

高级云架构

曹广杰

数据科学与计算机学院

授课教师：王昌栋

2017/12/31

高级云架构

分析比较云服务容错的几种机制

思考云服务负载均衡和存储负载均衡两种负载均衡需求及对应机制的差别

储存负载均衡

云服务负载均衡

SaaS环境中IT 资源分割通常不发生在其基础设施层次上，为什么？

分析比较云服务容错的几种机制

云服务容错架构的设计是用于应对云服务不可用的情况，比如：

1. 运行时服务器过载；
2. 维护更新暂时中断服务；
3. 云服务迁移；
4. 物理机的宕机失效；

为此设计实现不中断服务重定位架构，授课的时候对三种架构进行了举例：

- 第一种不中断服务重定位架构

1. 在云用户与云服务端之间有automated scaling listener；
2. 在服务端设置了预先定义的阈值；
3. 一旦到达了这个阈值，服务端与客户端之间的监听器就会开始向VIM请求；
4. VIM对过载的服务进行重定位操作；

- 第二种不中断服务重定位架构

与第一种重定位架构不同的是，在这个方案中用于处理过载的请求的虚拟机是被物理服务器上的hypervisor所复制而得。

- 第三种不中断服务重定位架构

1. 有两个同步的虚拟服务器实例；
2. 如果其中一个崩溃了，则用户的请求会全部转移到另一个服务器实例上。

思考云服务负载均衡和存储负载均衡两种负载均衡需求及对应机制的差别

储存负载均衡

实现负载均衡的粒度（即可用考虑的三个因素）：

1. 用户请求，前端负载均衡把用户请求分配给不同的云服务实例；
2. 虚拟服务器实例，把虚拟服务器放置到不同的物理机；
3. 云负载均衡，数据中心间的负载均衡；

实现负载均衡需要基于虚拟机监控器集群，核心是容量看门狗——动态地计算虚拟服务器实例以及工作负载；

- 容量看门狗

云使用监控器

1. 追踪物理和虚拟服务器使用情况；
2. 向容量计划器报告波动；

容量计划器

1. 动态计算物理服务器的可用能力和虚拟服务器的容量要求

VM在线迁移程序

云服务负载均衡

云服务负载均衡架构模型：一个特殊的架构模型，在多个云之间进行负载均衡。

目的：提高服务能力、可扩展性；提高可用性、可靠性；改进负载均衡和资源优化；

基础：自动伸缩监听器；故障转移系统；

以上两者都有故障转移系统，都配备有容量动态监听器，以及动态计算以及规划负载均衡的监听器。但是云负载均衡是在多个云之间进行负载均衡，而储存系统的负载均衡是在物理服务器以及虚拟机之间实现负载均衡。

SaaS环境中IT 资源分割通常不发生在其基础设施层次上，为什么？

对于云环境的搭建分为三个层次：

$$IaaS < PaaS < SaaS$$

以上三者从左到右

1. 可用性递增；
2. 权限递减；

3. 配置的程度递增；

所以到SaaS的环境配置的时候，已经是配置得比较完全比较完善了。此时的资源管控大多是基于IaaS或者PaaS的基础上，在这个配置中，已经很少涉及基础设施的层次了，或者说，那个层次上的资源已经配置过了，在SaaS的环境下不能再配置。