

大数据系统引擎技术简介

高光荣

ACM Fellow、IEEE Fellow

中国计算机学会 (CCF) 海外杰出贡献奖获得者

特拉华大学终身教授

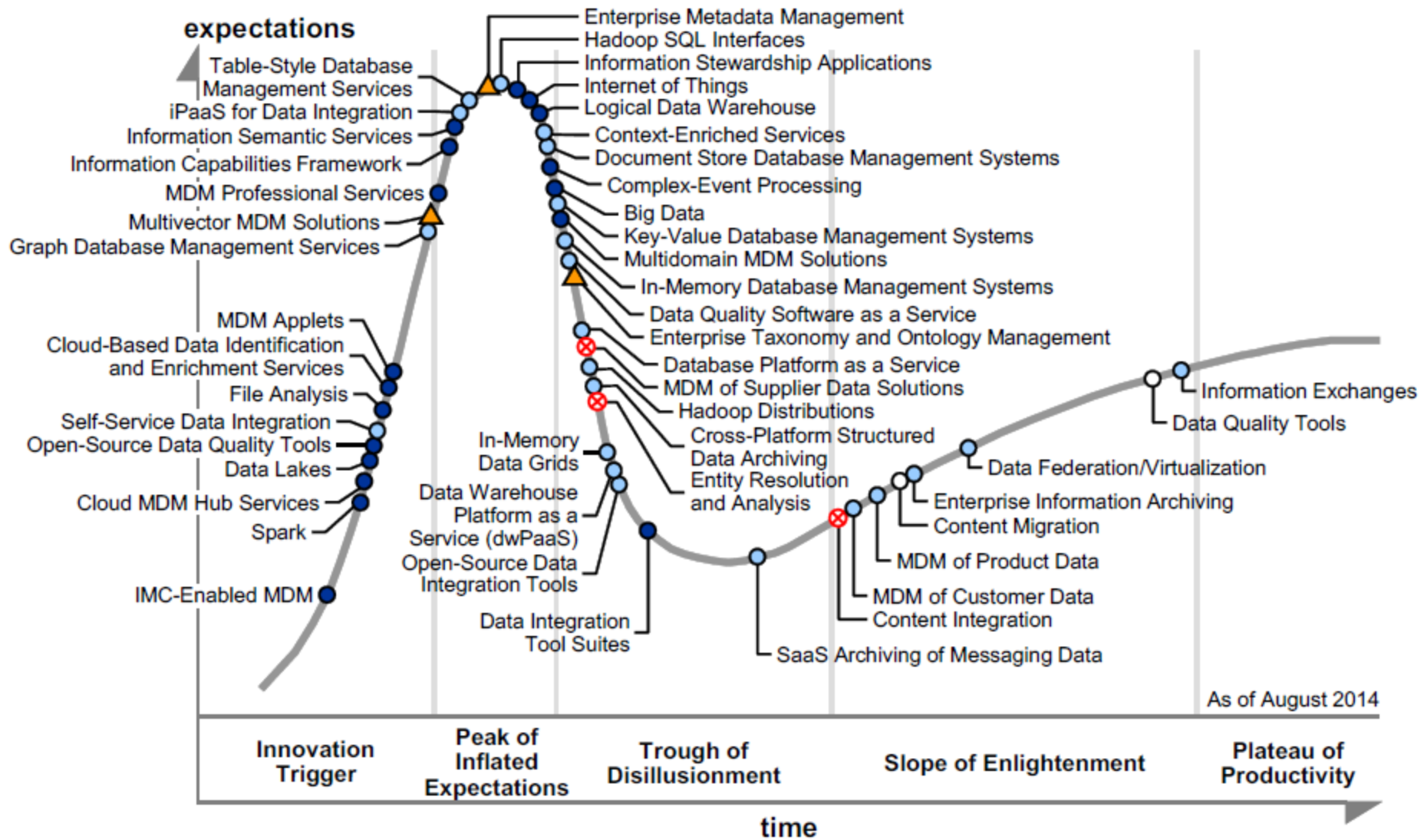
CAPSL 实验室主任

ET International 公司 创始人

Outline

- **Motivation:**
 - 大数据Hype Cycle的最新预测
 - Symbiotic trend between big data and big compute
- Background:大数据系统面临的严重挑战
- 大数据系统核心技术简介
- 数据流与大数据引擎的创新
- 大数据系统发展在中国的机遇与挑战
- 总结

Gartner Report



Plateau will be reached in:

- less than 2 years
- 2 to 5 years
- 5 to 10 years
- ▲ more than 10 years
- ⊗ obsolete before plateau

Hype Cycle and Big Data

- This Hype Cycle *sits mainly on the **Peak of Inflated Expectations***. And big-data has just passed the peak.
- We encourage ***department-level experimentation without enterprise commitment*** over the next three to five years.

Challenges and Opportunities

Symbiotic of Bigdata and Big compute

- **Symbiotic HPC computing and data-Intensive processing**
- Heterogeneity: including sensors, controllers, mobile devices, etc.
- Massive opportunity of concurrency
- Vast dynamic and distributed environment
- Asynchronous stream processing
- Realtime *continuous* interaction with environment
- Energy efficiency, resiliency and security challenges

[Supercomputing 2014 特邀论坛, Gao, 11/19/2014]

Outline

- **Motivation:**
 - 大数据Hype Cycle的最新预测
 - Symbiotic trend between big data and big compute
- **大数据系统引擎面临挑战**
- 大数据系统核心技术简介
- 数据流与大数据引擎的创新
- 大数据系统发展在中国的机遇与挑战
- 总结

新的需求需要新的技术

多

基础需求

- 能从大数据中挖掘出有价值的信息

快

性能需求

- 不仅关注要如何挖掘数据，更关注这个过程有多快

好

功能需求

- 对实时数据的复杂分析正成为最普遍的需求

省

成本需求

- 大数据技术需要的大量的处理资源对成本造成压力

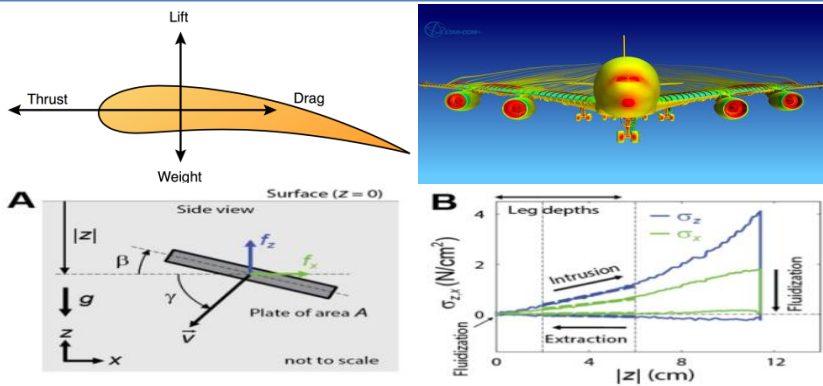
进一步提高性能，降低成本，让“大象”飞起来

让大象飞起来的关键

适合飞行的引擎



新的模型：从地面动力学模型到空气动力学模型



新的结构：从腿到翅膀、热气球、滑翔伞？

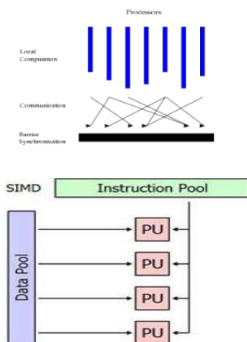


Outline

- **Motivation:**
 - 大数据Hype Cycle的最新预测
 - Symbiotic trend between big data and big compute
- **Background:大数据系统面临的严重挑战**
- **大数据系统核心技术简介**
- **数据流与大数据引擎的创新**
- **大数据系统发展在中国的机遇与挑战**
- **总结**

大数据引擎的核心技术

大数据引擎



大数据引擎**执行**
模型和结构技术



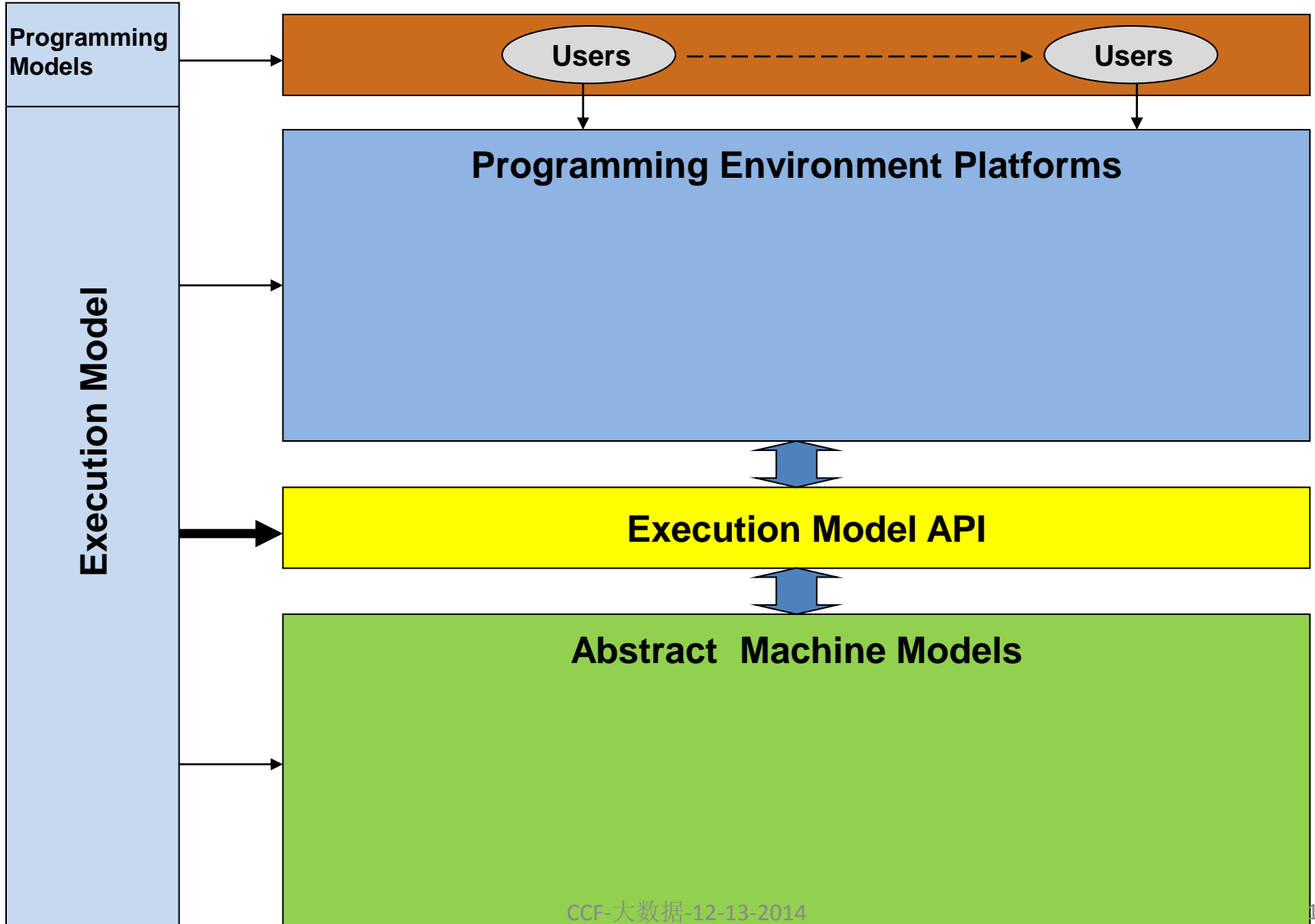
大数据引擎**系统**
软件技术

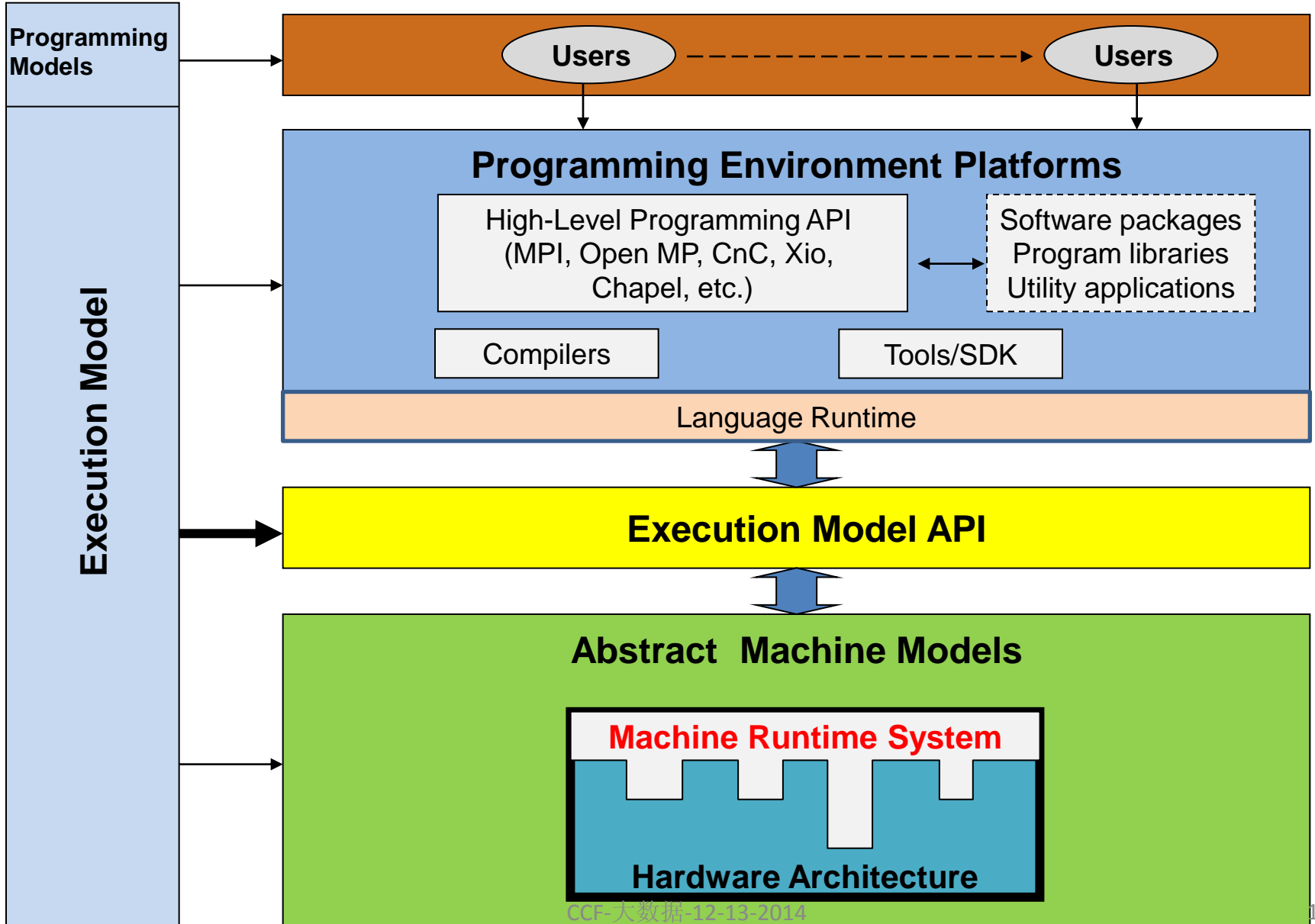


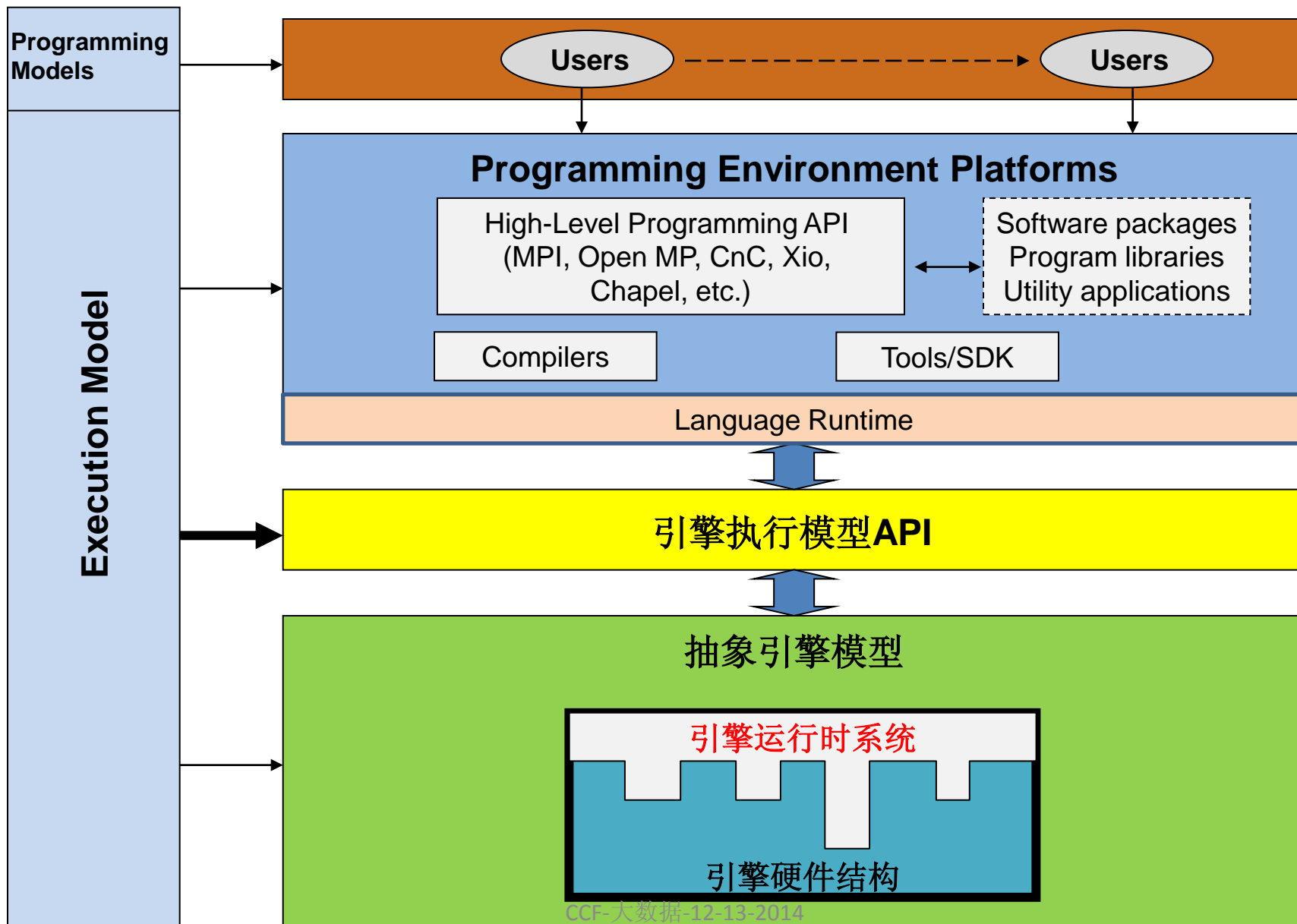
大数据引擎**编程**
模型和优化技术

Terminology Clarification

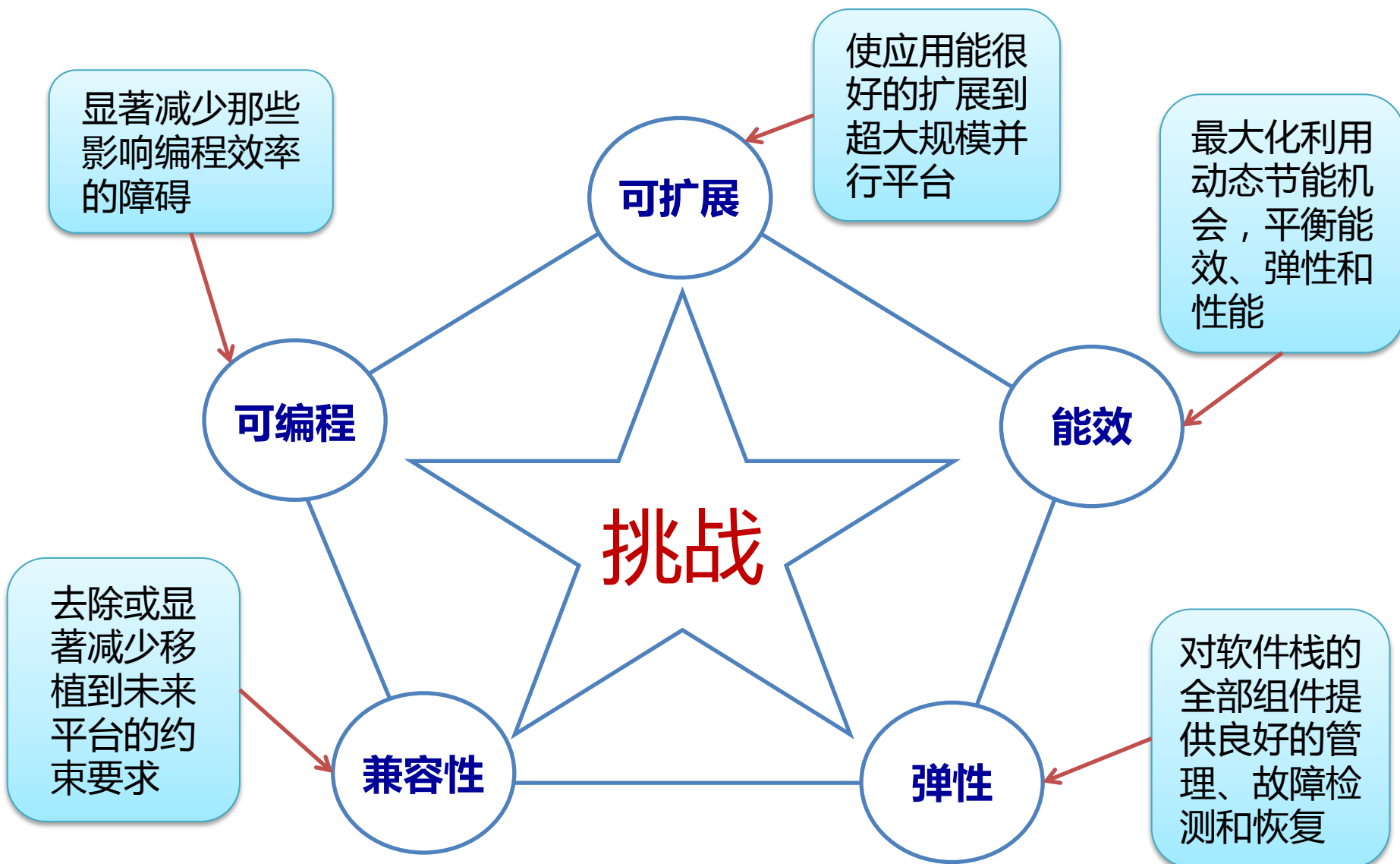
- Parallel Model of Computation
 - Parallel Models for Algorithm Designers
 - Parallel Models for System Designers
 - Parallel Programming Models
 - **Parallel Execution Models**
 - Parallel Architecture Models







并行执行模型及结构技术—挑战



系统软件—多核时代的挑战



Intel 8核处理器



AMD 12核处理器

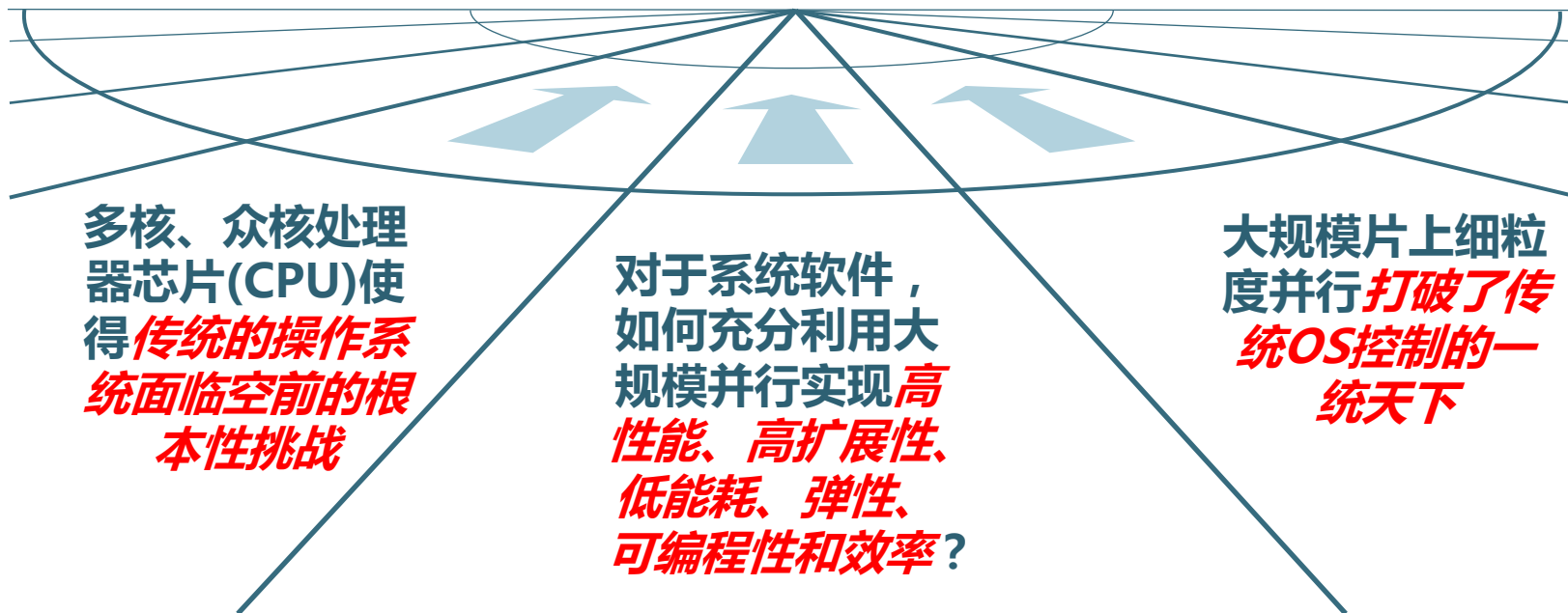


英伟达448核
众核处理器



由多核众核处理器
组成的多核、众核系统

多核、众核系统引发的挑战



系统软件—美国和西方技术走势

避免研发误区。

以实用性为前提，全面考虑高性能、高扩展性、低能耗、弹性、可编程性和效率的需求。

面向应用在实用案例中摸索前进。

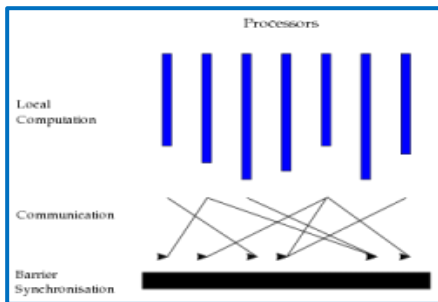
系统软件在
美国和西方的
最新技术走势

对于并行多核系统软件构思，**必须打破操作系统一统天下的格局!**

执行模型及结构技术的创新：**运行时系统软件独立技术和学科的兴起**和迅速发展。

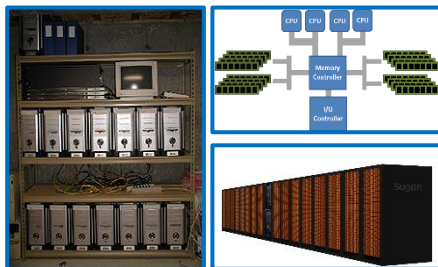
动态细粒度执行模型、结构的运行时系统技术的研发。

并行编程模型和优化技术—概述



BSP执行模型

计算机系统结构



当前优化技术集中在静态优化方法



~~目标机器动态调度~~

~~自适应并发和资源管理~~

~~能耗自感知和弹性控制~~

并行编程模型和优化技术—挑战

多核时代为并行编程模型和优化技术带来巨大挑战

多核芯片对于片上并发和其它资源管理是**空前的机会**

静态的并行编程模型和优化技术**不能适应**多核时代大规模并发资源的编程和管理

并行编程模型和优化技术如何**动态**利用这样的机会？

并行编程模型和优化技术—最近走势

并行编程模型和优化技术的创新，正在**与运行时系统软件技术密切交叉前进**

编程模型和优化技术主要**针对动态细粒度执行模型及结构**

细粒度并行编程模型

与系统软件技术交叉前进

避免研发误区

面向应用在实用案例中摸索前进

避免盲目跟进，以实际应用为导向，避免研发误区

解决**重大实际应用问题**的能力

综合考虑**多方面的性能需求**

以实用性为前提考虑高性能、高扩展性、低能耗、弹性、可编程性和效率的需求

我们的有关工作举例

一、核心技术基地

建成**数据流为背景动态细粒度**
多线程引擎核心技术的基地
(1996-2010)

二、巨型计算机

承担全套以**细粒度多线程系统**
软件总体设计和工程实现- 成
功用于世界领先采用众核芯片
技术的巨型计算机 (ETI获投资
总额超过3千万USD, 2004-
2011)。

我们的工作

四、研发超并行引擎

承担以**数据流为基础的**
runtime系统软件重大研发课
题: (DART/Dynax/SWARM,
2013-2015)

三、研发超并行执行模型

承担超大型以**数据流为背景的**
引擎执行模型 (**codelets**) 的
重大研发课题 (2010-2015)

Outline

- **Motivation:**
 - 大数据Hype Cycle的最新预测
 - **Symbiotic trend between big data and big compute**
- **Background:大数据系统面临的严重挑战**
- **大数据系统核心技术简介**
- **数据流与大数据引擎的创新**
- **大数据系统发展在中国的机遇与挑战**
- **总结**

Inspiration: Jack Dennis

General purpose parallel machines based on a dataflow graph model of computation

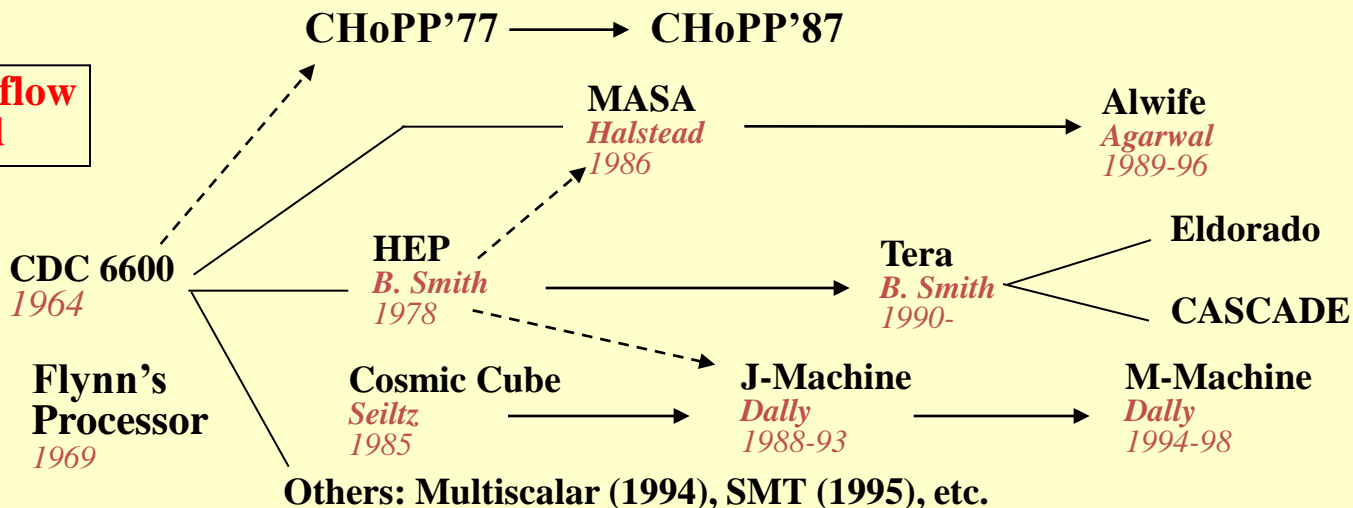


Inspired all the major players in dataflow during seventies and eighties, including Kim Gostelow and I @ UC Irvine

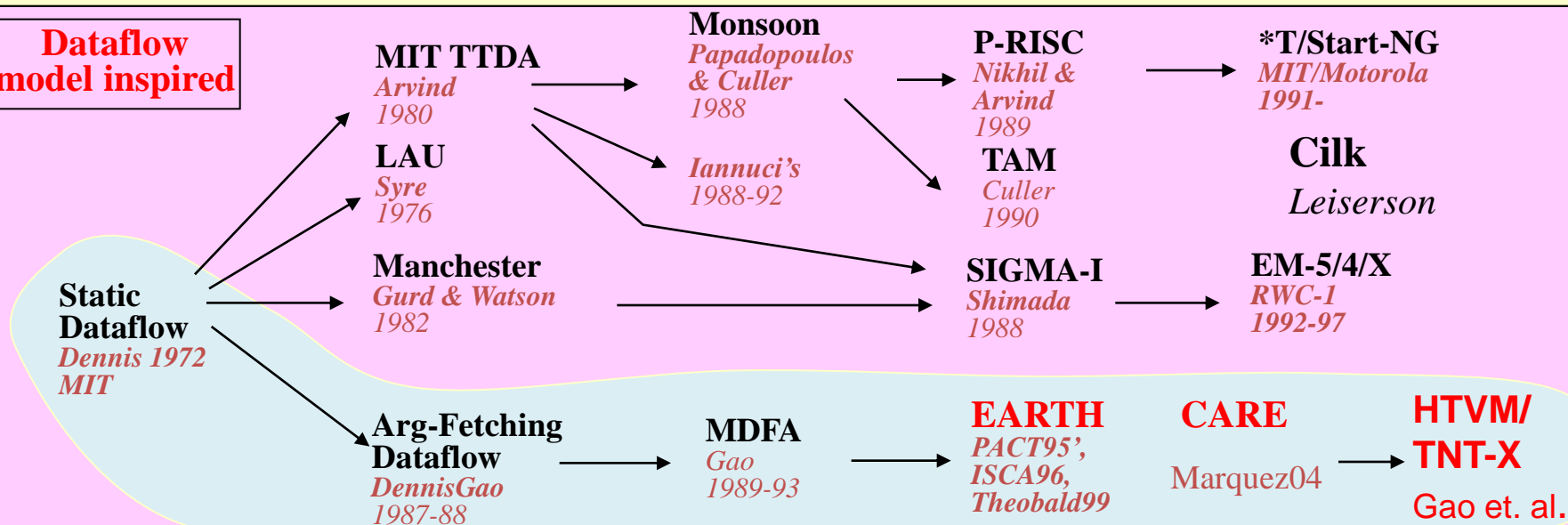
[By Arvind: ISCA 2006 Keynote]

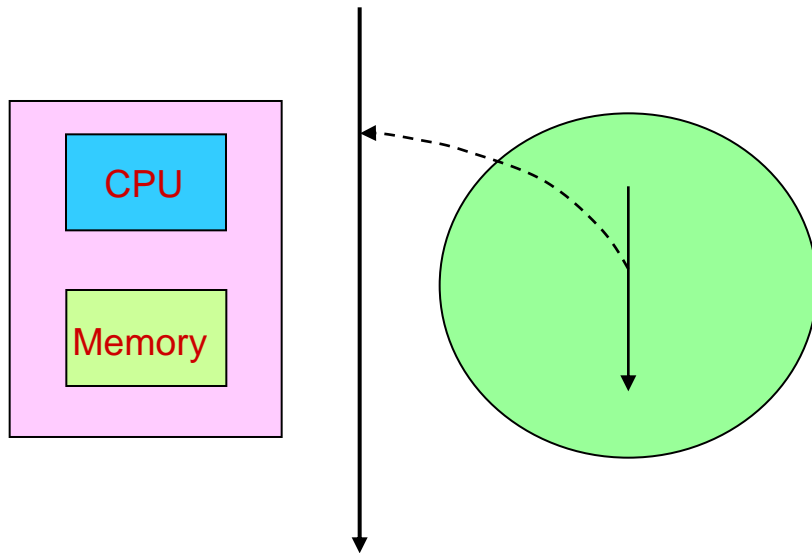
Evolution of Multithreaded Execution and Architecture Models

Non-dataflow based



Dataflow model inspired



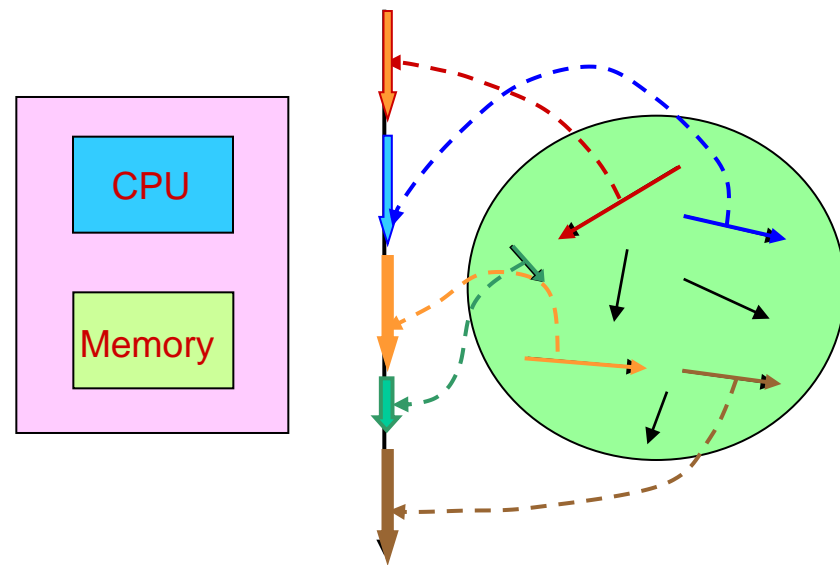


Thread Unit

Executor Locus

A Single Thread

**Coarse-Grain thread-
The family home model**



Thread Unit

Executor Locus

A Pool Thread

**Fine-Grain *non-preemptive* thread-
The “hotel” model**

Coarse-Grain vs. Fine-Grain Multithreading

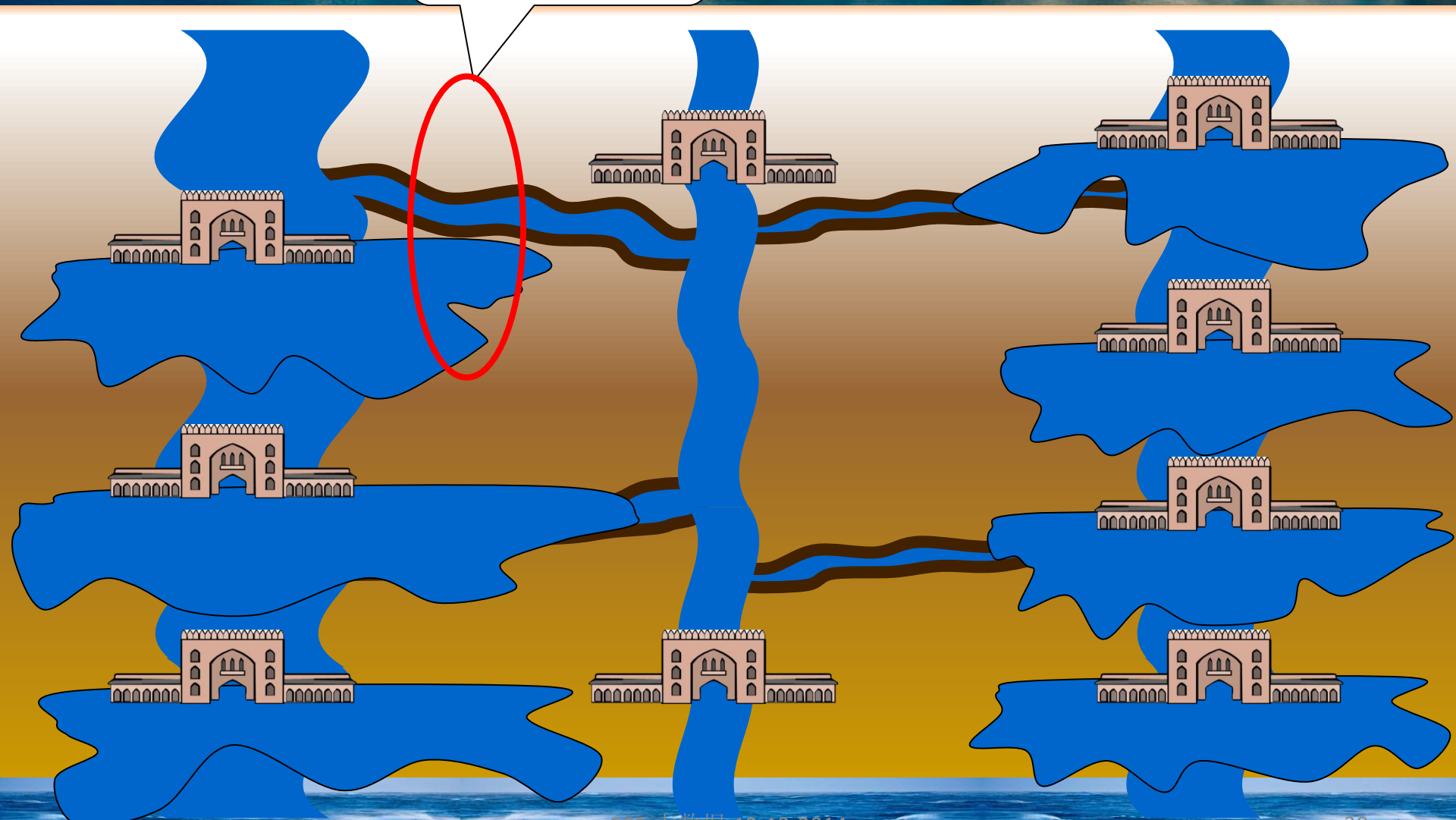
[Gao: invited talk at Fran Allen’s Retirement Workshop, 07/2002]

以堵为主？还是以疏为主？ 数据流关键技术之一





避开洪水的同时
保持数据流畅



数据流关键技术之二—流水线并行

大数据的初级解决思路



每次将一杯清水



每次倒出一杯

这对有处理海量数据需求的用户而言，这的确是雪中送炭；但这种方法难道就是十全十美的吗???

每次调制一杯糖水



流水线的方式调制糖水提供了一种更好的大数据处理解决方案，虽然对“调制糖水”的工艺要求更高、更精确但是所带来的性能提升亦令人满意。

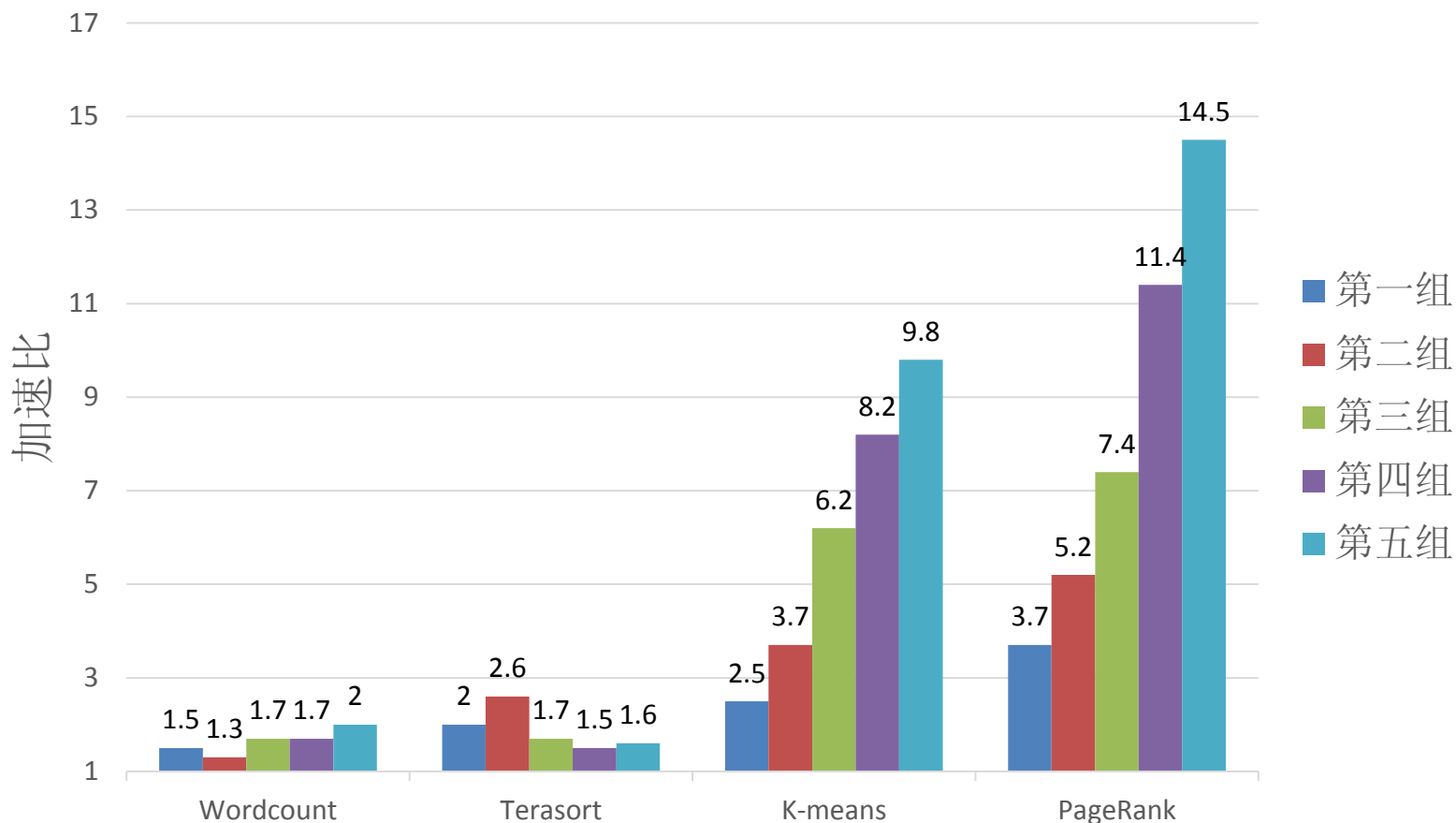
数据流的解决思路

在烧杯中不断地
调制糖水

数据流-大数据技术优势的初步例证

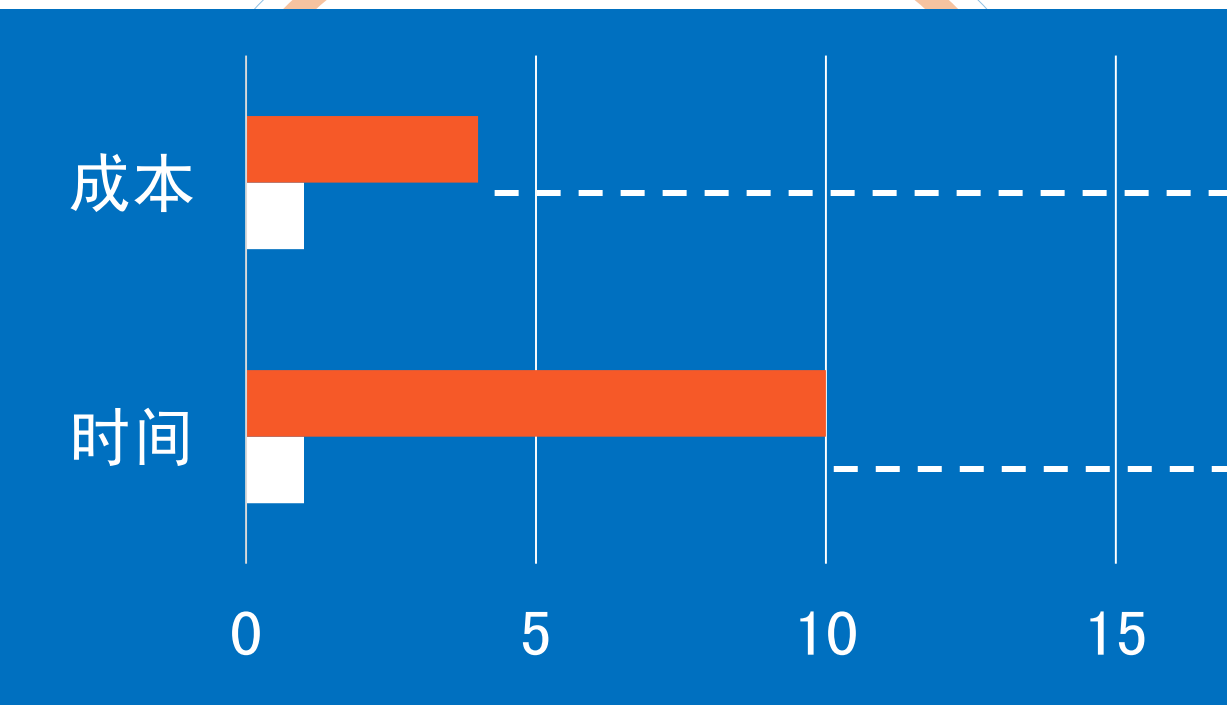
HT vs. Spark的加速比

加速比统计结果



图注：在各个测试用例中，测试数据量大小由第一组至第五组逐渐增大

HT vs. MapReduce



比MapReduce

成本下降**1/4**

比MapReduce

提速**十倍**

Outline

- **Motivation:**
 - 大数据Hype Cycle的最新预测
 - **Symbiotic trend between big data and big compute**
- **Background:大数据系统面临的严重挑战**
- 大数据系统核心技术简介
- 数据流与大数据引擎的创新
- **大数据系统发展在中国的机遇与挑战**
- 总结

数据流大数据技术在中国落地和起飞的可行方案

- **他山之石可以攻玉**

引进国外最先进的数据流大数据引擎技术，实现高起点的大数据产业，“消化吸收再创新”，避免陈旧技术的革新负担。

- **跨越式发展**

从 *中国制造到中国创造*，再到中国标准，建立新的大数据技术标准，实现跨越式发展，最终技术上反超国外。

大数据在中国的产业应用
正化蛹成蝶，呼之欲出！

BIG
DATA