国家重点研发计划“云计算与大数据”专项

面向云计算的网络化操作系统

项目/课题实施方案

【课题名称】

参与单位

二零一七年五月XX日

目 录

【三级目录即可，小四，1.5倍行距】

# 1. 项目/课题目标

# 2. 项目/课题主要研究内容

容器的在线迁移的研究内容主要分为三个部分：

（1）Docker Daemon与容器之间父子嵌套关系的迁移

由于Docker目前是结合CIRU来实现对容器状态信息的Checkpoint/ Restore操作，而CRIU不能支持对嵌套命名空间的C/R，故对容器检查点备份并传送到目的机恢复后，Docker不能再对该容器的生命周期进行管理。我们需要对CRIU展开研究，能够在维持Docker Daemon与容器之间父子嵌套关系的基础上实现对容器的Checkpoint/Restore。

（2）Docker与CRIU PageServer的结合实现端到端的无缝迁移

目前 Docker原生只支持对容器简单的Checkpoint/Restore，还不能实现端到端的无缝迁移，而端到端的无缝迁移需要CRIU PageServer技术的支持，PageServer是CRIU项目的一个组件，用于实现在进程热迁移的过程中内存的传输，这个步骤是容器热迁移中最关键的步骤，所以研究Docker如何整合PageServer来实现对容器热迁移过程中内存的传输非常重要。

（3）Docker迁移策略的研究

在研究内容(1)(2)完成的基础上，在技术层面已经可以有效的实现Docker容器的端到端的无缝热迁移。然而在实际的生产环境中，我们需要提出合理的迁移策略来对整个云平台下的热迁移做出决策，其中就包括容器何时需要迁移，及迁移目的机的选择等，从而提高云平台下计算资源的利用率。

# 3. 项目/课题的技术方案与技术路线

根据课题的三个部分的研究内容，该课题的技术方案也将根据这三大部分展开：

（1）第一个研究内容为Docker与容器之间的父子嵌套关系不能跨主机迁移，主要原因是CRIU目前不能支持对嵌套命名空间的C/R,因为仅从CRIU迁移进程的角度，当CRIU对进程进行检查点备份后，将检查点文件传输到目的主机，最后会以操作系统Init进程为父进程来恢复容器。然而理想情况是根据传输到目的机的检查点文件并以目的机上的Docker Daemon为父进程来恢复容器从而再次建立Docker与容器之间的关系。所以我们需要了解CRIU对容器进行检查点备份及恢复容器进程树的原理，然后修改相关源码来以指定的Docker Daemon为父进程恢复容器的运行。

（2）第二个部分研究内容是Docker对CRIU中的PageServer的整合，这个部分的研究是建立在Docker容器可以顺利实现端到端的Checkpoint/Restore的技术基础之上，在此基础上通过修改docker源码来调用PageServer模块并结合热迁移最常用的Pre-Copy迁移机制有条理地传输Docker容器所指定的用户态内存并迭代传输，直到达到Pre-Copy所设定的Stop-and-Copy条件再结束内存的迭代传输，最后利用Docker Restore能力来恢复容器的运行，从而实现Docker容器端到端的无缝迁移。这整个技术实现要求我们修改Docker源码并结合Pre-Copy的迁移机制，然后调用PageServer模块的相关命令实现端到端的Pre-Copy无缝迁移。

（3）Docker迁移策略同样是在Docker容器端到端无缝迁移的技术基础上实现，通过研究现有容器编排工具的一些迁移策略和容器编排思想，例如K8S或Docker Swarm等的编排算法，然后在此基础上，与阿里交流容器实际生产应用方面在容器编排的过程中所需要考虑的资源侧重点或者是资源调度所存在的一些问题，最后调整已有的迁移策略实现针对阿里容器应用的容器调度迁移系统。

# 4. 项目/课题成果、考核指标与评测方法

【量化、可考核】

# 5. 项目/课题进度计划

【至少半年一个里程碑】