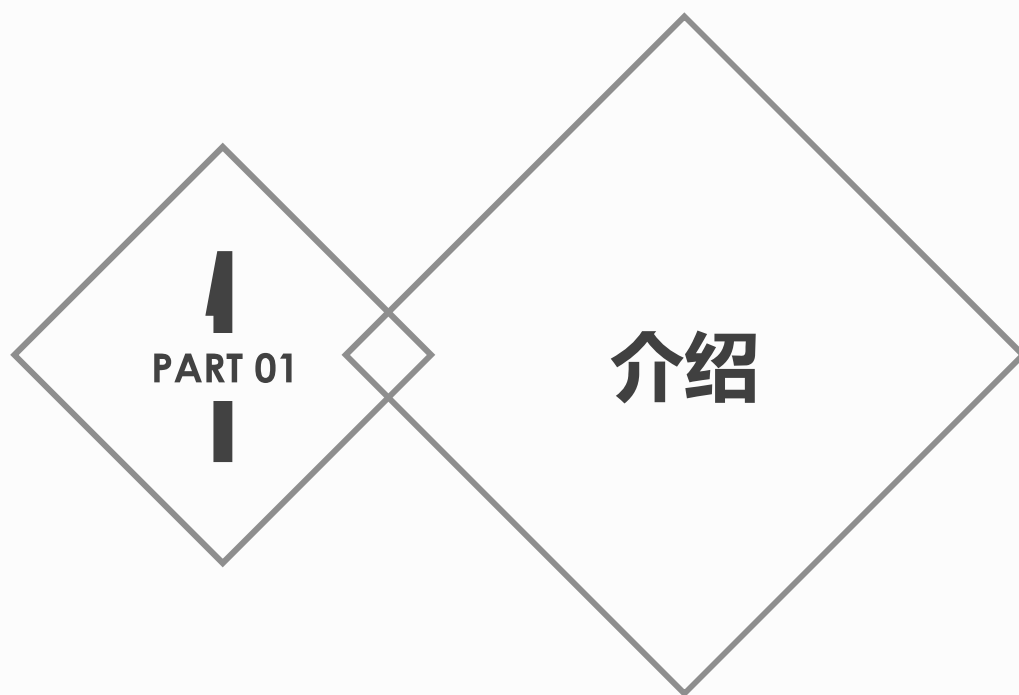


# Milkyway-2 Supercomputer System And Application

汇报人：贾朝阳

成 员：陶应娟 贾朝阳 汪浩

指导老师：龚春叶、甘新标、杨博



## 01 介绍

高性能计算（HPC）  
技术发展非常迅速

HPC被视为满足工业  
和社会要求的战略基  
础设施



第41届TOP500排行榜  
上被评为全球最快超  
级计算机

亿亿次级计算面的体  
系结构，编程模型，  
应用程序和许多其他  
方面的巨大挑战



  
PART 02

# 软硬件系统概述





02

## 软硬件系统概述：硬件系统

计算：计算机架、计算框架、计算节点

通信：胖树拓扑结构、专有的高速互连芯片



存储：层次存储结构，大容量、高带宽和低延迟的共享存储

监控诊断：集中式管理架构，实现实时监控

服务：加速新兴的信息服务应用程序，充当前端节点





## 软硬件系统概述：软件系统

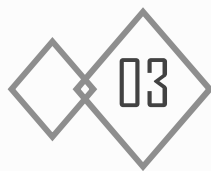




PART 03

## 系统重要组成部分





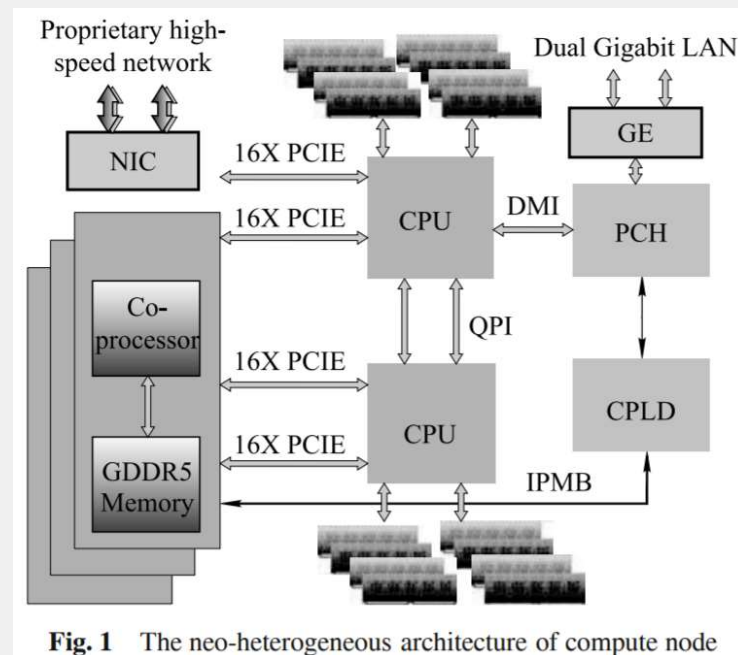
## 系统重要组成部分：异构计算节点

两个基于Ivy Bridge架构的英特尔至强处理器

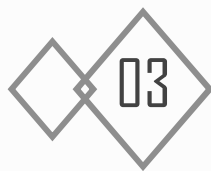
三个基于Intel MIC架构的英特尔至强协处理器

Intel QPI、PCI-E

统一编程模型的异构架构：具有相同编程模型，简化开发和优化过程







## 系统重要组成部分：TH Express-2网络

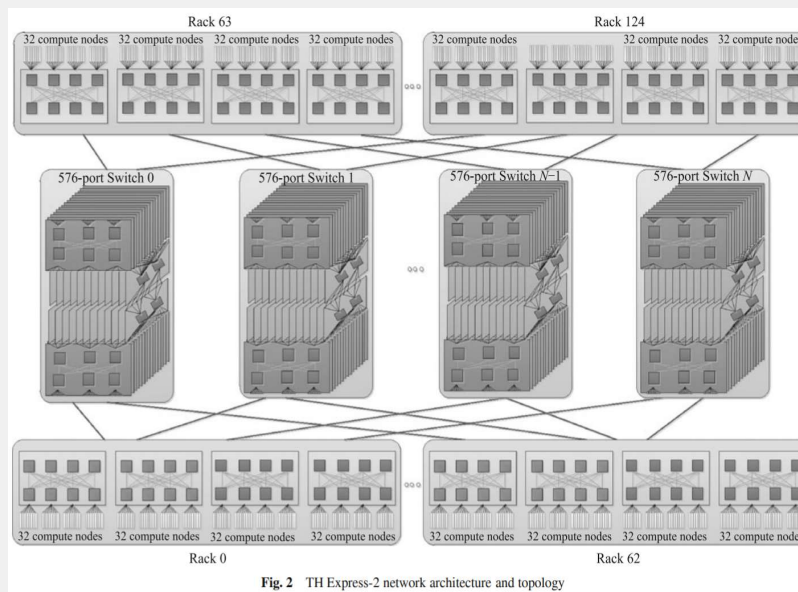
TH Express-2网络用于处理器间通信

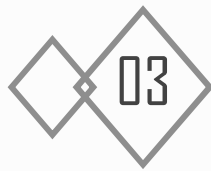
高基数路由器芯片和网络接口芯片

高基数路由器芯片用作基本构建块，  
用于创建各级交换机

galaxy express (GLEX2) 通信层，  
支持其他软件系统或流行编程模型的  
实现

MPI实现通过MPICH中Nemesis内的  
高性能网络模块 (Netmod)





## 系统重要组成部分：FT1500 CPU

16个核心被组织为四个基本单元

每个内核拥有一个专用的L1指令缓存一个L1数据缓存和一个512 KB的L2缓存

L3缓存分为四个库，每个库独立工作

每个单元内部：7×7的交叉开关网络、1×4线状网络、芯片间接口、I / O接口

片上系统处理器

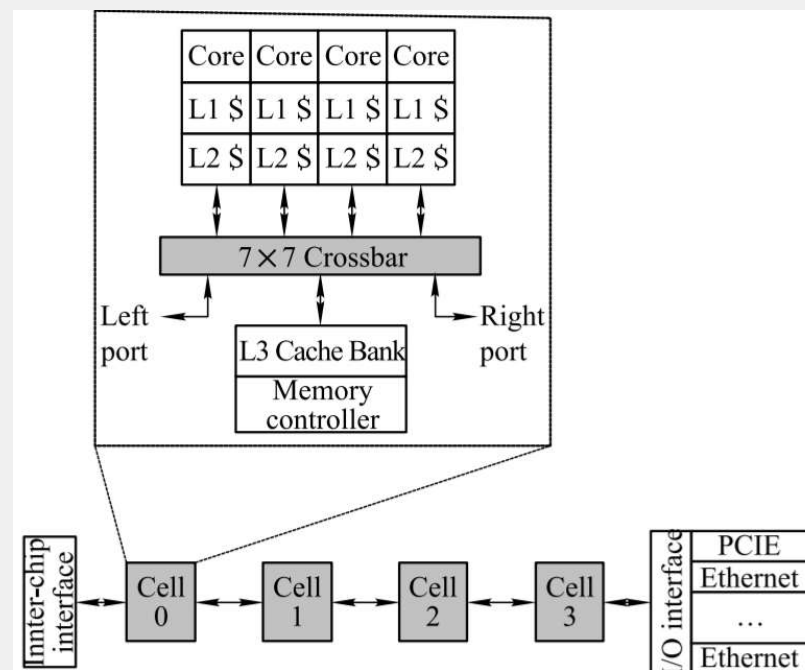
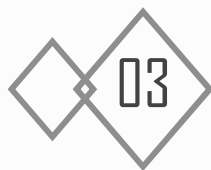


Fig. 3 Architecture of FT1500 processor

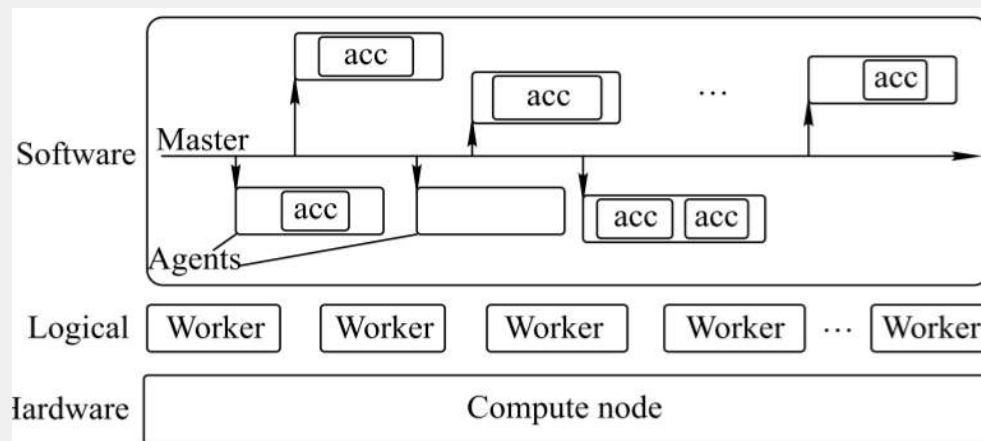


## 系统重要组成部分：OpenMC编程模型

银河机上的混合编程模型：MPI和OpenMP + X

基于指令的节点内编程模型：  
OpenMC

把计算节点中硬件资源的统一抽象作为工作者

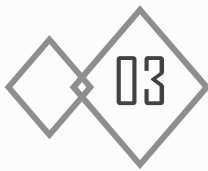


**Fig. 5** Execution model of OpenMC

所有硬件资源都可作为逻辑级别的一组工作者用于OpenMC程序

OpenMC程序中的任务根据主线程在软件级别进行组织





## 系统重要组成部分：智能系统管理

智能管理系统监控MilkyWay-2的整个系统，包括硬件和软件

实时错误探测、精确故障诊断、及时维护提示

自动管理所有故障的生命周期：自主管理。

自我意识、自我诊断、自我修复、自我保护、自我演化



**4.**  
**PART 04**  
**T**

**基准测试**





## 基准测试：LINPACK

LINPACK关注高性能计算系统的系统级性能。MilkyWay-2 LINPACK基准评估中每个进程使用一个CPU作为主机，使用1.5个Xeon Phi加速器作为协处理器。主机负责初始化、控制、通信和部分计算任务，协处理器处理大多数计算任务。

基准测试在19 452秒（大约5小时24分钟）完成，性能为33.86 PFLOPS。

T/U	<i>N</i>	<i>NB</i>	<i>P</i>	<i>Q</i>	Time	Gflops
WC15C2C4	9 960 000	1 008	248	128	19 452.04	3.386 27e+07
HPL_pdgesv()	start	time	Mon Jun 3	00:25:03	2013	
HPL_pdgesv()	end	time	Mon Jun 3	05:49:16	2013	
$\ Ax-b\ _{\infty}/(\epsilon \cdot (\ A\ _{\infty} \cdot \ x\ _{\infty} + \ b\ _{\infty}) \cdot N) =$						0.001 966 1 ... PASSED

**Fig. 7** Output of LINPACK benchmark





## 基准测试：Graph500

Graph500旨在根据计算机处理数据密集型任务的能力对计算机进行排名，性能主要受内存访问、通信和计算的影响。

关键性能指标是每秒遍历的边（TEPS）

弱缩放通过增加每个节点的负载进行测试，每个节点的内存使用量随着规模变大而增加

强缩放测试全局问题大小（即给定图中的顶点数）保持不变，而处理器的数量增加

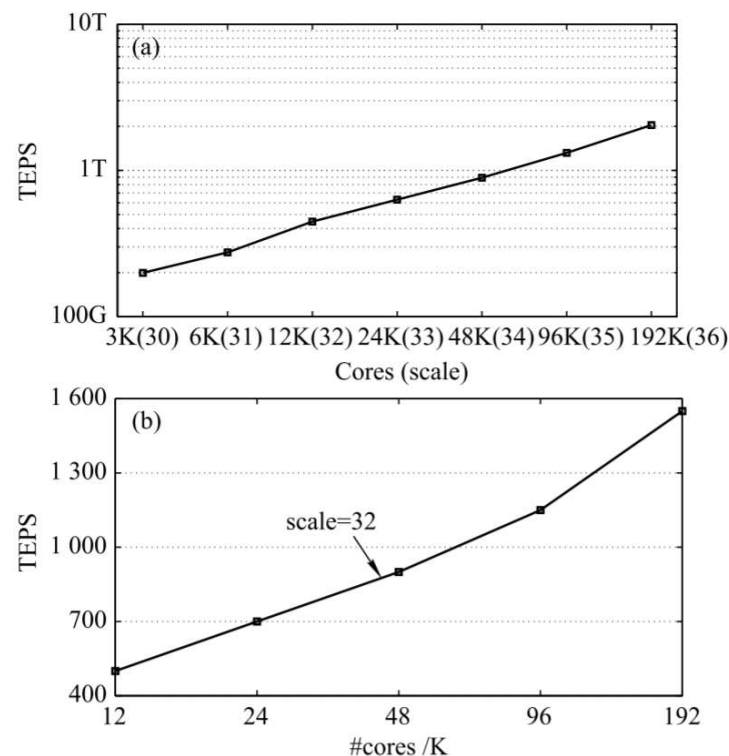
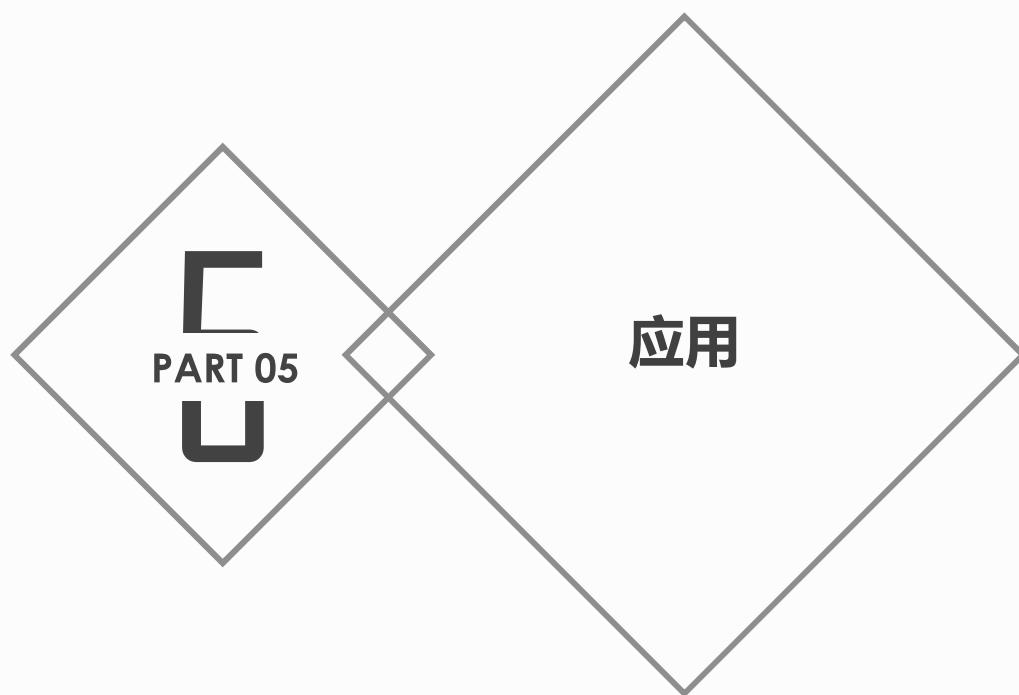


Fig. 8 Graph500 benchmark on MilkyWay-2. (a) Parallel weak scaling; (b) Parallel strong scaling







## 应用

### **Gyrokinetic toroidal gode**

一种三维粒子模拟系统，用于研究托卡马克聚变装置中的微湍流。

### **Parallel molecular docking**

虚拟高吞吐量筛选 (vHTS)  
分子对接  
使用超算计算并分析计算结果

### **全球大气模拟**

科学计算中最具挑战性的问题之一  
适合配备有处理器和加速器的异构系统进行模拟计算

