大数据研究的技术层面与主要研究内容

黄　宜　华
yhuang@nju.edu.cn

南　京　大　学

计算机科学与技术系
软件新技术国家重点实验室

主 要 内 容

第一部分：大数据处理的基本特点

主要介绍大数据处理的主要特点和研究原则

第二部分：大数据研究技术层面和主要研究内容

主要介绍大数据研究所涉及的各个技术层面以及各层面下主要的研究内容和热点问题

第三部分：大数据并行处理技术研究

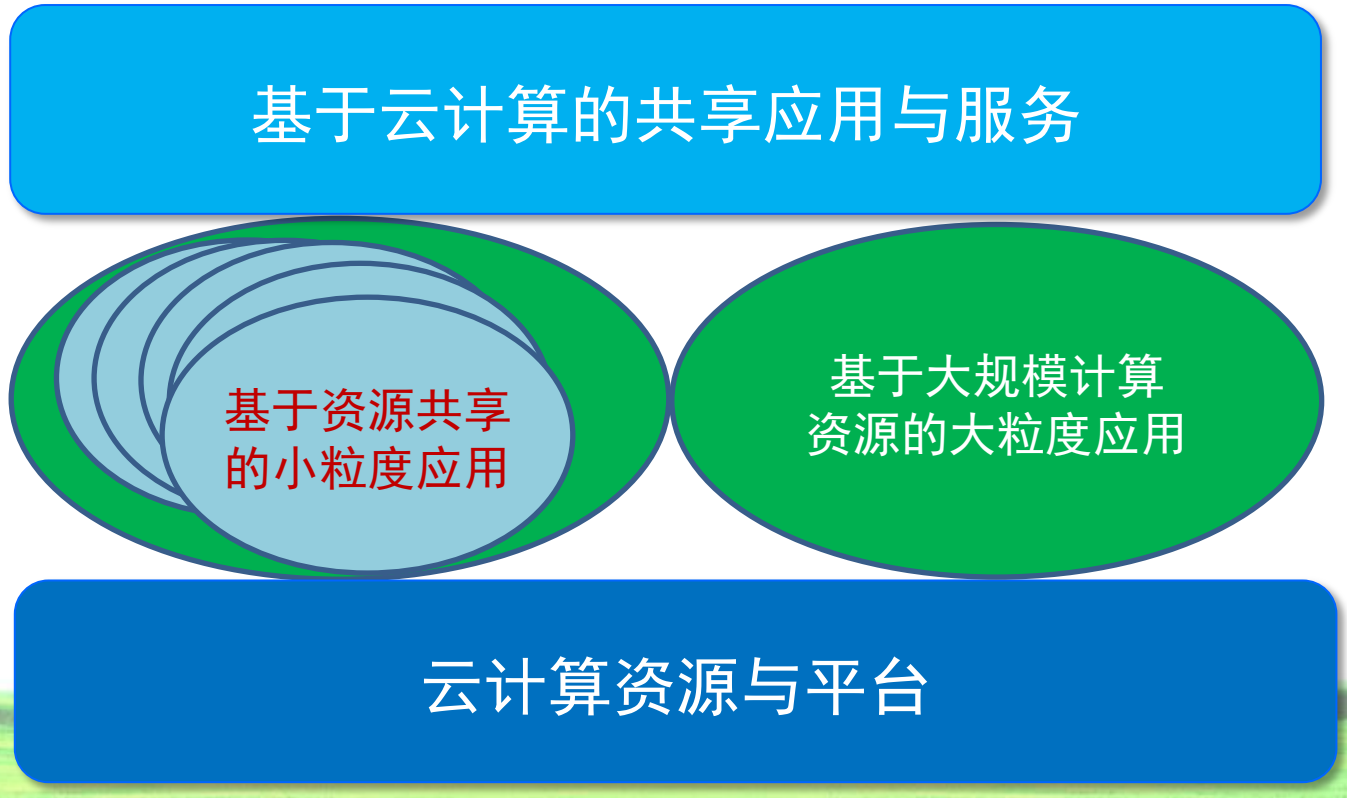
简要介绍本课题组在大数据方面所开展的一些工作

大数据是云计算的两大核心内容之一

云计算的主要目标是：用集中管理的巨大计算资源和计算能力

1）为小粒度应用提供资源共享；

2）为大粒度应用(主要是大数据应用)提供大规模计算能力



大数据是云计算的两大核心内容之一

什么是大数据？

- Wiki百科： big data is a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools

  大数据意指一个超大的、难以用现有常规的数据库管理技术和工具处理的数据集

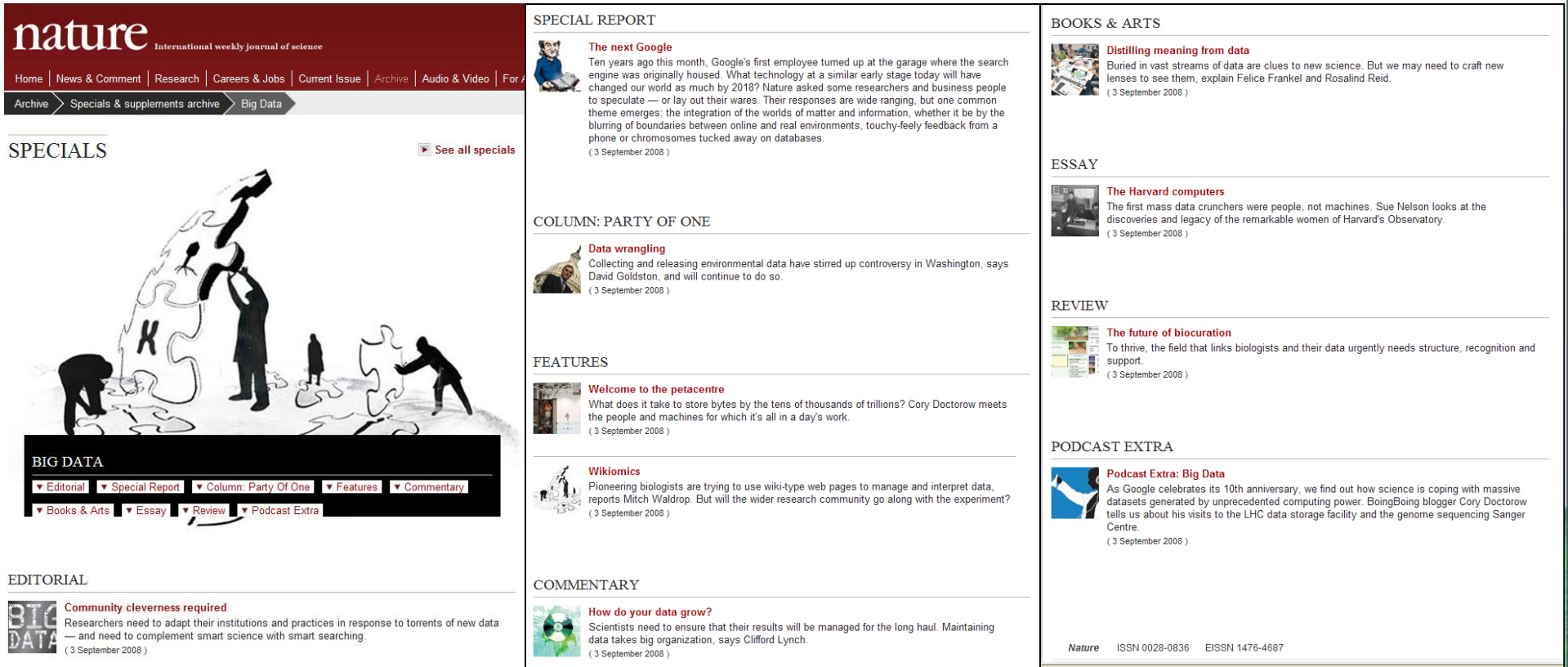
- IDC报告： Big data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis.

  大数据技术描述了一种新一代技术和构架，用于以很经济的方式、以高速的捕获、发现和分析技术，从各种超大规模的数据中提取价值

大数据处理技术的重要性

大数据(Big Data)应用需求

出现越来越多的大数据应用和行业需求。2008年，在Google成立10周年之际，《Nature》杂志出版一期专刊专门讨论未来的大数据（Big Data）处理相关的一系列技术问题和挑战。



nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For

Archive > Specials & supplements archive > Big Data

SPECIALS

See all specials

BIG DATA

▾ Editorial ▾ Special Report ▾ Column: Party Of One ▾ Features ▾ Commentary ▾ Books & Arts ▾ Essay ▾ Review ▾ Podcast Extra

EDITORIAL

Community cleverness required
Researchers need to adapt their institutions and practices in response to torrents of new data — and need to complement smart science with smart searching.
(3 September 2008)

SPECIAL REPORT

The next Google
Ten years ago this month, Google's first employee turned up at the garage where the search engine was originally housed. What technology at a similar early stage today will have changed our world as much by 2018? Nature asked some researchers and business people to speculate — or lay out their wares. Their responses are wide ranging, but one common theme emerges: the integration of the worlds of matter and information, whether it be by the blurring of boundaries between online and real environments, touchy-feely feedback from a phone or chromosomes tucked away on databases.
(3 September 2008)

COLUMN: PARTY OF ONE

Data wrangling
Collecting and releasing environmental data have stirred up controversy in Washington, says David Goldston, and will continue to do so.
(3 September 2008)

FEATURES

Welcome to the petacentre
What does it take to store bytes by the tens of thousands of trillions? Cory Doctorow meets the people and machines for which it's all in a day's work.
(3 September 2008)

Wikiomics
Pioneering biologists are trying to use wiki-type web pages to manage and interpret data, reports Mitch Waldrop. But will the wider research community go along with the experiment?
(3 September 2008)

COMMENTARY

How do your data grow?
Scientists need to ensure that their results will be managed for the long haul. Maintaining data takes big organization, says Clifford Lynch.
(3 September 2008)

BOOKS & ARTS

Distilling meaning from data
Buried in vast streams of data are clues to new science. But we may need to craft new lenses to see them, explain Felice Frankel and Rosalind Reid.
(3 September 2008)

ESSAY

The Harvard computers
The first mass data crunchers were people, not machines. Sue Nelson looks at the discoveries and legacy of the remarkable women of Harvard's Observatory.
(3 September 2008)

REVIEW

The future of biocuration
To thrive, the field that links biologists and their data urgently needs structure, recognition and support.
(3 September 2008)

PODCAST EXTRA

Podcast Extra: Big Data
As Google celebrates its 10th anniversary, we find out how science is coping with massive datasets generated by unprecedented computing power. BoingBoing blogger Cory Doctorow tells us about his visits to the LHC data storage facility and the genome sequencing Sanger Centre.
(3 September 2008)

Nature ISSN 0028-0836 EISSN 1476-4687

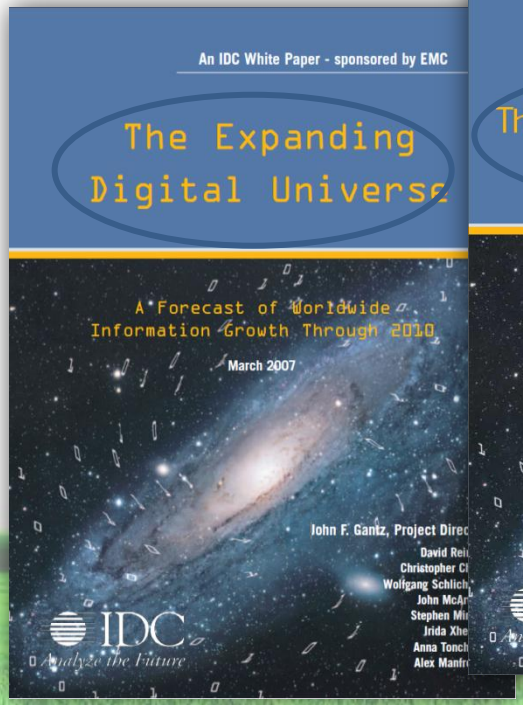
大数据处理技术的重要性

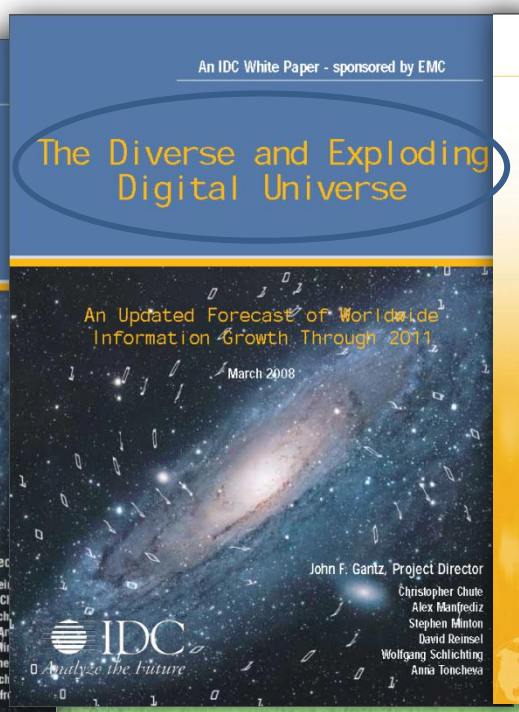
未来10多年数据将急剧增长

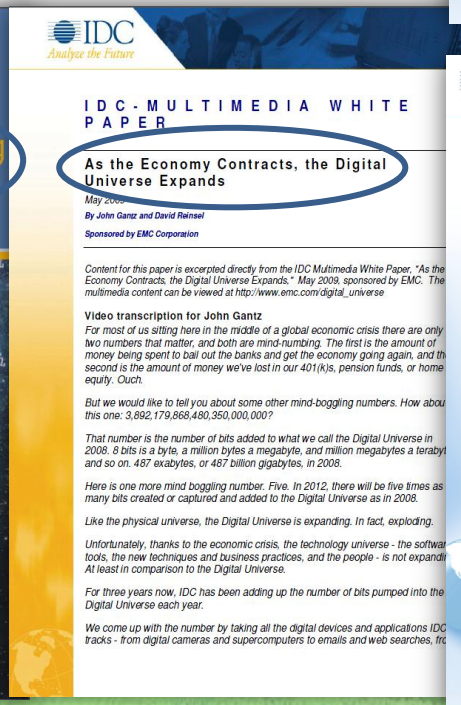
IDC研究报告《Data Universe Study》

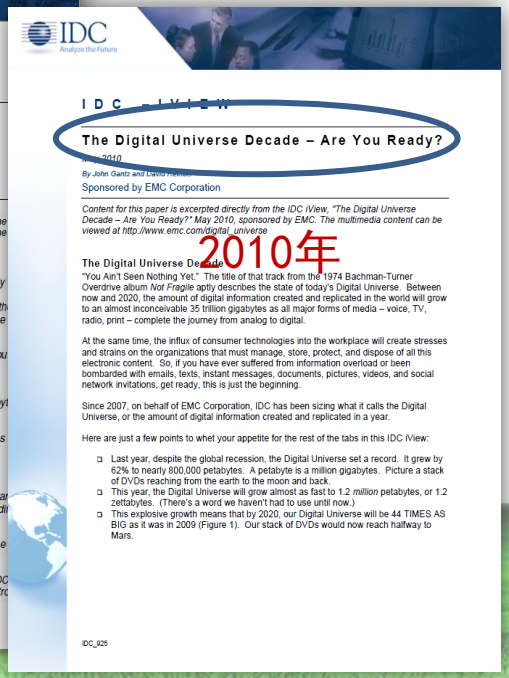
提出“数据宇宙”的说法描述海量数据

2007年　　2008年　　2009年　　2010年　　2011年









大数据处理技术的重要性

未来急剧增长的数据迫切需要寻求新的处理技术手段

IDC报告《Data Universe Study》

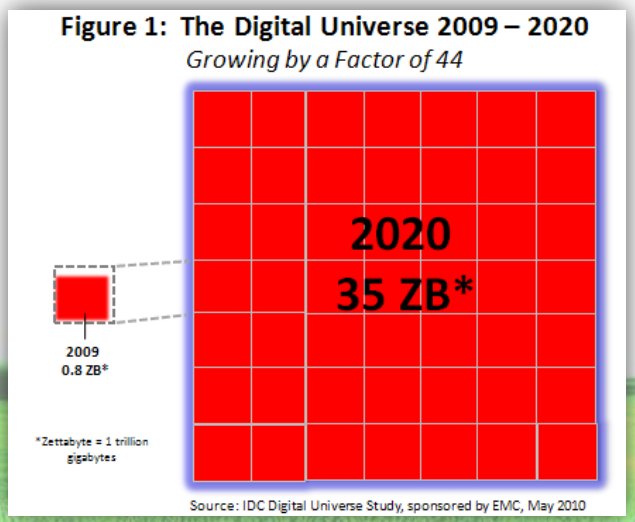
全世界权威IT咨询公司研究报告预测:

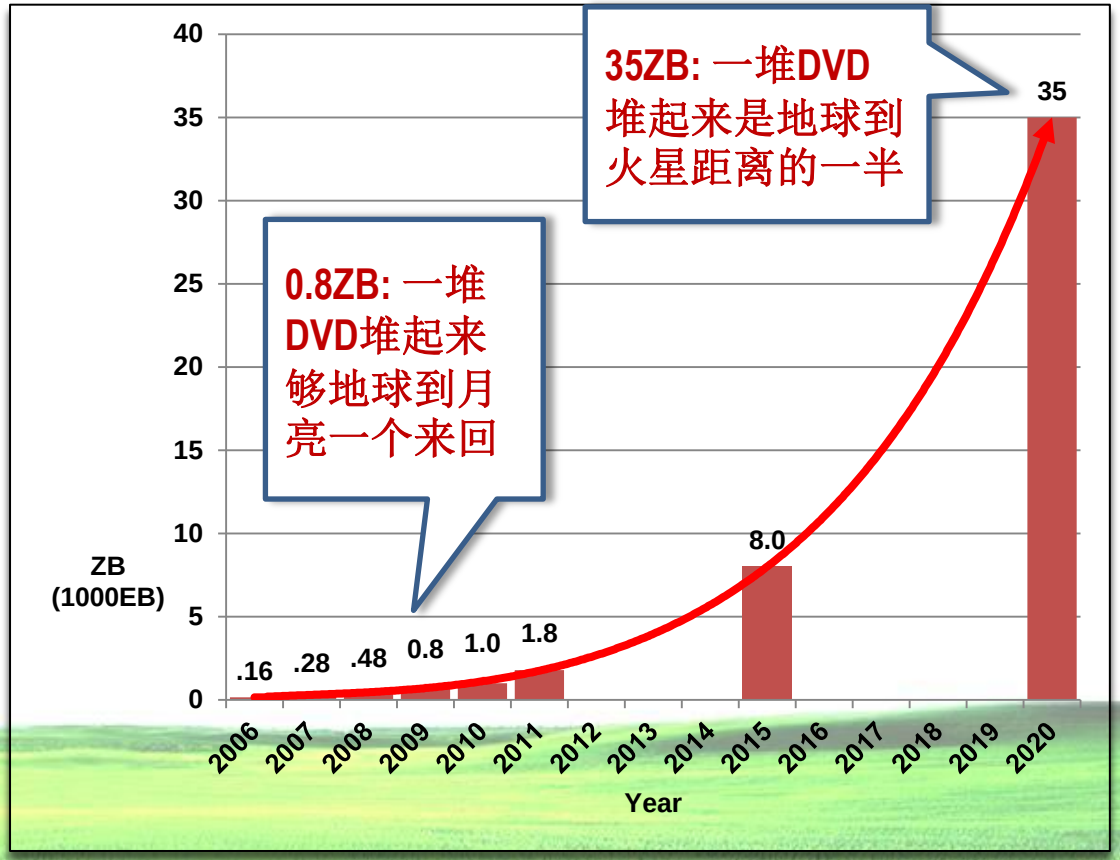
全世界数据量未来10年将从

2009年的0.8ZB增长到

2020年的35ZB,增长44倍!

年均增长率>40%!

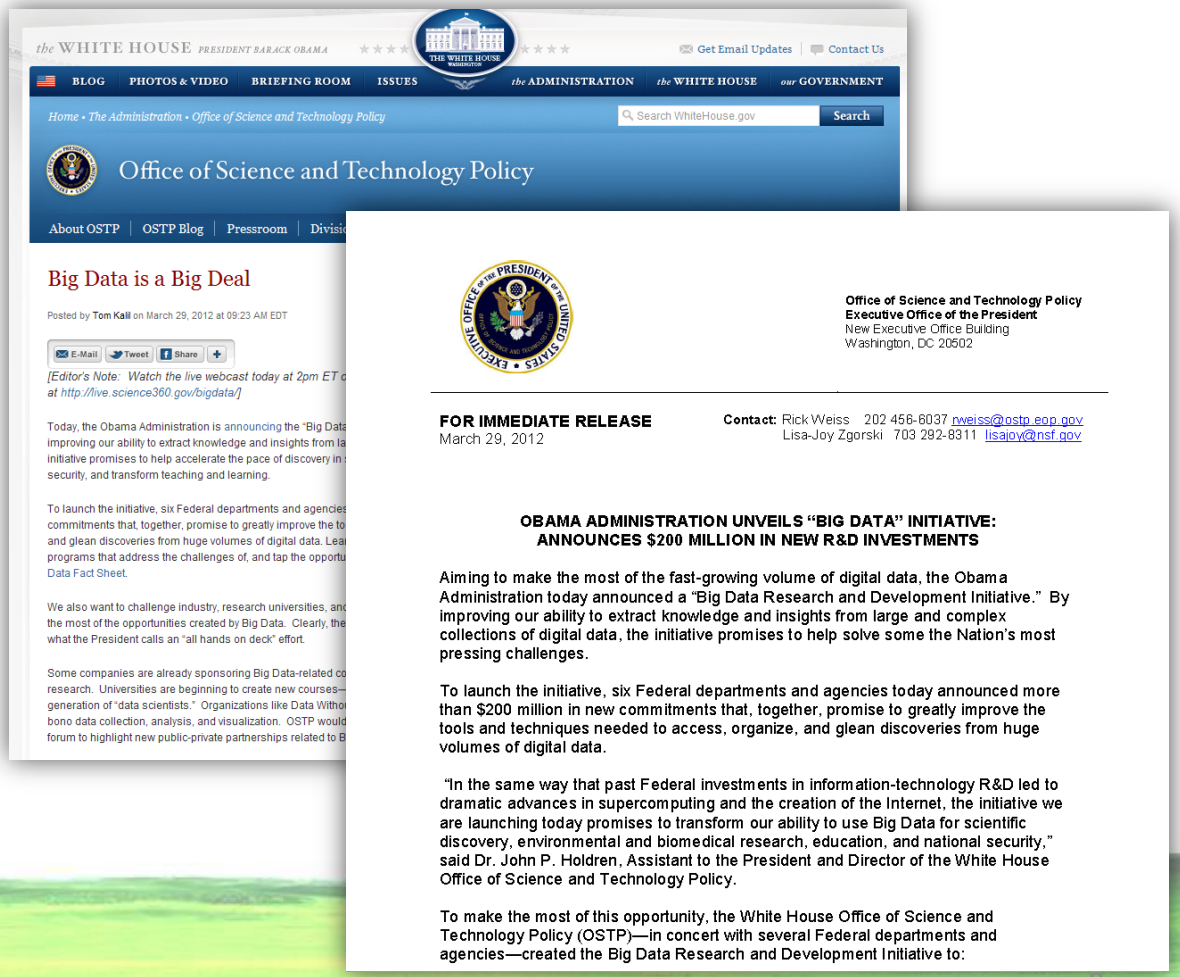




大数据处理技术的重要性

美国联邦政府发布大数据研发专项研究计划

美国联邦政府下属的国防部、能源部、卫生总署等7部委联合推动，于2012年3月底发布了大数据研发专项研究计划 (Big Data Initiative)，拟投入2亿美元用于研究开发科学探索、环境和生物医学、教育和国家安全等重大领域和行业所急需的大数据处理技术和工具，把大数据研究上升到为国家发展战略。



the WHITE HOUSE PRESIDENT BARACK OBAMA

Get Email Updates | Contact Us

BLOG | PHOTOS & VIDEO | BRIEFING ROOM | ISSUES | the ADMINISTRATION | the WHITE HOUSE | our GOVERNMENT

Home • The Administration • Office of Science and Technology Policy

Search WhiteHouse.gov

Search

Office of Science and Technology Policy

About OSTP | OSTP Blog | Pressroom

Big Data is a Big Deal

Posted by Tom Kalil on March 29, 2012 at 09:23 AM EDT

E-Mail | Tweet | Share

[Editor's Note: Watch the live webcast today at 2pm ET

at http://live.science360.gov/bigdata/]

Today, the Obama Administration is announcing the "Big Data

improving our ability to extract knowledge and insights from la

initiative promises to help accelerate the pace of discovery in

security, and transform teaching and learning.

To launch the initiative, six Federal departments and agencies

commitments that, together, promise to greatly improve the to

and glean discoveries from huge volumes of digital data. Lea

programs that address the challenges of, and tap the opportu

Data Fact Sheet.

We also want to challenge industry, research universities, and

the most of the opportunities created by Big Data. Clearly, the

what the President calls an "all hands on deck" effort.

Some companies are already sponsoring Big Data-related co

research. Universities are beginning to create new courses—

generation of "data scientists." Organizations like Data Witho

bono data collection, analysis, and visualization. OSTP would

forum to highlight new public-private partnerships related to B

EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES

Office of Science and Technology Policy
Executive Office of the President
New Executive Office Building
Washington, DC 20502

FOR IMMEDIATE RELEASE
March 29, 2012

Contact: Rick Weiss 202 456-6037 rweiss@ostp.eop.gov
Lisa-Joy Zgorski 703 292-8311 lisajoy@nsf.gov

OBAMA ADMINISTRATION UNVEILS "BIG DATA" INITIATIVE: ANNOUNCES $200 MILLION IN NEW R&D INVESTMENTS

Aiming to make the most of the fast-growing volume of digital data, the Obama Administration today announced a "Big Data Research and Development Initiative." By improving our ability to extract knowledge and insights from large and complex collections of digital data, the initiative promises to help solve some the Nation's most pressing challenges.

To launch the initiative, six Federal departments and agencies today announced more than $200 million in new commitments that, together, promise to greatly improve the tools and techniques needed to access, organize, and glean discoveries from huge volumes of digital data.

"In the same way that past Federal investments in information-technology R&D led to dramatic advances in supercomputing and the creation of the Internet, the initiative we are launching today promises to transform our ability to use Big Data for scientific discovery, environmental and biomedical research, education, and national security," said Dr. John P. Holdren, Assistant to the President and Director of the White House Office of Science and Technology Policy.

To make the most of this opportunity, the White House Office of Science and Technology Policy (OSTP)—in concert with several Federal departments and agencies—created the Big Data Research and Development Initiative to:

大数据处理技术的重要性

“大数据研究的科学价值”

李国杰，《中国计算机学会通讯》，vol. 8，no.9，2012.9

2012年3月，美国奥巴马政府宣布投资2亿美元启动“大数据研究和发展计划”，这是继1993年美国宣布“信息高速公路”计划后的又一次重大科技发展部署。美国政府认为大数据是“未来的新石油”，将“大数据研究”上升为国家意志，对未来的科技与经济发展必将带来深远影响。一个国家拥有数据的规模和运用数据的能力将成为综合国力的重要组成部分，对数据的占有和控制也将成为国家间和企业间新的争夺焦点。

大数据处理技术的重要性

数据科学(Data Science)

国内外出现了“数据科学”的概念

- 图灵奖获得者Jim Gray: 2007年最后一次演讲中提出“数据密集型科学发现(Data-Intensive Scientific Discovery)”将成为科学研究的第四范式

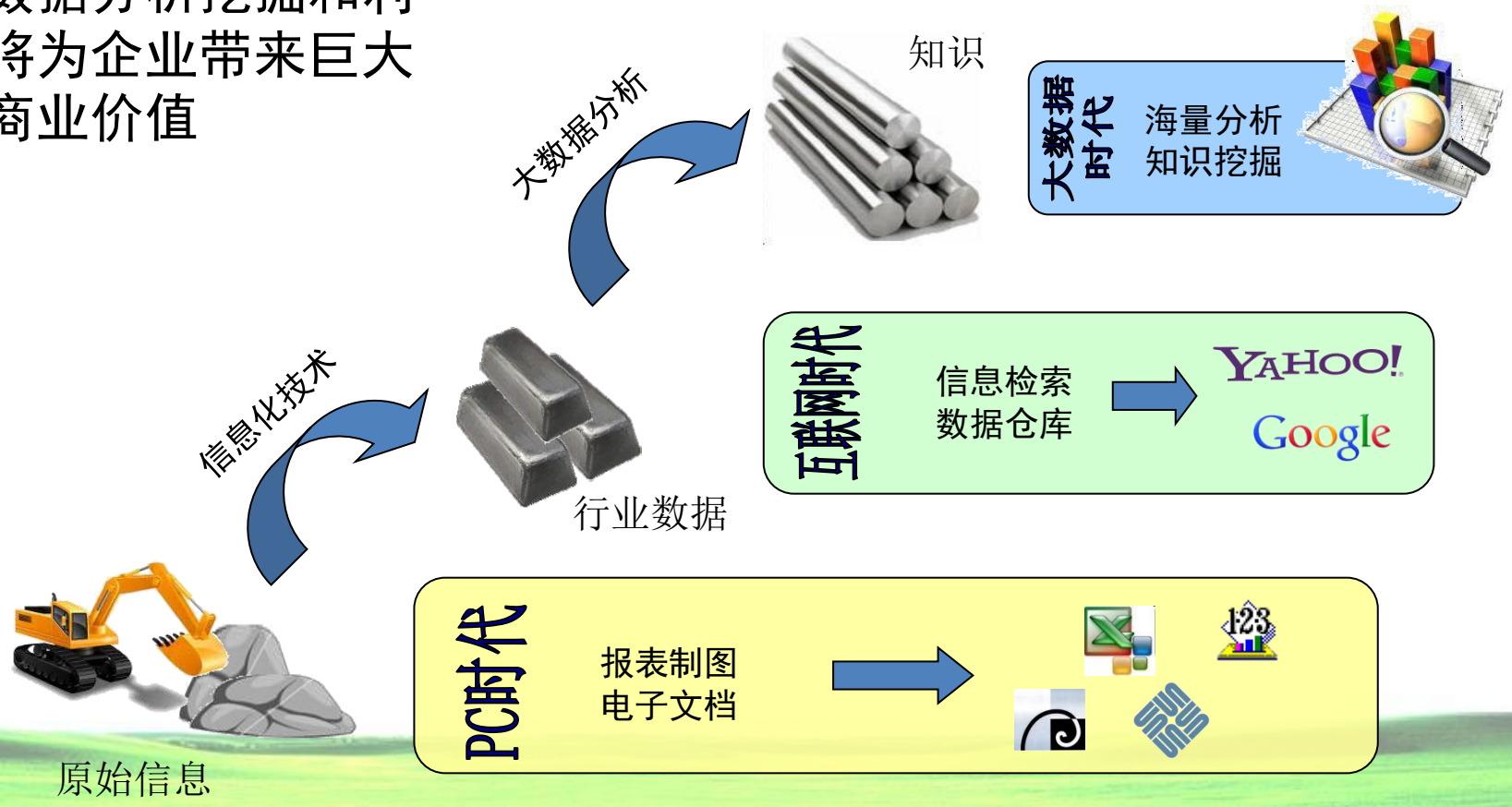
实验科学→ 理论科学 → 计算科学 → 数据科学

- 世界著名存储技术公司EMC：提出了“Data Science”的概念，“Data Science teams will become the driving force for success with big data analytics”
- 李国杰院士：“数据科学”研究的对象是什么？计算机科学是关于算法的科学，数据科学是关于数据的科学。

大数据处理技术的重要性

大数据将带来巨大的技术和商业机遇

- 大数据分析挖掘和利用将为企业带来巨大的商业价值



大数据处理技术发展的驱动力

应用数据规模急剧增加，传统计算面临严重挑战

- 中国移动一个省电话通联记录(CDR)数据每月可达0.5-1PB，而整个中国移动每月则高达7-15PB数据；如此巨大的数据量使得Oracle等数据库系统已经难以支撑和应对
- 南京市公安局320道路监控云计算系统，数据量为三年200亿条、总量120TB的车辆监控数据
- 百度存储数百PB数据，每天处理数据10PB
- 淘宝存储14PB交易数据，每天新增数据40-50TB

百度

数据总量	100~1000PB
数据处理量	10~100PB/天
网页	千亿~万亿
索引	百亿~千亿
更新量	十亿~百亿/天
请求	十亿~百亿/天
日志	100TB~1PB/天

淘宝网

商品总量	30亿注册商品 10亿在线商品
交易量	每天千万量级
注册用户	4亿
单日成交峰值	19.5亿元
数据	40-50TB/天
累积数据	14PB
处理集群	1500节点

大数据处理技术发展的驱动力

大规模数据处理和行业应用需求日益增加和迫切

出现越来越多的大规模数据处理应用需求，传统系统难以提供足够的存储和计算资源进行处理，云计算技术是最理想的解决方案。调查显示：目前，IT专业人员对云计算中诸多关键技术最为关心的是大规模数据并行处理技术

大数据并行处理没有通用和现成的解决方案

对于应用行业来说，云计算平台软件、虚拟化软件都不需要自己开发，但行业的大规模数据处理应用没有现成和通用的软件，需要针对特定的应用需求专门开发，涉及到诸多并行化算法、索引查询优化技术研究、以及系统的设计实现

大数据处理技术发展的驱动力

海量数据隐含着更准确的事实

研究发现：大数据量可显著提高机器学习算法的准确性；训练数据集越大，数据分类精度越高；大数据集上的简单算法能比小数据集上的复杂算法产生更好的结果，因此数据量足够大时有可能使用代价很小的简单算法来达到很好的学习精度。

例如，2001年,一个基于事实的简短问答研究，如提问:Who shot Abraham Lincoln？在很大的数据集时,只要使用简单的模式匹配方法,找到在“shot Abraham Lincoln”前面的部分即可快速得到准确答案：John Wilkes Booth

大数据的基本特点

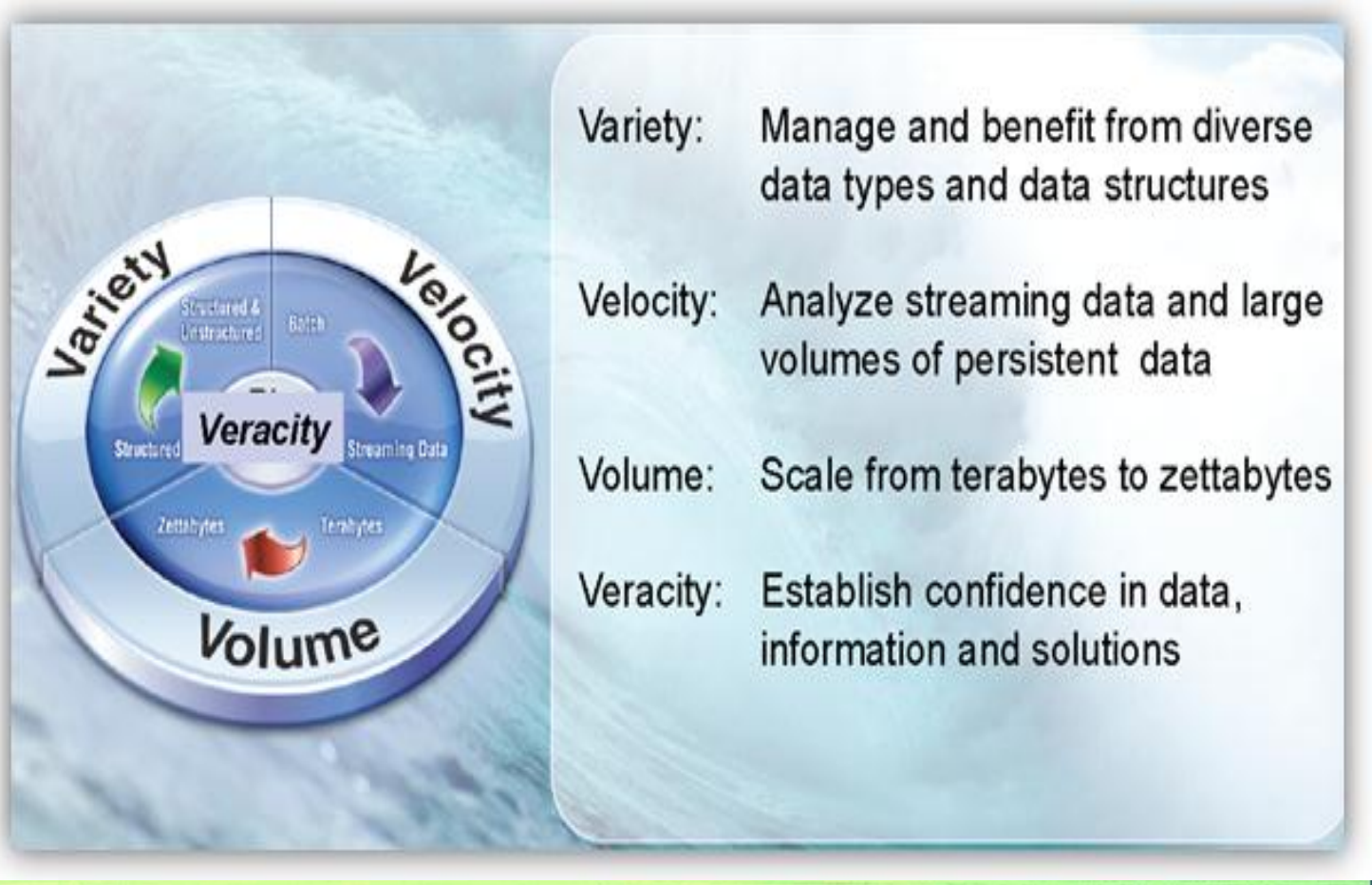
大数据特点

Volume: 大容量,TB-ZB

Variety: 多样性

Velocity: 时效性

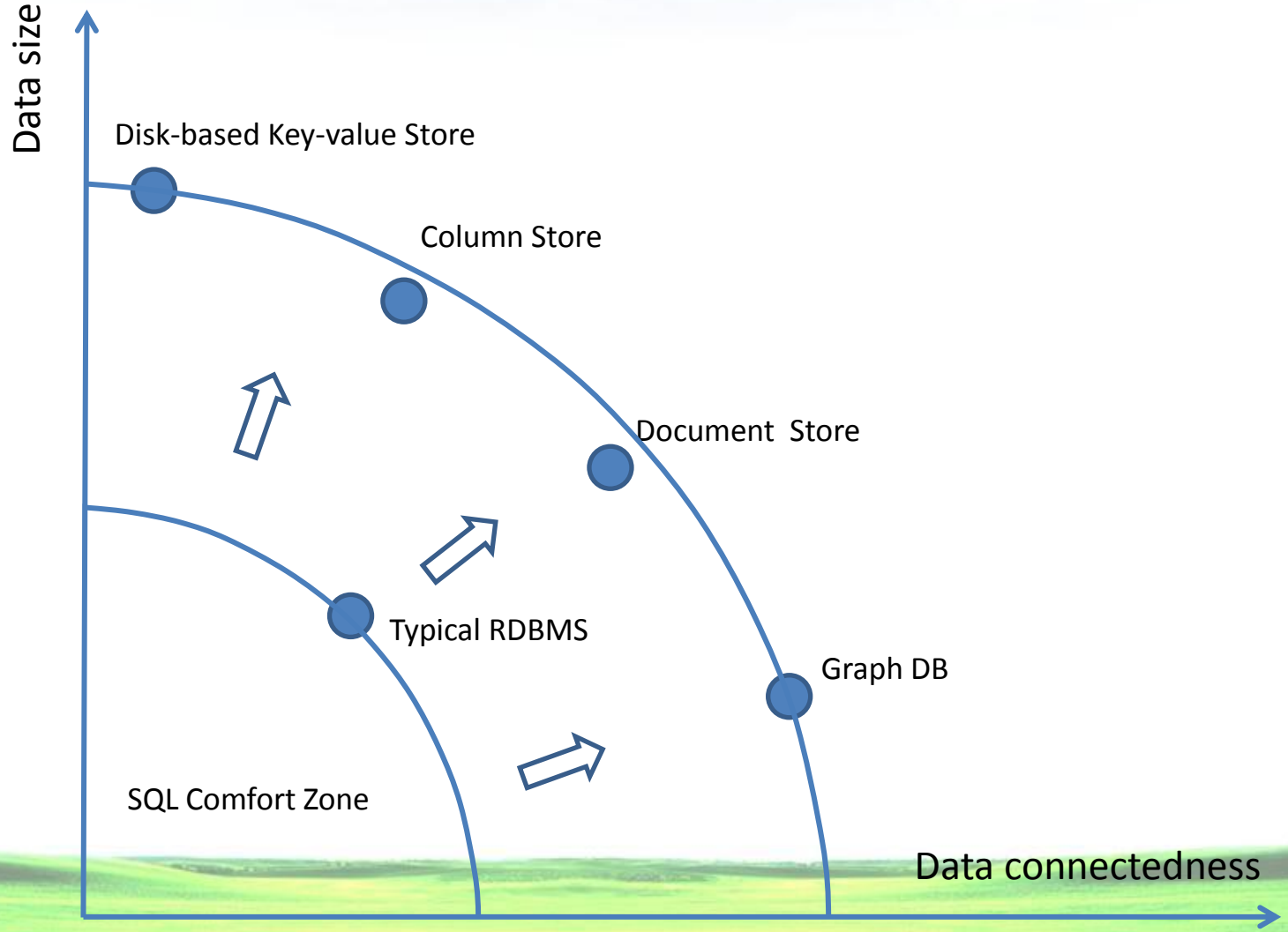
Veracity: 准确性



大数据的类型

- 结构特征
  - 结构化数据
  - 非结构化/半结构化数据
- 获取和处理方式
  - 动态(流式/增量式/线上)/实时数据
  - 静态(线下数据)/非实时数据
- 关联特征
  - 无关联/简单关联数据(键值记录型数据)
  - 复杂关联数据(图数据)

数据尺度和关联度空间



引自微软亚洲研究院，王海勋

大数据问题的特点和研究原则

- 大数据问题的基本特点
  - 大数据来自应用行业，具有极强的行业应用需求特性
  - 数据规模极大，达到PB甚至EB量级，超过任何传统数据库系统的处理能力
  - 大数据处理给传统计算技术带来极大挑战，大多数传统算法在面向大数据处理时都面临问题，需要重写
- 大数据研究的基本原则
  - 应用需求为导向
  - 领域交叉为桥梁
  - 计算技术为支撑

大数据研究的挑战和基本途径

- 大数据研究的挑战
  - 数据规模导致难以应对的存储和计算量
  - 数据规模导致传统算法失效
  - 大数据复杂的数据关联性导致高复杂度的计算

- 大数据研究的基本途径

  三个基本途径：
  - 继续寻找新算法降低计算复杂度
  - 降低大数据尺度，寻找数据尺度无关算法
  - 大数据并行化处理

大数据研究的挑战和基本途径

- 大数据研究的基本途径

降低尺度
寻找数据
尺度无关
近似算法

新算法
寻找新算法降低计算复杂度

并行化
分而治之
并行化处理

第 二 部 分

大数据研究层面与主要研究内容

大数据研究层面和主要内容

		研究层面	角色
电信/公安/商业/金融/遥感遥测/勘探/生物医药……	应用层	大数据行业应用/服务层	行业用户
领域应用/服务需求和计算模型			领域专家
行业应用系统开发		应用开发层	应用开发者
社会网络,排名与推荐,商业智能,自然语言处理,生物信息媒体分析检索, Web挖掘与检索, 3维建模与可视化计算…	算法层	应用算法层	计算技术研究和开发者
并行化机器学习与数据挖掘算法		基础算法层	
MapReduce, BSP, MPI, CUDA, OpenMP, 定制式, 混合式 (如MapReduce+CUDA, MapReduce+MPI)	系统层	并行编程模型与计算框架层	
大数据查询(SQL, NoSQL, 实时查询, 线下分析) 大数据存储(DFS, HBase, MemDB, RDB) 大数据预处理		大数据存储管理层	
集群, 多核, GPU, 混合式构架 (如集群+多核, 集群+GPU) 云计算资源与支撑平台	平台层	并行构架和资源平台层	

大数据行业应用与服务层

- 行业应用系统和服务
  - 行业应用系统

    电信、公安、商业、金融、遥感遥测、地质勘探、生物医药 ……
  - 行业应用公共服务中间件
- 领域应用/服务需求和计算模型
  - 领域应用问题和需求
  - 领域应用问题计算模型

大数据行业应用开发层

- 行业应用系统和服务
  - 大数据应用开发环境和工具
  - 大数据应用和服务集成框架和接口
  - 大数据应用测试环境和工具
  - 大数据应用发布和运行环境

应用算法/技术层研究内容

- 社会网络
- 排名与推荐系统
- 商业智能
- 媒体分析检索
- Web挖掘与搜索
- 3维建模与科学计算可视化
- 生物信息
- 自然语言处理
- ……

应用算法/技术层研究内容

- 搜索引擎综合应用案例

未来的搜索引擎，将不再是基于简单关键词检索的网页聚合，而是基于精准化和智能化搜索的信息和知识的聚合，能够分析用户的意图，信息将以更精准、更智能化方式提供给用户

Google Knowledge Graph

基于搜索关键词语义理解和信息关联性的智能化搜索功能，可提供搜索对象相关的综合性和多样化信息（文字和媒体信息）。涉及到前述大多数应用技术的综合性应用：

- 一种深度搜索技术
- 基于语义分析理解
- 基于信息关联网络分析
- 多样化排名与推荐
- 基于图片内容的搜索

目前 Google Knowledge Graph 已经有五亿个信息“对象”包括 35 亿个属性和相互关系；但目前只支持英文，不支持中文

应用算法/技术层研究内容

- 综合应用案例

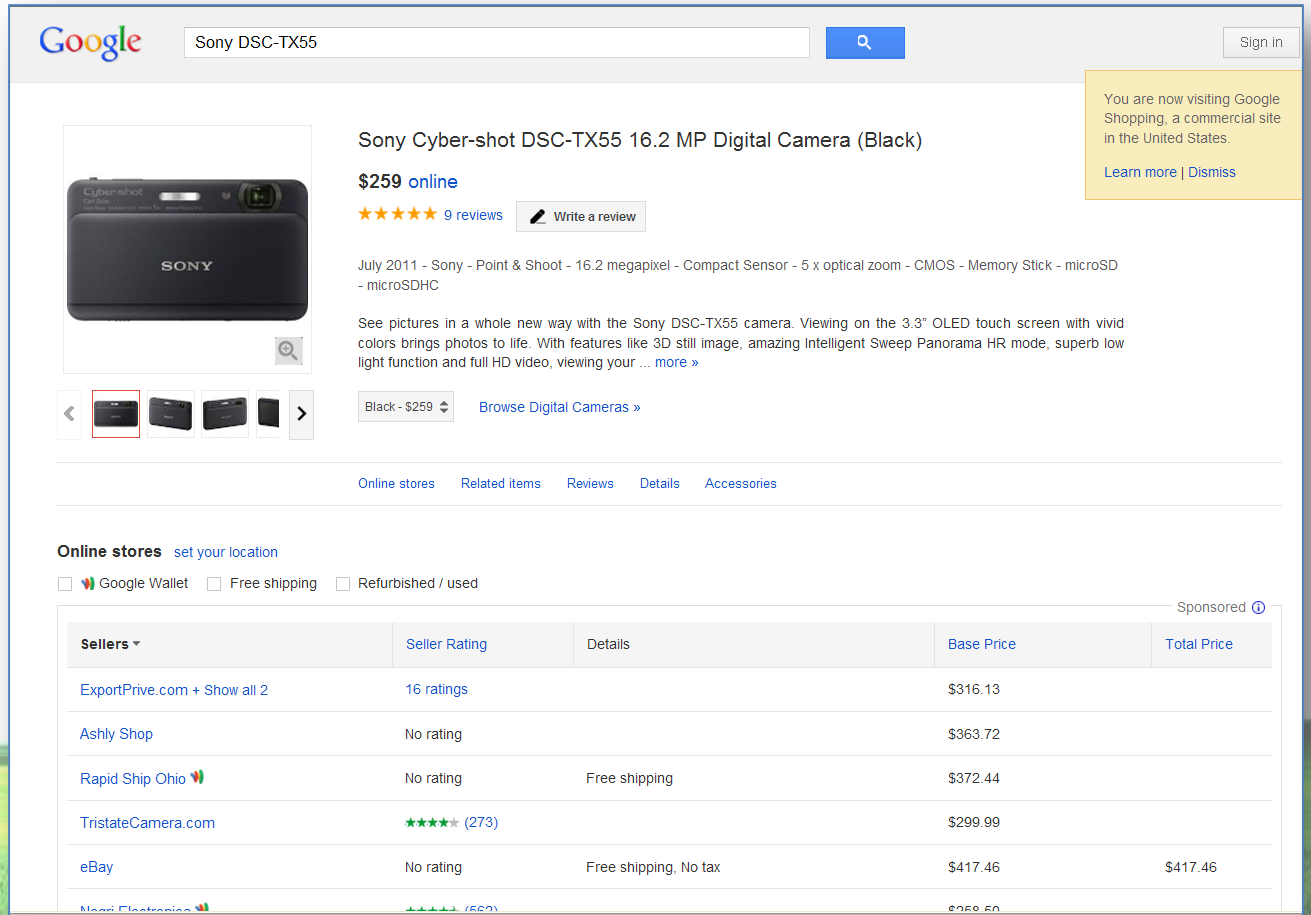
  Google Knowledge Graph

应用算法/技术层研究内容

- 综合应用案例

Google 商品搜索

提供商品信息垂直搜索功能



应用算法/技术层研究内容

- 综合应用案例

  搜狗“知立方”

第一个具有中文知识图谱功能的搜索引擎

应用算法/技术层研究内容

- 社会网络
  - 社团发现 (Community Detection)
  - 网络建模 (Network Modeling)
  - 中心分析和影响力建模 (Centrality Analysis and Influence Modeling)
  - 分类推荐 (Classification and Recommendation)
  - 隐私安全等 (Privacy, Spam and Security)

应用算法/技术层研究内容

- 排名与推荐系统
  - 常规排名(Ranking)
  - 多样性排名(Diversified Ranking)
  - 基于内容的推荐 (Content-based Recommendation)
  - 基于标签的推荐 (Tag-based Recommendation)
  - 协同过滤推荐 (Collaborative Filtering Recommendation)

应用算法/技术层研究内容

- 媒体分析检索
  - 大规模图像检索 (Image Retrieval)
  - 大规模图像分类 (Image Classification)
  - 目标检测 (Object Recognition)
  - 视频异常行为检测 (Abnormal Event Detection)

应用算法/技术层研究内容

- Web搜索与数据挖掘
  - 深度Web搜索(Deep Web Search，精确化、智能化、综合化信息搜索)
  - 页面分类 (Document Classification)
  - 页面聚类 (Document Cluster)
  - 网页摘要 (Document Automatic Summarization)

应用算法/技术层研究内容

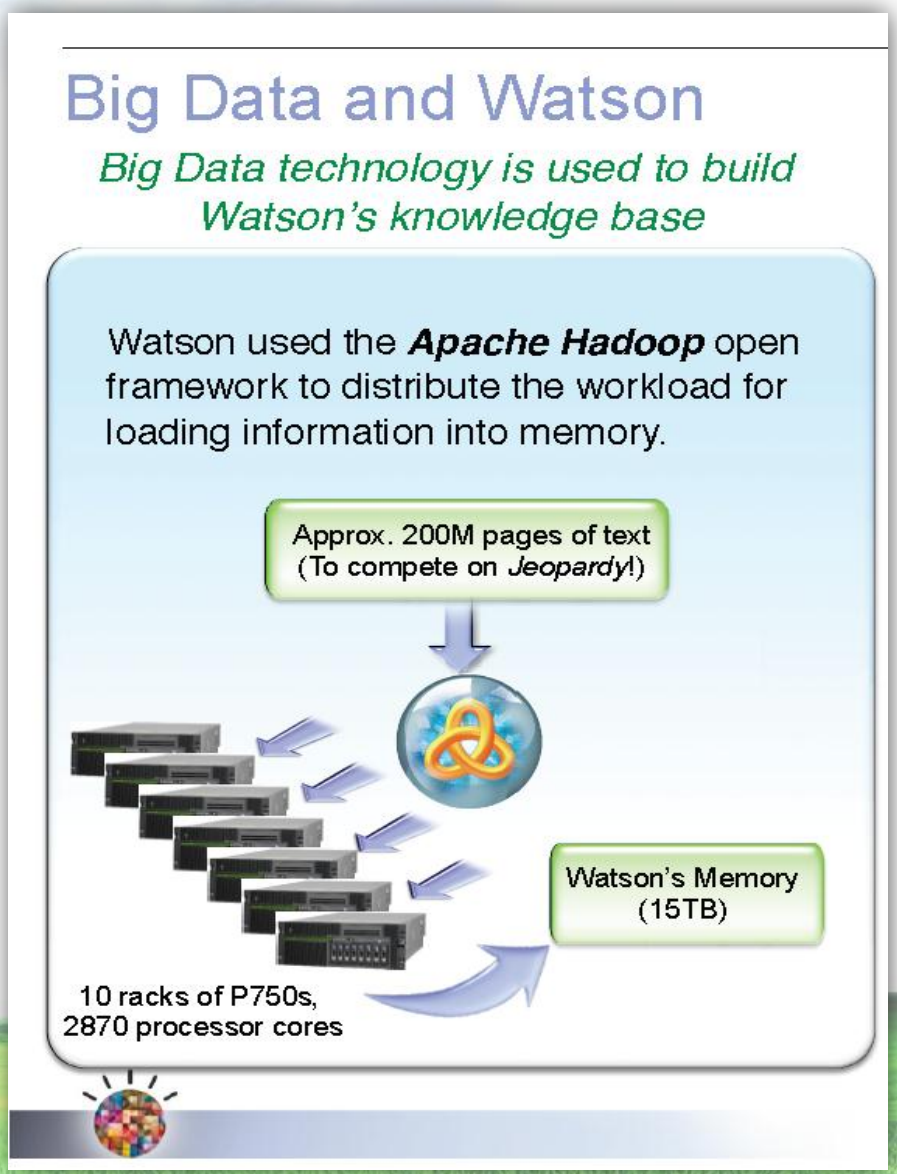
- 自然语言处理
  - 机器翻译 (Machine Translation)
  - 情感分析 (Sentiment Analysis)
  - 舆情分析 (Public Opinion Analysis)
  - 智能输入 (Smart Input)
  - 问答系统 (QA)

IBM智力问答机器人Watson

IBM 智力竞赛机器人Watson是一个基于MapReduce数据并行处理和统计模型自然语言处理的成功应用。

IBM智力问答机器人Watson

Watson收集了2亿页知识文本数据，并基于Hadoop MapReduce并行处理集群进行数据分析，采用了优化的并行体系结构和优化的知识和自然语言处理算法，可在1秒内完成对大量非结构化信息的检索，并实时回答知识竞赛问答题。



应用算法/技术层研究内容

- 3维建模与大数据可视化计算
  - 地质建模与分析 (Geological Modeling and Analysis)
  - 电影渲染 (Movie Rendering)
  - 大规模数据可视化计算与分析 (Scale Visual Analytics)

应用算法/技术层研究内容

- 生物信息处理
  - 高通量基因序列拼序 (High-Throughput Genomic Sequence Assembly)
  - 高通量基因序列比对 (High-Throughput Genomic Sequence Alignment)
  - 生物网络建模与分析 (Biological Network Modeling and Analysis)

基础算法/技术层研究内容

- 大数据并行化机器学习和数据挖掘算法
  - 大数据处理并行化学习和挖掘算法
  - 不同并行模型下并行化学习和挖掘算法
  - 并行化机器学习和数据挖掘工具和平台

研究表明：基于大数据集的机器学习会取得更好的学习效果，这已经成果目前机器学习领域的共识

基础算法/技术层研究内容

大数据并行化机器学习和数据挖掘算法

2001, 微软研究院的Banko and Brill*等发表了一篇自然语言理解领域的经典研究论文，探讨训练数据集大小对分类精度的影响，发现数据越大，精度越高；更有趣的发现是，他们发现当数据不断增长时，不同算法的分类精度趋向于相同，使得小数据集时不同算法在精度上的差别基本消失！

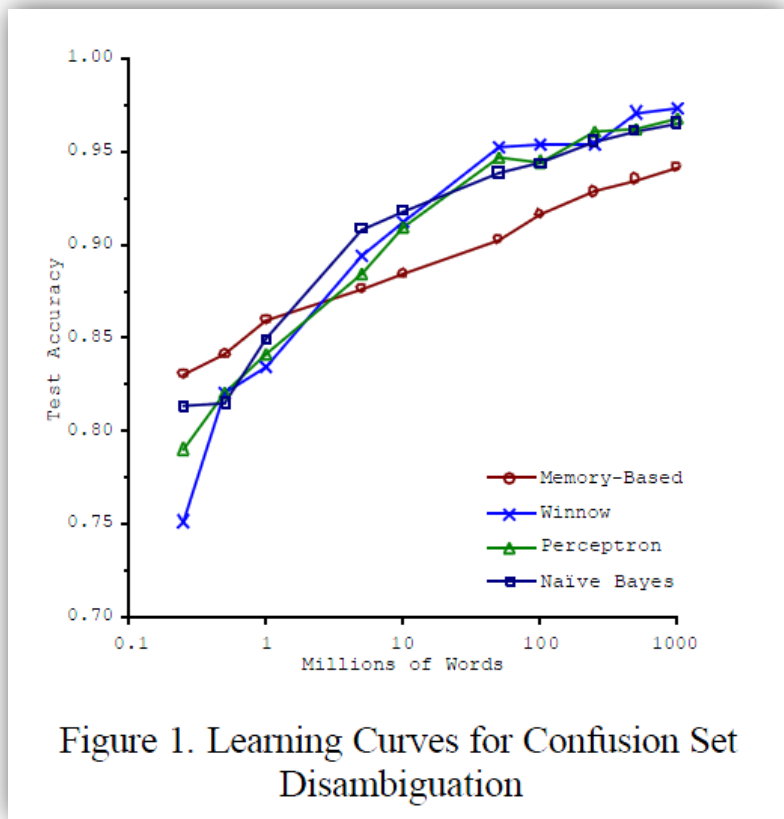


Figure 1. Learning Curves for Confusion Set Disambiguation

结论引起争论：看似算法不再要紧，数据更重要！看似不再需要研究复杂算法，找更多数据就行了

* M. Banko and E. Brili (2001). **Scaling to very very large corpora for natural language disambiguation**. *ACL 2001*

基础算法/技术层研究内容

大数据并行化机器学习和数据挖掘算法

2007, Google公司Brants *等基于MapReduce研究了一个基于2万亿个单词训练数据集的语言模型，比较了当时最先进的Kneser-Ney smoothing 算法与他们称之为“stupid backoff “的简单算法, 最后发现, 后者在小数据集时效果不佳，但在大数据集时，该算法最终居然产生了更好的语言模型！

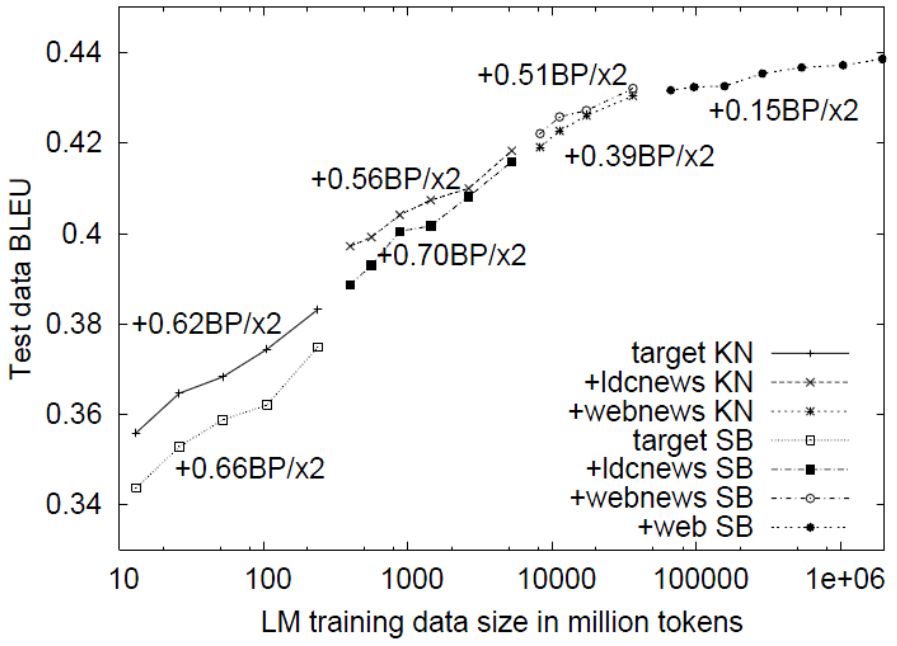


Figure 5: BLEU scores for varying amounts of data using Kneser-Ney (KN) and Stupid Backoff (SB).

结论：大数据集上的简单算法能比小数据集上的复杂算法产生更好的结果！

* T. Brants, A. C. Popat, et al. **Large Language Models in Machine Translation**. In EMNLP-CoNLL 2007 - Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning

基础算法/技术层研究内容

大数据并行化机器学习和数据挖掘算法

- 分类算法 (Classification)
  - 大规模支持向量机 (Large Scale SVM)
  - 神经网络 (Neural Network)
  - 朴素贝叶斯 (Naïve Bayes)
  - 决策树 (Decision Trees)
- 聚类 (Clustering)
- 关联规则挖掘
- 参数估计 (Parameters Estimation)

基础算法/技术层研究内容

大数据并行化机器学习和数据挖掘算法

- 高维度数据降维 (Dimension Reduction)
- 集成学习 (Ensemble Learning)
- 图数据算法
  - 图聚类
  - 图分类
  - 图模式匹配(子图同构、最大公共子图...)
- ...

并行编程模型与计算框架层研究内容

- MapReduce
  - Hadoop性能优化
    - 针对作业、任务和Slot资源的调度优化 (IBM的AMapReduce，Facebook的Corona）
    - 针对I/O的优化、针对充分利用内存的优化 (Berkeley的Spark)
    - 针对流程的优化 (优化Shuffle过程 、SHadoop)
  - MapReduce并行计算框架改进
    - 迭代式MapReduce执行框架 (Twister, HaLoop)
    - 流式MapReduce执行框架 (Hadoop Online)

并行编程模型与计算框架层研究内容

- MapReduce
  - MapReduce在不同构架上的实现
    - 基于众核构架的MapReduce (Stanford的Phoenix, 上海交大)
    - 基于GPU的MapReduce（香港科大、上海交大）

并行编程模型与计算框架层研究内容

- BSP (Bulk Synchronized Parallel)
  - 基于BSP模型的并行处理框架
- 大规模图数据并行处理框架和系统
  - Google的Pregel
  - 微软的Trinity
- CUDA、MPI、OpenMP
  - 提升可编程性

并行编程模型与计算框架层研究内容

- 定制式并行计算框架
  - 全内存集群计算 (Spark)
  - 大规模流式数据处理 (S4)
  - 特定应用问题的定制式并行计算框架
- 混合式并行计算模型和框架*
  - MapReduce+CUDA并行计算框架的设计与优化
  - MapReduce+MPI和MapReduce+BSP并行计算框架设计与优化

* A Survey of Parallel Programming Models and Tools in the Multi and Many-Core Era Javier Diaz, Camelia Mun˜oz-Caro, and Alfonso Nino. IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 23, NO. 8, AUGUST 2012

大数据存储管理层研究内容

- 大数据预处理技术
  - 大数据的采集和传输
  - 大数据的清洗过滤和质量管理技术
  - 大数据的压缩技术
- 记录型大数据索引和查询技术
  - 静态记录型大数据索引技术
  - 流式/增量式记录型大数据索引技术
  - 大数据表的高效关系型操作 (如查询连接)
  - 大数据并行化查询技术

大数据存储管理层研究内容

- 图数据表示和查询技术
  - 静态图数据的表示和存储
  - 静态图数据的查询
  - 流式/增量式图数据的表示和存储
  - 流式/增量式图数据的查询
  - 图数据并行化查询技术
- SQL/NoSQL查询语言接口与技术
  - SQL/NoSQL查询语言接口
  - 并行化查询执行机制

大数据存储管理层研究内容

- 混合式数据表示和存储管理模型
  - NoSQL数据库技术
  - 结构化/半结构/非结构化数据混合存储管理模型
  - 混合式数据下的数据关系和查询操作技术

大数据存储管理层研究内容

- 分布式数据库
  - HBase性能优化
  - 基于HBase的大数据索引和查询技术
  - 分布式内存数据库存储技术
- 分布式文件系统
  - HDFS的优化

并行构架和计算平台层研究内容

- 共享内存构架
  - 多核，GPU
- 分布内存构架
  - 集群
- 混合式构架
  - 集群+多核
  - 集群+GPU

并行构架和计算平台层研究内容

- 大数据应用/服务云计算支撑平台
  - 云计算支撑平台和框架
  - 大数据云存储技术
  - 大数据并行计算系统可靠性与容错恢复技术
  - 数据访问隐私保护和安全技术

大数据十个热点研究问题

一、大数据存储管理和索引查询问题
二、Hadoop性能优化问题
三、图数据并行计算模型和框架
四、并行化机器学习和数据挖掘算法
五、社会网络分析
六、排名和推荐
七、Web信息挖掘和检索
八、媒体分析检索
九、自然语言处理
十、大数据可视化计算与分析

系统层
基础算法
应用算法
应用技术

大数据研究方式

横向方式：
集中研究解决某个层面的问题

例如：

- 典型应用算法
- 并行机器学习数据挖掘算法
- 并行化编程模型和框架
- 大数据索引查询技术

电信/公安/商业/金融/遥感遥测/勘探/生物医药……

领域应用/服务需求和计算模型

行业应用系统开发

社会网络,排名与推荐,商业智能,自然语言处理,生物信息
媒体分析检索, Web挖掘与搜索, 3维建模与可视化…

并行化机器学习和数据挖掘算法

MapReduce, BSP, MPI, CUDA, OpenMP, 定制式,
混合式 (如MapReduce+CUDA, MapReduce+MPI)

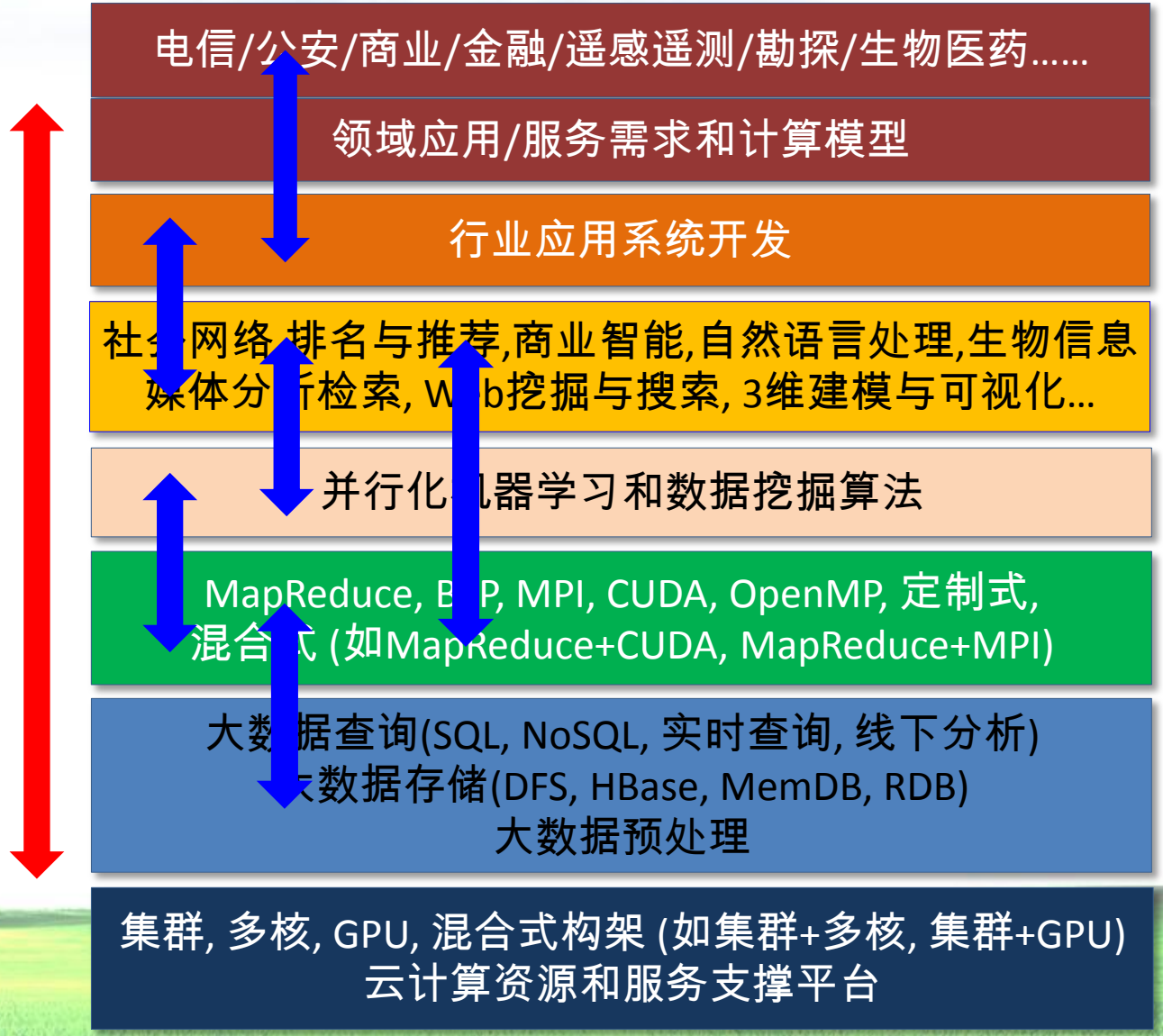
大数据查询(SQL, NoSQL, 实时查询, 线下分析)
大数据存储(DFS, HBase, MemDB, RDB)
大数据预处理

集群, 多核, GPU, 混合式构架 (如集群+多核, 集群+GPU)
云计算资源和服务支撑平台

大数据研究方式

纵向方式：上下层交叉组合

单一层面的研究往往难以获得理想的综合解决方案，上下层交叉组合既可以获得理想的综合解决方案，又能组合产生很多不同的研究点



大数据研究方式

上下层交叉组合，应用、算法、系统研究兼顾

- 从应用问题和需求出发

- 验证性原型应用系统

向上形成验证性原型应用系统

- 围绕典型和热点大数据应用算法和技术问题
- 解决并行化应用算法和基础算法问题

向下研究并行计算模型框架和数据存储系统层技术

- 研究混合式/定制式并行计算模型和框架

- 研究大数据存储和查询问题

第 三 部 分

大数据并行处理技术研究

大数据索引和查询技术

问题背景

大数据使得传统的关系数据库已经难以胜任，在存储能力和查询性能上都难以满足大数据存储和查询管理的需求。因此，需要针对应用需求研究大数据的索引和查询技术

- Oracle海量数据库系统Exadata，每个定制集群系统2千万元，存储100TB高性能数据
- IBM基于数据库DB2构建了定制的数据仓库集群系统，每集群存储数据60TB，价格5百万元
- 这些定制的分布式关系数据库系统价格过于昂贵，而数据存储处理能力仍然难以满足大数据处理要求，且系统难以扩充

大数据索引和查询技术

主要研究问题

大数据索引和查询技术主要研究非结构化或半结构化大数据的快速索引和查询优化技术，尤其是面向特定应用领域的大数据索引机制和管理技术、以及流式或增量式实时/准实时数据的索引和查询优化技术

目前本课题组正在研究基于分布式混合树索引的大数据索引和快速查询技术和算法。

大数据索引和查询技术

大规模移动电话通联记录索引和查询技术

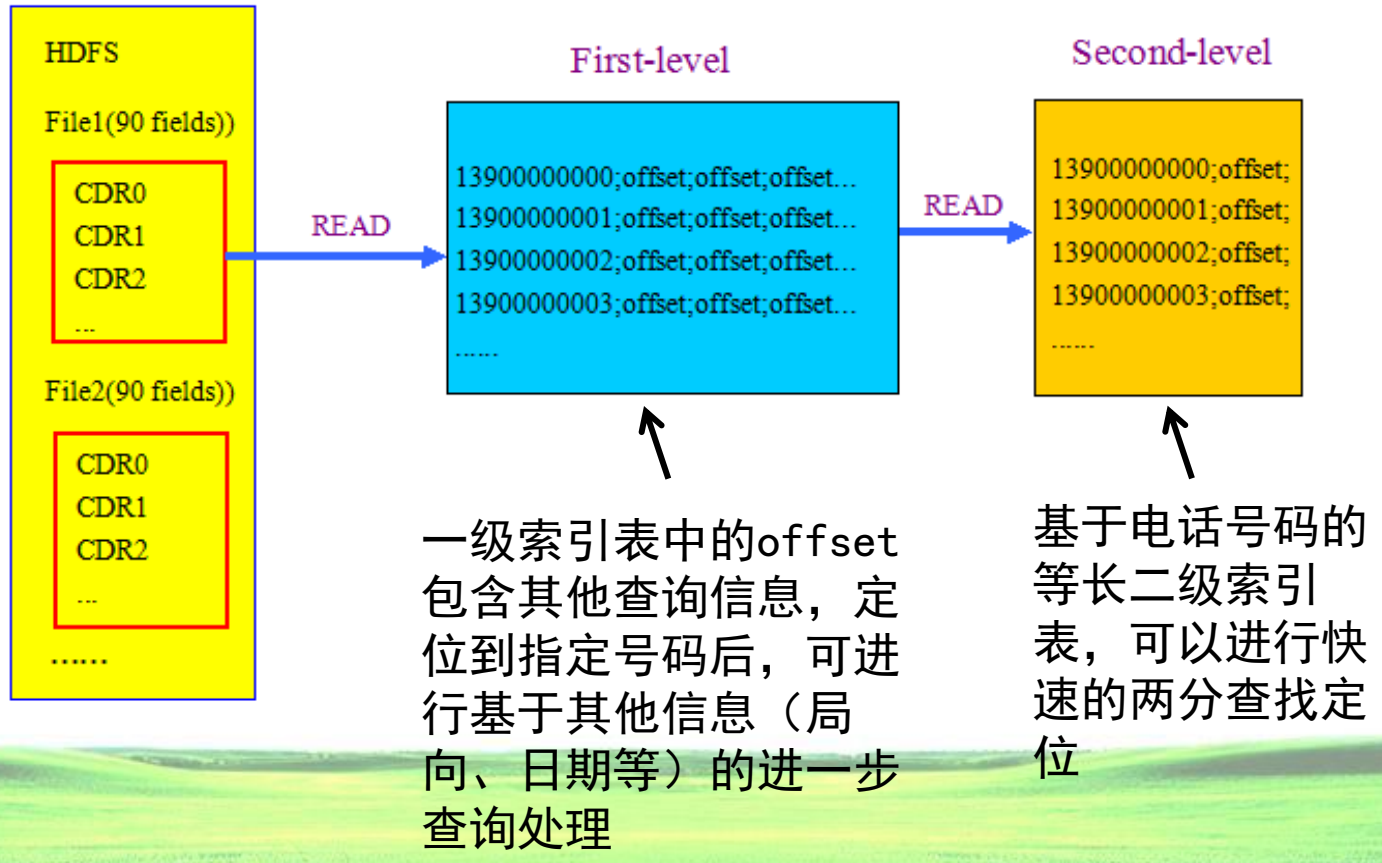
移动电话通联记录（CDR）数据量巨大，关系数据库已经越来越难以承受和胜任大量电话记录的管理和查询处理，为此，需要考虑基于Hadoop的分布式CDR数据存储和查询技术。

例如，在移动电话公司内部，最常使用的查询是依据电话号码（一个指定号码或者一个屏蔽了最后4位数字的万字段号码查询），加上其他查询信息（如局向、拨打或接受时间等）。为此提高查询速度，我们可以基于电话号码建立专门的快速查询索引表，然后使用两分快速查找方法，即可快速查询到指定号码的CDR数据记录。

大数据数据索引和查询技术

大规模移动电话通联记录索引和查询技术

CDR两级查询索引

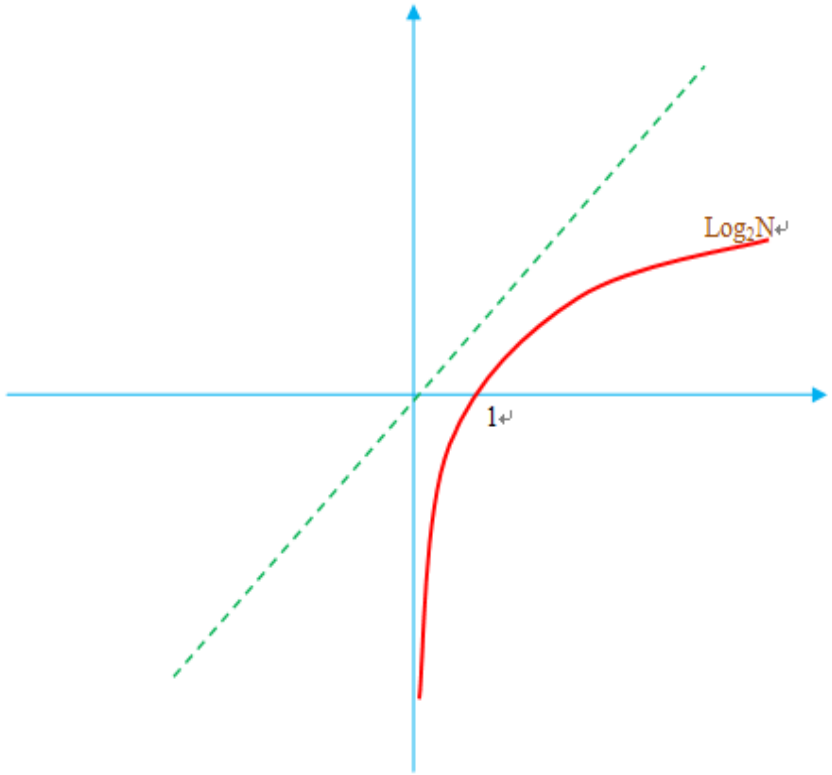


大数据数据索引和查询技术

大规模移动电话通联记录索引和查询技术

CDR两级查询索引

20亿个号码的CDR电话记录最多只需要比较大约31次即可完成！



Hadoop系统改进与优化研究

面向在线查询类任务的Hadoop系统性能优化

- Hadoop系统设计时重点考虑了高吞吐率大数据的处理，在作业执行性能上不够理想，对实时响应要求较高的查询类作业难以满足要求。

我们进行的工作：

1. 基于短作业任务调度的性能优化
2. 基于动态slot调度的性能优化

研究论文：BigDataMR2012，计算机研究与发展，IPDPS2013

SHadoop: Optimizing Execution Performance of Short MapReduce Jobs

Rong Gu, Xiaoliang Yang, Jinshuang Yan, Chunfeng Yuan, and Yihua Huang

Performance Optimization for Short MapReduce Job Execution in Hadoop

Jinshuang Yan, Xiaoliang Yang, Rong Gu, Chunfeng Yuan, and Yihua Huang

Hadoop系统改进与优化研究

基于短作业任务调度的Hadoop系统性能优化

现有标准MapReduce作业初始化和结束时需要花费十几秒的常数时间，作业执行时，map和reduce任务的调度都依赖于心跳机制进行任务调度时的消息传递和通信，因而任务调度时间开销较大，效率较低

解决方案：

1. we optimize the *setup* and *cleanup* tasks to reduce the time cost during the initialization and termination stages of a job

   我们优化了作业初始化和作业结束阶段的setup和cleanup两个特殊任务的调度，去除了以前所有作业都需要花费的十几秒常数时间

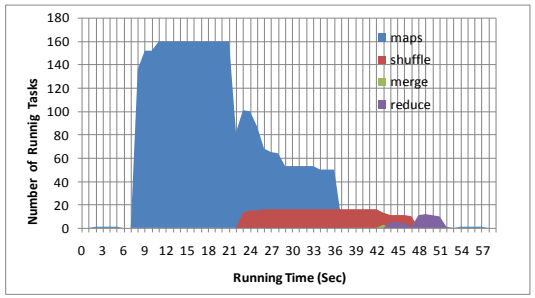
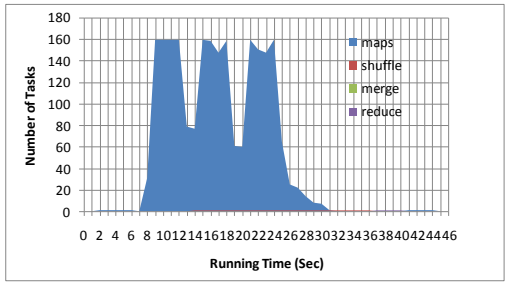
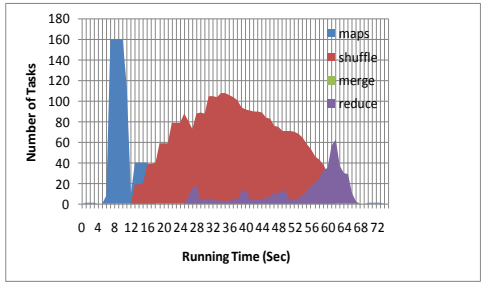
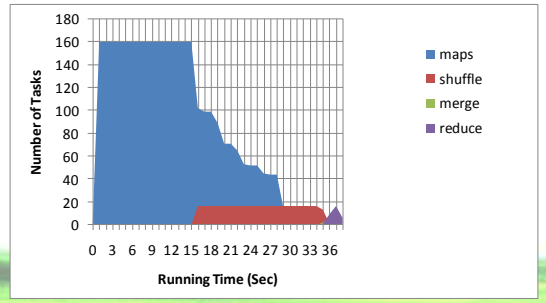
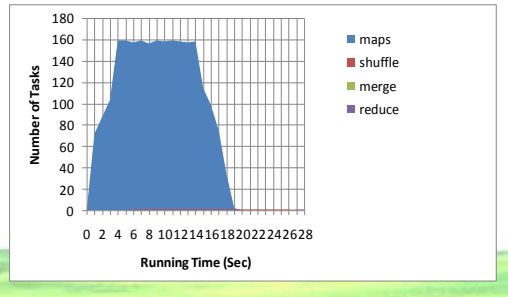
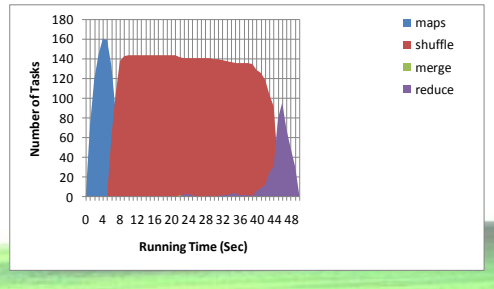
2. we design and implement an instant messaging model into the standard Hadoop for task scheduling event notifications between the JobTracker and TaskTrackers, instead of using the original heartbeat-based communication mechanism

   我们在JobTracker和TaskTracker之间设计实现了一种即时消息传递机制，去除了原有的心跳通信机制，显著缩短了作业内任务的调度时间

Hadoop系统改进与优化研究

基于短作业任务调度的Hadoop系统性能优化

实验结果：对WorldCount, Grep 和 TeraSort等MapReduce的标准Benchmark程序执行性能提升达到35%

	WorldCount	Grep	TeraSort
优化前			
优化后			

Hadoop系统改进与优化研究

基于动态slot调度的Hadoop系统性能优化

现有标准MapReduce作业执行时其底层的Slot调度是通过静态的配置文件设置的，作业执行过程中即使有空闲的Slot也无法为忙绿的任务所使用，map任务与reduce任务间的Slot也不能互换使用，因而系统的Slot计算资源使用率不高，也导致作业执行性能不高

解决方案：

基本解决方案是，我们在作业执行过程中及时收集Hadoop系统Slot资源分配使用的动态信息，并根据这些信息为作业动态分配和调度Slot资源

此项工作目前正在编码实现和调试阶段

基础性大数据并行算法

机器学习与数据挖掘基础算法

- 并行化聚类算法
- 并行化分类算法
- 并行化关联规则挖掘算法
- 神经网络并行化算法
- 图比对并行化算法
- ……

基础性大数据并行算法

频繁项集挖掘并行化算法

本研究组进行了基于MapReduce的频繁项集挖掘算法研究

基本思路是基于传统的Apriori算法和SON算法，提出并实现了一个并行化的频繁项集挖掘算法PSON，用两轮MapReduce实现了大规模频繁项集挖掘并行计算

研究论文，已发表于PAAP2011国际会议

PSON: A Parallelized SON Algorithm with MapReduce for Mining Frequent Sets

Tao Xiao, Shuai Wang, Chunfeng Yuan, Yihua Huang

The Fourth International Symposium on Parallel Architectures, Algorithms and Programming (PAAP 2011)， Tianjin， Dec. 9-11, 2011

Frequent sets

- Suppose I is an itemset consisting of items from the transaction database D
  - Let N be the number of transactions D
  - Let M be the number of transactions that contain all the items of I
  - M/N is referred to as the *support* of I in D

TID	Items
T100	I1, I2, I5
T200	I2, I3, I4
T300	I3, I4
T400	I1, I2, I3, I4

Example

Here, N = 4, let I = {I1, I2}, than M = 2

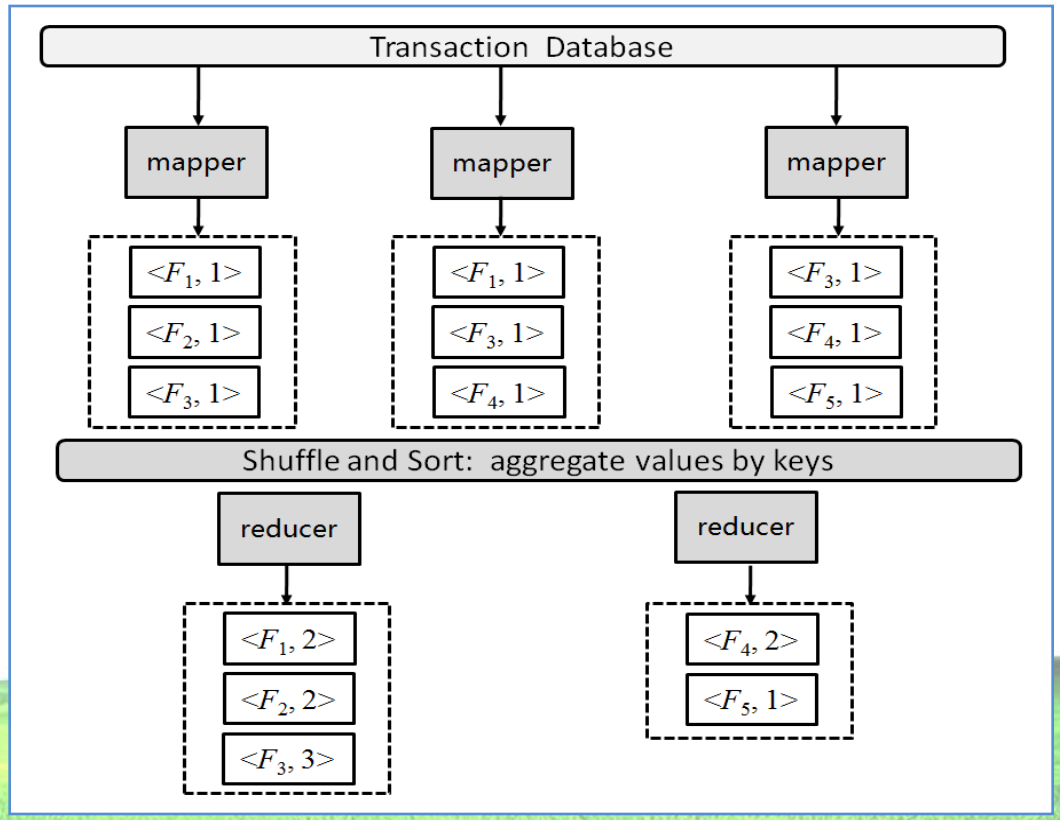
because I = {I1, I2} is contained in transactions T100 and T400

so the support of I is 0.5 (2/4 = 0.5)

- If $sup(I)$ is no less that an user-defined threshold, then I is referred to as a frequent itemset
- Goal of frequent sets mining
  - To find all frequent k-itemsets from a transaction database (k = 1, 2, 3, ...)
- 枚举计算的时间复杂度是：$O(2^n*N*t)$， n是Item的总数，N是Transaction总数，t是每个Transaction平均包含的Item数

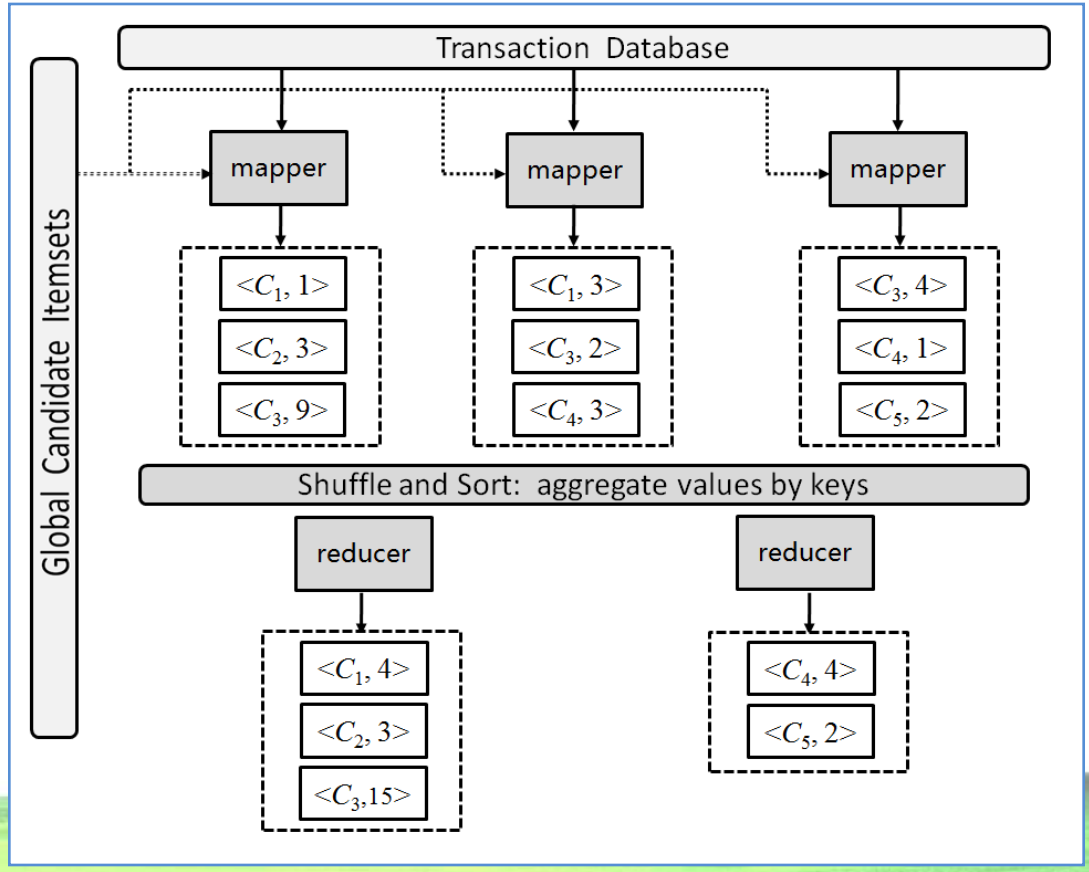
The 1st MapReduce Job

- **Map phase**
  - Each map node takes in one partition and generates local frequent itemsets for that partition using Apriori algorithm.
  - For each local frequent itemset *F*, emits key-value pair <*F*, 1>. Here, the value 1 is only to indicate that *F* is a local frequent itemset for that partition.
- **Shuffle and Sort phase**
  - The same local frequent itemsets are sent to one reduce node.
- **Reduce phase**
  - Each reduce node emits one and only one key-value pair <F, 1> to DFS
- **Finally**
  - Merging all the pairs in DFS gives us all global candidate itemsets



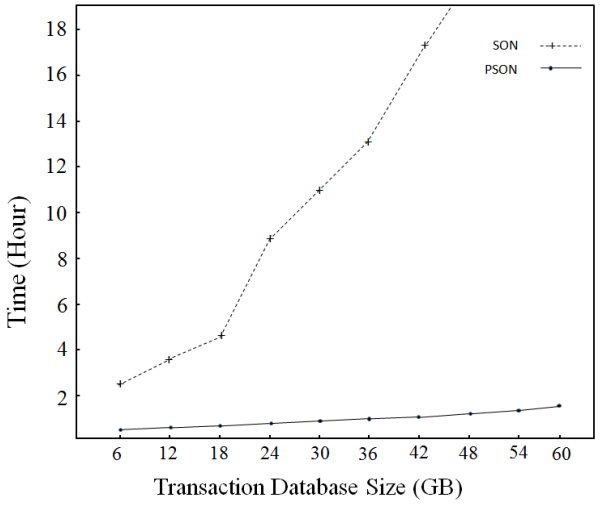
The 2nd MapReduce Job

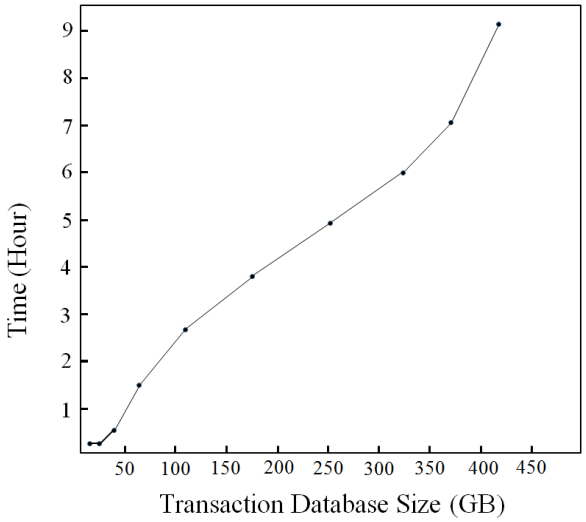
- **Assumption**
  - Each node is given a full duplicate of the global candidate itemsets generated by the 1st MapReduce job beforehand
- **Map phase**
  - Each map node counts for each of the global candidate itemsets in the partition the map node is assigned
  - Then emits pairs like <*C*, *v*> where *C* is a global candidate itemset and *v* is the count of it in that partition
- **Shuffle and Sort phase**
  - Each global candidate itemset and its counts in all the partitions are sent to one reduce node
- **Reduce phase**
  - For each global candidate itemset *C*, reduce node adds up all the associative counts for *C* and emits only the actual global frequent itemsets to DFS

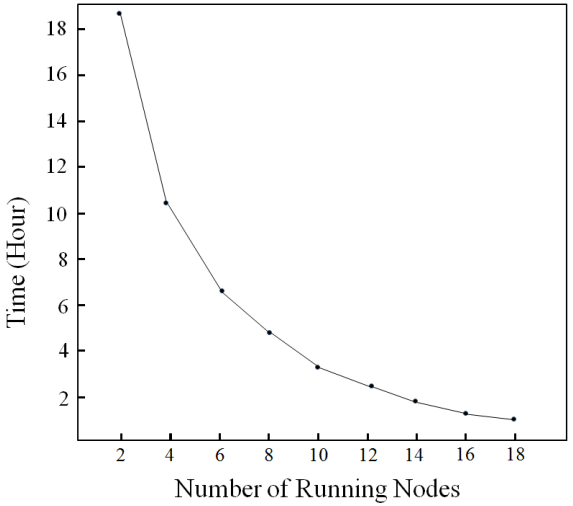


Experimental Results

- The transaction database size varies from 6GB to 60GB， with the number of transactions varies from 1 million to 500 billion







Conclusion:

- When the size of the database reaches a threshold of hundreds of GB, PSON can finish running in an acceptable period of time， achieving a good performance in scale-up
- PSON can achieve a good performance in speed-up

基础性大数据并行算法

查询推荐QUBIC并行化算法

本研究组进行了基于MapReduce的查询推荐QUBIC并行化算法。基本思路是基于用户日志设计查询推荐算法，首先挖掘用户日志中Query与URL之间的关系，寻找Query中若干关联性较大的组，最后基于MapReduce并行构造Query-URL二部图和查询亲和图QAG，在此基础上最终完成查询词的聚类，并以此为基础推荐查询关键词

基础性大数据并行算法

短文本多分类并行化算法

本研究组进行了基于MapReduce的查询短文本分类并行化算法研究。原始数据有1000万条查询短文本，需要分为500个类别，其中1万条已经标注出所属类别作为训练样本，需要对其他短文本进行分类。

研究成果：
本项目为本系研究生组队参加2012年中国第一届“云计算与移动互联网大奖赛”的指定的4个大数据并行处理赛题之一，经过角逐获得1、2、3等奖各一名。

基础性大数据并行算法

神经网络并行化算法

本研究组基于Hadoop的Hbase和底层RPC远程过程调用通信，采用分布内存式数据缓存机制，为经典的BackPropagation神经网络算法研究设计了一个定制的轻量级专用并行化计算框架，并最终设计实现了并行化的BackPropagation神经网络算法。神经网络需要经过数万至数十万轮的迭代计算，计算量巨大。由于实现了并行化算法，大大缩短了训练时间，本课题进行了从未有人做过的800万超大训练样本的神经网络训练并行化计算测试。

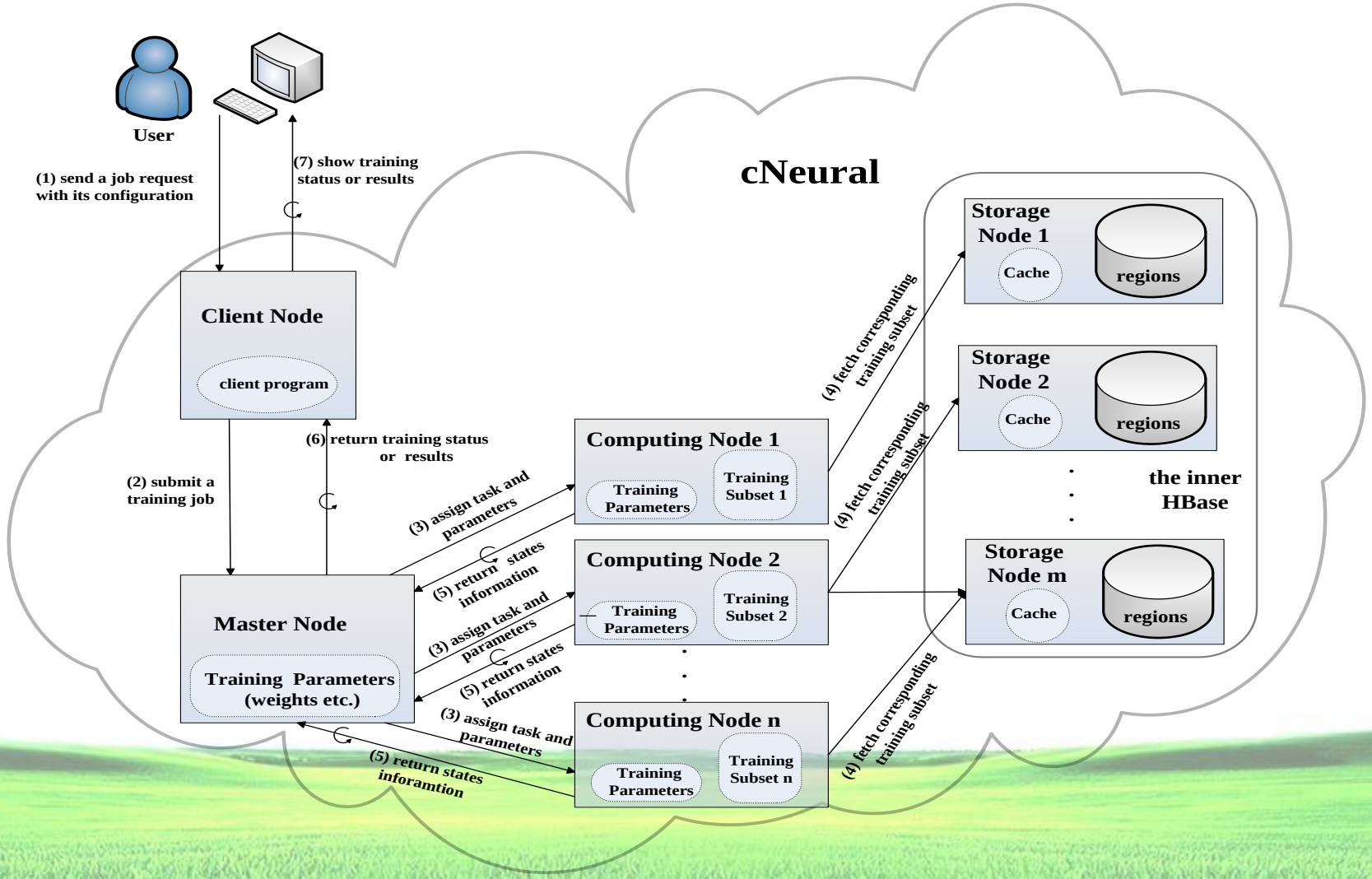
研究论文：投稿IPDPS2013,审稿中

A Parallel Computing Platform for Accelerating Large Scale Neural Network Training

Rong Gu, Furao Shen, Yihua Huang

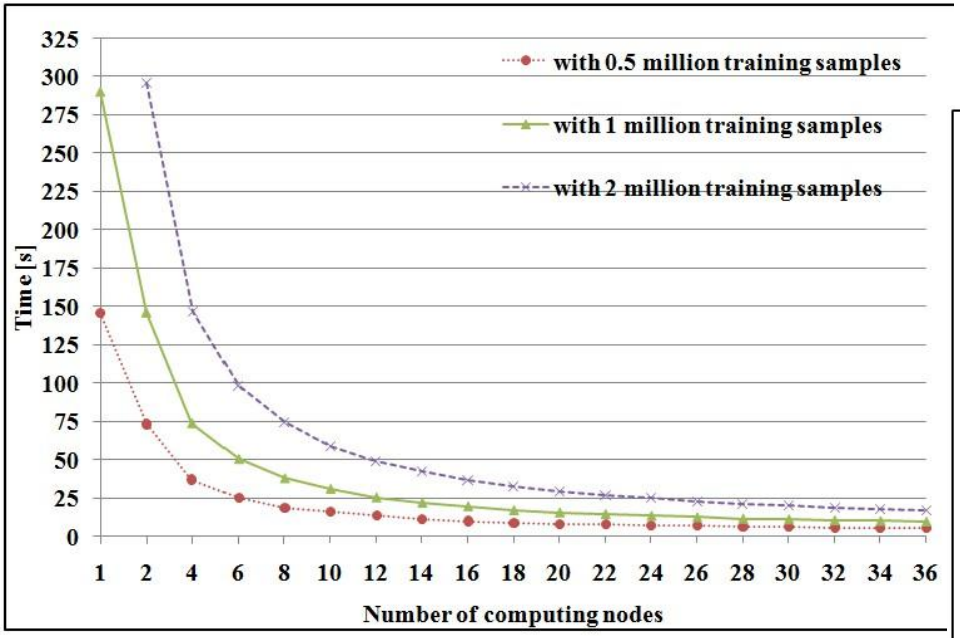
基础性大数据并行算法

神经网络并行化算法

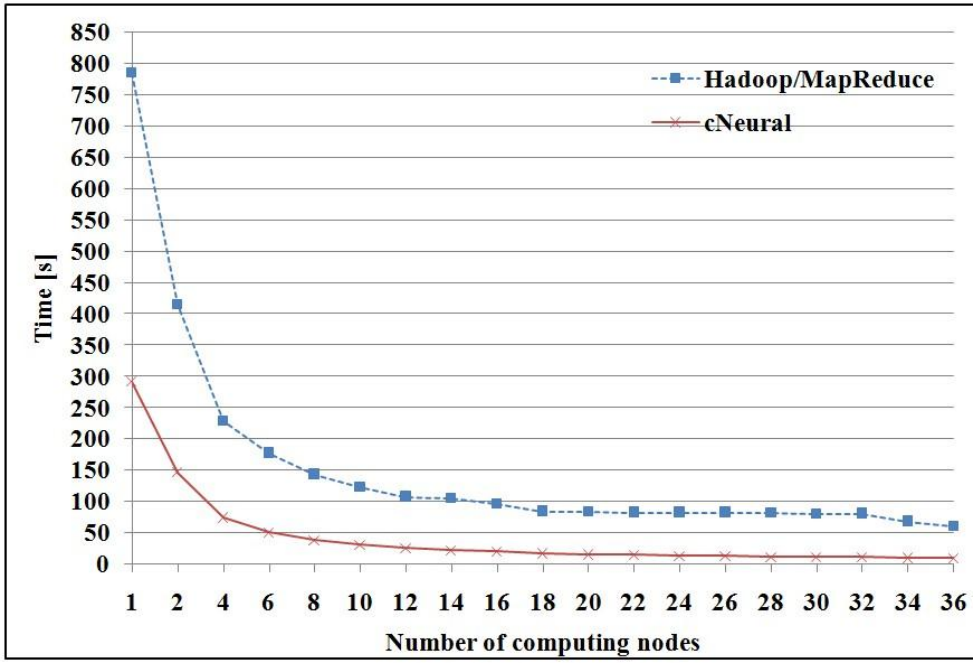


基础性大数据并行算法

神经网络并行化算法



Performance of each epoch's training time cost in cNeural with various numbers of computing nodes and various sizes of training samples



Comparison of each epoch's training time cost in cNeural and Hadoop with different number of computing nodes on 1 million training samples.

基础性大数据并行算法

图比对并行化算法

本研究组进行了基于MapReduce的图相似性比对并行化算法研究， 基于多趟复杂的MapReduce计算，研究实现了并行化的经典图比对算法Similarity flooding algorithm，取得显著的并行化性能提升，可以有效地解决大规模子图的相似性比对问题。

研究论文：投稿PDCAT2012国际会议，已接收

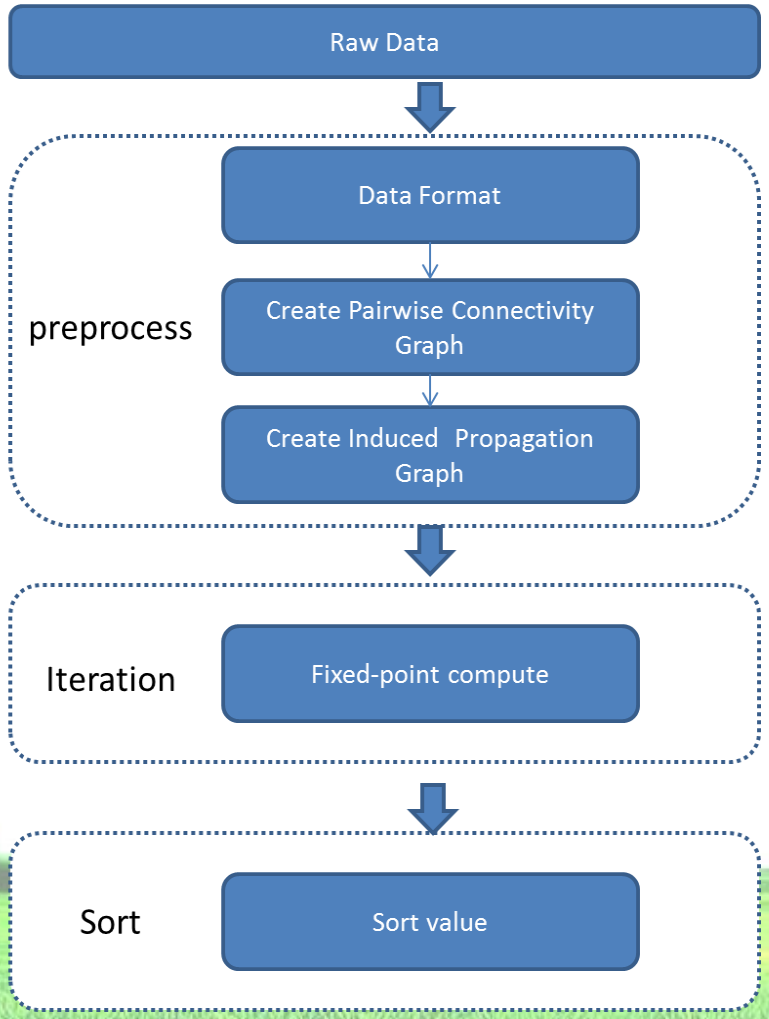
Parallelized Similarity Flooding Algorithm for Processing Large Scale Graph Datasets with MapReduce

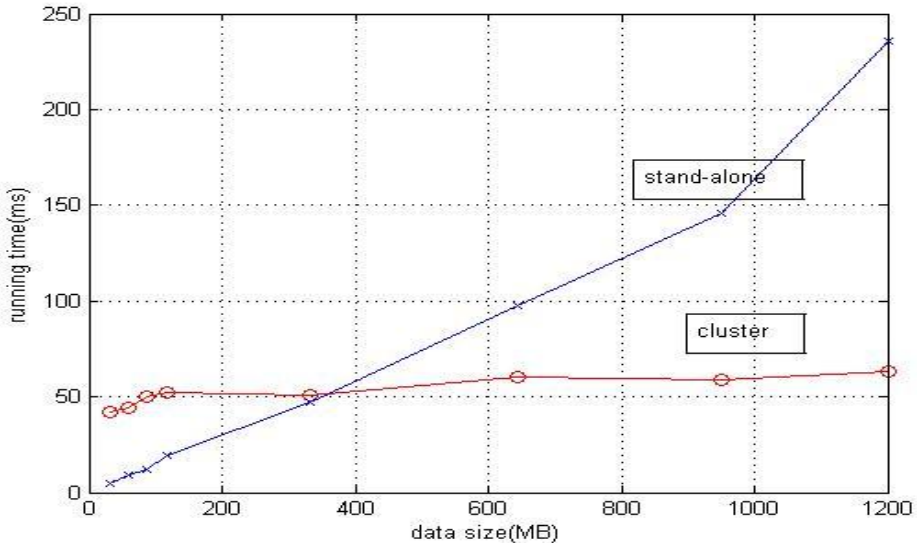
Jian Zhang, Chunfeng Yuan, and Yihua Huang

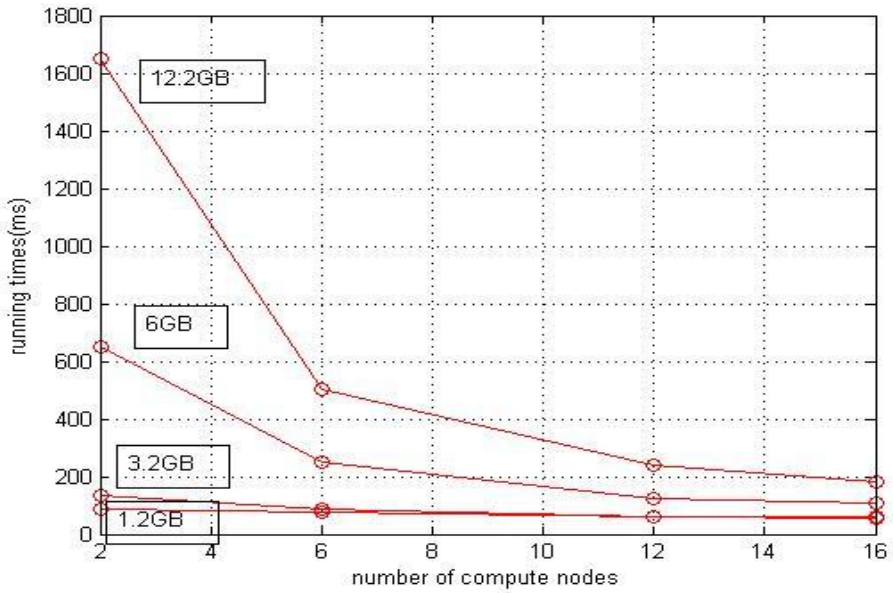
基础性大数据并行算法

图比对并行化算法

The Whole Process







特定应用问题的大数据并行算法

重复文档检测算法（Duplicate Document Detection）

本研究组进行了重复文档检测算法研究

问题：
搜索引擎的结果中包含大量重复文档链接

Numerous copies of web documents creating a serious problem for search engines:

a. enlarge the *space* to store index
b. increase the *cost* of crawling, ranking, clustering…
c. unbeneficial information on the *first page* in search result

特定应用问题的大数据并行算法

重复文档检测算法（Duplicate Document Detection）

特定应用问题的大数据并行算法

重复文档检测主要处理过程

根据Shingling算法具有的计算高效的特点以及IMatch算法具有的准确度较高的特点，基于Shingling算法，并借助IMatch算法中强化语义特征的处理思想，研究并提出一种改进的重复网页检测和过滤算法CoreMatch，该算法针对目前现有的英文文档重复检测方法在处理效果和适用性方面的不足，研究并提出一种适用于中文文档重复检测和过滤的方法，并基于多趟执行的MapReduce程序设计实现了大规模并行化重复文档检测算法

研究论文：PDCAT2012 国际会议，已接收

Parallelized Near-Duplicate Document Detection Algorithm for Large Scale Chinese Web Pages

Yongzhuang Wei, Shuai Wang, Chunfeng Yuan, and Yihua Huang

算法结果比较

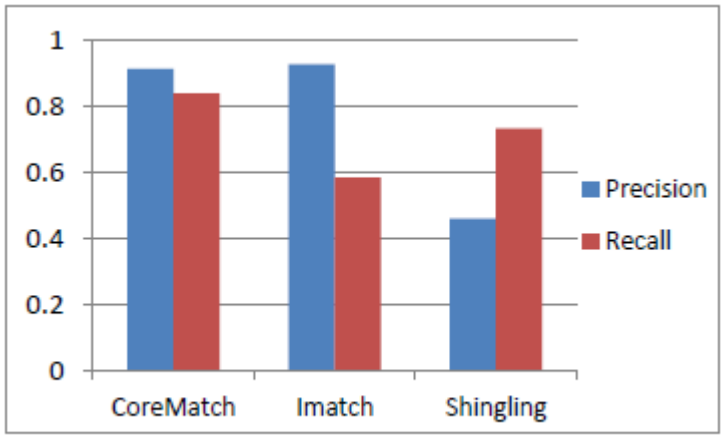


Figure 7 The precision and recall of each algorithm.

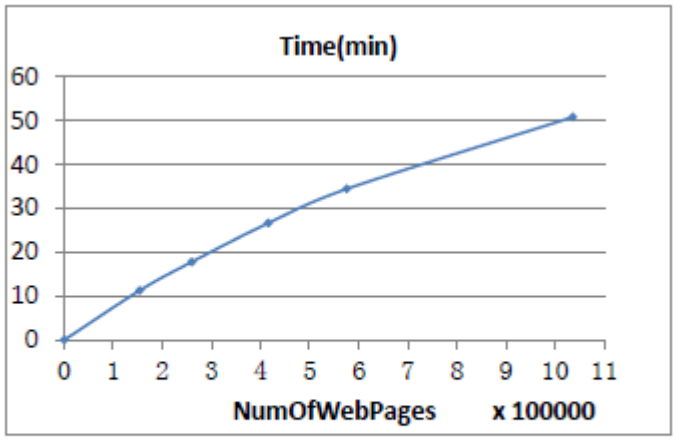


Figure 8 The time of our improved approach in processing huge number of Web pages range from 100,000 to 1,000,000.

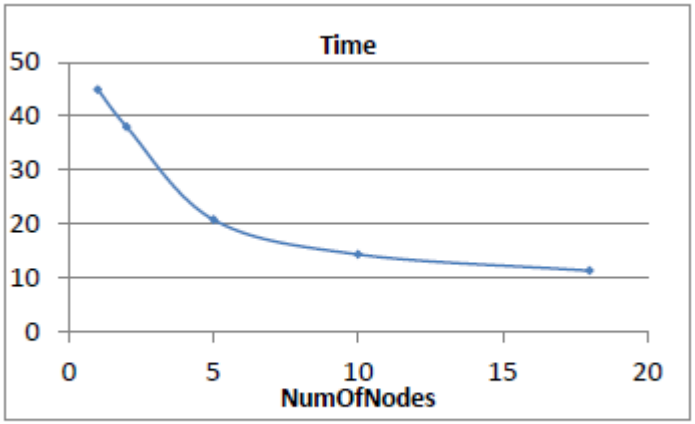


Figure 9 The processing time of Hadoop system with datanodes range from 1 to 18.

特定应用问题的大数据并行算法

具有重复网页检测和聚类功能的中文新闻搜索引擎

英文重复网页检测已有很多相关的研究工作，然而迄今为止，还未见相关的研究文献专门用于解决中文网页的重复检测问题。虽然已有算法在一定程度上可以解决中文网页的重复检测问题，但是由于中文与英文之间在语法和语义上存在的显著差别，使得中文处理方法与英文处理方法有着很大的不同，尤其在新闻网页的处理上。因此，针对中文新闻网页的特点，本文提出一种基于“句号”特征来提取新闻网页特征的方法CCDet。该方法首先提取新闻网页中的句号特征，并定义一种新的网页相似度度量方法称为“一般包含相似度”，该方法可以有效的度量网页之间的重复关系和包含关系。同时CCDet会对具有重复关系和包含关系的网页进行聚类。由于重复网页检测的网页数据量和计算量巨大，因此，我们进一步研究实现了基于MapReduce的CCDet算法和中文新闻搜索引擎。

特定应用问题的大数据并行算法

具有重复网页检测和聚类功能的中文新闻搜索引擎

实验结果显示，CCDet在检测网页重复关系和包含关系上的精确度和召回率均达到很好的效果，比现有算法在精度上有显著的提高。

算法	重复对的个数	正确个数	精确度
CCDet	393	392	0.997
IMatch	131	53	0.405
SpotSigs	1030	47	0.045

研究成果：

本项目为本研究组研究生组队参加2012年中国第一届“云计算与移动互联网大奖赛”的创意赛题，经过角逐已经成为10个优胜创意项目之一进入复赛，并获得二等奖。

特定应用问题的大数据并行算法

大规模长基因序列比对算法

本研究组进行了基于MapReduce的大规模基因序列比对并行化算法研究，设计实现了两种并行化比对算法map side extension BLAST和reduce side extension BLAST

研究论文，已发表于PAAP2011国际会议：

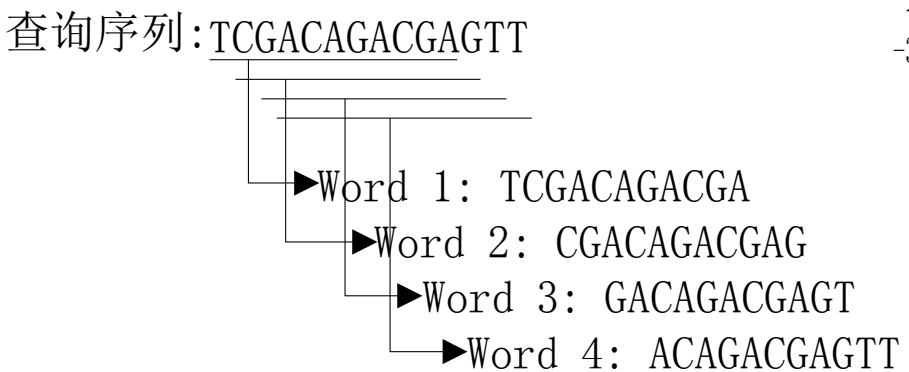
Parallization of BLAST with MapReduce

Xiaoliang Yang, Chunfeng Yuan, Yihua Huang

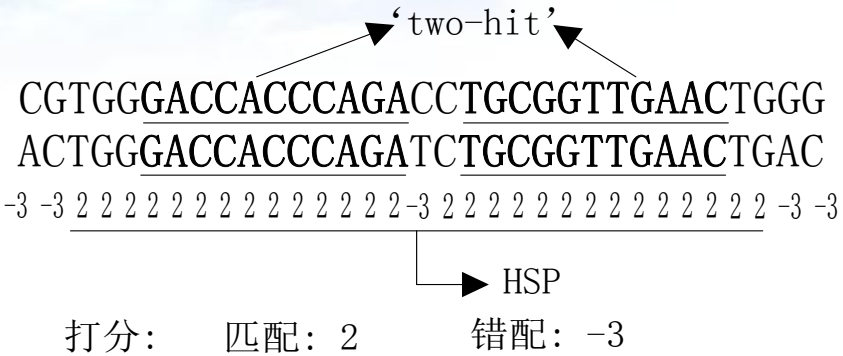
The Fourth International Symposium on Parallel Architectures, Algorithms and Programming (PAAP 2011)， Tianjin， Dec. 9-11, 2011

特定应用问题的大数据并行算法

基因比对处理方法



1. 划分单词片段



2. 用查询序列中的单词片段到已知基因序列中比较，找到两个相邻的单词片段匹配
3. 以此为基础，向序列两侧扩展，找到一个最高分的匹配串

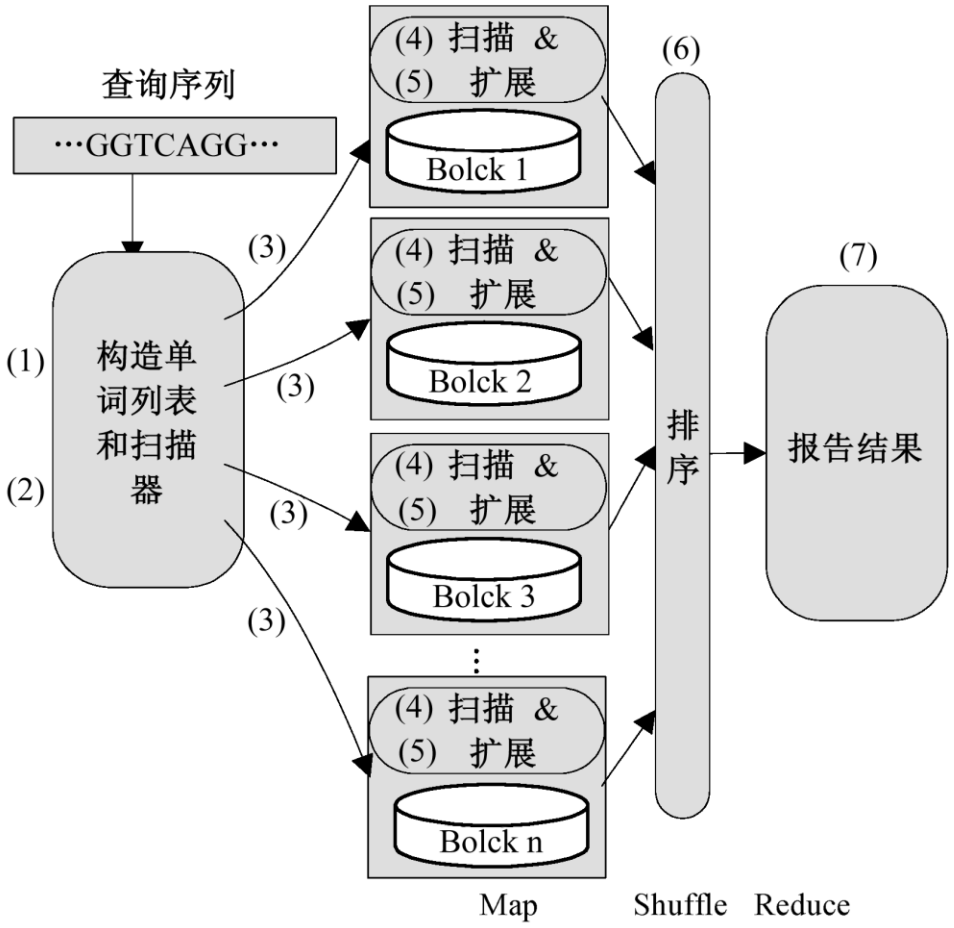
4. 当这个最高分匹配串达到一定的分值时，触发一个对查询序列与已知基因序列进行精确匹配比较的过程，该过程用动态规划方法完成

1-3步进行初步的筛选，快速过滤掉大量不可能匹配的序列，以此大大减少比对数量，第4步对筛选出的可能匹配的序列进行精确比对

特定应用问题的大数据并行算法

基因比对并行化算法

（1）由查询序列构造单词列表；
（2）从单词列表构造一个扫描器
（3）利用Hadoop的Distributed Cache将查询序列和扫描器发送到每个节点上，然后启动MapReduce Job进行序列比对；
（4）在Map阶段，每个map task从Distributed Cache文件中读取查询序列并加载扫描器，然后在本地的数据块上扫描单词匹配（word hit）；满足two-hit条件的匹配会被保留下来进行扩展；
（5）在每个节点上，扫描完成后，对保留下来的单词匹配先后做精确匹配扩展和允许空位的扩展（动态规划方法）

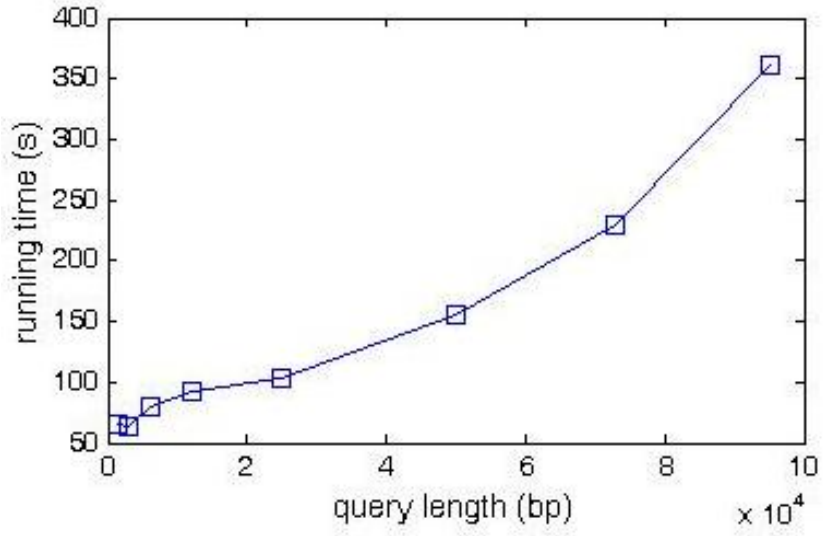


MapReduce BLAST overview

特定应用问题的大数据并行算法

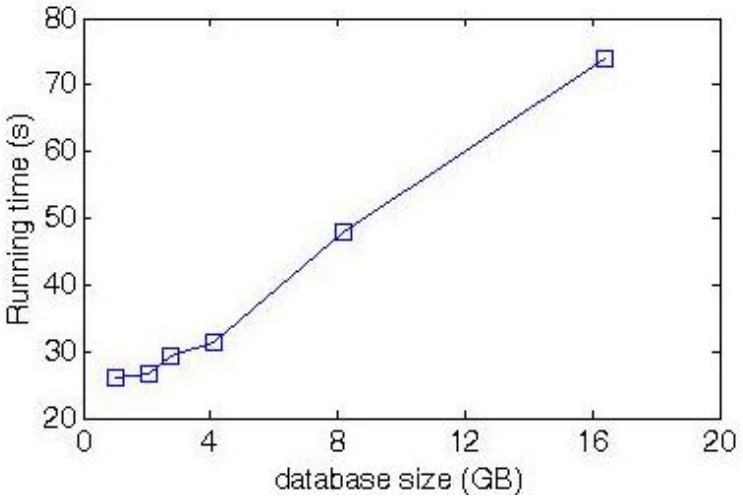
基因比对并行化算法实验结果

The running time grows nearly linearly as the query length increases



Fragments of increasing length from a 95kb neucleoside sequence were aligned with the 16GB nt sequence database.

The running time scales nearly linearly as the size of the sequence database increases

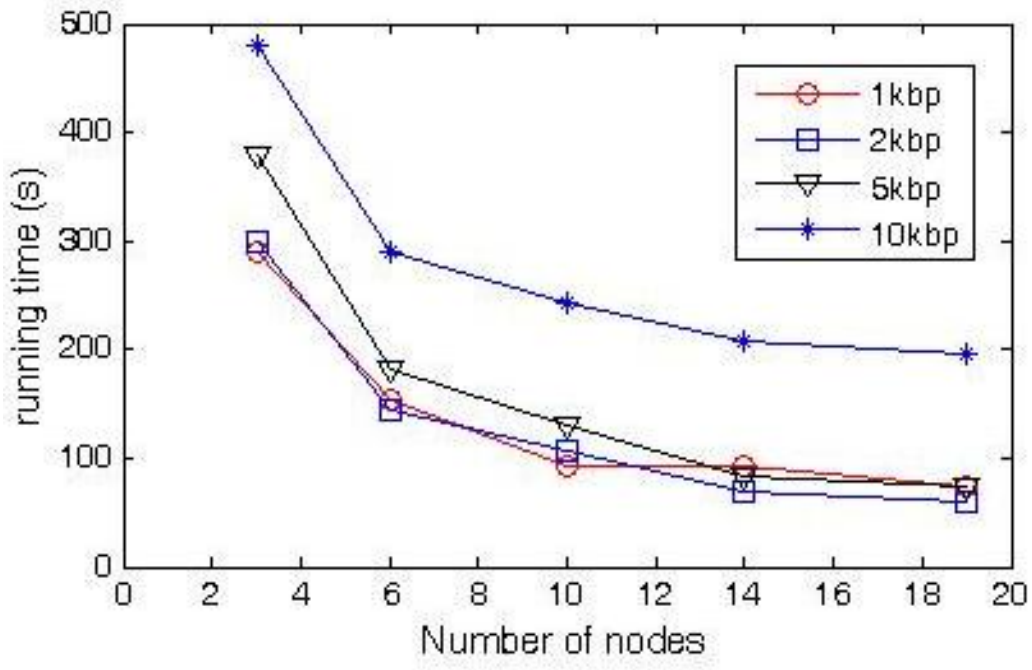


A set of queries were aligned with sequence databases of different sizes

特定应用问题的大数据并行算法

基因比对并行化算法实验结果

- The running time of searching a sequence database dropped quickly as the number of compute nodes increases.



Four nucleoside sequences of 1kbp, 2kbp, 5kbp, and 10kbp in length respectively were aligned with the 16GB nt database on different cluster configurations(3 to 19 nodes, 24 to 152 CPU cores).

特定应用问题的大数据并行算法

统计机器翻译并行化算法和系统

本课题组进行了基于大语料库统计机器翻译并行化算法研究。研究显示，语料库数据越大，翻译精度越高，但计算量太大，速度太慢。统计机器翻译主要包括生成翻译模型，目标语言模型和翻译decode解码三个主要过程。

翻译模型生成需要处理大量的平行预料，一般都是TB级别，算法未并行化时，训练速度太慢，单机处理需要1周左右；

目标语言模型生成需要处理大量目标语言，生成N-gram语言模型，算法未并行化时，训练速度太慢，单机处理需要好几天左右

翻译Decoder解码算法对翻译的精准度和速度都有重要影响，算法涉及到大量的查表操作（每翻译一个句子，几百万次），处理不好会导致速度很慢。

特定应用问题的大数据并行算法

统计机器翻译并行化算法和系统

本研究组研究实现了基于MapReduce的翻译模型和语言模型生成并行化处理算法，并分别基于Hbase和分布式内存数据库系统Redis设计实现了快速并行化Decode解码算法。

翻译Decode解码算法实验效果

我们采用了多机多线程的方式进行查询，同样的100条句子针对语言模型表的查询，在线程数和机器数量选择合适的情况下，我们耗时是能从原来的360s降低到25s

- 比基于HBase的并行算法性能提升8.8倍！
- 比原来基于胖节点的并行算法性能提升14.4倍！

研究成果：

目前该项工作整体性系统优化还在进行中，即将撰写研究论文

特定应用问题的大数据并行算法

大规模图像检索并行化算法

本研究组进行了基于MapReduce的大规模图像搜索并行化算法和系统研究，针对数十万、数百万量级的图片，通过颜色特征值粗选方式快速筛选出最接近的数千张图片，然后再采用二级细粒度比对的方法，完成快速的图像比对和搜索。

研究成果：
本项目为本系研究生组队参加2012年中国第一届“云计算与移动互联网大奖赛”的指定的4个大数据并行处理赛题之一，经过角逐获得1、3等奖各一名。

特定应用问题的大数据并行算法

城市路径规划并行化算法

本研究组进行了基于MapReduce的大规模城市路径规划并行化算法和系统研究，针对北京市的实际城市道路和交通时空信息，完成优化的交通路径规划。

研究成果：
本项目为本系研究生组队参加2012年中国第一届“云计算与移动互联网大奖赛”的指定的4个大数据并行处理赛题之一，经过角逐获得2、3等奖各一名。

大规模数据并行处理应用研究与开发

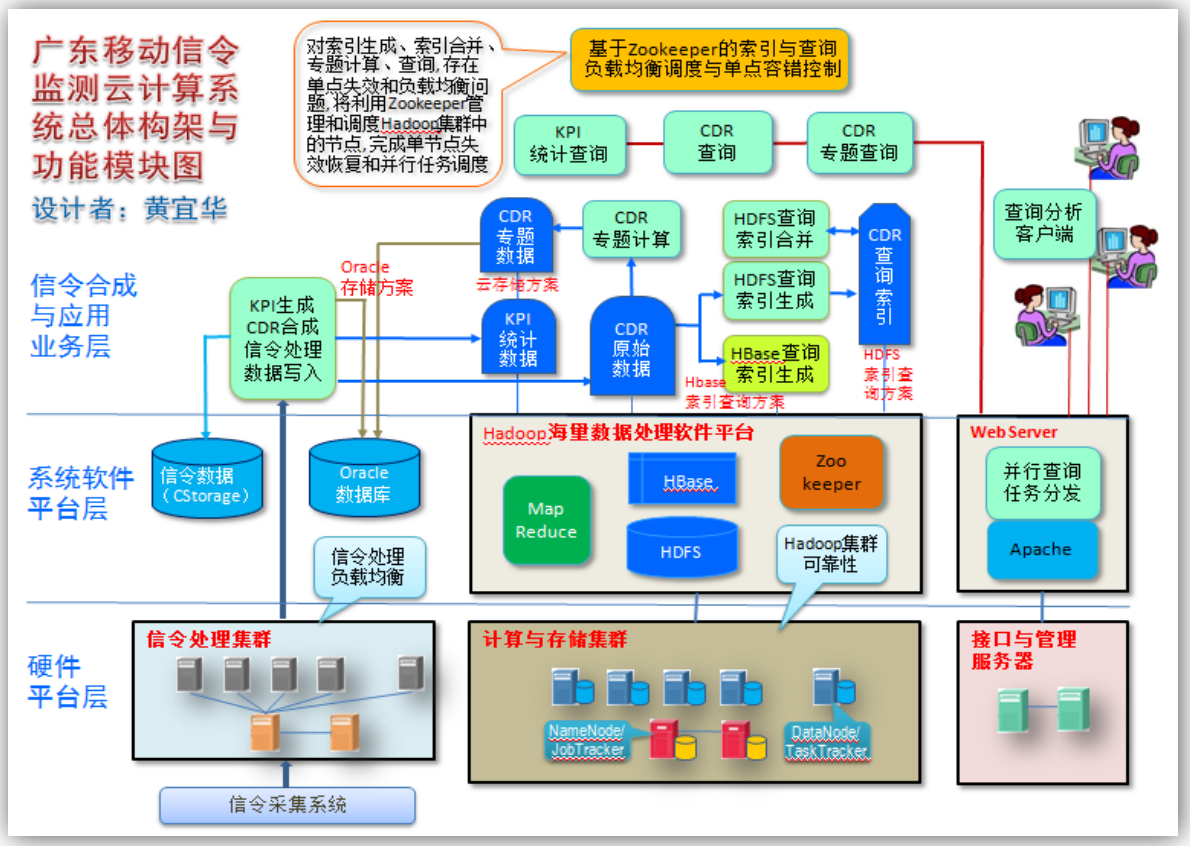
大规模数据处理将可以应用于各种需要处理海量数据的行业和应用

- 电信数据信息处理与挖掘
- 电网数据信息处理与挖掘
- 警务云应用系统（道路监控、视频监控、网络监控、智能交通、反电信诈骗、指挥调度等公安信息系统）
- 大规模基因序列分析比对
- Web信息挖掘
- 多媒体数据并行化处理
- 其他各种行业的云计算和海量数据处理应用

大数据应用系统开发

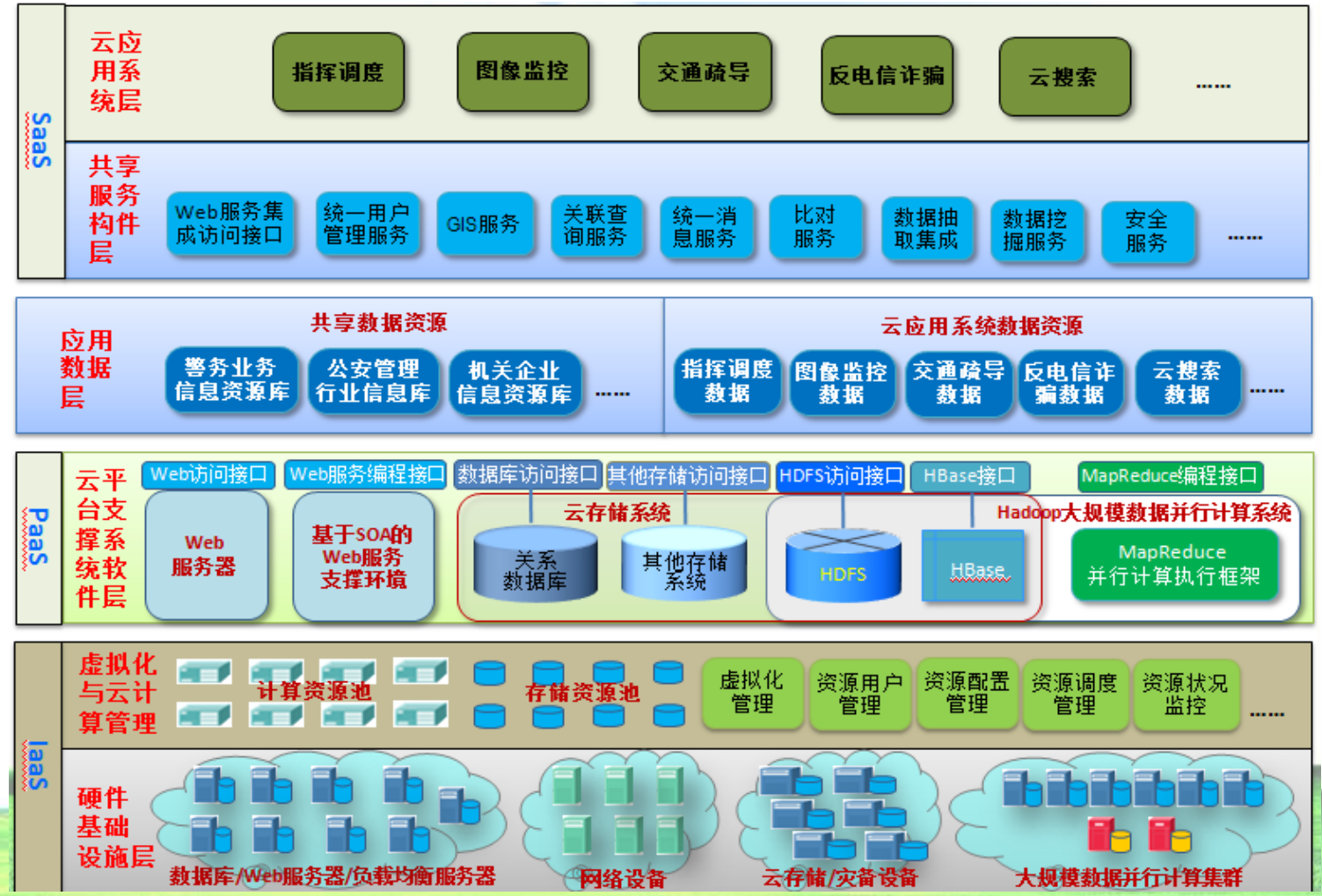
移动信令查询监测云计算系统

本课题组参加了广东移动信令检测大数据云计算系统项目，并负责设计了基于Hadoop的云计算系统构架和软件框架，研究设计了基于MapReduce的信令数据查询分析并行化算法和系统。



大数据应用系统开发
公安警务云计算平台与系统设计

本课题组与南京市公安局合作，帮助进行警务云计算平台和应用系统的规划和设计



大数据并行处理技术教学

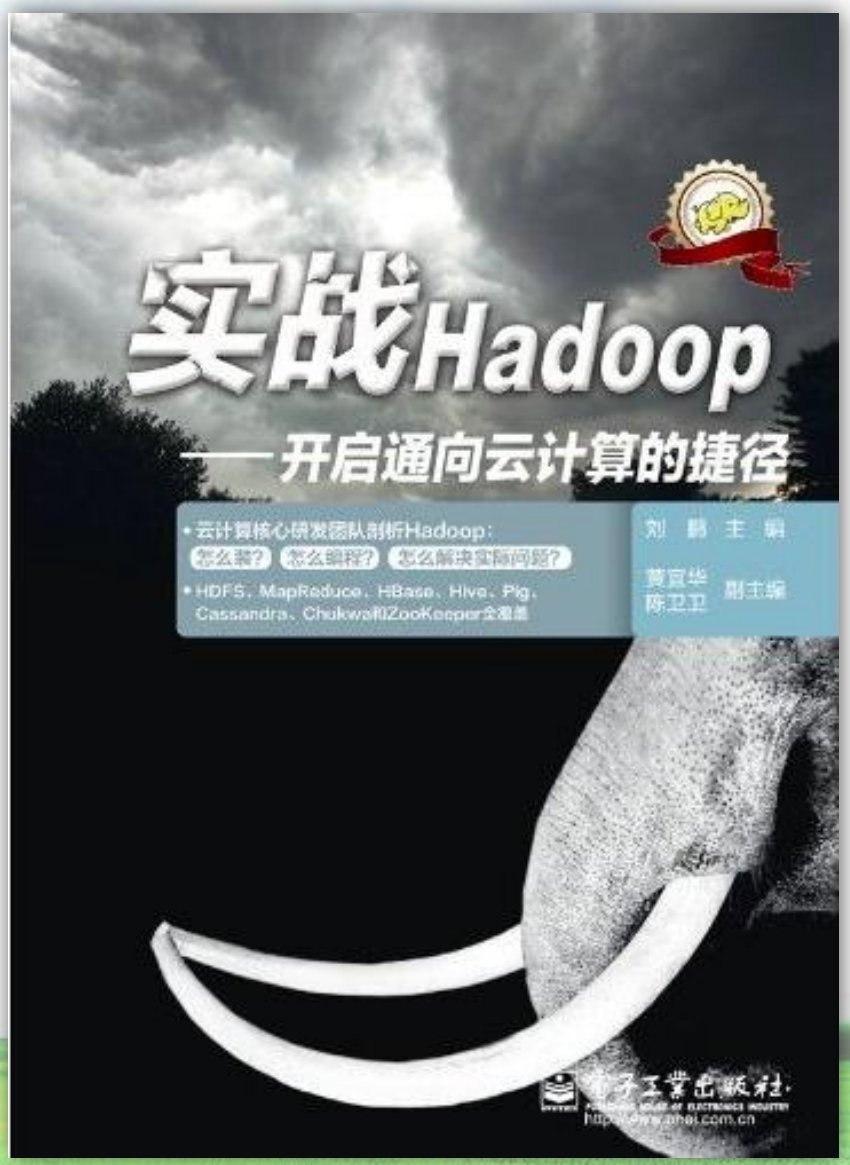
Google技术培训

2009年12月Google在清华大学举办的MapReduce技术培训班

大数据并行处理技术教学

教材出版

2011年7月合著编写《实战Hadoop》，有关Hadoop技术第一本具有原著性质的书籍，456页，9月电子工业出版出版发行。



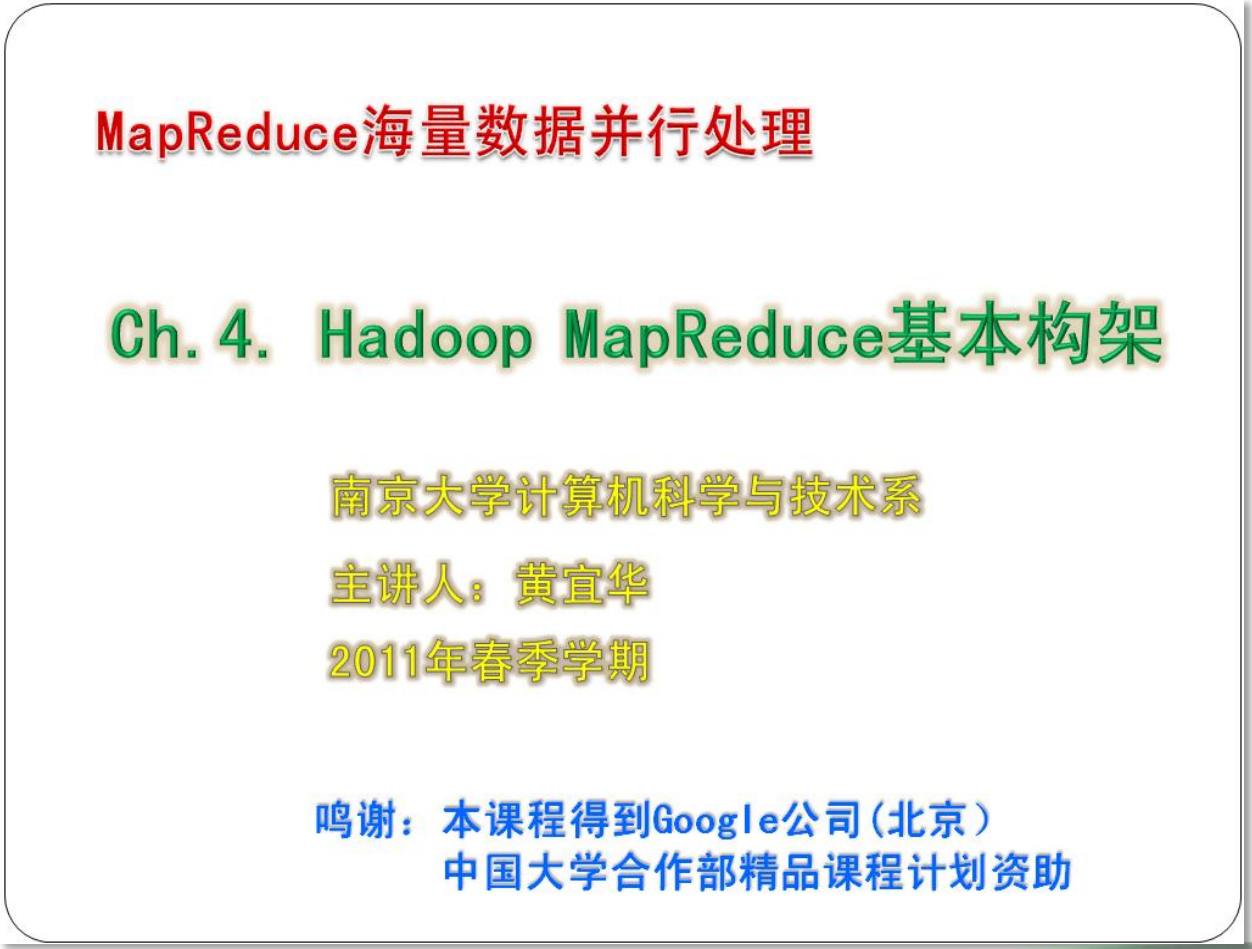
大数据并行处理技术教学

《实战Hadoop》

大数据并行处理技术教学

课程建设和教学

2009年参加了Google公司MapReduce技术培训班，后在Google公司资助下开设了“MapReduce大规模数据并行处理”课程，是目前为止江苏省唯一开设该课程的教师和院系



MapReduce海量数据并行处理
课 程 简 介

鸣谢：本课程得到Google公司（北京）
中国大学合作部精品课程计划资助

南京大学计算机科学与技术系

主讲人：黄宜华

2012年春季学期

课程简介

教学内容简介

本课程将系统介绍目前业界和学术界最新的并行计算和大规模海量数据并行处理技术和方法。课程首先介绍并行计算技术的基本概念、原理、方法和技术，在此基础上，介绍基于集群的大规模海量数据并行处理技术原理和方法，重点介绍MapReduce并行计算集群的构架、用于海量数据存储和计算的分布式文件系统、以及基于MapReduce集群的大规模海量数据并行处理技术和编程方法，MapReduce并行化算法设计技术、并行化算法应用研究案例。

教学目标

课程的主要目标是通过介绍多处理器并行处理技术、以及基于集群的大规模海量数据并行处理技术和MapReduce并行编程模型和方法，要求学生理解和掌握并行处理技术的基本概念、原理和构架、以及基于集群的大规模海量数据并行处理与编程技术方法，并能够用MapReduce实际设计和编写具体的大数据处理应用问题的算法和程序。

选课要求

具有Java程序设计能力，除课堂听课外需要完成编程实验；研究生还要求在学期结束时自选课题完成一个课程设计

课程内容

Ch.1 并行计算技术简介

简要介绍并行计算技术的概况，基本分类，主要技术问题，MPI并行程序设计，大规模并行数据处理技术

Ch.2 MapReduce简介

简要介绍MapReduce技术的由来，基本构思，编程模型，主要设计思想和技术特征，基本应用

Ch.3 Google MapReduce的基本构架

介绍Google MapReduce并行计算框架的基本结构、工作原理，Google分布式文件系统GFS的基本构架与工作原理，Google结构化数据管理系统BigTable的基本结构与工作原理

Ch.4 Hadoop 的基本构架

介绍开源MapReduce系统Hadoop 的基本结构、工作原理，Hadoop分布式文件系统HDFS的基本构架与工作原理，Hadoop数据管理系统的基本结构与工作原理

课程内容

实验1：Hadoop的安装与配置(在个人的电脑上安装一个单机版本）

Ch.5 Hadoop系统安装运行与程序开发

介绍单机和集群Hadoop系统安装方法和步骤，以及程序开发环境与开发过程

实验2：莎士比亚文集词频统计(Word Count)实验

Ch.6 MapReduce算法设计

介绍排序算法、文档倒排索引、文档共现算法、专利文献数据分析应用

实验3：莎士比亚文集倒排序实验

Ch.7 高级MapReduce编程技术

介绍复杂I/O数据表示、用复合键值对完成特殊处理、程序员定制的I/O格式、Partitioner、Combiner，基于迭代的MapReduce求解方法、数据相关MapReduce任务计算、链式MapReduce计算、多数据源连接、访问关系数据库等高级技术

课程内容

Ch.8 基于MapReduce的搜索引擎算法

介绍网页排名算法PageRank，搜索引擎文档倒排索引算法，以及全文检索系统的设计实现

实验4：Wikipedia网页PageRank实验

Ch.9 基于MapReduce的数据挖掘基础算法

介绍机器学习和数据挖掘中的聚类算法、分类算法、频繁项集挖掘等算法的MapReduce并行化设计技术方法

Ch.10 基于MapReduce的并行化算法应用研究案例

介绍基于MapReduce的DNA序列比对算法、重复Web文档检测算法、统计机器翻译算法的并行化设计和实现

Ch.11 云计算技术简介

介绍云计算技术基本概念、发展现状、关键技术与云计算应用

课程设计大作业(研究生)：自选具有一定难度和工作量的题目，鼓励结合导师的研究工作自选课程设计题目，完成课程设计

课程开课情况

- 2011-2012学年分别开设了两个学期，每周2学时，研究生和本科生，本系选修人数200多人
- 安排4次又简到难的课程实验，从Hadoop安装到编程实验
- 要求结合导师研究课题或自行选题完成一个大的课程设计
- 课程结束后，要求研究生结合导师课题选题或自主选题，分小组完成一个具有一定难度的课程项目设计。一共有50多个小组提交了开题报告，最后出现一批相当出色的课程设计项目。

课程开课情况

课程项目设计（开题报告和评审意见）

分组		成员	题目	难度（	工作量	可行性与开题报告质量	开题报告评审意见
st52	MF1033037	殷昆燕	基于Hadoop的影片推荐系统	4.5	4	可行，4.5	选题有较大的技术难度（4.5）和设计实现工作量（4），技术方案可行，达到课程设计要求，同意按开题报告进行，请按开题报告的设计目标认真完成课题。
	MF1033011	金国平					
	MG1033057	余宗桥					
st53	MF0933002	陈光鹏	基于Mapreduce的频繁闭项集挖掘算法研究及其实现	4	3	可行，4	选题"基于Mapreduce的频繁闭项集挖掘算法研究及其实现"有一定的难度(4)和设计实现工作量(3)，技术方案可行，达到课程设计要求，同意按开题报告进行，请按开题报告的设计目标认真完成课题。但开题报告中主要技术难点挖掘和分析不足，请在最终设计报告中加强这部分的讨论。
	MF0933009	黄刚					
st54	MG0933035	王团团	基于Map-Reduce框架的SQL语句解析及执行系统	5+	5+	偏难，4	选题类似于Hadoop的子项目Hive，难度（5+）和工作量（5+）很高，具有很大的技术挑战性，难度和工作量都达到并超过课程设计的要求，同意按开题报告进行。但选题目标过大，可能无法按期完成，且开题报告中对基本解决方法和设计实现思路缺少讨论，可行性分析不足，请在这方面进行一定的讨论和可行性分析，在此基础上确定一个适当难度和工作量、可以如期完成的设计目标，最终按设计目标认真完成课题。
	MG1033080	江凯					
	MG1033088	陆瑶					
	MG1033075	顾小东					
st55	MG1033060	张航	基于MapReduce的本体匹配技术	4	3.5	可行,4	选题具有较好的应用问题背景，有一定的技术难度和工作量，技术方案可行，达到课程设计的要求，同意按开题报告进行，请按开题报告的设计目标认真完成课题。但开题报告对具体的本体匹配技术的并行化处理的难点分析讨论不足，对研究问题以及引入MapReduce并行化处理的必要性缺少清晰的描述和讨论，也缺少参考文献。请在最终设计报告中补充这些方面的内容。
	MG1033052	杨琬琪					
	MF1033023	陶承恺					
st56	MG1033015	李文凯	汽车推荐系统	4	4	基本可行,3.5	选题新颖有趣，具有较好的应用前景。选题具有较大的难度（4）和工作量（4）,达到课程设计的要求，同意按开题报告进行，请按开题报告的设计目标认真完成课题。但开题报告内容较为单薄，报告中对课题的主要技术难点及其并行化处理的难点和必要性缺少足够的分析讨论，尽管原型系统可能数据量不会太大，但作为课程设计，要体现出对大数据量并行处理的特点。在基本技术方案中对具体的汽车评估模型及其并行化处理也缺少基本的讨论和描述。请在最终的设计报告中就以上问题提供足够的内容叙述。
	MF1033014	李若冰					
	MF1033041	赵靓					
st57	mf1033015	刘敏	NBA球员数据分析工具	4	4	基本可行，3	选题新颖有趣，具有一定的潜在应用价值。选题具有较大的难度（4）和工作量（4），达到课程设计的要求，同意按开题报告进行，请按开题报告的设计目标认真完成课题。但开题报告过于简单，报告中对主要技术问题及其难点、课题的可行性等缺少足够的分析讨论；基本技术方案中对数据爬取的MapRedcuce并行化处理设计可行性不足，对基本的球员评估模型缺少基本的讨论和描述。请详细研究以上问题，确定课题的可行性，并在最终的设计报告中就以上问题提供足够的内容叙述。
	mf1033019	鲁林					
	mf1033017	刘振兴					
st59	MF1033001	鲍慧慧	Netflix电影推荐	4.5	3.5	可行,3.5	选题有较大的技术难度（4.5）和一定的设计实现工作量（3.5），达到课程设计要求，同意按开题报告进行，请按开题报告的设计目标认真完成课题。但主要评估算法的有效性和可行性研究和分析不足，请详细研究并确定所选算法的有效性和可行性，如果所选算法不是足够有效，需要考虑更为复杂和有效的算法；另外开题报告内容较为单薄，最终设计报告中请注意保证有足够和完整的内容。
	MF1033009	蒋慧					
	Mf033035	杨丽					
	MF1033033	薛艳					
st60	MF1033002	蔡希辉	？？推荐系统			重新提交	选题没有题目，主要研究内容是什么？问题关键点在哪里？与MapReduce有什么关系？都没有交待。如果是做推荐系统，需要说明具体是哪个应用领域、什么样的具体推荐系统问题。开题报告简单粗糙，内容不清，不符合要求，请重新提交开题报告。
	MF1033036	杨阳					
	MF1033030	徐佳一					
	MG1033102	朱正文					

优秀课程项目设计示例

陈虎，笪庆小组：基于内容的图像搜索引擎EagleEye
—MapReduce海量数据并行处理项目

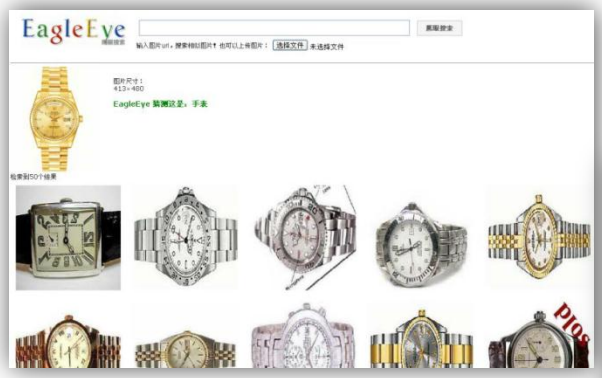
主要研究内容：

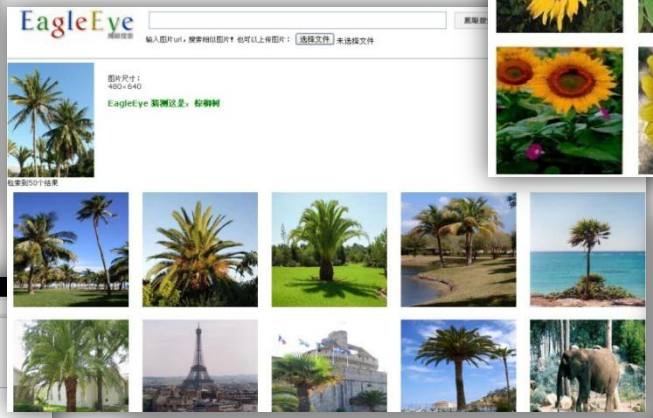
1、研究解决了有效的图像特征表示和快速提取方法：表示和提取图像的特征使其在基于内容的图像检索中能够更准确地表征不同图像之间的相似程度。
2、研究解决了基于MapReduce的海量图像特征索引和图像搜索算法
3、完成了一个基于内容的图像搜索EagleEye原型系统的设计实现

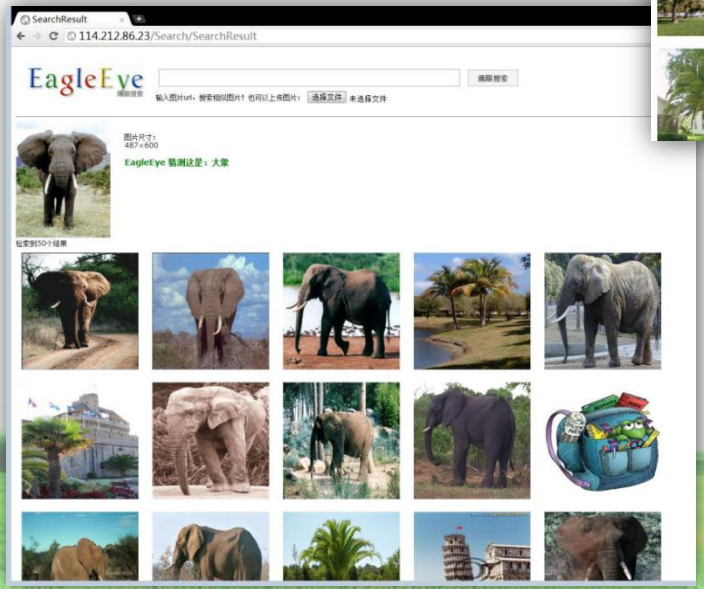
优秀课程项目设计示例

陈虎，笪庆小组：基于内容的图像搜索引擎EagleEye

搜索结果示例





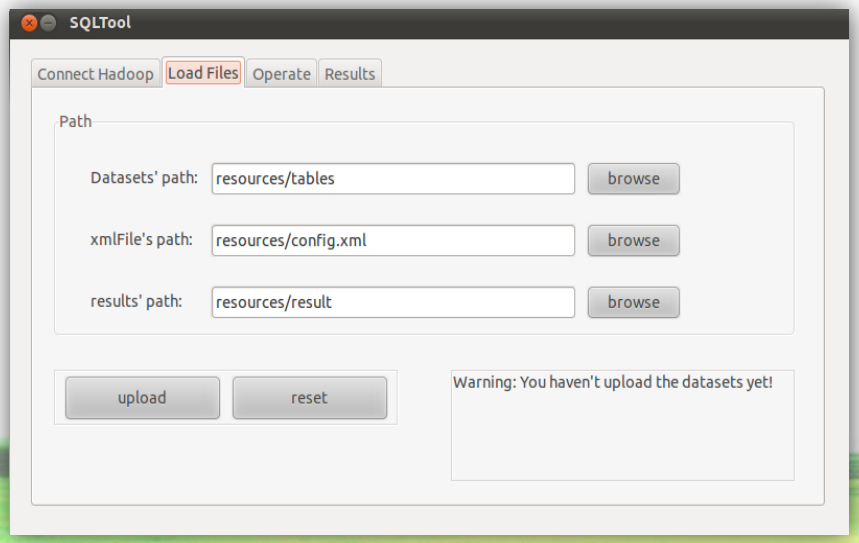


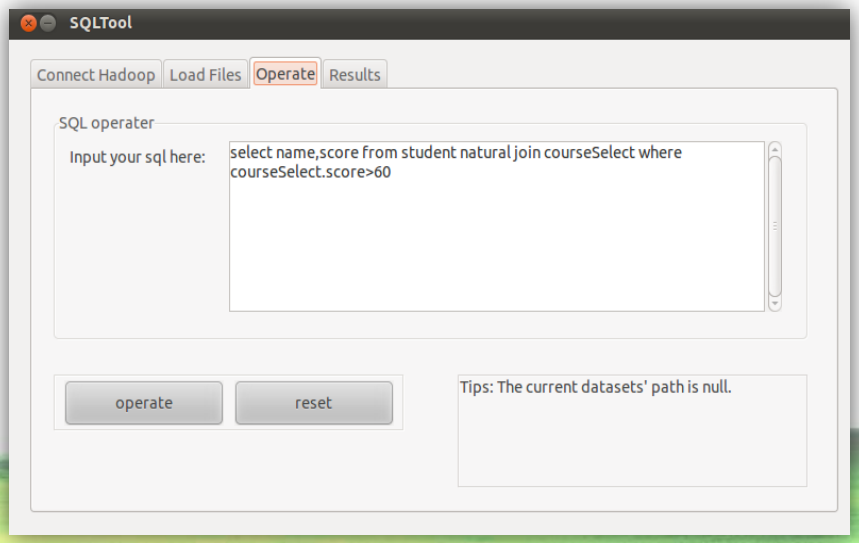
优秀课程项目设计示例

江凯, 顾小东, 陆瑶, 王团团小组：**基于Hadoop的SQL查询工具**

主要研究了在Hadoop分布式文件系统环境下设计和模拟一个管理和查询结构化数据的原型数据库系统，主要技术内容包括：

- 设计了基于XML的数据库Schema的描述和处理方法
- 设计了基本的SQL查询语言
- 完成SQL语句的解析处理
- 完成SQL到关系代数的转换处理
- 基于MapReduce并行计算框架完成关系代数的并行化处理，提高计算效率
- 设计实现了一个原型的查询工具





优秀课程项目设计示例

梁亚澜, 李杰, 钮鑫涛：Hadoop平台下覆盖表生成遗传算法参数配置启发式演化工具

主要研究内容：

1. 采用启发式演化方法对遗传算法的种群规模、进化机制、交叉概率、变异概率及其变种算法5个因素进行取值组合演化，系统地探索各个因素对遗传算法覆盖表生成效果的影响程度和性质，并以覆盖表规模和消耗时间为依据寻找出最佳配置
2. 遗传算法生成覆盖表的计算量极大，设种群规模为100，进化代数为1000，则完整的进化过程需运行遗传算法100*1000 = 100,000次，以一次生成覆盖表的时间为1分钟为例，采用串行计算共需100000分钟，约71天。课题研究实现了基于Hadoop MapReduce的并行化遗传算法生成覆盖表算法, 大大缩短了计算时间

表 3：各待测实例的最终最优配置和覆盖表生成结果

	Algorithm	m	T	Pc	Pm	Size	Time		Algorithm	m	T	Pc	Pm	Size	Time
4^{10}	GAr climb	100	100	0.2	0.2	28	0.234	6^{30}	GA climb	100	1100	0.2	0.2	87	52.6
3^{13}	GAr climb	100	1100	0.8	0.2	17	2.28	10^{11}	GA climb	100	1100	0.8	0.2	154	19.8
6^{10}	GA climb	6100	1100	0.2	0.2	58	402	$7^6 6^7 5^6$	GAr climb	100	1100	0.8	0.2	82	23.5
4^{20}	GAr climb	100	1100	0.8	0.2	35	10.1	$8^2 7^2 6^2 5^2$	GA- climb	2100	600	0.8	0.6	70	277
8^{10}	GA climb	2100	600	0.6	0.2	98	604	$6^1 5^1 4^6 3^8 2^3$	GAr climb	4100	1100	0.8	0.4	36	568.1
3^{20}	GA- climb	100	600	0.2	0.2	21	3.31	6^4	GAr climb	100	100	0.6	0.2	41	0.03
6^{20}	GA climb	100	1100	0.8	0.2	74	22.9	$5^1 3^8 2^2$	GAr climb	100	100	0.8	0.2	20	0.43
4^{30}	GAr climb	100	600	0.2	0.2	40	12.4								

基于本课程设计项目的研究成果作者和导师发表了两篇学术论文

1. 梁亚澜，聂长海，覆盖表生成的遗传算法配置参数优化 2011年6月，计算机学报已录用.
2. Liang Yalan, Changhai Nie, Jonathan M. Kau_man, Gregory M. Kapfhammer, and Hareton Leung. **Empirically identifying the best genetic algorithm for covering array generation**. In Proceedings of the 3rd International Symposium on Search Based Software Engineering, Szeged, Hungary, September 2011

课程项目设计

梁亚澜, 李杰, 钮鑫涛:Hadoop平台下覆盖表生成遗传算法参数配置启发式演化工具

李袁奎, 刘文杰, 王姜: 使用Mapreduce框架进行软件代码分析

软件工程

黄刚, 陈光鹏: 一种基于MapReduce的频繁闭项集挖掘算法研究及其实现

王苏琦, 金龑, 罗爱宝, 王灵江: 基于模型的协同过滤并行化算法

胡昊然，冯子陵，窦文科，刘晶晶：面向新浪微博的关注推荐系统

段轶：Netflix电影数据聚类分析

机器学习
数据挖掘

孙道平：基

刘敏, 刘

刘正, 朱小

王尧, 苏宗

社会网络
分析

化的研究

的分析实验

- 选题覆盖了我系大多数研究方向
- 随着研究问题数据规模越来越大，越来越多的研究领域都需要使用并行计算技术提供新的计算方法
- 本课程的开设对推动我系各方向的研究将起到积极的作用

金惠益, 刘

基于短

机器翻译

式设计

张旭，何良朋：P2P流媒体中的结点分簇与最短路径构造

网络通信

陈虎，笪庆小组：基于内容的图像搜索引擎EagleEye

多媒体检索

张航, 杨琬琪, 陶承恺: 基于MapReduce的本体匹配技术

Web本体

江凯, 顾小东, 陆瑶, 王团团小组：基于Hadoop的SQL查询工具

数据库

第一届“中国云/移动互联网创新大奖赛”

本课程开设后，我系机器学习与数据挖掘研究所和云计算与大数据并行计算课题组学习了MapReduce技术的同学组织了4支研究生代表队在“中国云产业联盟”组织的首届“中国云·移动互联网创新大奖赛”中参赛并荣获9项优胜奖（一等奖2项，二等奖4项，三等奖3项）和4项优秀领队奖，并获得大赛奖金20万元！占据大赛全部30个奖项中的9项，4道大数据赛题全部17个奖项中的8项！

第一届“中国云/移动互联网创新大奖赛”

技术类赛题 1：调色板搜图—在百万图片中搜索与指定调色板相近的图片

技术类赛题 2：多快好省的速递员 — 动态路况环境下的物流规划

技术类赛题 3：你不知道我知道 — 互联网问答系统用户行为分析

技术类赛题 4：难舍难分 — 大规模搜索关键字（短文本）分类

技术类赛题 5：麻雀级云数据中心 — 规定时间内在小规模硬件环境上部署大量虚拟机

创意类竞赛说明： 创意类赛题没有具体的问题约束。

第一届“中国云/移动互联网创新大奖赛”

我系4支研究生代表队荣获9项优胜奖和4项优秀领队，获得奖金20万

第一届“中国云/移动互联网创新大奖赛”

“中国云·移动互联网创新大奖赛”是由“中国云产业联盟”和百度、阿里巴巴、腾讯、用友等国内著名企业和北航、北大等著名高校于2012年5月联合发起组织的第一届全国云计算和互联网创新技术大赛。这次大赛由北航的怀俊鹏院士与中国云产业联盟联合倡议并发起，来自国内多所著名高校和著名企业的十多位专家学者共同参与，是目前为止国内规模和影响最大、级别最高的云计算和互联网创新技术大赛。颁奖仪式上，中国科学院院士怀俊鹏教授、微软集团副总裁陆奇博士到会做了关于云计算的主题报告，百度总裁李彦宏、宽带资本董事长田溯宁、用友软件董事长兼总裁王文京、中国联通总裁陆益民等嘉宾也到会并做了云计算主题对话

联系信息

单位：南京大学计算机科学与技术系
姓名：黄宜华
Email：yhuang@nju.edu.cn
电话：189-5167-9127

谢 谢！