

中国人民大学物理系高性能计算集群“Dirac”介绍

王璟琨、刘凯

2016.11

1 Dirac集群简介

1.1 硬件介绍

目前Dirac集群共有72个节点（1个管理节点，2个存储节点，64个瘦节点，5个胖节点），共2288个计算核心，约6.6 TB内存，87 TB集中存储空间，理论计算峰值为76.88万亿次。管理网络采用千兆以太网，计算网络采用56Gb/s FDR Infiniband网络无阻塞链接。并行存储架构采用2台机架式I/O服务器，通过冗余的56Gb/s FDR IB 线缆连接后端DDN SFA7700X双控制器存储系统。

管理节点1个，型号为HP DL380 G9服务器，配置2个16 核Intel Xeon E5-2683 v4系列CPU（主频2.1GHz，L3缓存40MB，Intel QPI 9.6GT/s）；64GB内存；2块600G 10K转SAS硬盘，硬盘阵列Raid 1；1块56Gb/s FDR Infiniband HCA卡。

存储节点2个，型号为HP DL380 G9服务器，配置2个10 核Intel Xeon E5-2630 v4系列CPU（主频2.2GHz，L3缓存25MB，Intel QPI 8GT/s）；64GB内存；2块300G 10K 转SAS硬盘，硬盘阵列Raid 1；2块56Gb/s FDR Infiniband HCA卡。

计算节点69个，分布于5个HP c7000刀片机箱中，其中64个双路计算节点为HP BL460c G9刀片，5个四路胖节点为HP BL660c G9刀片。具体参数如下：

数量	型号	CPU规格	内存	硬盘
64	BL460c G9	Intel Xeon E5-2683 v4 2.1GHz 16核*2	64GB DDR4	300GB
5	BL660c G9	Intel Xeon E5-4650 v3 2.1GHz 12核*4	512GB DDR4	2*1TB

69个计算节点之间通过3台36口Mellanox Infiniband（56Gb FDR）交换机无阻塞连接，管理节点和存储节点也接入了56Gb FDR Infiniband网络。

1.2 软件介绍

集群部署和作业调度系统采用联科CHESS软件，编译器版本管理采用Modules Enviroment模块化环境。

操作系统： Linux，内核版本2.6.32-573

编译器： 系统已默认加载Intel C/C++/Fortran 编译器和MKL数学库，用户可以在管理节点编译程序：Fortran为ifort；C为icc；C++为icpc

并行环境： Fortran为mpiifort；C为mpiicc；C++为mpiicpc

数学库： 推荐使用Intel MKL（包括Intel fftw，位于/opt/intel/mkl/interfaces/文件夹下），比如fortran 程序链接方法为

```
ifort code.f90 -L/opt/intel/mkl/lib/intel64/ -lmkl_intel_lp64 -lmkl_sequential -lmkl_core -lpthread
```

其他数学库全部安装在/opt/libs文件夹中，主要包括以下开源数学库：

Lapack: 开源BLAS和LAPACK库

FFTW: 快速傅里叶变换；

开源软件： GNU PLOT（开源绘图软件），用户可直接输入命令gnuplot。

1.3 系统性能

1.3.1 峰值

节点数量	型号	CPU规格	内存	理论峰值 (TFlops)	实际峰值 (TFlops)	Linpck 效率
64	BL460c G9	Intel E5-2683 2.1GHz 16核*2	64GB DDR4	68.813	61.940	90.0%
5	BL660c G9	Intel E5-4650 2.1GHz 12核*4	512GB DDR4	8.064	6.842	84.8%

注：由于节点硬件类型不同，分别做了Linpck测试，实测数据由联智科技(北京)有限公司工程师完成。

1.3.2 存储带宽

存储系统IOZONE测试，多节点多线程聚合带宽实测，以512 KB、1MB、4MB、8MB文件块大小测试，读带宽5.71-5.86 GB/s，写带宽5.06-5.22 GB/s。由DataDirect Networks公司工程师完成。