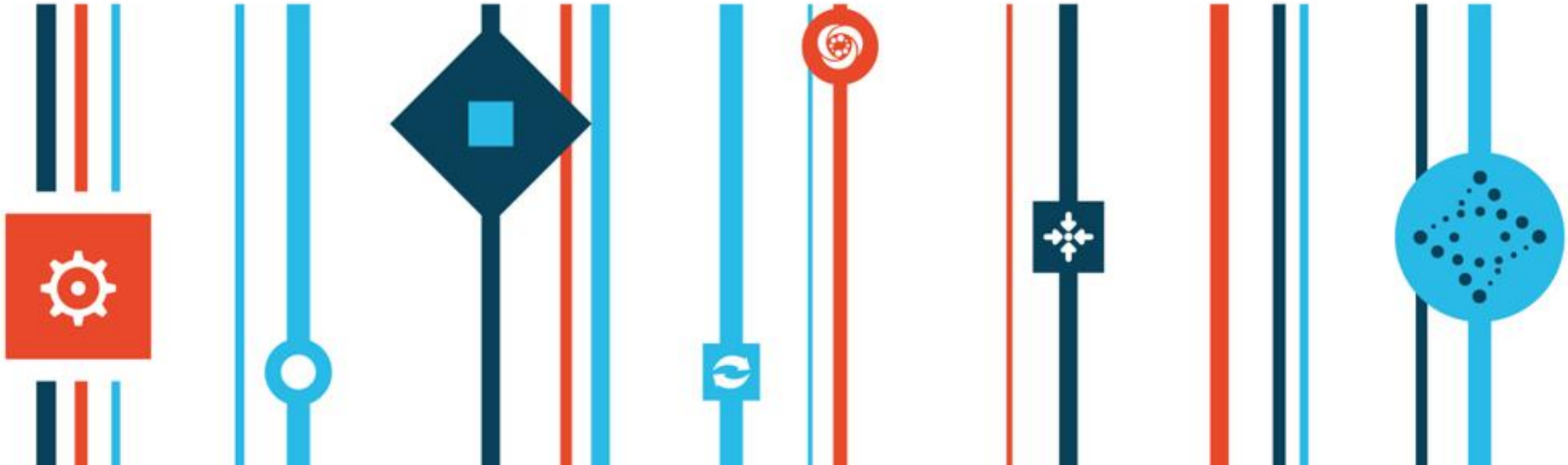
# 云计算在高性能计算中心的最佳实践

曹凡 IBM云计算中心架构师
梁毅 北京工业大学计算机学院副教授



# IBM全球云计算中心

- IBM把云计算视为一项重要战略，成立专门的云计算部门，并已在全球建立**十一个云计算中心**，为客户提供云计算技术支持和服务，提供全面的**端到端云计算解决方案**。
  - 现场设计实施云计算中心的基础架构
  - 提供云计算的高技能的人力资源支持
  - 提供下一代数据中心服务的培训
  - 快速部署和实施云计算的概念验证及试运行


美国 爱尔兰 中国 中国香港 新加坡 南非 越南 日本 巴西 韩国 印度

# IBM大中华区云计算中心

## 计世资讯的调查显示：IBM是国内客户第一想到的云计算解决方案提供商

- 以IBM中国研究院和软件开发中心的数千研发人员为后盾
- 以推动云计算在国内的发展为目标
- 结合国内客户的实际需求和业务特点
- 为客户提供端到端解决方案，充分整合IBM各产品线的优势

中国移动通信 CHINA MOBILE

北京工业大学 BEIJING UNIVERSITY OF TECHNOLOGY

中国中化集团公司 SINOCHEM CORPORATION

WUXI · CHINA

中国·东营 www.dongying.gov.cn

### 2010年

- 进一步深化云计算与行业应用的紧密结合
- 积极探索云计算在商业智能和物联网方面的应用前景

### 2009年

- 推出面向行业的云计算“6+1”解决方案
- 实现云计算在国内各行业的全面开花
- 加入由电子学会主办的云计算专家委员会

### 2008年成立

- 在中国建立第一个云计算中心——无锡

# 目录

- 高性能计算中心发展
- 云架构的高性能计算中心
- 北工大高性能计算中心的成功应用

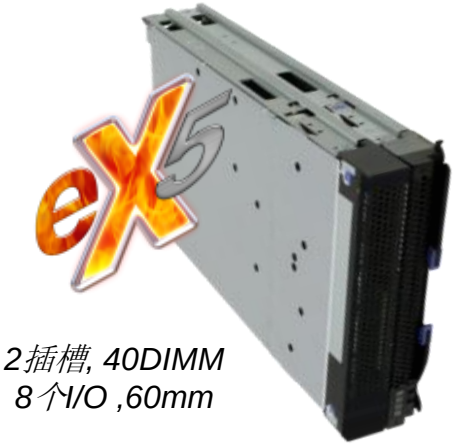
# 高性能计算应用不同/需求不同

## 侧重处理能力的高性能计算机 (CPU Intensive)

IBM BladeCenter HS22

15分钟搜索10亿多文件
IO吞吐量：134+GB每秒
实测最大的文件系统大小为~2PB
*川庆物探：1PB, 4GB每秒*

## 侧重内存的高性能计算机 (Memory Intensive)



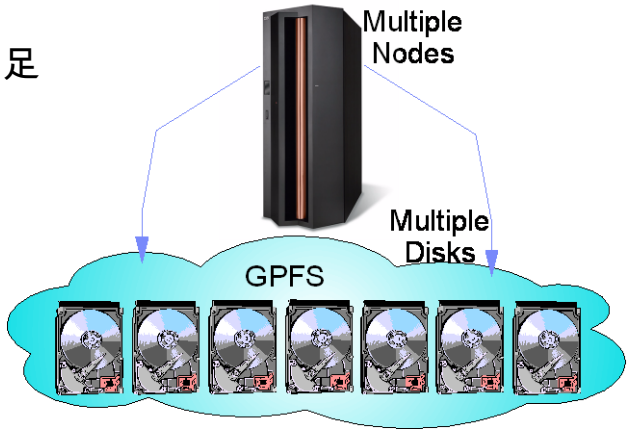

*2插槽, 40DIMM*
*8个I/O ,60mm*

IBM BladeCenter HX5

**看到：**
32cores 2.53 128GB内存，内存不足
SWAP频繁导致宕机
局部仿真20-30小时完成
可MPI扩展

*侧重内存的高性能计算*

## 侧重I/O的高性能计算机 (I/O Intensive)




IBM General Parallel FileSystem (GPFS)

# 高性能计算中心的挑战

- 应用模式复杂
  - 商用、开源、自开发
  - 多种平台需求：Linux、Windows、UNIX
  - 串行、并行
- 管理复杂
  - 作业管理
  - 应用平台管理
  - 数据管理
  - License管理
  - 用户管理
- 资源共享的需求
  - 开发、生产
  - 网格平台的需求
  - 跨WAN的资源共享
- 能源消耗巨大

# 什么是云计算

- 云计算是一种通过计算管理分配的方式共享资源的计算，计算资源可以动态部署、动态调优、动态收回。
- 在云计算基础设施中，各种计算资源被连接在一起形成统一的资源池，这些资源会被动态的分配给不同的应用和服务，满足它们在不同时刻的需求。

云的优点:

云 = 更少的投资
云 = 动态的规模
云 = 灵活和高效率

# 高性能计算云方案架构


IBM 云基础架构

# 高性能计算云服务模式

各研究人员

公众使用人群

云计算服务Portal

专用计算软件
并行运算支撑平台
操作系统平台

项目一

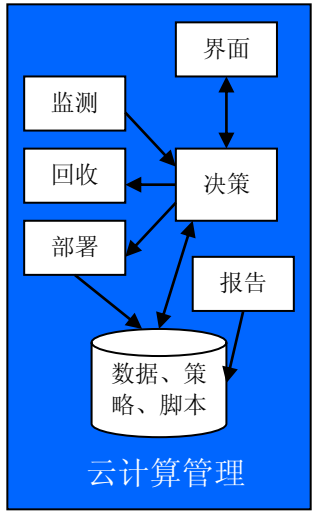
专用计算软件
并行运算支撑平台
操作系统平台

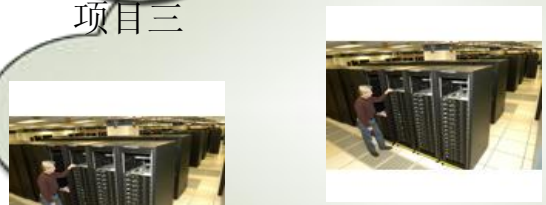
项目二

专用计算软件
并行运算支撑平台
操作系统平台

项目三

动态产生



北京

上海

其它地区

所有计算资源

# 云计算在HPC应用的优势

更多的服务模式

应用服务

平台服务

架构服务

项目用户

独享计算资源
便利的资源扩展
更多更灵活的平台应用支持
自动化服务体系

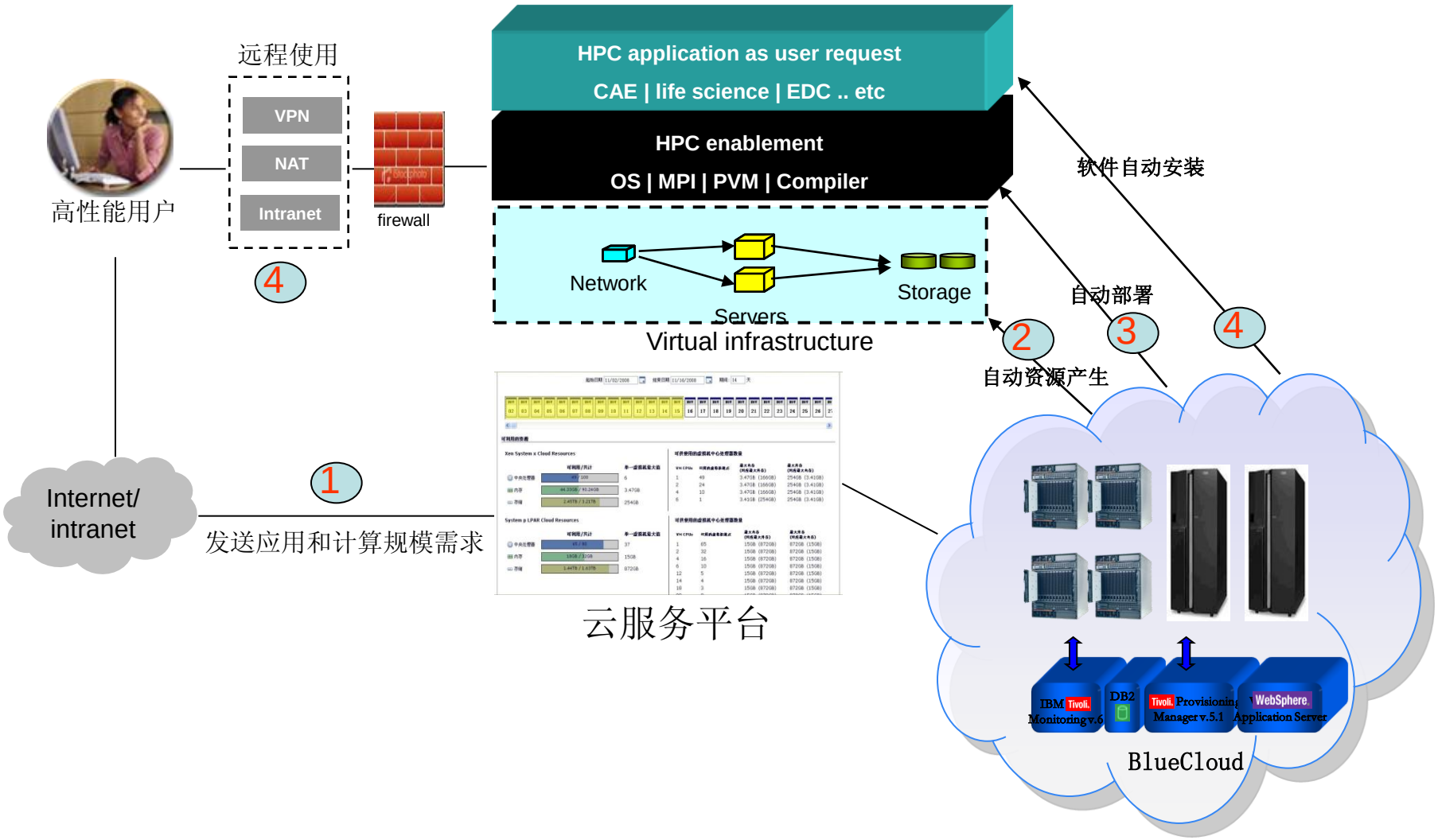

维护成本降低
更高的资源利用率
降低能源消耗

管理员

快速响应
便利的管理
安全的环境

临时用户

# 服务流程


远程使用
VPN
NAT
Intranet
高性能用户
firewall
4
HPC application as user request
CAE | life science | EDC .. etc
HPC enablement
OS | MPI | PVM | Compiler
Network
Servers
Storage
Virtual infrastructure
软件自动安装
自动部署
2
3
4
自动资源产生
Internet/
intranet
1
发送应用和计算规模需求
云服务平台
IBM Tivoli Monitoring v.6
DB2
Tivoli Provisioning Manager v.5.1
WebSphere Application Server
BlueCloud

# 计算云特点

- 适合对象
  - 高性能计算中心
  - 多用户、多学科研究平台
  - 研究院集中数据中心
- 服务模式
  - 公众服务平台
  - 基于研究课题动态分配
  - 基于研究团队资源分配
- 优点
  - 为研究者提供独立环境
  - 快速响应各种研究需求
  - 支撑多种研发平台

# IBM云计算技术在北京工业大学高性能计算平台的应用

报告人：梁毅

2010-06-11

# 提纲

1. 北京工业大学高性能计算平台简介
2. IBM云成功案例---
   北京工业大学并行计算大赛

# 北京工业大学高性能计算平台

- 建设目标

☆ 服务教学科研

☆ 开展科学研究

☆ 支撑服务北京

# 北京工业大学高性能计算平台

- 高性能计算平台建设规模及技术路线

☆ 遵循分区规划、统一管理的建设思路

☆ 采用基于云计算技术的高性能计算中心新一代解决方案

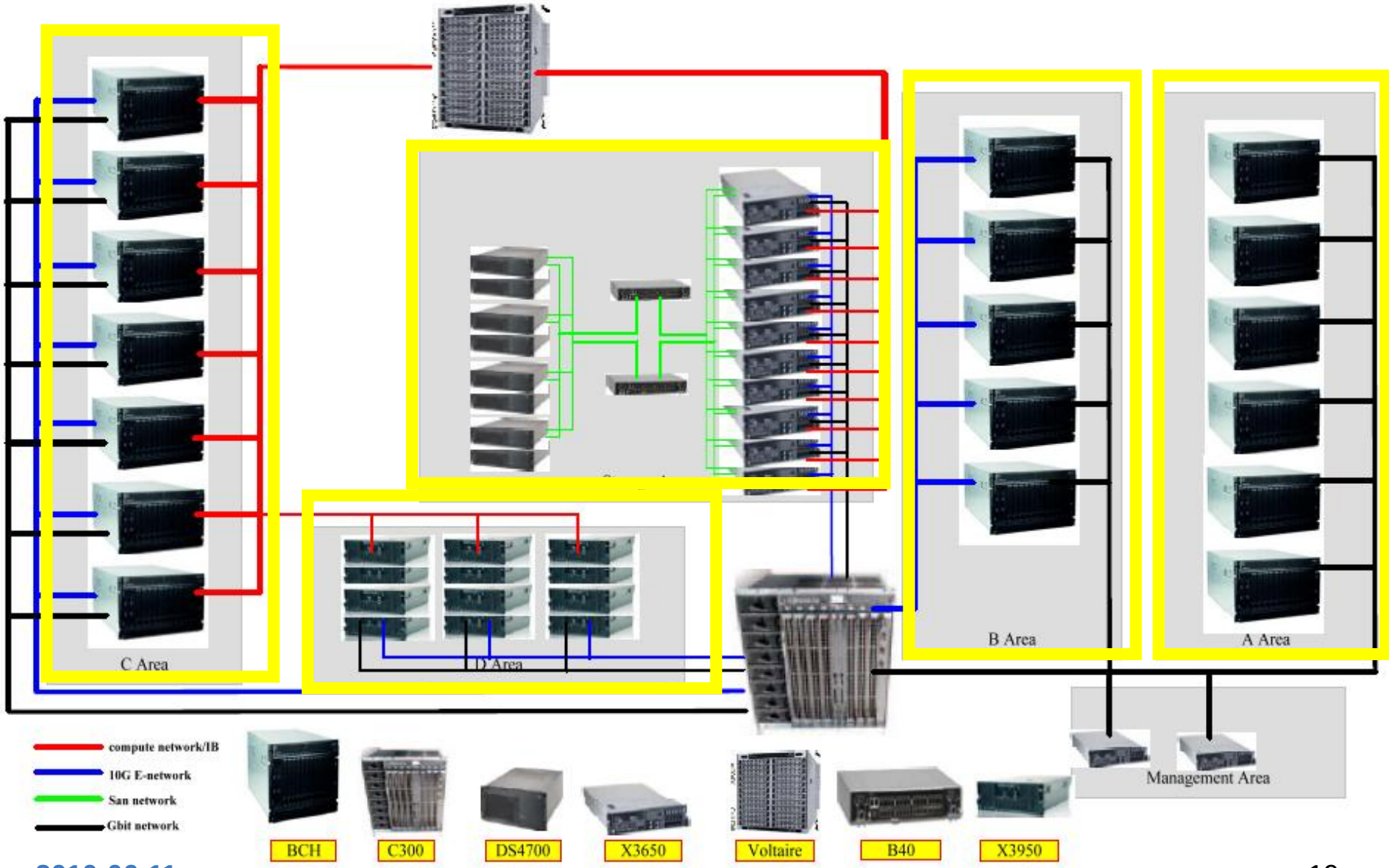
☆ 总计算能力23TFlops，总存储能力40TB，位居全国高校第三。

# 北京工业大学高性能计算平台

## 硬件基础设施概况

| | |
|---|---|
| 计算资源 | 252台IBM HS21刀片服务器 |
| | 3台IBM X3950M2大内存机架服务器 |
| 存储资源 | 4台IBM TotalStorage DS4700-70A磁盘阵列 |
| | 基于SAN架构的存储网络系统 |
| 网络资源 | 一套 Voltaire 20Gb/s Infiniband高性能网络 |
| | 一套Force10 万兆以太网 |
| | 一套Force10 千兆以太网 |

# 北京工业大学高性能计算平台



# 北京工业大学高性能计算平台

# 提纲

1. 北京工业大学高性能计算平台简介

2. IBM云成功案例---
   北京工业大学并行计算大赛

# 北京工业大学并行计算大赛

- 共10个学院，29个参赛小组
- 自选课题，86%来源于实际项目需求
- 应用类型多样，并行环境需求异构
- 所需的机群节点规模各异

# 北京工业大学并行计算大赛

- 并行环境需求

| 学院 | 应　　用 | 软 件 环 境 |
|---|---|---|
| 电　控 | 基于Hadoop的图像检索与地理信息查询系统的并行算法设计与实现 | Linux + MPI + NetBean + JDK + Hadoop + Hbase +Zookeeper + tomcat + geoserver + postgresql |
| 生　命 | 生物分子相互作用复杂网络的并行模块划分方法 | Linux + MPI + Boost_1_34_1 |
| 激　光 | 飞秒激光与等离子的相互作用机制的数值模拟 | Linux + MPI + Pvm |
| 建　工 | 建筑工程有限元分析软件OpenSees的并行优化 | Linux + MPI + OpenMPI + OpenMP + Opensees |
| 机　电 | 多孔介质化-力耦合问题的并行算法研究 | Linux + MPI + Fortran90 |
| 机　电 | 喷涂中液滴形成和撞击的并行计算模拟 | Linux + MPI |
| 机　电 | 变速箱减振降噪优化设计程序的并行化 | Windows XP + MPI |
| 数　理 | 并行计算在识别飞行物着陆点的中应用 | Windows xp + MPI + Vc++6.0 |
| 材　料 | 大规模锂电池生产管理中优化组合解决方案 | Linux + MPI + Gcc |
| 计算机 | 基于大规模数据库的人脸识别研究 | Linux + MPI + Opencv |

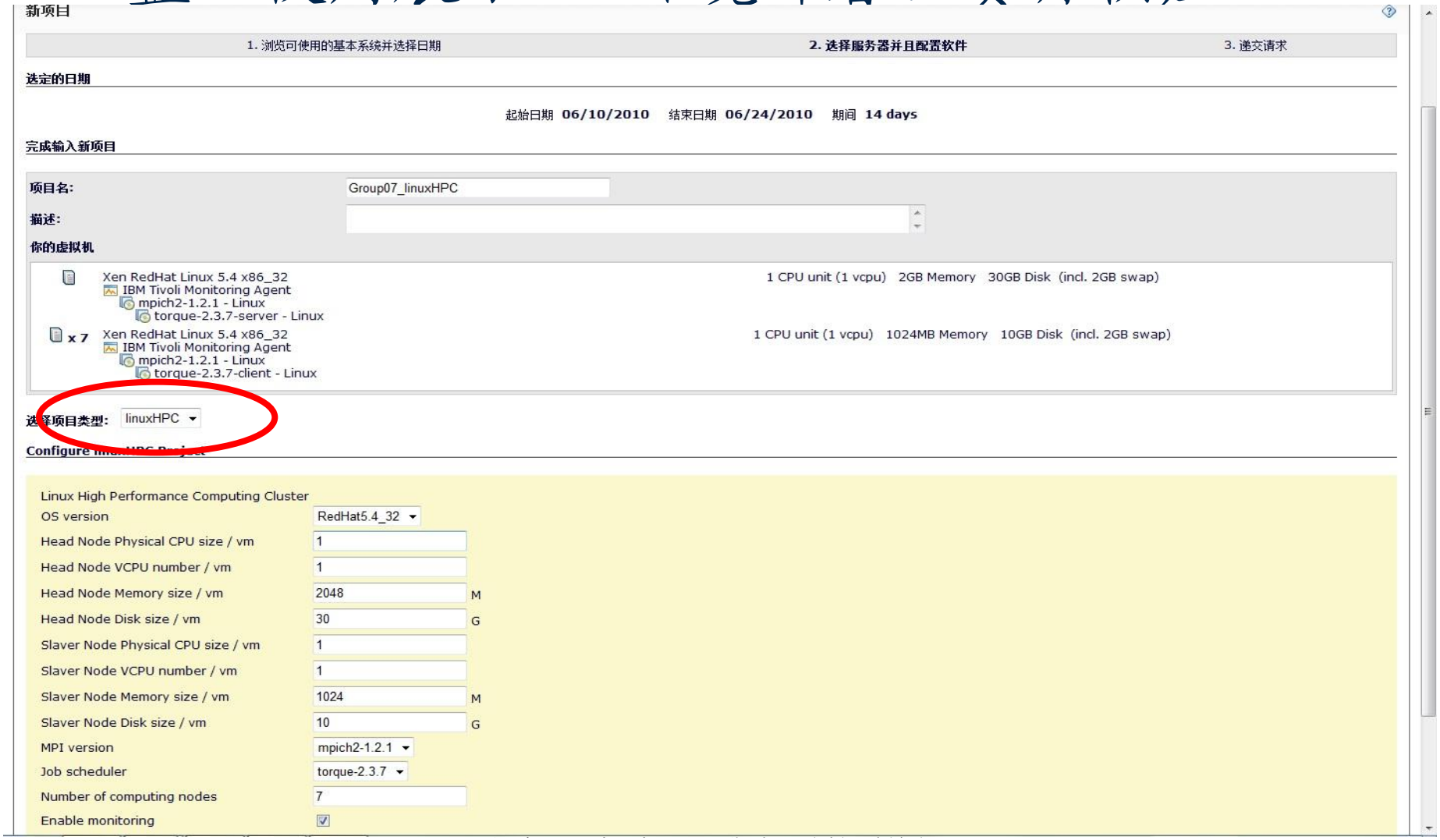
# 北京工业大学并行计算大赛

- 基于高性能平台A区，采用IBM云计算技术，提供并行应用调试/运行环境。

| | 虚拟机个数 | 单个虚拟机硬件配置 | | | 单个虚拟机软件配置 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | CPU | 内存 | 硬盘 | 操作系统 | 集群配置 | 并行环境 | 监控 | 作业调度 |
| 头节点 | 1 | 1*2.83 GHz | 2GB | 30 GB | Windows xp/ Windows 2003/ RHEL 5.4-32bit/ RHEL 5.4-64bit/ RHEL 5.2-32bit | ssh/nfs | Mpich 2/ OpenMP /Hadoop | ITM Monitoring Agent | Torque--server |
| 计算节点 | 7 | 1*2.83 GHz | 1GB | 10 GB | Windows xp/ Windows 2003/ RHEL 5.4-32bit/ RHEL 5.4-64bit/ RHEL 5.2-32bit | ssh/nfs | Mpich2 / OpenMP /Hadoop | ITM Monitoring Agent | Torque--client |

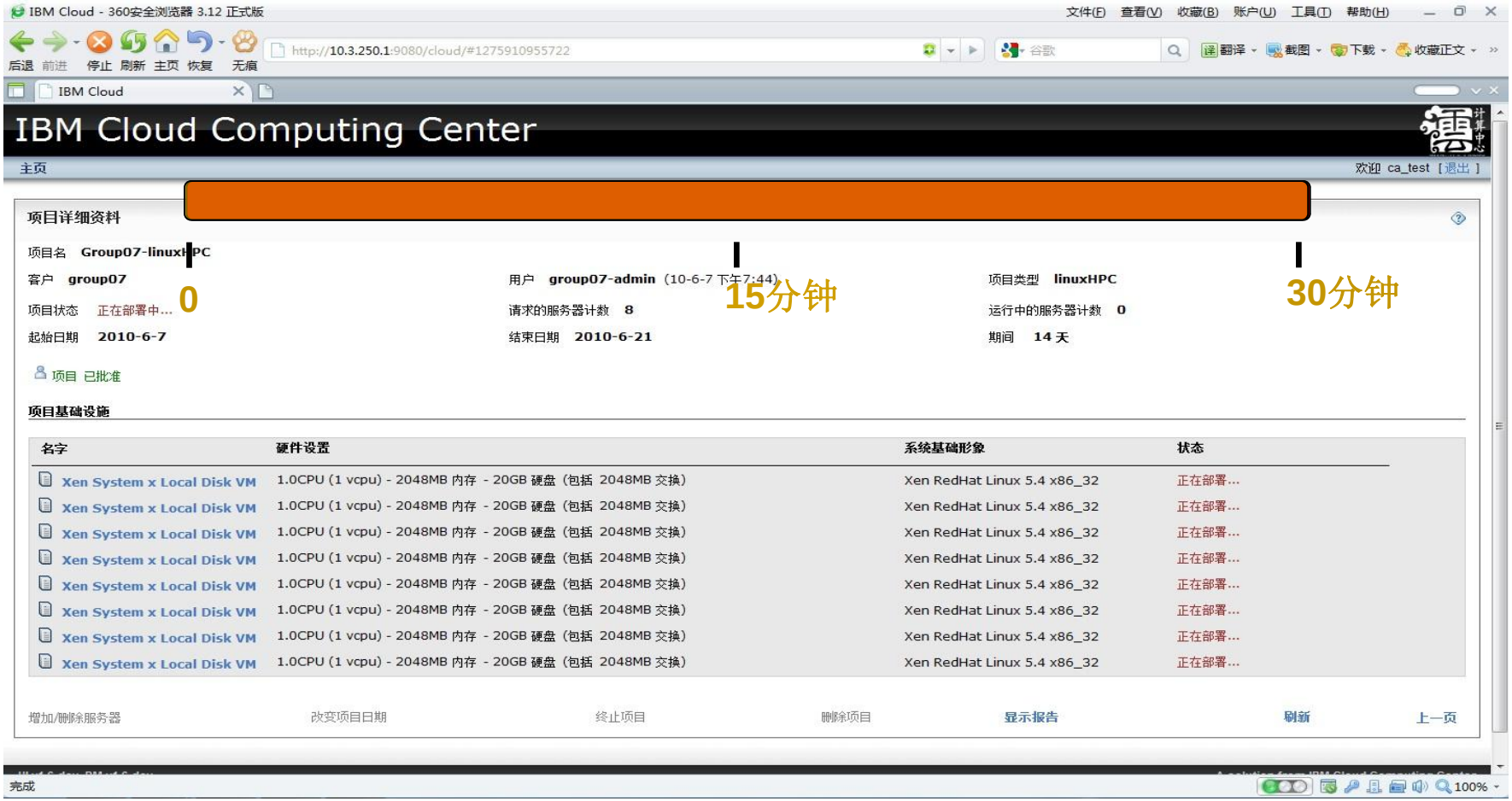
# 北京工业大学并行计算大赛

- 蓝云使用展示---环境部署、资源供应



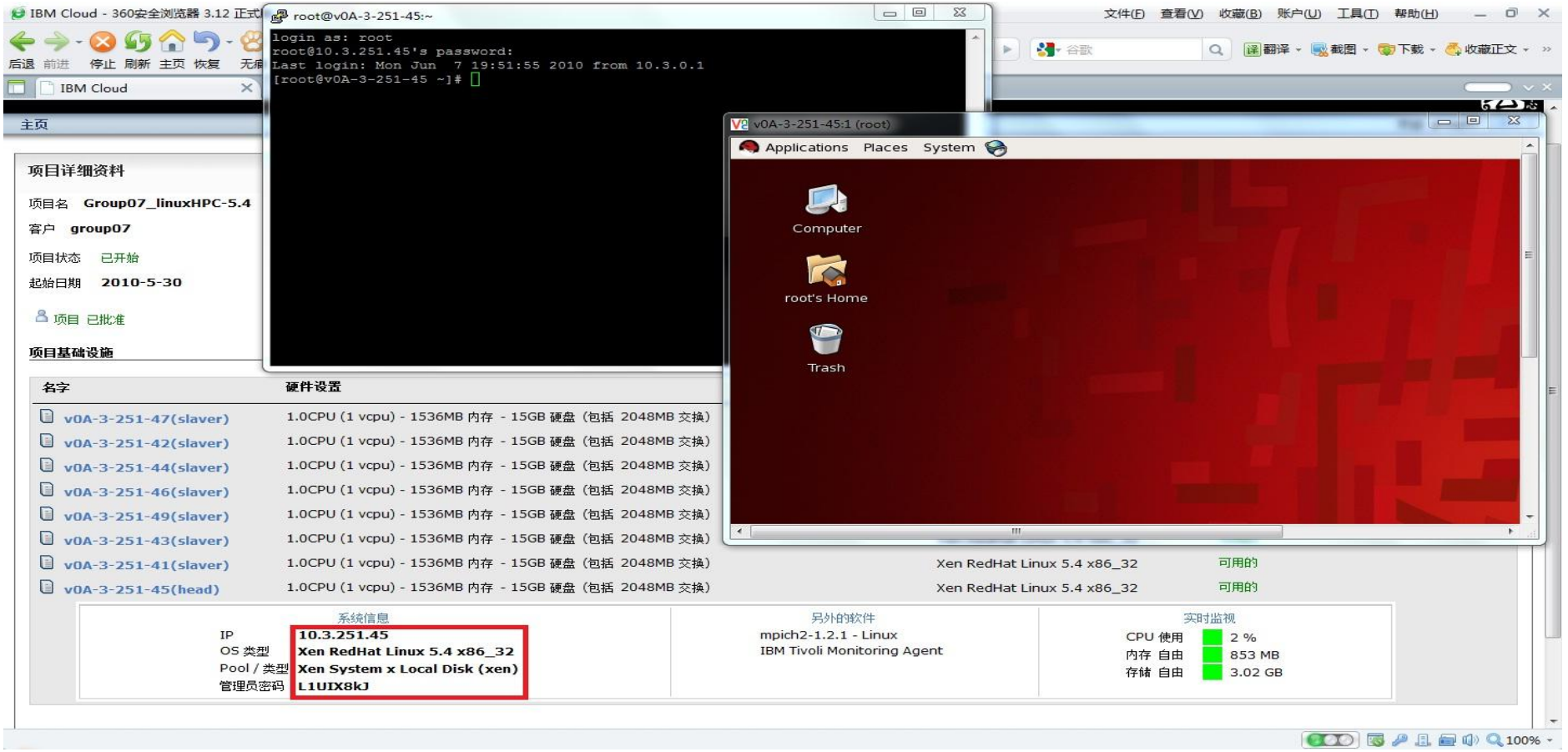
# 北京工业大学并行计算大赛

- IBM云使用展示---环境部署、资源供应



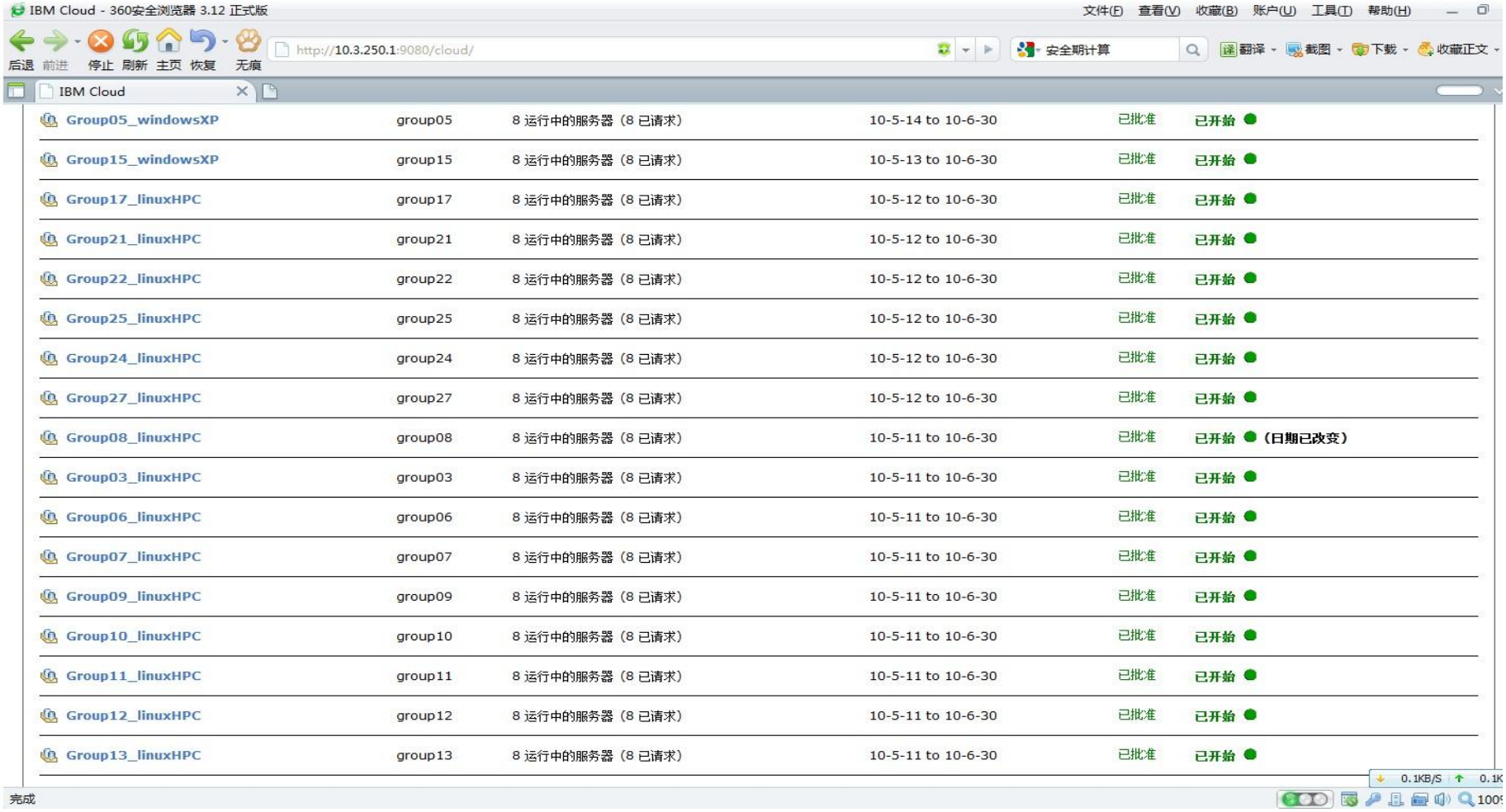
# 北京工业大学并行计算大赛

- IBM云使用展示--- 并行环境使用



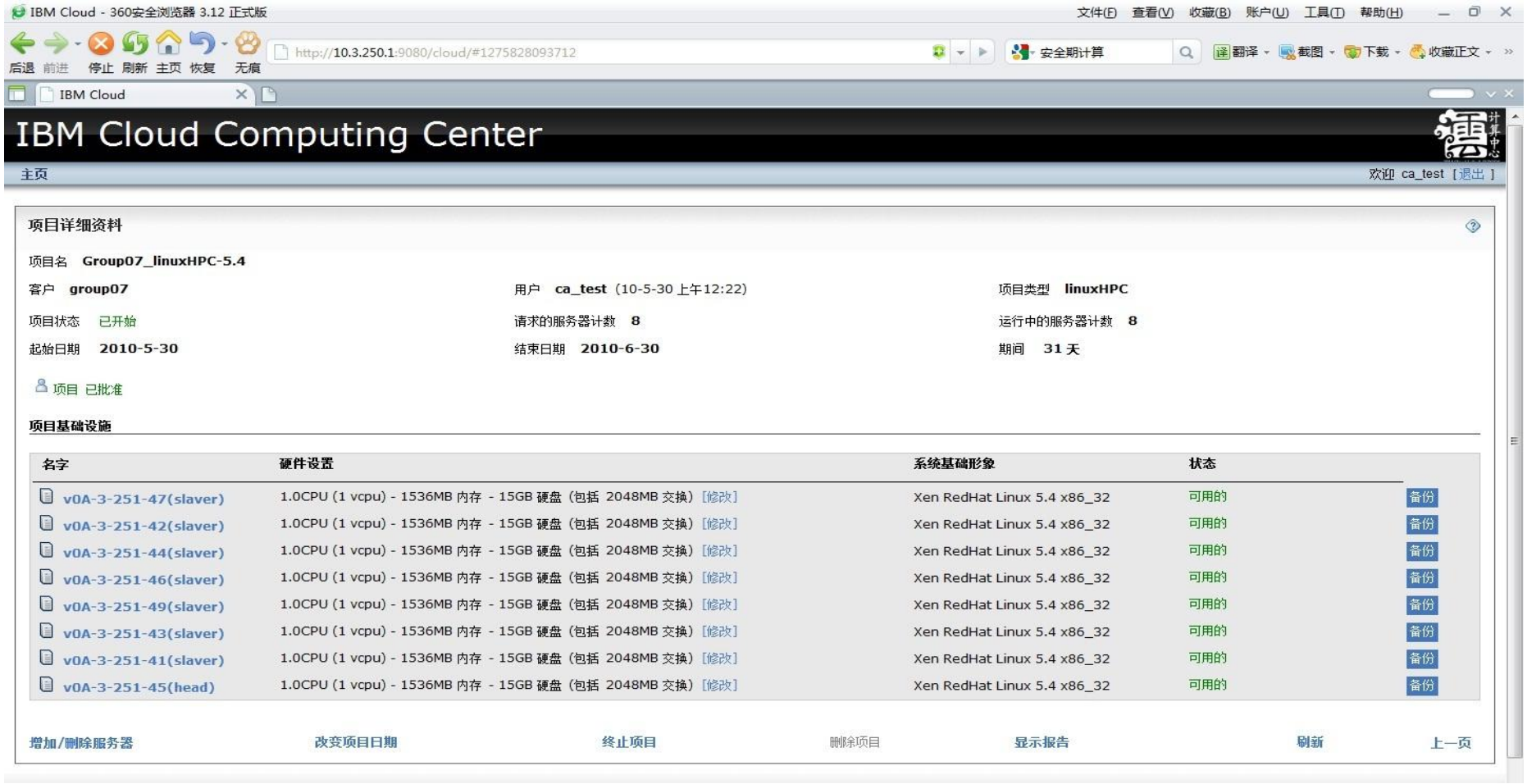
# 北京工业大学并行计算大赛

- IBM云使用展示—在线项目统计和管理



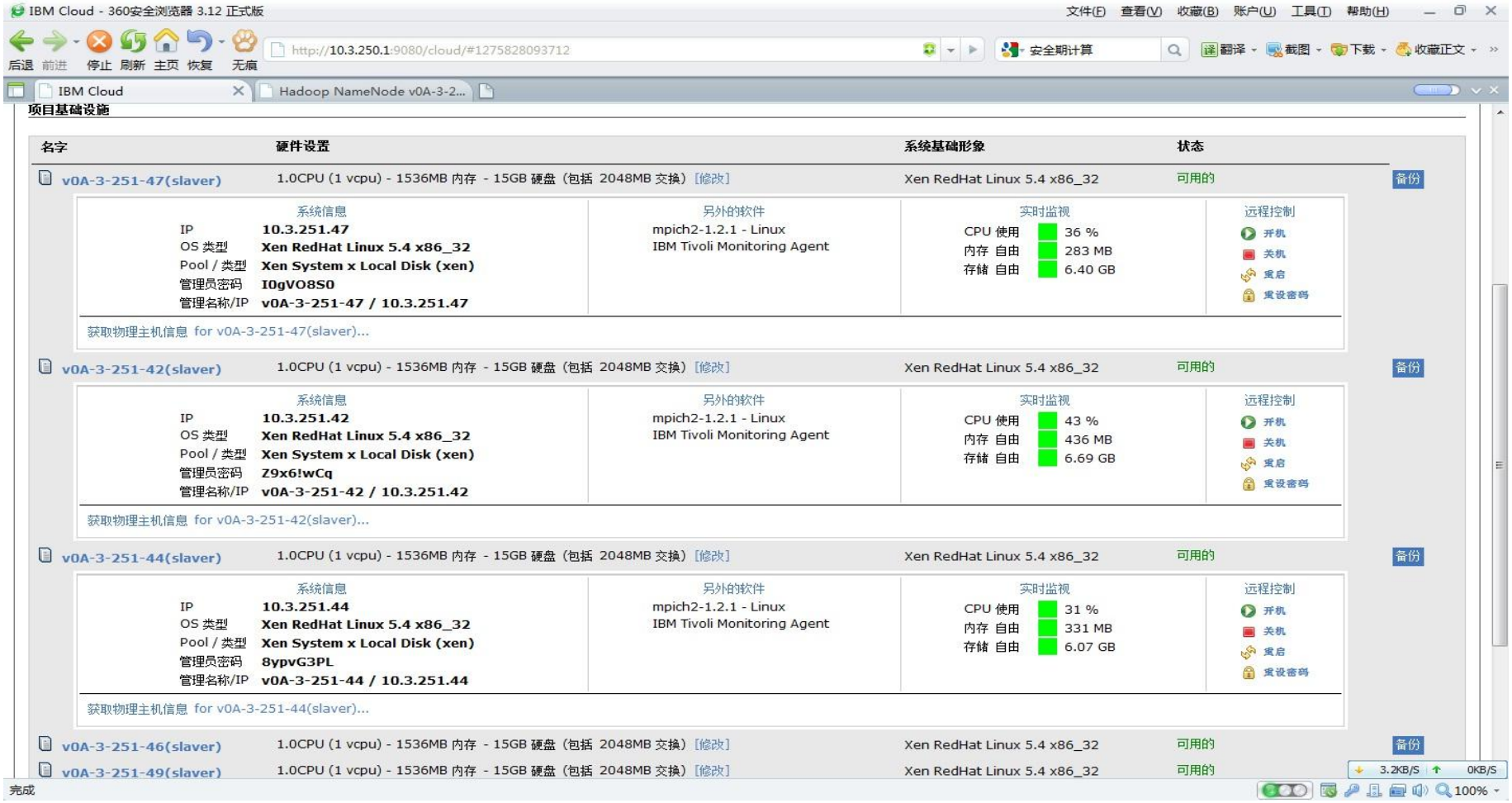
# 北京工业大学并行计算大赛

- IBM云使用展示—基于项目组的资源配置明细



# 北京工业大学并行计算大赛

- IBM云使用展示—资源状态实时监控



# 北京工业大学并行计算大赛

- 总结
  - 共计部署于84个刀片服务器
  - 虚拟机群规模8-100个节点
  - 在线同时管理虚拟机群数最大为32个
  - 虚拟机群部署时间约为30分钟
  - 通过IBM云提供的细粒度资源供给功能，仅使用了50-65%的硬件资源，满足所有29个参赛小组的高性能资源需求。

# THANK YOU!

2009-12-25
2010-06-11

北京工业大学2008级硕士开题答辩