

文章编号: 1006-4354 (2005) 01-0025-04

# 华云神箭高性能集群计算机系统管理与维护

夏巧利, 高 宇, 燕东渭, 邸永强

(陕西省气象台, 陕西西安 710014)

中图分类号: P409

文献标识码: B

华云神箭高性能集群计算机系统于 2003 年 6 月投入业务运行以来, 在全省的精细化预报业务中发挥着重要作用。根据系统的运行情况, 总结了系统管理维护技术方法, 希望能对业务人员有所帮助。

## 1 系统安装

### 1.1 操作系统安装

并行计算机的操作系统为 Linux RedHat 8.0, 以 custom 方式选全部 (everything) 安装。1 至 9 号节点机均安装 Linux RedHat 8.0。业务处理服务器的操作系统为 Windows2000, 以标准方式安装。

### 1.2 初始引导

初始化引导主要用于构建各并行计算模块间的数据交换功能。引导方法: 将安装光盘插入 9 号计算节点的光驱, 在 9 号机上运行命令:

```
% mount /mnt/cdrom
% cd /mnt/cdrom
% rcp-r InsDir $ {目的计算机名}: /
切换到目的计算机上运行命令:
% cd /InsDir
% ./netconf
```

按屏幕提示操作: 当提示 New password: 时输入 parusr。提示 Retype new password: 时输入

收稿日期: 2004-09-27

作者简介: 夏巧利 (1963-), 女, 陕西户县人, 高级工程师, 从事气象信息网络系统管理。

一分类中的云的液水含量可能有数倍的差异。云的宏观观测与辐射计的探测相结合, 可以更准确地描述云的特征。

在层状云降水产生之前, 水汽、液水含量有一低值 (横坐标 32 h 左右), 这一特征在过去的辐射计探测空中水的研究也有发现, 可作为降水前的特征之一, 用于降水临近预报的参考, 比降水开始时间提前 5 h。

## 4 小结

利用双波长 (0.95 cm 和 1.25 cm) 地基微波辐射计探测空中水汽、液水含量的变化, 反演水汽、液水含量值为大气水的研究提供参考。在降水开始前 5 h 可提供降水产生的临近预报。今后还需进一步累积探测数据量, 更好地为人工影响天气工作服务。

## 参考文献:

- [1] H D Toong, D H staelin. Passive Microwave Spectrum Measurement of Atmospheric Water Vapor and Cloud [J]. J Atmos Sci, 1970, 27 (5): 781-784.
- [2] 周秀骥. 大气微波辐射及遥感原理 [M]. 北京: 科学出版社, 1982. 420.
- [3] David H Staelin. Measurement and Interpretation of the Microwave Spectrum of the terrestrial Atmosphere Near 1-cm Wavelength [J]. J G R, 1966, 71 (12): 2875-2881.
- [4] 段 英, 吴志会. 利用地基遥感方法监测大气中气态、液态水含量分布特征的分析 [J]. 应用气象学报, 1999, 10 (1): 34-40.
- [5] 梁 谷, 李 燕. 层状云的降水潜力 [J]. 陕西气象, 2003, (5): 28-30.

parusr。提示 Please input the computer name want to install. [node1] 时, 要求输入机器名, 1 至 9 号计算节点机器名依次为 node1, node2, … node9。

最后测试网络是否已经连通, 在业务化处理服务器中点“开始 → 运行”, 输入“ping 30.0.0.X”(各并行计算机节点 IP), 如果不通检查各并行节点与交换机、业务化处理服务器与交换机间的连接。

### 1.3 业务处理服务器软件安装

1.3.1 安装加密狗 安装业务处理服务器软件前, 先将业务处理服务器正常关机, 再将该系统所带加密狗安装在服务器背面的并口上。

1.3.2 安装软件 将“华云神箭高性能数值天气预报业务系统安装盘”插入管理机光驱中, 运行光盘目录中的 Setup.exe。屏幕显示授权宣告, 选择同意按钮 (Y)。系统安装的默认路径为 C:\AUTOMM5, 该路径不可更改。系统安装完毕, 接着安装加密狗驱动, 点“安装”, 当显示“是否安装 16 位驱动”时点“是 (Y)”, 安装完成后“退出”。接着安装程序自动调用路径初始化配置, 在华云神箭高性能数值天气预报业务系统运行过程中使用目录: 原始数据和回传结果所在的备份文件目录 F:\mm5bak; 下载美国资料存放目录 D:\sour; 接收 T213、地面、高空报文存放目录 C:\Automm5\datatran。至此, 业务化服务器软件系统安装完毕。

### 1.4 MM5 并行环境安装

方法 1: 将“华云神箭高性能数值天气预报业务系统安装盘”插入管理机光驱中, 将安装光盘中的整个 InsDir 目录通过网络分别拷贝到每个并行计算节点中的/mm5setup 目录下, 分别在每个并行计算节点中执行命令:

```
% mv /mm5setup/InsDir /
% cd /InsDir
% ./insmm5
```

方法 2: 将“华云神箭高性能数值天气预报业务系统安装盘”插入 9 号计算节点的光驱, 在 9 号机上运行命令:

```
% mount /mnt/cdrom
```

```
% cd /mnt/cdrom
```

```
% rep-r InsDir $ {目的计算机名}: /
```

切换到目的计算机上运行命令:

```
% cd /InsDir
```

```
% ./insmm5
```

安装时检查计算模块中的/mm5setup/和 home/parusr/路径下是否存在有用数据, 如果有则将这些数据移动到根 (/) 目录下。

### 1.5 验证 MM5 并行环境

所有软件安装完成后, 在业务处理机上运行业务系统软件, 执行“系统操作”→“系统检测”→“系统自测试”确定软件是否安装正确。如果提示信息显示“测试正常完成”, 则系统已经正常安装完成, 否则需重新安装。

## 2 NAS 服务器的管理

### 2.1 接口地址设置

接口 1 接内网: IP 地址 172.23.65.10 按 LCD 流程图进行设置; 网关在 NAS 的网页“NAS 管理→接口→路由表”中设置。接口 2 接机群: IP 地址 30.0.0.30。

### 2.2 进入 NAS 管理界面

方法 1, 在业务处理机中的业务管理软件中选择“资源管理”菜单中的子菜单“NAS 管理”, 给出用户名和密码; 方法 2, 在浏览器 (Explorer) 中给出 NAS 网址, http://10.0.0.30/nas; 或 http://172.23.65.10, 选择“NAS 管理”图标, 给出用户名和密码。

### 2.3 建用户

进入 NAS 管理界面后, 点击“用户”选项下的“新建用户”, 给出新建用户的用户名、密码、用户组名、UID/GID 和容量限制等。

### 2.4 初始化 NAS (第一次安装时)

在处理机上进入//30.0.0.30; 进入“NAS 管理→系统→Write Default”; 选第二项“Erase all datas on disk”; 进入“NAS→服务→Windows→卷”新建卷 mm5bak。

## 3 HYS-G Monitor 集群管理软件

HYS-G Monitor 集群管理软件, 可基于 Web 有效地对高性能计算系统的负载及资源使用情况实时监测管理, 包括处理机忙闲程度, 存

存储器使用情况、进程数目和列表等。

集群管理软件调用方法: 按 KVM 切换开关, 切换至业务处理服务器; 在 Windows 2000 界面, 调用 IE; 输入网址: HTTP://30.0.0.9/ 调用该系统。

#### 4 系统开机和关机

开、关机步骤正确与否直接关系到集群计算机系统能否正常运行, 应严格按照开、关机步骤操作。

##### 4.1 开机步骤

开机顺序: UPS→NAS→各计算节点机和服务节点机→业务处理服务器。

UPS 开机: 接上 UPS 电源, 将 UPS 后板电源开关置于 ON; 按 UPS 前面板上的开机按钮。

NAS 开机: 按压 NAS 后面板的开机开关, 启动过程约 3 min 左右。当前面板液晶窗口出现 172.23.65.10 及日期信息时, NAS 启动完毕。

计算节点开机: 依次按压计算节点 Computer1—Computer9 机箱右上方的红色按钮, 完成计算节点开机和 Linux 系统的启动。可通过显示屏监视节点是否启动成功, 按压机箱右边 KVM 开关(上)1—8 和(中)1, 若出现 login: 信息时, 节点启动完毕。

业务处理机开机: 按压业务处理机前面板的开机开关; 按压机箱右边 KVM 开关(中)的第 3 个接上业务处理机, 屏幕显示启动信息, 启动完成出现 Windows2000 界面。

做系统自测试以确定系统是正常的。测试步骤见 1.5 节。

##### 4.2 正常关机步骤

关机顺序: 各计算节点机和服务节点→NAS→业务处理服务器→UPS。

计算节点关机: 业务系统软件中, 打开“系统操作”, 点击“关机”; 在用户登陆界面输入用户名和口令后, 从屏幕的列表中选择要关机的计算机, 点击“确定”, 所选计算机开始关机; 按压机箱右边 KVM 开关(上)1—8 和(中)1, 待出现 Power down 信息时, 按压计算节点

Computer1—Computer9 机箱右上方的红色按钮使节点电源关闭, 完成关机。

NAS 关机: 业务系统软件中, 打开“资源管理”, 点击“NAS 管理”, 输入用户名和密码, 进入 NAS 管理。选择“系统”中的“关闭 NAS”完成关机。

业务处理机关机: 按 Windows2000 正常关机方法关机。

UPS 关机: 按 UPS 前面板上的关机按钮; 将 UPS 后板电源开关置于 OFF。

#### 4.3 非正常关机方法

计算节点关机: 将显示屏连接关机节点, 按压机箱右边 KVM 开关(上)1—8 和(中)1, 出现 login: 提示信息时, 输入 root; 出现 Linux 提示符 # 后, 输入 Halt, 开始关机, 出现 Power down 信息时, 按压计算节点机箱右上方的红色按钮使节点电源关闭, 完成关机。若上述方法不行, 强行按压计算节点机箱右上方的红色按钮使节点电源关闭, 完成关机。

NAS 关机: 强行按压 NAS 前面板黑色按键直至完成关机。

业务处理机: 强行按压业务处理机机箱前面板黑色按键直至完成关机。

## 5 业务处理系统日常维护与故障对策

### 5.1 日常维护步骤

检查是否将 supernas 机上的 mm5bak 目录映射为 F 盘。检查 NAS 服务器、处理机的 D 盘是否有空间。每天检查 T213 资料、AVN 资料和 MRF 资料是否齐全。

T213 资料检查: 进入到 F:\1\T213\日期(date-1)\次数(1或2)\inputdata 目录, 查看是否有 \*.3dy 和 \*.DAT 和 ok 文件。如果有 \*.3dy 而没有 ok, 则添加一个 ok 文件, ok 文件在 C:\AutoMM5 目录中有备份。如果没有 \*.3dy, 打开 C:\AutoMM5\m\_\_setup\t213nwp\_\_1.ftp (20 点资料), 或 C:\AutoMM5\m\_\_setup\t213nwp\_\_2.ftp (08 点资料), 文件中列出了所有所缺的 T213 报文, 先从相关途径收齐所缺资料, 再运行 C:\AutoMM5 目录下的 pred0.exe。如没有 \*.DAT 文件, 则先查看 C:\AutoMM5\m\_\_setup\nuns.ftp, 该文件列出所有未收到的地面、高空实况报文, 先从相关途径收齐所缺资

料,再运行 C:\AutoMM5 目录下的 pred0.exe。

AVN 资料检查:进入到 F:\1\AVN\日期(date-1)\次数(1或2)\inputdata 目录,查看是否有 \*.3dy 和 ok 文件。如果有 \*.3dy 而没有 ok,则添加一个 ok 文件。如果没有 \*.3dy,则先检查 D:\Sour\日期(date-1)\1或2\目录下有无 AVN 原始文件,即是否有 gblav.t12z.pgrbf00(12,24,36,48,60,72)7 个文件或 gblav.t00z.pgrbf00(12,24,36,48,60,72)7 个文件,如果软件界面显示全部下载完成,但以上 7 个文件不全,则需重新下载。

## 5.2 常见故障处理

### 5.2.1 业务软件任务表中出现多个“运行”状态

先停止所有任务,分析任务号,检查该号任务是否有结果;有结果则删除该号任务,没有结果则恢复该号任务;恢复当天任务。

### 5.2.2 业务软件任务表中有很多天的业务化任务

删除所有非当天的任务;重新启动所有计算节点和业务处理机。

### 5.2.3 如果有资料,运算时提示没有原始数据

在主控机的/mm5bak 目录下使用 ls 命令,看能否列出 NAS 服务器中 mm5bak 目录下的所有内容,其中包括有 1,AT0301GR,TT0301GR,RF021GR 等目录,如果没有则重启主控机,是由于 NAS 没有映射到主控机上。

### 5.2.4 计算过程中出错

MM5 的计算分 3 个阶段并用 3 个信号灯表示,无论那个阶段出错(信号灯为红色)均无计算结果。这时需要重新运行软件并建立任务提交作业。任务建立时必须给出资料来源(用 T213 资料还是美国 AVN、MRF 资料计算)、计算方式(用同化还是非同化方式计算)、资料时间(用 MM5 计算的资料时间)、资料时次(用 20 时资料还是 08 时资料)、初始场形成方案、MM5 运行方案(共 GR, SQ, KF, BM, B 五种方案)、开始时间(指定新选择增加业务化作业开始计算的时间)、作业调度优先级(1 为最

高优先级,2 为业务化作业、3 级较高、4 级为普通任务)、是否容错、最小分辨率、网格嵌套、结果处理等。

### 5.2.5 系统出错处理

(1)测试 MM5 运行出错,再运行一次系统自测试,若此错误消失,则通过;若三次测试都依然出错,则需要重新安装 MM5 程序。

(2)同化出错转非同化,仍无法完成 MM5 第一步,错误原因一般为资料不可用,以 3 号级别任务再提交一次或更换主控机再提交一次,若还是不能计算,重新启动机群再试,若还出现同样错误必须更换资料。

(3)MM5 运行时间在非容错状态下超过 1.5 h,或容错状态下超过 3 h,用户界面显示仍然在运行 MM5 第 2 步,系统可能已死机,按如下步骤操作:检查 NAS 是否正常(如:能否在客户端连接并进行文件操作)。如无法完成,则重启 NAS 及整个机群。如 NAS 正常且与服务器的连接正常(查看当前任务状态,如不出现“服务器端无返回”错误,则连接正常),则可以通过业务机提供的命令重新启动机群。如 NAS 正常而与服务器的连接不正常,则需手工重启整个机群。

(4)计算节点故障,一般计算节点故障不影响机群的计算,但会降低整体计算性能,应查找原因尽快恢复。若是当前设置的主控(计算)节点故障,可在软件中更换主控节点即可。

(5)千兆交换机故障,需要尽快维修,若要在维修期间计算,可采用百兆网来实现计算:将插入千兆交换机 1—9 插口的网线,接入百兆交换机对应的位置,即可重新启动系统计算。

(6)百兆交换机故障:需要尽快进行维修,若要在维修期间计算,可采用下列步骤仅应用千兆网实现计算:将插入百兆交换机 10—11 插口的网线,接入千兆交换机对应的位置,即可重新启动系统计算。