



Libvirt分布式锁管理插件 DLM-Corosync的实现

2018.06

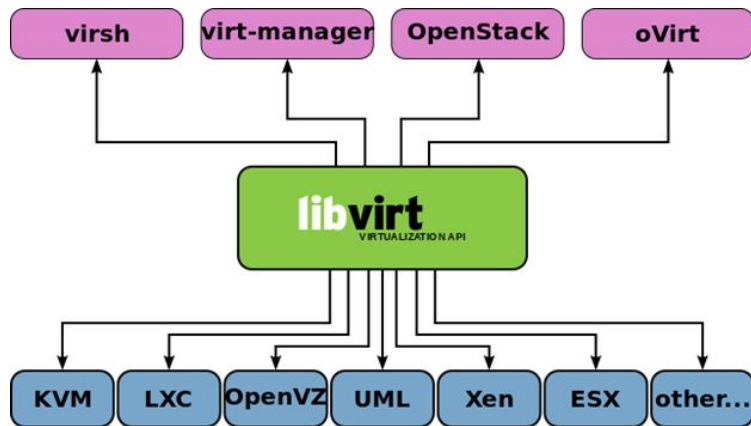


目录

- 研究背景
- 插件需求分析
- 插件设计与实现

研究背景 - libvirt简介

- 一个开源软件的集合 :API + libvirtd + virsh
- 虚拟化软件的管理工具:致力于提供在不同的操作系统平台之上以单一的方式管理不同的虚拟化方案以及虚拟机





研究背景 - libvirt 分布式锁管理框架

提出背景

- 集群环境中普遍使用共享存储
- 多个VM存在同时使用同一个磁盘镜像的可能性
- 多个VM同时向同一个磁盘镜像写入可能会导致不可逆转的数据损坏



研究背景 - libvirt 锁管理框架(续)

现存的插件

nop:

- 什么都不做

sanlock:

- 使用disk-paxos算法读写共享存储以实现对分布式锁的获取、释放和超时
- 带来了大量的磁盘IO



研究背景 - libvirt 锁管理框架(续)

现存的插件(续)

lockd:

- 极少的磁盘IO使用
- 使用POSIX Standard文件锁
- 需分布式文件系统支持POSIX Standard文件锁



提出 - libvirt第三种锁管理插件DLM-Corosync

技术背景

- DLM与Corosync在高可用套件中使用, 不依赖共享存储
- DLM是一种分布式锁管理器, 用于控制资源访问。Linux Kernel中有相关的实现, 用户态程序可通过libdlm间接使用
- Corosync是集群成员之间通信的基础设施, 使用TCP/STCP协议

插件需求分析

达成的目标

在以下场景中实现资源互斥：

- 两个不同的VM运行在同一个磁盘上。
- 因管理员失误导致的一个VM在不同的宿主机上多次启动
- 因libvirt的驱动错误导致的一VM在同一个hypervisor上多次启动

插件需求分析

提出的要求

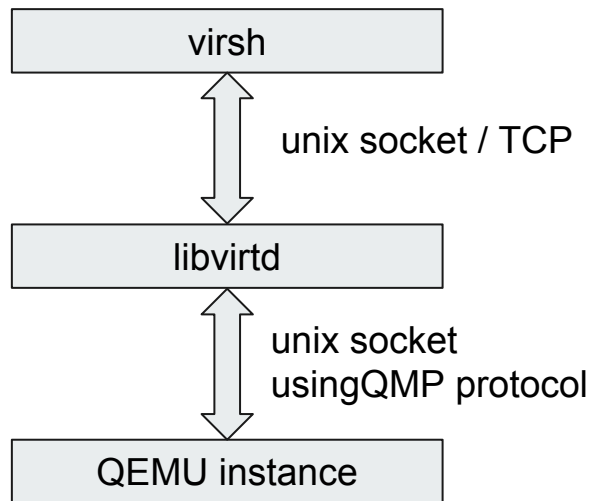
- QEMU 进程打开磁盘的时候, 这个磁盘镜像必须被加锁。
- 插件必须实现只读、共享写、互斥写三种锁的模式。
- 在 QEMU 实例迁移的过程中, 锁的所有权转移必须可行。
- 如果某个资源的锁被移除, 锁管理器必须能够识别并杀死使用这个资源的进程实例。
- 通过使用管理程序可以对与虚拟机有关的任意资源进行加锁。

插件需求分析

因libvirt与被管理的虚拟机进程相互独立

带来的额外要求：

- 锁与相关联的虚拟机进程组有着相同的生命周期。
- 锁不会因 libvirt 守护进程的重启、退出而失效。



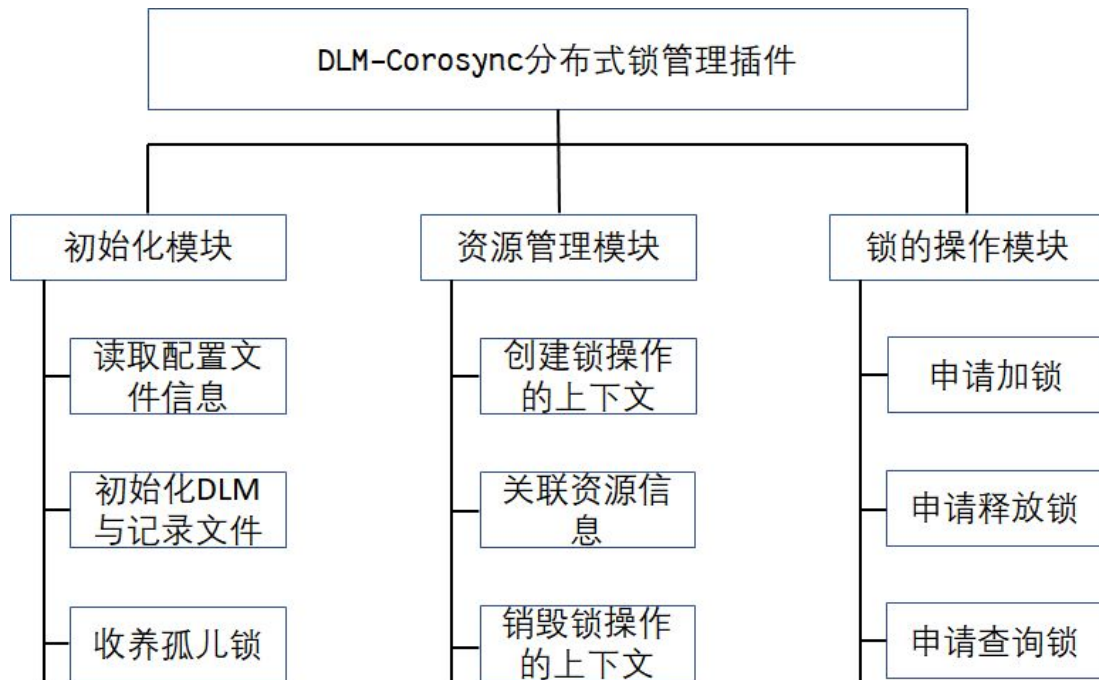
插件设计与实现

设计的操作

- Register object: 注册一个用于获取锁的对象, 该对象的标识符是独一无二的。
- Add resource: 把锁的对象与资源相互关联, 用于之后对锁的获取/释放操作。
- Acquire locks: 为与锁的对象相关联的所有资源进行加锁。
- Release locks: 释放与锁的对象相关联的所有资源的锁。
- Inquire locks: 查询与该锁的对象所关联的锁的状态。

插件设计与实现

功能结构图



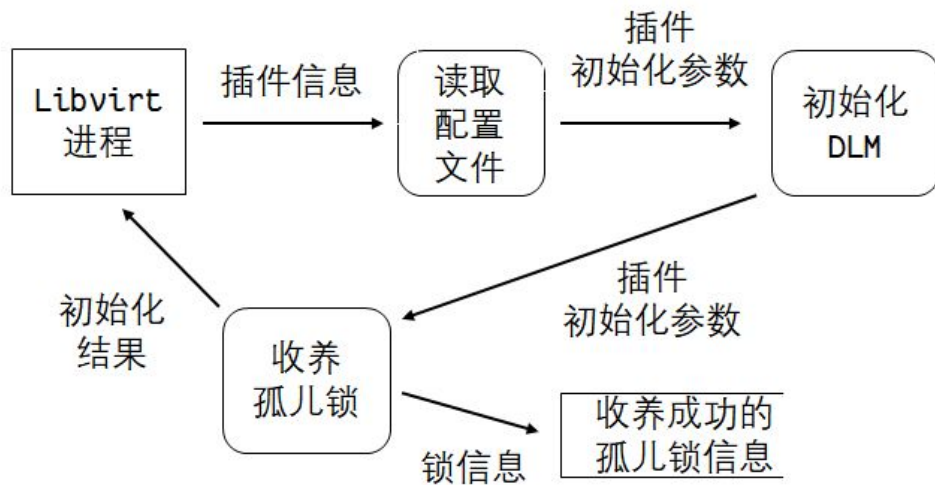
插件设计与实现

```
virLockDriver virLockDriverImpl =  
{  
    .drvInit = virLockManagerDLMInit,           初始化模块  
    .drvDeinit = virLockManagerDLMDeinit,  
  
    .drvNew = virLockManagerDLMNew,  
    .drvFree = virLockManagerDLMFree,           资源关联模块模块  
    .drvAddResource = virLockManagerDLMAAddResource,  
  
    .drvAcquire = virLockManagerDLMAcquire,  
    .drvRelease = virLockManagerDLMRelease,     资源关联模块  
    .drvInquire = virLockManagerDLMInquire,  
};
```

插件设计与实现

初始化模块

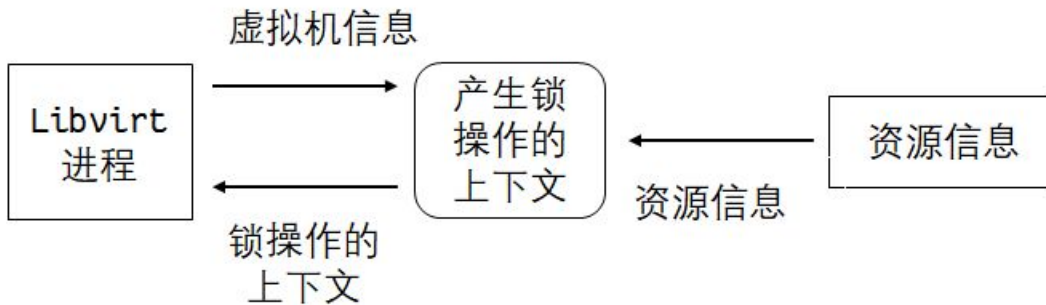
- virLockManagerDLMInit



插件设计与实现

初始化模块

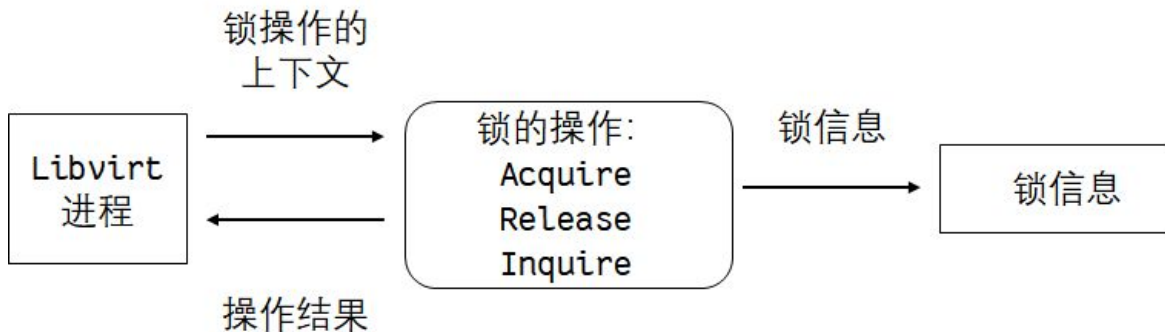
- virLockManagerDLMNew
- virLockManagerDLMAAddResource



插件设计与实现

锁的操作模块

- virLockManagerDLMAcquire
- virLockManagerDLMRelease
- virLockManagerDLMinquire





結束

谢谢指导老师与在座的各位老师

Q & A