

爱维 Linux 公开课

企业运维监控平台架构设计与实现

(Ganglia 篇)

主讲人：南非蚂蚁

课程安排

- 开源监控软件的选择与思考
- 统一运维监控平台设计思路
- Ganglia 数据收集平台的构建
- Ganglia 的优势与经验汇总

一、Cacti/Nagios/Zabbix/centreon/Ganglia 之抉择

1、cacti

Cacti 是一套基于 PHP,MySQL,SNMP 及 RRDTool 开发的网络流量监测图形分析工具。简单的说 Cacti 就是一个 PHP 程序。它通过使用 SNMP 协议获取远端网络设备和相关信息, (其实就是使用 Net-SNMP 软件包的 snmpget 和 snmpwalk 命令获取) 并通过 RRDTOOL 工具绘图, 通过 PHP 程序展现出来。我们使用它可以展现出监控对象一段时间内的状态或者性能趋势图。

2、nagios

Nagios 是一款开源的免费网络监视工具, 能有效监控 Windows、Linux 和 Unix 的主机状态, 交换机路由器等网络设置, 打印机等。在系统或服务状态异常时发出邮件或短信报警第一时间通知网站运维人员, 在状态恢复后发出正常的邮件或短信通知。

3、zabbix

zabbix 是一个基于 WEB 界面的提供分布式系统监视以及网络监视功能的企业级的开源解决方案。zabbix 能监视各种网络参数, 保证服务器系统的安全运营; 并提供柔软的通知机制以让系统管理员快速定位/解决存在的各种问题。

zabbix 由 2 部分构成, zabbix server 与可选组件 zabbix agent。zabbix server 可以通过 SNMP, zabbix agent, ping, 端口监视等方法提供对远程服务器/网络状态的监视, 数据收集等功能, 它可以运行在 Linux, Solaris, HP-UX, AIX, Free BSD, Open BSD, OSX 等平台上。

4、ganglia

Ganglia 是一款为 HPC（高性能计算）集群而设计的可扩展的分布式监控系统，它可以监视和显示集群中的节点的各种状态信息，它由运行在各个节点上的 gmond 守护进程来采集 CPU、内存、硬盘利用率、I/O 负载、网络流量情况等方面的数据，然后汇总到 gmetad 守护进程下，使用 rrdtool 存储数据，最后将历史数据以曲线方式通过 PHP 页面呈现。

Ganglia 监控系统有三部分组成，分别是 gmond、gmetad、webfrontend。

5、centreon

Centreon 是一款功能强大的分布式 IT 监控系统，它通过第三方组件可以实现对网络、操作系统和应用程序的监控：首先，它是开源的，我们可以免费使用它；其次，它的底层采用 nagios 作为监控软件，同时 nagios 通过 ndoutil 模块将监控到的数据定时写入数据库中，而 Centreon 实时从数据库读取该数据并通过 Web 界面展现监控数据；最后，我们可以通过 Centreon 管理和配置 nagios，或者说 Centreon 就是 nagios 的一个管理配置工具，通过 Centreon 提供的 Web 配置界面，可以轻松完成 nagios 的各种繁琐配置。

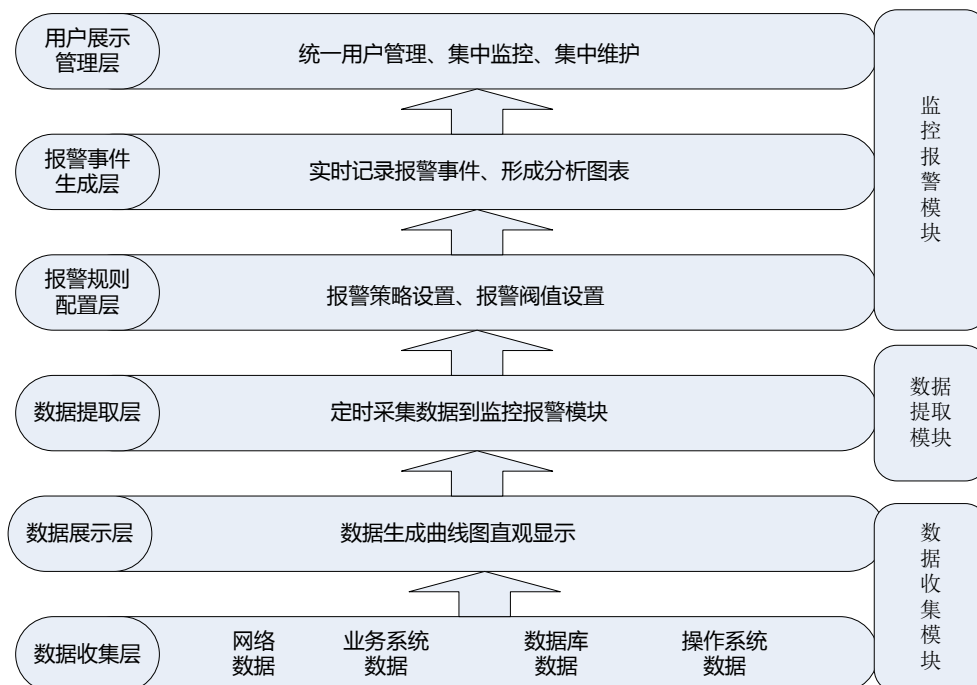
6、对比图

功能	nagios	cacti	zabbix	ganglia+centreon
管理方式	文本编辑	web界面	web界面	web界面
数据展示	PNP插件绘图	RRD模式绘图	PHP GD绘图	RRD模式绘图
数据采集	插件模式	SNMP/Scripts	自身函数/插件	ganglia专业收集/插件收集/支持C/PHP/SHELL/python接口
告警方式	Email/自定义	Email	命令/邮件	邮件/手机/第三方接口
监控方式	手动建立	手动建立	自动发现/手动建立	自动发现/手动建立
管理功能	无	无	命令下发	命令下发/API
监控结构	集中式/分布式	集中式	集中式/分布式	集中式/分布式
用户权限	简单	灵活的权限分配	灵活的权限分配	灵活的权限分配/支持角色/组
维护时间	有	无	有	有/支持多种维护模式
审计功能	无	无	精细化审计	精细化审计/支持历史查阅
汇总状态图	无	复杂，但效果好	简单	灵活/精确/支持自定义汇总
性能、扩展	差	很差	一般	优秀

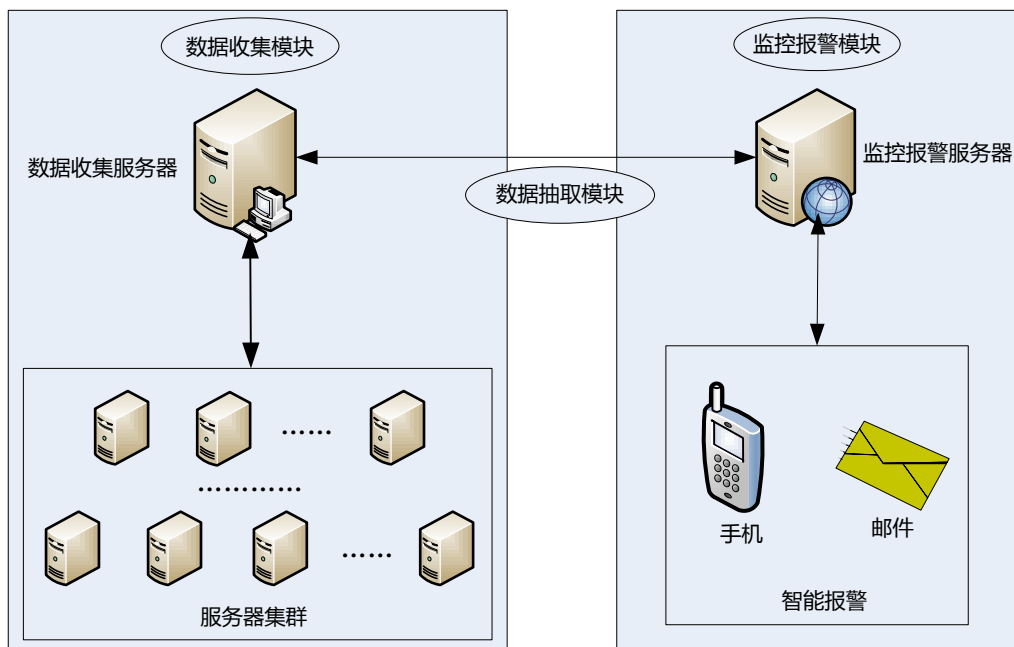
二、 统一运维监控平台设计思路

构建一个智能的运维监控平台，必须以运行监控和故障报警这两个方面为重点，将所有业务系统中所涉及的网络资源、硬件资源、软件资源、数据库资源等纳入统一的运维监控平台中，并通过消除管理软件的差别，数据采集手段的差别，对各种不同的数据来源实现统一管理、统一规范、统一处理、统一展现、统一用户登录、统一权限控制，最终实现运维规范化、自动化、智能化的大运维管理。

智能的运维监控平台，设计架构从低到高可以分为 6 层，三大模块，如下图：



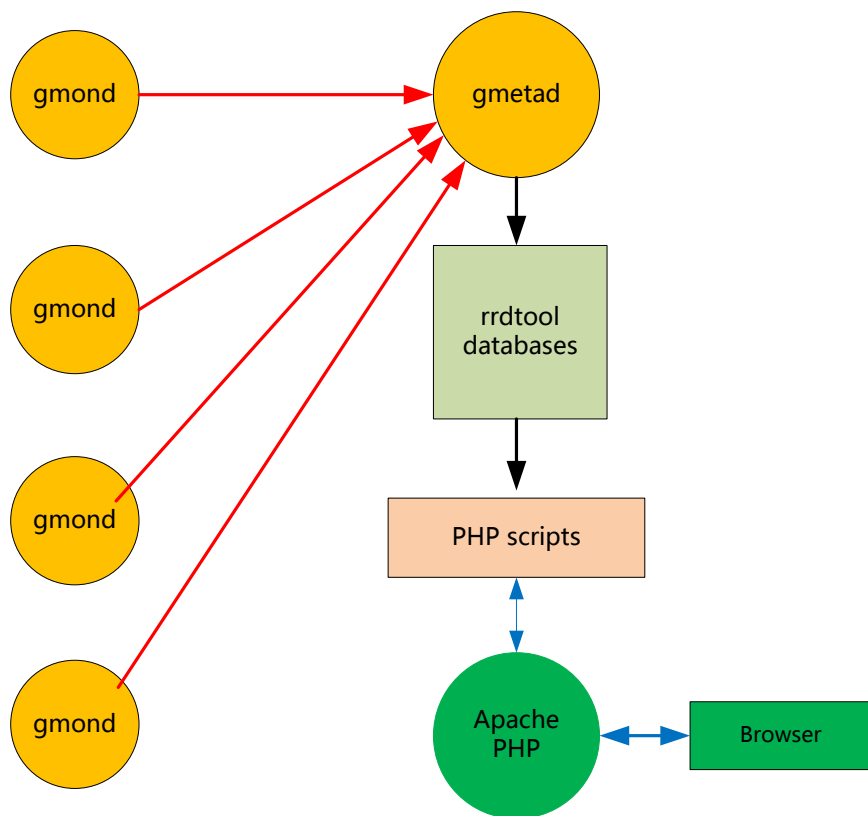
运维监控平台实现拓扑图，请看下图：



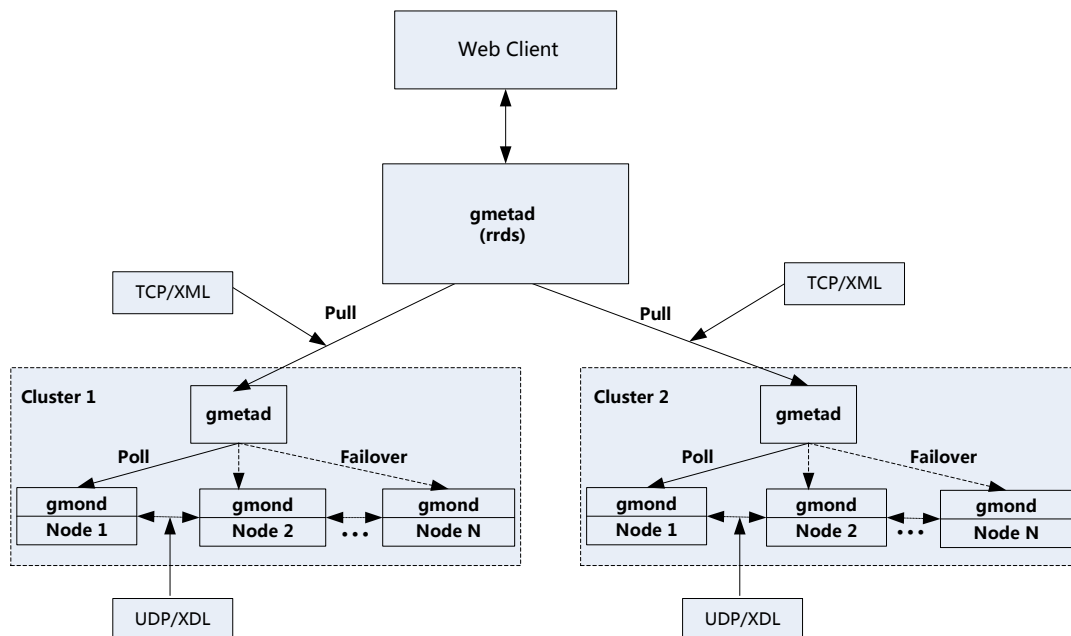
三、Ganglia 的安装

1、ganglia 的常用架构

Ganglia 监控系统有三部分组成，分别是 gmond、gmetad、webfrontend，如下图所示：



同时，Ganglia 支持多种监控架构，这是由 gmetad 的特性决定的，gmetad 可以周期性地到多个 gmond 节点收集数据，这就是 ganglia 的两层架构。同时，gmetad 不但可以从 gmond 收集数据，也可以从其他的 gmetad 得到数据，这就形成了 Gnaglia 的三层架构。多种架构方式也体现了 Ganglia 作为分布式监控系统的灵活性和扩展性。



2、yum 方式安装 ganglia

CentOS 系统中默认的 yum 源并没有包含 Ganglia，所以我们必须安装扩展的 yum 源。从下面这个地址下载 Linux 附加软件包（EPEL），然后安装扩展 yum 源：

```
[root@node1~]# wget http://dl.fedoraproject.org/pub/epel/5/i386/epel-release-5-4.noarch.rpm
[root@node1~]# rpm -ivh epel-release-5-4.noarch.rpm
```

完成 yum 源安装，就可以直接通过 yum 方式安装 Ganglia 了。

Ganglia 的安装分为两个部分，分别是 gmetad 和 gmond，gmetad 安装在监控管理端，gmond 安装在需要监控的客户端主机，对应的 yum 包名称分别为 ganglia-gmetad 和 ganglia-gmond。

下面介绍通过 yum 方式安装 Ganglia 的过程。

以下操作是在监控管理端进行的，首先通过 yum 命令查看下可用的 Ganglia 安装信息：

```
[root@monitor~]# yum list ganglia*
```

安装 gmetad 需要 rrdtool 的支持，而通过 yum 方式，会自动查找 gmetad 依赖的安装包，自动完成安装，这也是 yum 方式安装的优势。

最后在需要监控的所有客户端主机上安装 gmond 服务：

```
[root@node1~]# yum -y install ganglia-gmond.x86_64
```

这样，Ganglia 监控系统就安装完成了。通过 yum 方式安装的 Ganglia 默认配置文件位于 /etc/ganglia 中。

3、Ganglia 监控管理端配置

监控管理端的配置文件是 gmetad.conf，这个配置文件内容比较多，但是需要修改的配置仅有如下几个：

```
data_source "Cluster1" cloud0cloud2
```

```
gridname "IIVEYGrid"
```

```
xml_port 8651
```

```
interactive_port 8652
```

```
rrd_rootdir "/var/lib/ganglia/rrds"
```

data_source: 此参数定义了集群名字，以及集群中的节点。**Cluster1** 就是这个集群的名称，**cloud0** 和 **cloud2** 指明了从这两个节点收集数据，**Cluster1** 后面指定的节点名可以是 IP 地址，也可以是主机名，由于采用了 **multicast** 模式，每个 **gmond** 节点都有本 **Cluster1** 集群节点所有监控数据，因此不需要把所有节点都写入 **data_source** 中。但是建议写入不低于 2 个，这样，在 **cloud0** 节点出现故障的时候，**gmetad** 会自动到 **cloud2** 节点采集数据，这样就保证了 **Ganglia** 监控系统的高可用性。

上面通过 **data_source** 参数定义了一个服务器集群 **Cluster1**，对于要监控多个应用系统的情况，还可以对不同用途的主机进行分组，定义多个服务器集群，分组方式可以通过下面的方法定义：

```
data_source "mycluster" 10localhost my.machine.edu:8649 1.2.3.5:8655
```

```
data_source "mygrid" 50 1.3.4.7:8655 grid.org:8651 grid-backup.org:8651
```

```
data_source "another source" 1.3.4.7:8655 1.3.4.8
```

可以通过定义多个 **data_source** 来实现监控多个服务器集群，而每个服务器集群在定义集群节点的时候，可以采用主机名或 IP 地址等形式，也可以加端口，如果不加端口，默认端口是 8649，同时可以设定采集数据的频率，如上面的“10localhost、50 1.3.4.7:8655”等，分别表示每隔 10 秒钟、50 秒钟采集一次数据。

gridname: 此参数是定义一个网格名称。一个网格有多个服务器集群组成，每个服务器集群由“**data_source**”选项来定义。

xml_port: 此参数定义了一个收集数据汇总的交互端口，如果不指定，默认是 8651，可以通过 **telnet** 这个端口得到监控管理端收集到的客户端的所有数据。

interactive_port: 此参数定义了 Web 端获取数据的端口，这个端口在配置 **Ganglia** 的 Web 监控界面时需要指定。

rrd_rootdir: 此参数定义了 **rrd** 数据库的存放路径，**gmetad** 在收集到监控数据后会将其更新到该目录下的对应的 **rrd** 数据库中。

4、Ganglia 的客户端配置

Ganglia 监控客户端 **gmond** 安装完成后，配置文件位于 **Ganglia** 安装路径的 **etc** 目录下，名称为 **gmond.conf**，这个配置文件稍微复杂，如下所示：

```
globals{
```

```
daemonize=yes #是否后台运行，这里表示以后台的方式运行
```

```
setuid=yes #是否设置运行用户，在 Windows 中需要设置为 false
```

```
user=nobody #设置运行的用户名称，必须是操作系统已经存在的用户，默认是 nobody
```

```
debug_level = 0 #调试级别，默认是 0，表示不输出任何日志，数字越大表示输出的日志越多
```

```
max_udp_msg_len = 1472
```

```
mute=no #是否发送监控数据到其他节点，设置为 no 表示本节点将不再广播任何自己收集到的数据到网络上
```

```
deaf=no #是否接受其他节点发送过来的监控数据，设置为 no 表示本节点将不再接
```

收任何其他节点广播的数据包

`allow_extra_data=yes` #是否发送扩展数据

`host_dmax = 0 /*secs */` #是否删除一个节点, 0 代表永远不删除, 0 之外的整数代表节点的不响应时间, 超过这个时间后, Ganglia 就会刷新集群节点信息进而删除此节点

`cleanup_threshold=300 /*secs */` #gmond 清理过期数据的时间

`gexec=no` #是否使用 gexec 来告知主机是否可用, 这里不启用

`send_metadata_interval = 0` #在单播协议中, 新添加的节点在多长时间响应一下以表示自己的存在, 0 代表仅在 gmond 启动时通知一次, 单位秒

}

cluster{

`name = "Cluster1"` #集群的名称, 是区分此节点属于某个集群的标志, 必须和监控服务端 `data_source` 中的某一项名称匹配

`owner="junfeng"` #节点的拥有者, 也就是节点的管理员

`latlong="unspecified"` #节点的坐标, 经度、纬度等, 一般无需指定

`url="unspecified"` #节点的 URL 地址, 一般无需指定

}

host{

`location="unspecified"` #节点的物理位置, 一般无需指定

}

`udp_send_channel{` #udp 包的发送通道

`mcast_join = 239.2.11.71` #指定发送的多播地址, 其中 239.2.11.71 是一个 D 类地址。如果使用单播模式, 则要写 `host=host1`, 在单播模式下也可以配置多个 `udp_send_channel`

`port = 8649` #监听端口

`ttl = 1`

}

`udp_rcv_channel{` #接收 udp 包配置

`mcast_join=239.2.11.71` #指定接收的多播地址, 同样也是 239.2.11.71 这个 D 类地址

`port = 8649` #监听端口

`bind = 239.2.11.71` #绑定地址

}

`tcp_accept_channel{`

`port = 8649` #通过 tcp 协议监听的端口, 在远端可以通过连接到 8649 端口得到监控数据

}

在一个集群内, 所有客户端的配置是一样的。完成一个客户端配置后, 将配置文件复制到此集群内的所有客户端主机上即可完成客户端主机的配置。

5、Ganglia web 端配置

Ganglia 的 web 监控界面是基于 PHP 的, 因此需要安装 PHP 环境。

有两种方式安装 Ganglia 的 web 监控界面, 一种是 yum 直接安装, 另一种是通过源码安装。PHP 环境的安装这里不做介绍, 大家可以在 <http://sourceforge.net/projects/ganglia/files/>

下载 ganglia-web 的最新版本，然后将 ganglia-web 程序放到 Apache Web 的根目录即可，这里我们下载的版本是 ganglia-web-3.7.1。

配置 Ganglia 的 Web 界面比较简单，只需要修改几个 php 文件即可。首先是 conf_default.php，可以将 conf_default.php 重命名为 conf.php，也可以保持不变，Ganglia 的 Web 默认先找 conf.php，找不到会继续找 conf_default.php，需要修改的内容如下：

```
$conf['gweb_confdir']="/var/www/html/ganglia"; #ganglia web 的根目录
$conf['gmetad_root']="/opt/app/ganglia"; # ganglia 程序安装目录
$conf['rrds']="{$conf['gmetad_root']}/rrds"; #ganglia web 读取 rrd 数据库的路径，这里
是/opt/app/ganglia/rrds
$conf['dwoo_compiled_dir']="{$conf['gweb_confdir']}/dwoo/compiled"; #需要“777”权限
$conf['dwoo_cache_dir']="{$conf['gweb_confdir']}/dwoo/cache"; #需要“777”权限
$conf['rrdtool']="/opt/rrdtool/bin/rrdtool"; #指定 rrdtool 的路径
$conf['graphdir']=$conf['gweb_root'].'graph.d'; #生成图形模板目录
$conf['ganglia_ip']="127.0.0.1"; #gmetad 服务所在服务器的地址
$conf['ganglia_port']=8652; #gmetad 服务器的交互式提供监控数据端口发布
```

这里需要说明的是：“\$conf['dwoo_compiled_dir']”和“\$conf['dwoo_cache_dir']”指定的路径在默认情况下可能不存在，因此需要手动建立 compiled 和 cache 目录，并授予 Linux 下“777”的权限。另外，rrd 数据库的存储目录/opt/app/ganglia/rrds 一定要保证 rrdtool 可写，因此需要执行授权命令：

```
Chown -R nobody:nobody /opt/app/ganglia/rrds
```

这样 rrdtool 才能正常读取 rrd 数据库，进而将数据通过 Web 界面展示出来。其实 ganglia-web 的配置还是比较简单的，一旦配置出错会给出提示，根据错误提示进行问题排查，一般都能找到解决方法。

四、扩展 Ganglia 监控功能

1、通过 gmetric 接口扩展 Ganglia 监控

gmetric 是 Ganglia 的一个命令行工具它可以将数据直接发送到负责收集数据的 gmond 节点，或者广播给所有 gmond 节点。

在 Ganglia 安装完成后，会在 bin 目录下生成 gmetric 命令。下面通过一个实例介绍一下 gmetric 的使用方法：

```
[root@cloud1~]# /opt/app/ganglia/bin/gmetric \
> -n disk_used -v 40 -t int32 -u '% test' -d 50 -S '8.8.8.8:cloud1'
```

其中：

- n，表示要监控的指标名。
- v，表示写入的监控指标值。
- t，表示写入监控数据的类型。
- u，表示监控数据的单位。
- d，表示监控指标的存活时间。
- c，用于指定 ganglia 配置文件的位置。
- S，表示伪装客户端信息，8.8.8.8 代表伪装的客户端地址，cloud1 代表被监控主机的主

机名。

2、python 扩展插件

现成可用扩展插件：

https://github.com/ganglia/gmond_python_modules

五、Ganglia 的优势与注意事项

- 1、可以轻松监控上万台服务器，数据延时在 10s 以内。
- 2、分布式架构，扩展方面，非常适合多地跨机房部署。
- 3、与 centreon 无缝整合，实现监控、报警一体化。
- 4、数据存储磁盘 IO 可能成为瓶颈，需要高性能磁盘做支撑。

下期公开课预告

主题：智能化运维监控平台构建实战

内容：

- 1、Centreon 的快速安装和使用经验
- 2、如何实现 ganglia 与 centreon 的集成
- 3、ganglia 与 centreon 集成实现的几个脚本

参与方式：

加入 Linux 运维专家 (134896298), 然后关注群公告, 每周会定期进行公开课技术分享, 具体开课时间会进行提前通知, 欢迎大家届时参加。

参与方式：打开爱维 Linux 腾讯课堂：<http://ke.qq.com/course/117291>，点击报名，即可免费参加，如果你没来得及参加线上直播，那么可以通过访问爱维 Linux 交流论坛：<http://i.iivey.com/> 观看公开课的录播视频。

爱维 Linux 实战培训

爱维 Linux ,专注与 Linux 运维实战教育 ,课程涉及 Linux 运维、集群架构、大数据、虚拟化等 ;所有授课教师都来自企业一线 ,确保学员能够学习到最优质的课程和最好的学习体验 !

开班介绍

爱维 Linux , 专注 Linux 运维实战教育 , 我们开设了两个班级 :

- 高薪运维入门提高班 详情 : <http://www.iivey.com/666-2>
- 高薪运维实战提升班 详情 : <http://www.iivey.com/archives/66>

授课模式

课程汇聚了以**南非蚂蚁**领衔的行业顶尖技术专家 10 年一线工作经验和培训心得 ,课程由浅入深 ,循序渐进 ,能够帮助学员们系统学习 Linux 一线经验 ,并迅速掌握 Linux 的各种应用技能。

我们的授课方法 :

直播+录播+理论结合实际+实战技巧+经验分享+实时互动+专业学习教材

开课时间

入门提高班将在 3 月开课 ,而实战提高班将在 5 月份开课 ,5 个月的授课时间 ,现在入门提高班接受报名 ,有意向的朋友可通过如下方式联系我们 :

QQ : 397824870 (蚂蚁老师) 3335603751 (章老师) 18966929688 (王老师)

微信 : ixdba8

课程下载 :

百度网盘：

<http://yun.baidu.com/share/link?shareid=1471417839&uk=4245145857>

360 云盘：<https://yunpan.cn/OcY6DdsPd53vBI>（提取码：0ca8）

在线观看：<http://edu.csdn.net/agency/index/187>