

# 实验四 集群结构体验——MPI 编程

## 一、MPI 内容及 cannon 程序代码

见《多路处理器实验教学系统实验指导书-计算机系统结构.pdf》。

## 二、MPI 编程原理

(1) OpenMP 编程：多线程是一种便捷的模型，其中每个线程都可以访问其它线程的存储空间。因此，这种模型只能在共享存储系统之间移植。

(2) MPI 编程：集群计算机没有共享存储，当面向非共享存储系统开发并行程序时，程序的各部分之间通过来回传递消息的方式通信。要使得消息传递方式可移植，就需要采用标准的消息传递库。这就促成的消息传递接口(Message Passing Interface, MPI)的面世，MPI 是一种被广泛采用的消息传递标准。

MPI 标准定义了一组具有可移植性的编程接口。由于 MPI 提供了统一的编程接口，程序员只需要设计好并行算法，使用相应的 MPI 库就可以实现基于消息传递的并行计算。MPI 程序是基于消息传递的并行程序。消息传递指的是并行执行的各个进程具有自己独立的堆栈和代码段，作为互不相关的多个程序独立执行，进程之间的信息交互完全通过显式地调用通信函数来完成。

## 三、实验目标和内容

本实验用 MPI 完成 Cannon 算法的并行计算。在 4 路 4 核龙芯的配置环境下，调整 MPI 进程数量 (1~16)，计算加速比，分析通信在 MPI 程序中对加速比的影响。注意 MPI 中消息传递的使用，并对比实验 10 中与 SMP 结构中使用共享内存的差异。

## 四、实验步骤

### (1) 《多路处理器实验教学系统用户使用手册 2 版 1.1.pdf》实验箱开机准备

1) 打开实验箱，看到的电路板是 1 个主控电脑 (2H 处理器) 占左 1/3 部分，SSD 硬盘附近有并排 2 个小按键，左开机右复位。右 2/3 四个风扇为 4 个 3A 处理器 CPU，分别代表 4 个计算节点。每个 CPU 是 4 核的，最多可使用  $4 \times 4 = 16$  核。

2) 接电源，接 VGA 显示器，接 USB 口鼠标及 USB 键盘，其它连线按键都不要动。

3) 按开机键启动主机 (linux 系统)，在账号选择时不选默认，选其它—输入 root 回车—输入密码 loongson 回车 (注意区分大小写)，进入 linux 图形桌面。桌面上有 5 个图标，其中第 3 个为龙芯实验箱设置软件 Loongson Multi-node Teaching Control System 应用程序，龙芯实验箱设置软件是用来操作和设置多路处理单元的辅助工具，集成了模式切换功能，计算节点开关机和重启功能，串口展示和切换功能。同时软件用可视化图形实时显示各计算节

点的开关状态。具体见《多路处理器实验教学系统用户使用手册 2 版 1.1.pdf》3.2.2 软件启动多路处理系统。

4) 在软件启动多路处理系统软件主界面上点**单节点模式** (Single Node)，并点 4 个计算节点**开机按钮**使 4 个 3A 计算节点开机启动运行，启动后均运行于 linux 操作系统中。

5) 其它配置出厂已经配置好，不要动。记住以下信息：

4 个节点所有 root 账号均为 loongson，sdbox 账号均为 sdbox；

主机 IP: 198.168.100.250；

计算节点 IP 分别为：IP1=198.168.100.1；IP2=198.168.100.2；IP3=198.168.100.3；IP4=198.168.100.4

计算节点名称为：loongsonbox-n1, loongsonbox-n2, loongsonbox-n3, loongsonbox-n4

## (2) 多路处理器实验教学系统用户使用手册 2 版 1.1.pdf3.2.5 建立分布式系统

声明，本分布式系统可在节点 1 的 **root 用户** 下操作完成实验，也可在节点 1 的 **sdbox 用户** 下操作完成实验，root 权限高，且已经出厂配置成了 4 个节点 MPI 无密码访问，实现简单，但不要动与实验无关的操作；sdbox 是一般用户，实验说明书中提倡的操作方式，下面分两种用户下操作进行提示。

### **A———root@192.168.100.1 帐户下实验步骤：**

1) 分布环境须在计算节点 1 (主节点上) 操作，节点 1 带领其它 3 个节点组成一个 4CPU 的机群，在主控机的桌面上右键选**在终端中打开**，启动一个终端 (可同时启动多个窗口)，使用 SSH 命令 (**ssh root@192.168.100.1 回车，密码 loongson 回车**) 登入到节点 1 电脑进行后续操作。登录节点 1 后，可用 ls,pwd 等命令测试一下。

### **2) 建立 mpd.hosts 文件**

在节点 1 (loongsonbox-n1) 主处理单元 root 帐户的家目录 (~表示家目录) 下，查看是否存在 mpd.hosts 文件，如有这个文件查看并修改编辑里面的内容，如不存在就新建这个 mpd.hosts 文件，添加需要连接的计算节点名称，保存退出。使用 cat mpd.hosts 命令查看，使用 vi mdp.hosts 命令编辑修改。连接 loongsonbox-n1 ~loongsonbox-n4 4 台主机。mpd.hosts 内容应为：

```
loongsonbox-n1
loongsonbox-n2
loongsonbox-n3
loongsonbox-n4
```

### **3) 检查并追加修改 /etc/hosts 文件内容**

```
。 。 。
192.168.100.1 loongsonbox-n1
192.168.100.2 loongsonbox-n2
192.168.100.3 loongsonbox-n3
```

```
192.168.100.4 loongsonbox-n4
```

```
~
```

```
“/etc/hosts” 10L, 300C
```

#### 4) 启动分布式集群

在主处理单元节点 1 的家目录下启动分布式集群，采用如下命令：

```
mpdboot -n 4 -f mpd.hosts
```

-n 后的数字表示要启动的处理单元个数，一般是不大于 mpd.hosts 文件中的机器数，即 1~4。

```
mpdtrace 或 mpdtrace -l
```

测试命令，若输出联机机群的全部节点名，说明启动成功，这样就可以执行 MPI 程序。

**B———sdbox@192.168.100.1 帐户下实验步骤:如不存在 sdbox 用户可以自行创建 useradd -s /bin/bash sdbox, passwd sdbox,由于是共享, 节点 1 创建一次 4 个节点都有了, 然后按后续方法配置无密钥访问集群环境**

1) 分布环境须在计算节点 1 (主节点上) 操作, 节点 1 带领其它 3 个节点组成一个 4CPU 的机群, 在主控机的桌面上右键选在终端中运行, 启动一个终端 (可同时启动多个窗口), 使用 SSH 命令 (ssh sdbox@192.168.100.1 回车, 密码 sdbox 回车) 登入到节点 1 电脑进行后续操作。登录节点 1 后, 可用 ls,pwd 等命令测试一下。

#### 2) 建立 mpd.hosts 文件

在节点 1 (loongsonbox-n1) 主处理单元 sdbox 帐户的家目录 (~表示家目录) 下, 查看是否存在 mpd.hosts 文件, 如有这个文件查看并修改编辑里面的内容, 如不存在就新建这个 mpd.hosts 文件, 添加需要连接的计算节点名称, 保存退出。使用 cat mpd.hosts 命令查看, 使用 vi mdp.hosts 命令编辑修改。连接 loongsonbox-n1 ~ loongsonbox-n4 4 台主机。mpd.hosts 内容应为:

```
loongsonbox-n1
```

```
loongsonbox-n2
```

```
loongsonbox-n3
```

```
loongsonbox-n4
```

#### 3) 检查并追加修改 /etc/hosts 文件内容

```
192.168.100.1 loongsonbox-n1
```

```
192.168.100.2 loongsonbox-n2
```

```
192.168.100.3 loongsonbox-n3
```

```
192.168.100.4 loongsonbox-n4
```

#### 4) .mpd.conf 创建:

第 1 种方法, 节点 1 上 sdbox 用户家目录下创建.mpd.conf (带.的隐藏文件), 里面的内容与本机 root 用户/etc/mpd.conf (不带.) 抄写。内容为:

```
secretword=loongson
```

MPD\_SECRETWPRD=mr.chen

## 第 2 种复制方法:

root 用户下

```
cp /etc/mpd.conf /tmp/mpd.conf;
```

```
chmod a+r /tmp/mpd.conf
```

切换到 sdbox 用户下使用命令:

```
cat /tmp/mpd.conf > /home/sdbox/.mpd.conf (创建带隐藏文件, > 是覆盖, >> 是追加)
```

第 3 步, 修改.mpd.conf 权限为 600, `chmod 600 /home/sdbox/.mpd.conf`

## 5) 启动分布式集群

在主处理单元节点 1 的家目录下启动分布式集群, 采用如下命令:

```
mpdboot -n 4 -f mpd.hosts
```

-n 后的数字表示要启动的处理单元个数, 一般是不大于 mpd.hosts 文件中的机器数, 即 1~4。

```
mpdtrace 或 mpdtrace -l
```

测试命令, 若输出联机机群的全部节点名, 说明启动成功, 这样就可以执行 MPI 程序。

关闭机群用 `mpdallexit` 命令。

6) 如 5) 启动 `mpdboot` 过程报错还须建立无密钥访问集群的配置过程, 即在本机如节点 1 的 sdbox 家目录下使用 `ssh-keygen -t rsa` 生成 `id_rsa` 和 `id_rsa.pub`, 生成的密钥在 `~/.ssh` 中, 用 `ls -al` 可以查看。将公钥发给其它 3 个节点, 使用 `ssh-copy-id sdbox@192.168.100.2` (3 个机器 2, 3, 4), 并进行登入验证, 第一次可能需要密码。如此过程在节点 2-4 都做一遍, 即可实现无密码机群访问 (需要以 `ping -c 5 192.168.100.X` 为前提条件)。

## (3) 实验正式开始——“用 MPI 完成 Cannon 算法的并行计算”

1) 源程序 C 代码获取: 从指导书中的算法程序编辑成 `cannon.c` 文件, 本实验可从教师机上利用 U 盘 copy 到实验箱主机的共享文件夹中, 可使用 linux 桌面资源管理器完成。

2) 在主控服务机上操作复制 `cannon.c` 文件到机群主机节点 1 用户 (root 用户/sdbox 用户) 家目录中, 编译后将生成的可执行目标文件如 `cannon911` 复制到其它 3 个节点的对应帐户家目录中 (由于 4 个节点使用共享文件夹, 所以实验箱每个节点不用复制就都存在了)。

主控服务器机上使用: `scp cannon.c root@192.168.100.1` (拷到节点 1 root 用户家目录中)

或者: `scp cannon.c sdbox@192.168.100.1:/home/sdbox` (拷到节点 1 sdbox 用户家目录中)

节点 1 上使用编译命令为:

```
mpicc -lm -o cannon911 cannon.c
```

将目标文件复制到其它节点: (由于本机群共享文件系统, 以下 3 条复制命令无须执行)

运行程序命令 (节点 1 上运行):

```
mpirun -np 16 ./cannon911 4
```

注: (`np` 后面的参数为 1, 4, 9, 16 整数的平方, 最大 16, 最后的 4 须为 `sqrt(np)` 的整数倍)

**运行结果为输出 A，B，A\*B 矩阵的值**

《多路处理器实验教学系统实验指导书-并行算法实践.pdf》