Lime通过Collectd实时采集Lustre集群中节点的性能指标信息，以这些信息作为依据来调节Lustre NRS TBF的rate值，来提供QOS保障以及某些性能的限制。Lime的源码：<https://github.com/DDNStorage/Lime> 。

**Collectd的作用**

1. 采集Lustre集群的性能指标，主要通过filedata插件完成；
2. 上报采集信息到Lime，主要通过write\_http插件完成；

**Lime的作用**

1. 监听Collectd上报数据的特定端口，处理Collectd上报的性能指标信息；
2. 识别Lustre集群：有哪些节点以及节点的角色（MDS/OSS/Client）。做到这一点，只需要在配置文件中完整的配置Lustre集群节点的IP即可，Lime通过SSH执行命令“lctl dl”后，即可自动判别节点的角色；
3. 以Collectd上报的性能指标为依据，通过SSH远程执行Lustre NRS TBF的命令，调整rate值；

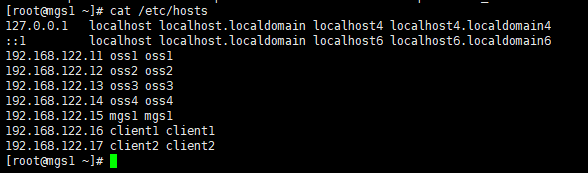
配置环境

在物理机上通过KVM虚拟7个虚拟机几点，作为Lustre集群，虚拟机节点和物理机节点互通，虚拟机、物理机均可以访问公网，Lime配置在物理机上。

|  |  |  |
| --- | --- | --- |
| **名称** | **版本** | **备注** |
| 物理机 | CentOS Linux release 7.2.1511 (Core) | CPU Intel(R) Xeon(R) CPU E5-2620 0 @ 2.00GHz 2x  内存容量 16384 MB  内存详情 16GB DDR3 RAM  硬盘容量 1168 GB  硬盘详情 2x SAS Disks (15000RPM, 146GB)  网卡信息 Mellanox InfiniBand QDR 40Gb/s NIC  IP 192.168.3.181 |
| KVM | qemu-img version 1.5.3 |  |
| 虚拟机 | CentOS Linux release 7.2.1511 (Core) | IP配置情况：192.168.122.11-192.168.122.17  192.168.122.11-192.168.122.14：OSS  192.168.122.15：MDS  192.168.122.16-192.168.122.17：Client |
| Collectd | collectd-5.7.2.564 | DDN提供的 |
| Lustre | 2.8.0 |  |
| Lime |  | <https://github.com/DDNStorage/Lime> |
| Python | 2.7.5 |  |

配置步骤

1. 配置Lustre集群
2. 配置物理机和虚拟机网络，使它们可以相互ping通，物理机、虚拟机均可以访问公网（后面需要使用yum安装依赖软件）。配置物理机可以SSH免密码登录任何一台虚拟机（非必须步骤，方便后面的操作）；
3. 修改物理机和虚拟机的hosts文件，如下：



1. 将所有虚拟机节点配置为一个简单的Lustre集群：

虚拟机节点角色

192.168.122.11-192.168.122.14：OSS

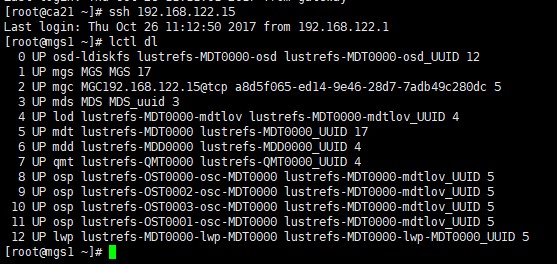
192.168.122.15：MDS

192.168.122.16-192.168.122.17：Client

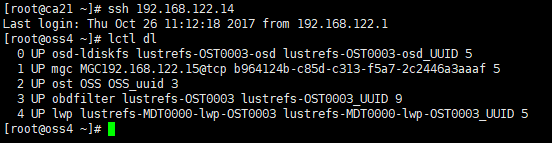
文件系统的名称：lustrefs #可以随意设置，只需要与Lime的配置文件lime\_config.json的名称保持一致即可。

配置完成后，登录一些节点做简单验证：

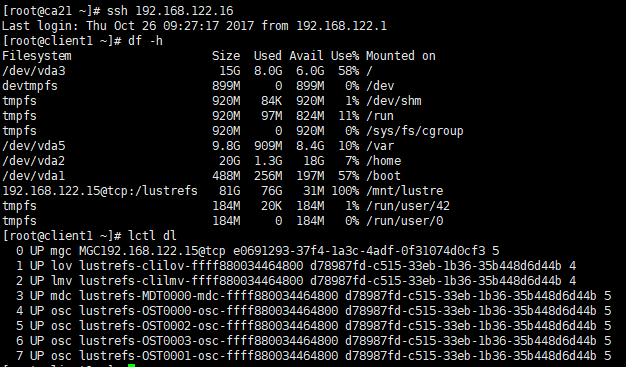
登录MDS执行lctl dl



登录其中任何一台OSS，执行lctl dl



登录任何一台Client，执行df -h 以及lctl dl



1. 安装并配置Collectd，需要在Lustre集群中的所有节点中安装

**安装**

rpm -ivh collectd-5.7.2.564.gced30bf.ddn-7.el7.x86\_64.rpm

rpm -ivh collectd-filedata-5.7.2.564.gced30bf.ddn-7.el7.x86\_64.rpm

rpm -ivh libcollectdclient-5.7.2.564.gced30bf.ddn-7.el7.x86\_64.rpm

rpm -ivh collectd-write\_http-5.7.2.564.gced30bf.ddn-7.el7.x86\_64.rpm

rpm -ivh xml\_definition-2.0.g370b24d.ddn1-1.noarch.rpm #相关的配置文件

备注：安装过程中可能需要安装一些依赖软件，根据提示安装即可，缺失什么依赖就安装什么依赖。若与旧版本软件冲突，先将旧版本软件卸载。

**配置**

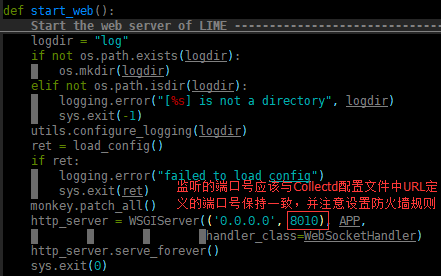
Collectd配置文件：/etc/collectd.conf

|  |
| --- |
| LoadPlugin logfile  <Plugin logfile>  LogLevel info  File "/var/log/collectd.log"  </Plugin>  Interval 1  LoadPlugin filedata  <Plugin "filedata">  <Common>  DefinitionFile "/etc/lustre-ieel-2.7\_definition.xml"  </Common>  <Item>  Type "ost\_jobstats"  <Rule>  Field "job\_id"  Match "dd\..+"  </Rule>  <Filter>  Field "read\_samples"  Field "min\_read\_bytes"  Field "max\_read\_bytes"  Field "sum\_read\_bytes"  Field "write\_samples"  Field "min\_write\_bytes"  Field "max\_write\_bytes"  Field "getattr"  Field "setattr"  Field "punch"  Field "sync"  Field "destroy"  Field "create"  Field "statfs"  Field "get\_info"  Field "set\_info"  Field "quotactl"  </Filter>  </Item>  # <Item>  # Type "mdt\_jobstats"  # </Item>  </Plugin>  LoadPlugin write\_http  <Plugin write\_http>  <Node "lime">  format "JSON"  #URL "http://ddnlab.kmdns.net:6004/metric\_post"  URL "http://192.168.3.181:8010/metric\_post"  SyncSend true  </Node>  </Plugin> |

说明：URL中的IP和端口，是Lime所在的节点的IP以及相应的监听的端口（注意防火墙要开放对应的端口），可以根据实际环境灵活配置。

安装完毕后：service collectd start 看是否能正常启动。

1. Lime配置和启动：在物理机上配置Lime，实际情况下可以在Lustre集群中的任何一台计算机上配置Lime，也可以在独立于Lustre集群的一台计算机上配置，只要它们在网络上是相通的即可。
2. Lime监听端口设置: vim lime\_web.py



1. Lime配置文件设置：vim static/lime\_config.json

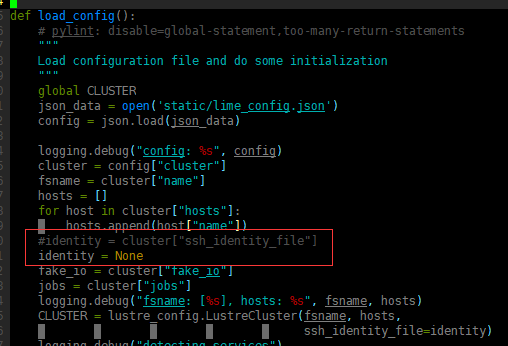
|  |  |
| --- | --- |
|  | {  "cluster": {  "name": "lustrefs",  "hosts": [  {  "name": "192.168.122.11"  },  {  "name": "192.168.122.12"  },  {  "name": "192.168.122.13"  },  {  "name": "192.168.122.14"  },  {  "name": "192.168.122.15"  },  {  "name": "192.168.122.16"  },  {  "name": "192.168.122.17"  }  ],  "fake\_io": false,  "ssh\_identity\_file": "~/.ssh/id\_dsa",  "policy": "priority",  "jobs": [  {  "job\_id": "dd.0",  "login\_name": "root",  "throughput": "10000"  },  {  "job\_id": "dd.1001",  "login\_name": "lime",  "throughput": "10000"  }  ]  }  } |

1. Lime源码修改(根据实际情况修改，非必须)

**SSH执行命令认证**

若前面配置过Lime所在的主机可以免密码登录Lustre集群中任意一台主机，可将lime\_web.py中load\_config()函数里关于SSH秘钥的代码做如图中修改：identity = None

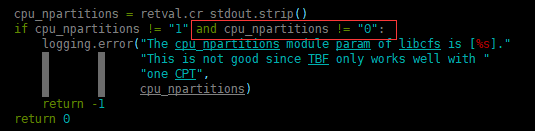
以避免lime\_web.py运行过程中多次输入密码，本步骤不是必要的。



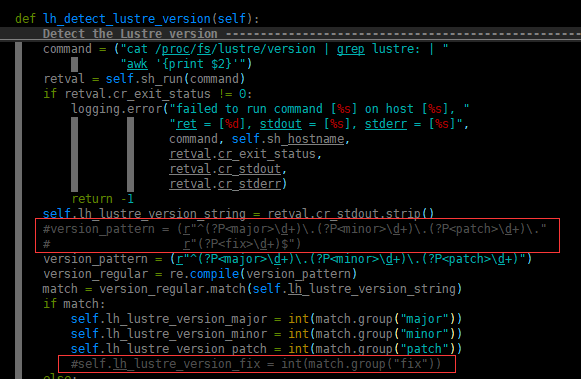
**根据实际情况可能需要修改的代码**

若在虚拟机中实验，需要修改lustre\_config.py: lh\_check\_cpt(self)函数的代码，修改判别条件为

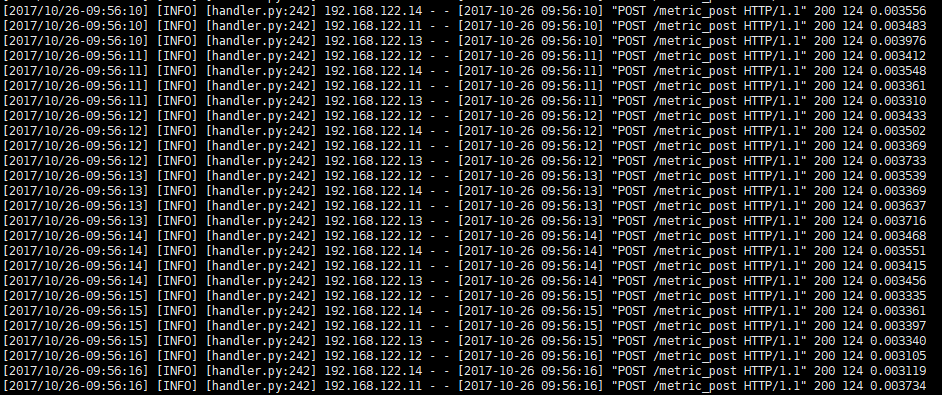
cpu\_npartitions != "1" and cpu\_npartitions != "0"



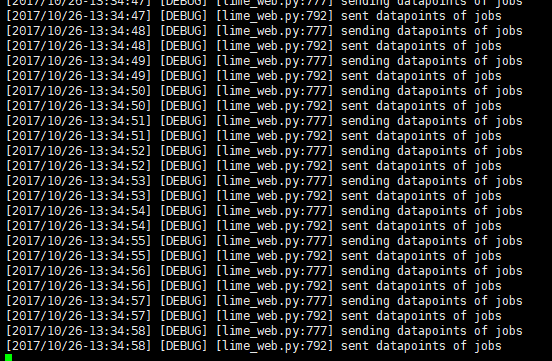
lh\_detect\_lustre\_version函数中需要注释这两行版本检测相关的代码（跟Lustre编译时设定的rpm包版本名称有关，若是按照ddn的规范生产的rpm包，则不需要修改这部分代码）。



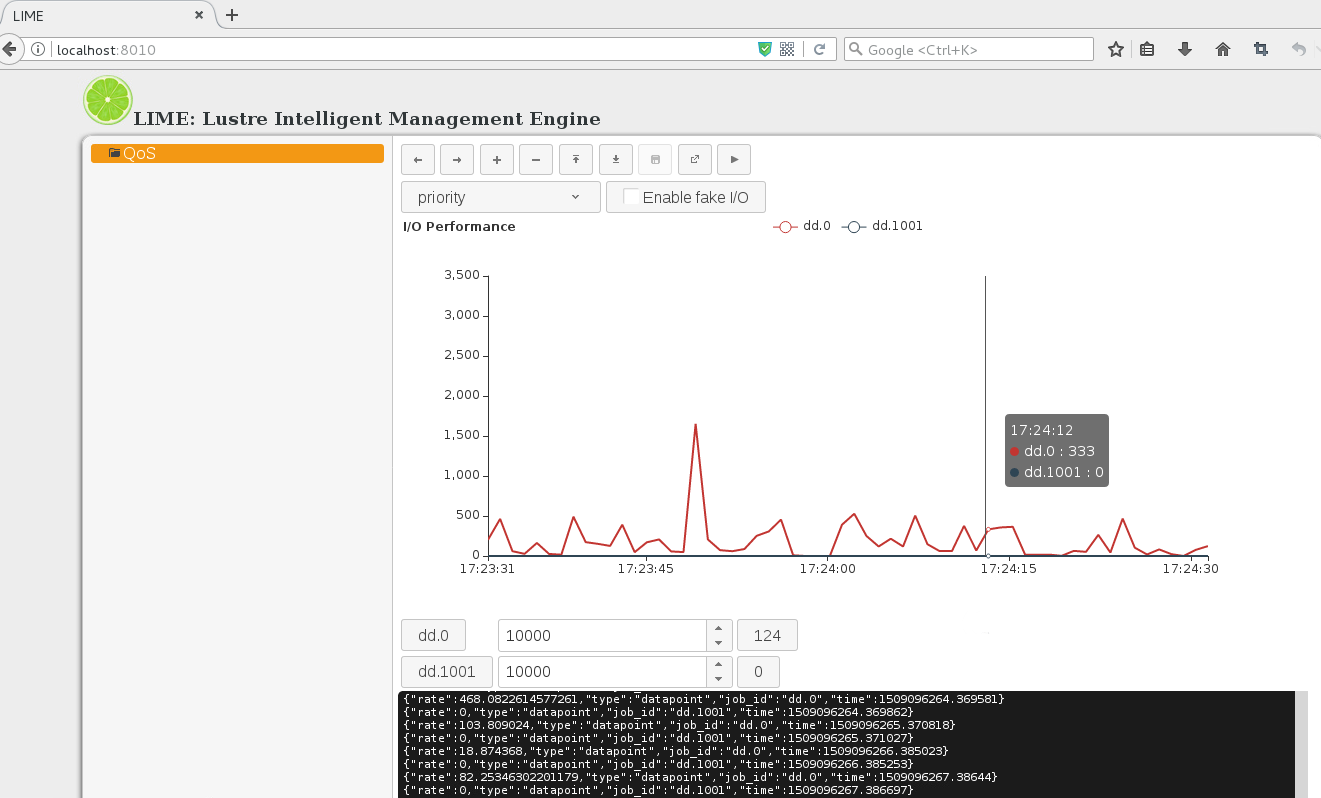
1. 执行python lime\_web.py，等待几秒钟，会有数据采集的log输出：



同时，可以tail –f log/debug.log 以及其它的log信息，查看相关输出以及运行情况：



访问：http:192.168.3.181:8080



注意：Lime使用了众多的Python包，根据不同的环境，可能会因为包依赖问题报错，这需要根据实际情况安装缺少的依赖包。如：pip install gevent && pip install gevent-websocket && pip install flask