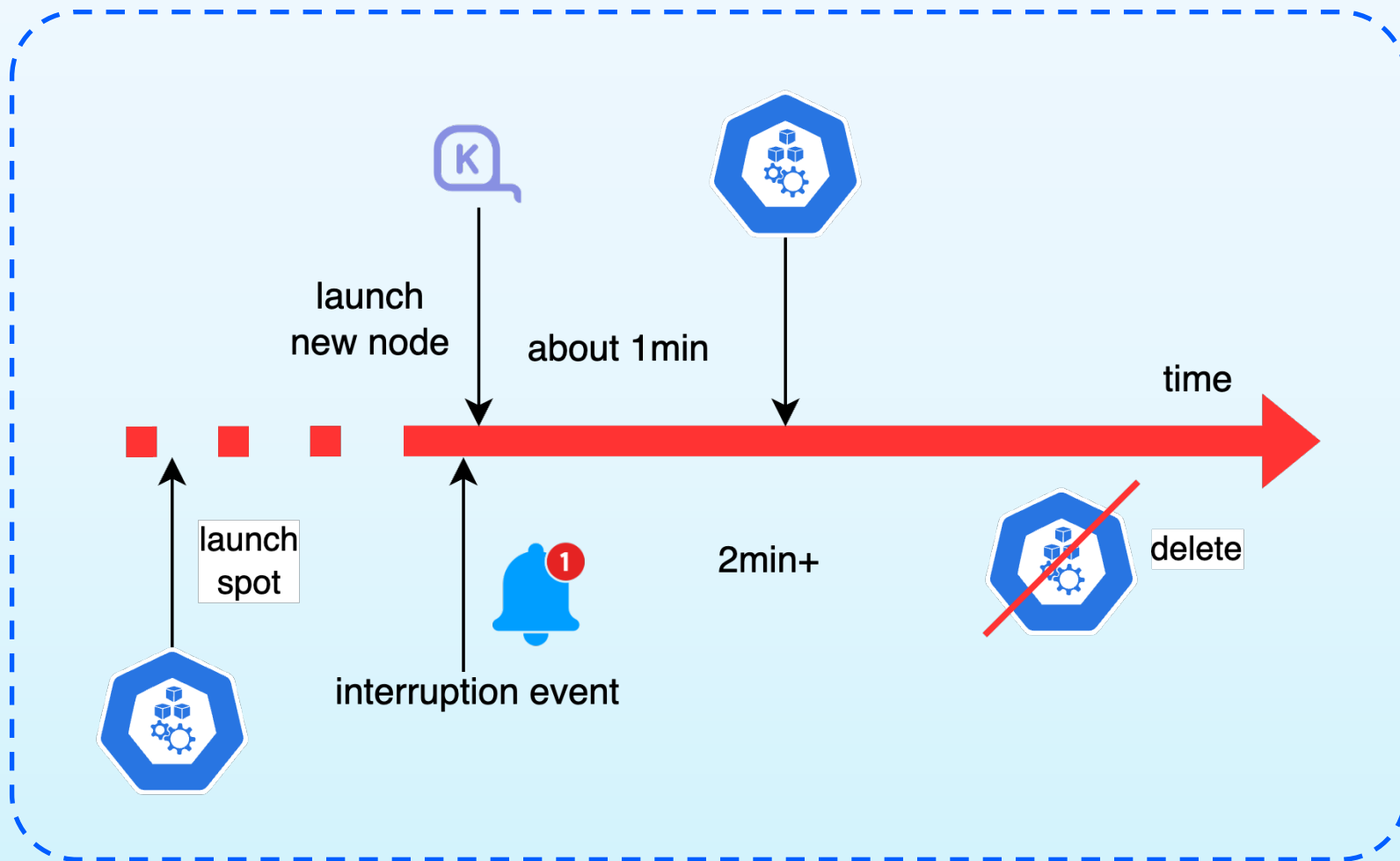
最大化节点利用率，
尽可能将 Pod 填充
到较少的节点上



Rightsizing

为 Pod 选择最合适
的实例类型，避免
资源过剩或不足

Karpenter介绍

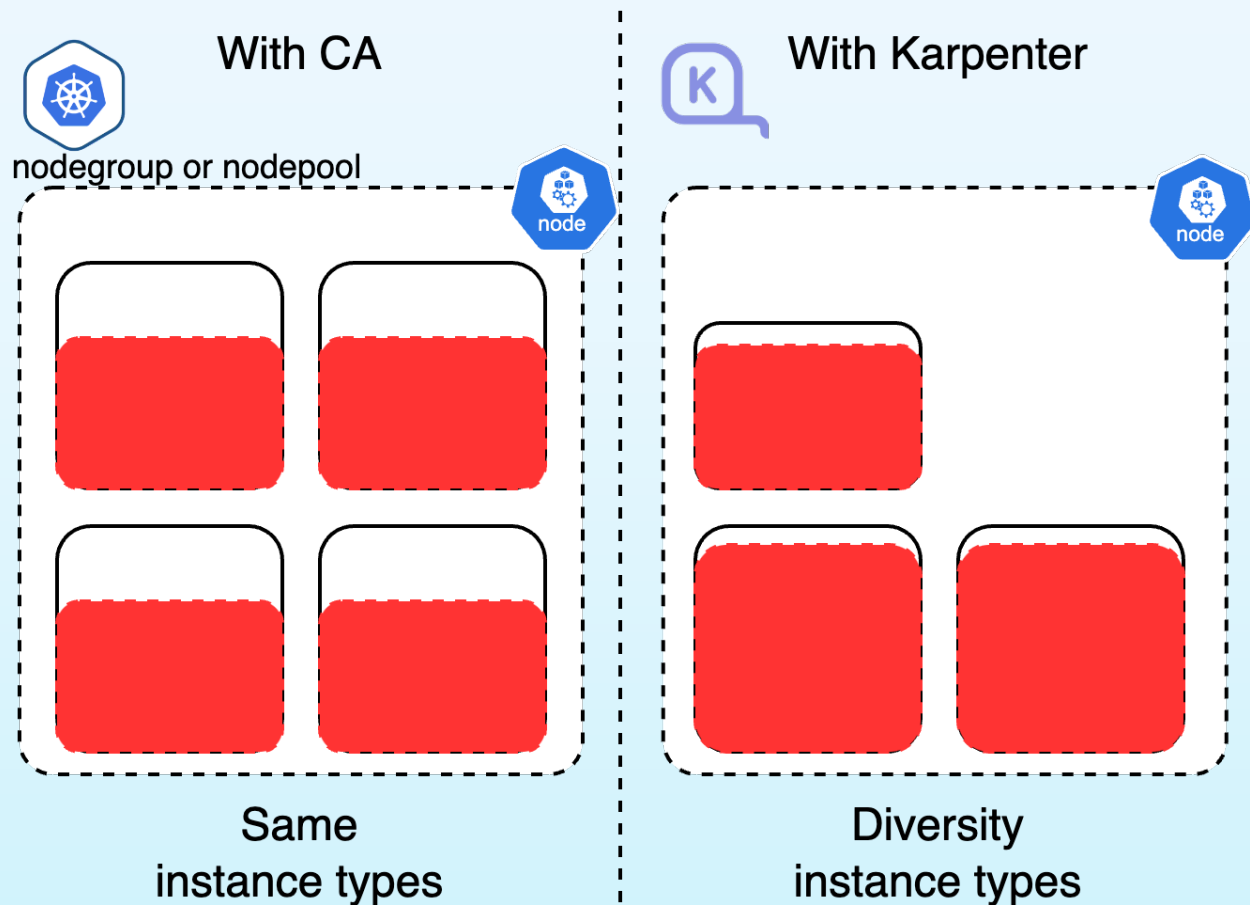


Spot 实例中断处理流程

- Spot 实例启动：
Karpenter 创建新节点。
- 中断通知：实例收到 2 分钟终止警告。（不同云厂商中断通知时间不一样）
- 迁移 & 替换：Karpenter 重新调度工作负载。
- 实例删除：释放即将中断的节点。

- 不同云厂商中断通知时间节点不一样，AWS中断提前2min通知，阿里云提前5min通知。CloudPilot AI能够选择中断率更低的实例，同时自动替换将被中断的实例。

Karpenter介绍



相比 Cluster AutoScaler , Karpenter有如下优势：

更灵活：

- NodePool 定义弹性节点类型
- NodeClass 设定节点属性（如镜像）
- 无需依赖节点组，配置更简单

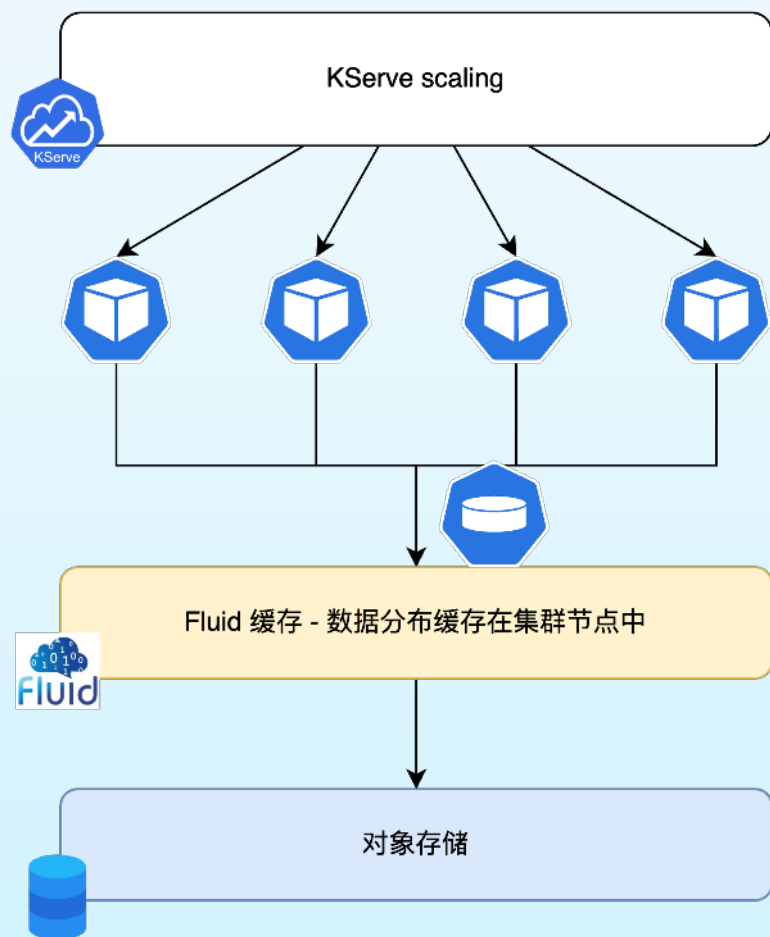
更快扩缩容：

- 直接调用云厂商 VM API
- 比 CA 多层 API 交互，响应更快

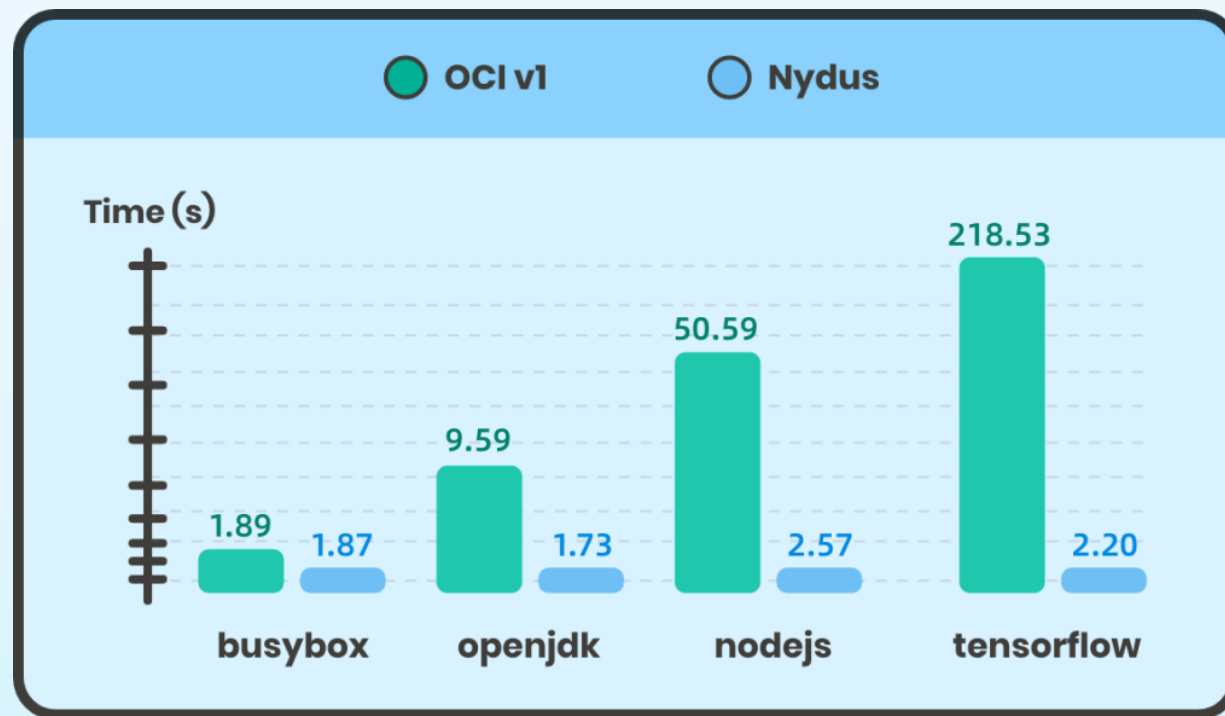
更高性价比：

- 稳定利用 Spot 实例
- 支持更多节点类型，优化成本与资源利用

存储 & 镜像加速



- 实测 9GiB Model Size / 从46s缩减到26s



Pod Creation to Container Start

Part 04

总结

AI



总结：高效AI云上基础设施



Instance Pricing

US East (N. Virginia)

On Demand
\$383.98

us-east-1

Spot Min
\$120.304

use1-az2

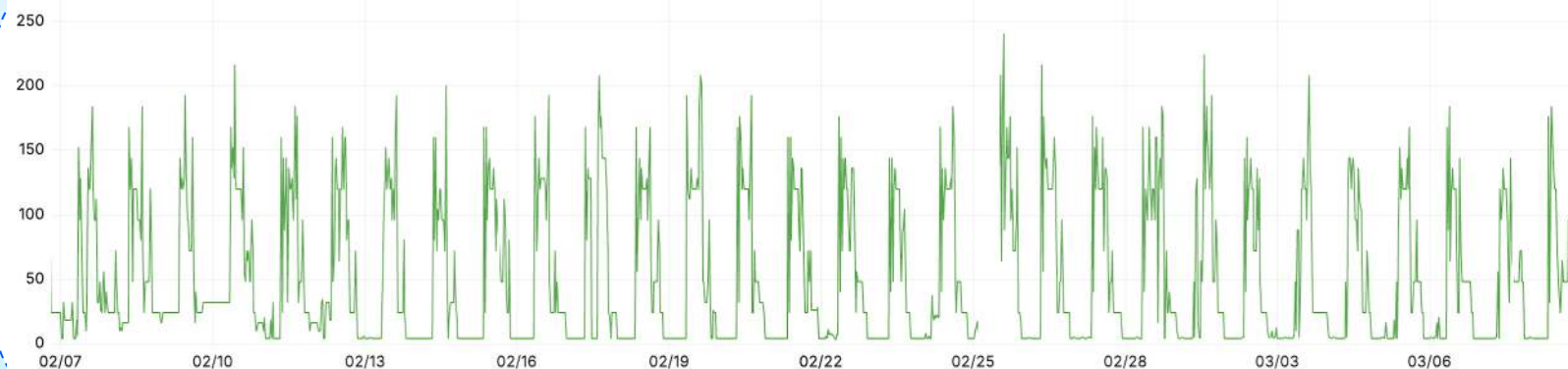
Spot Max
\$129.721

use1-az1

-68.7%

-66.2%

Per Month



- 利用Spot，实现接近**70%**的成本降低 (g4dn.xlarge, NVIDIA T4)
- 数据来源：Spot Insights (spot.cloudpilot.ai)
- 利用Karpenter，实现底层GPU资源按需弹性，实现计算弹性免运维

总结



Karpenter 负责云上节点资源的自动扩缩容，支持 **CPU** 和 **GPU**

节点资源，提供以下卓越特性：

高确定性的资源弹性

通过简化节点资源配置，支持将数百种实例类型纳入资源池。在特定实例类型资源紧张时，Karpenter 能自动切换到其他实例，确保计算资源弹性调度成功。

集群高资源利用率

通过主动 binpacking 和节点替换，提升集群的资源利用率，最大化节点的计算价值。

极致成本优化

支持使用 Spot 实例，针对可迁移节点的工作负载，实现极高的性价比，在保证服务稳定的同时大幅降低运行成本。

配置简化

通过 CR (Custom Resource , 自定义资源) ，快速定义弹性资源池属性，实现资源扩缩容的高效自动化。

Part 05

Q&A

AI



Q&A

Karpenter Cloud Provider Repo :

阿里云

<https://github.com/cloudpilot-ai/karpenter-provider-alibabacloud>

GCP

<https://github.com/cloudpilot-ai/karpenter-provider-gcp>

AWS

<https://github.com/aws/karpenter-provider-aws>



扫码关注公众号
获取更多
云成本优化实践



扫码添加小助手
加入「Karpenter 技
术交流群」

Thanks.

AI

