

裸机多计算可用区优化需求文档

需求概要	既有的多可用区实现及调度为前端实现, 在大规模的场景下不满足性能, 实时性需求, 需要修改为后端实现
隶属产品线	计算产品线
商务报价模式	待定
部署方式	现有组件在线升级
涉及组件/模块/服务	ironic、nova
产品版本	设计：Baremetal 6.1.2 ； 开发启动：Baremetal 6.1.2 ； 发布：Baremetal 6.1.2
文档状态	Draft
产品负责人	@李向军 @杨鹏毅 (Unlicensed)
界面设计师	@吴奇骏
技术架构负责人	@于尚斌
后端负责人	@王亚 (Unlicensed) @于尚斌
前端负责人	@翟元杰
测试负责人	@刘伟 @廖正勇
JIRA Link	
UE Link	暂定
UI Link	暂定
前置需求文档(如有)	

修订记录

修改日期	修改内容
2022.6.17	初版

一、需求背景

现状

目前裸机的多计算可用区及相关调度均由前端实现, 在大规模的场景下, 性能和实时性等不满足要求. 并且无法做到 nova-compute-ironic 组与 ironic-conductor 组的对应, 后端的 nova/ironic 间数据同步可能会出现非预期情景.

客户

邮储等重点客户都有裸金属大规模部署及多可用区划分的场景与需求

行业

二、价值综述

客户价值

商业价值

三、名词定义

四、使用场景

在环境中网络可用区划分,或需要依据裸金属特性进行划分时,可通过划分计算可用区来实现资源隔离

五、竞品&社区分析

社区现状

社区中通过 ironic conductor group 实现该功能,通过 ironic-conductor 中配置 conductor_group ,nova-compute 中配置 partition_key 和 peer_list 实现 ironic conductor 组与 nova compute 组的对应.再通过为裸金属节点指定 conductor_group 实现节点与 conductor 的映射,具体工作流程参见 [Openstack Master 版本实现](#)

竞品分析

六、业务流程

- 1. ironic-conductor-init 按照节点 label 生成 conductor_group 配置(现有逻辑)
- 2. ark-ironic 中 nova-compute-ironic statefulset 服务的生成方式与 ironic-conductor 相似,通过 .Values.pod.deploy.ironic_conductor.name 来生成若干个三副本的 statefulset,命名规则与 ironic-conductor 类似,形如 nova-compute-ironic-{conductor_group_name}
- 3. nova-compute-ironic-init 获取 pod 名称,并以中间的 {conductor_group_name} 作为配置项 partition_key 的值.同时获取该 statefulset 的副本数,按照形如 nova-compute-ironic-{conductor_group_name}-0, nova-compute-ironic-{conductor_group_name}-1 的形式作为配置项 peer_list 的值.
- 4. 注册/编辑裸金属节点时,可用区和节点组的选择有所联动
 - a. 注册节点到空(无nova-compute-ironic服务)可用区: 允许选择节点组,但可选择的节点组需要未与可用区对应
 - b. 注册节点到非空((有nova-compute-ironic服务))可用区: 不允许选择节点组,自动切换到可用区对应的节点组上
 - c. 编辑节点到空(无nova-compute-ironic服务)可用区: 需要重新选择节点组,可选择的节点组需要未与可用区对应
 - d. 编辑节点到非空((有nova-compute-ironic服务))可用区: 不允许选择节点组,自动切换到可用区对应的节点组上
- 5. 可用区管理界面中,允许向可用区添加 nova-compute-ironic 服务,并具有相同 partition_key 配置(即 conductor_group 名称)的 nova-compute-ironic 服务应当同时只能在一个可用区中.

七、功能需求

序号	一级功能	二级功能	三级功能	功能描述	备注
----	------	------	------	------	----

八、非功能性需求

升级与进化需求

保证可以通过产品版本升级获得该功能，并保持功能可迭代，同时兼容存量裸机。

要点在于：

1. 当前多可用区通过在裸金属节点的 `extra` 中写 `reserve_availability_zone` 实现, 升级后, 需要将该裸金属节点所在的 nova-compute-ironic 服务添加到对应的可用区中

能力需求

- License 需求：支持license控制
- 版本管理需求：支持版本升级
- 安装部署需求：支持云产品形态安装部署
- 扩容/缩容能力需求：支持
- 计费需求：裸金属当前版本不支持计费功能

高可用需求

- 云主机高可用：不涉及云主机高可用
- 控制平面高可用：裸金属服务多副本支持

接口需求

- API接口：提供功能API和说明文档
- 硬件接口：不涉及硬件接口

性能需求

- 时间特性要求：满足平台通用时间要求
- 存储性能要求：满足平台通用存储性能要求
- 网络性能要求：满足平台通用网络性能要求

安全性需求

- 密码：平台和裸金属主机登录需要密码
- 容灾：裸金属当前版本不支持容灾
- 备份：裸金属当前版本不支持备份

权限需求

- 权限：支持管理员和项目用户权限
- 多区域：裸金属当前版本不支持多区域
- 配额：裸金属已支持配额管理

可运维性需求

- 日志：平台统一日志支持
- 监控：平台统一监控支持
- 告警：平台统一告警支持

硬件需求 [↗](#)

不涉及特殊硬件需求

九、产品边界与限制 [↗](#)

功能边界 [↗](#)

1. 向可用区添加 nova-compute-ironic 服务时, 必须要将同一组(相同 `partition_key` 配置)的 nova-compute-ironic 加入到一个可用区中, 移除时同理

升级边界：

存量裸金属节点需要迁移到新的可用区使用方式上

十、附录 [↗](#)