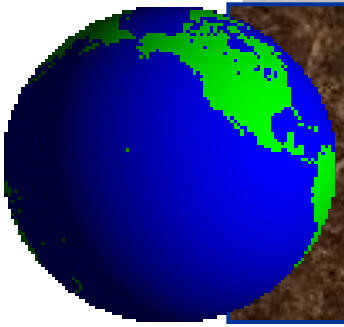
INSTITUTE OF COMPUTING TECHNOLOGY

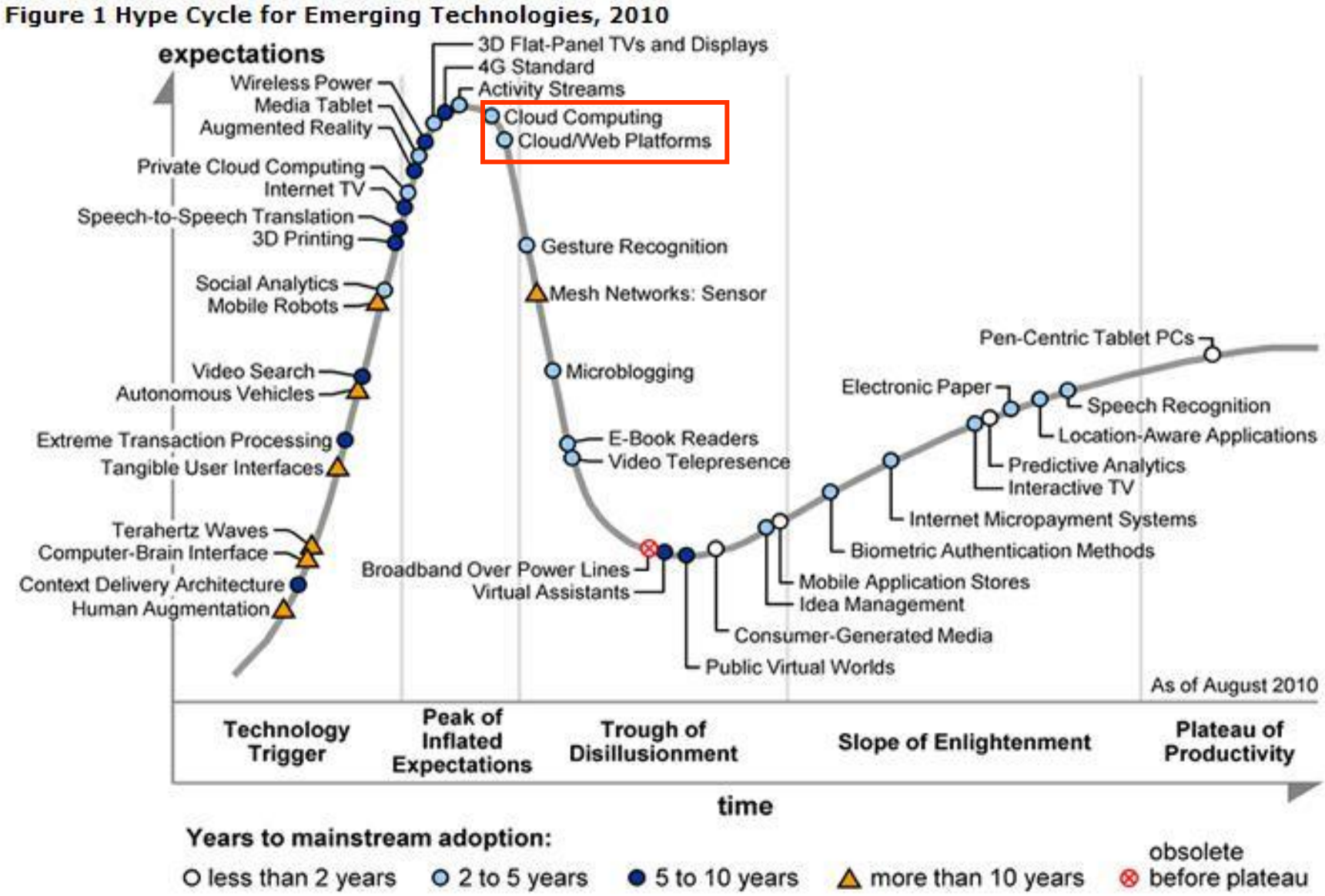
发展云计算应重视的几个问题



李国杰
中国科学院计算技术研究所
2011.04.28，成都

我国已进入云计算建设高潮

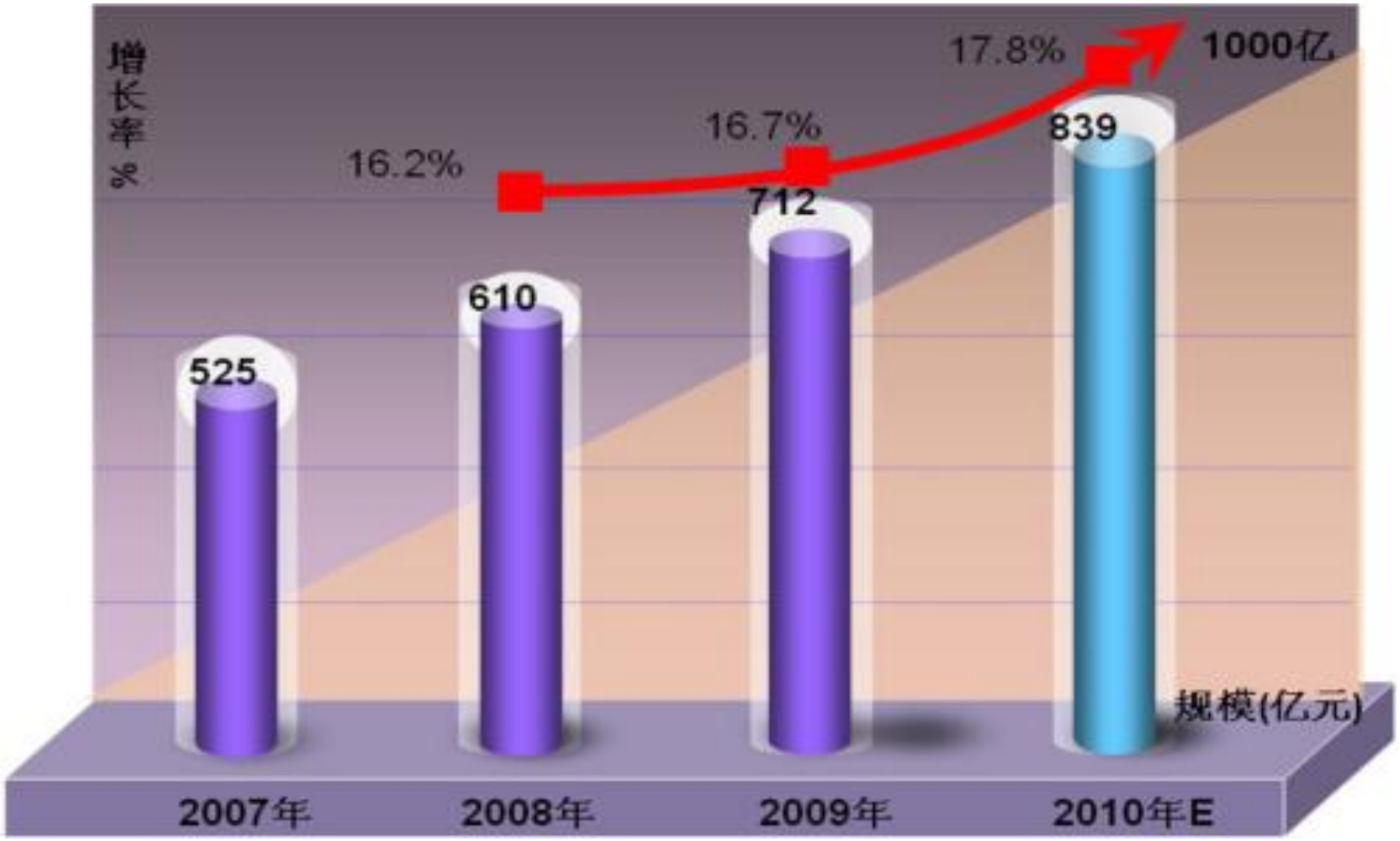
Gartner's 2010年 Hype Cycle



云计算的市场很大

- IDC预测，未来4年全球云计算服务市场平均每年将增长26%。目前全球云计算服务市场规模为172亿美元，2013年将增长至442亿美元。存储将是增长最快的云计算服务。
- 赛迪顾问发布的《中国云计算产业发展白皮书》显示，2010年中国云计算市场规模达到167.31亿元，比2009年的92.23亿元同比增长81.4%，预计到2012年将突破600亿元。
- 有预计称，2015年全国“云计算”产业链规模可能达到7500亿至一万亿人民币,有望占到2015年战略性新兴产业15%以上的产值规模
- 目前全球有1000多万台数据中心的服务器还不是以云的方式工作，还有3000多万台服务器不在托管中心，发展空间很大。

国内数据中心市场

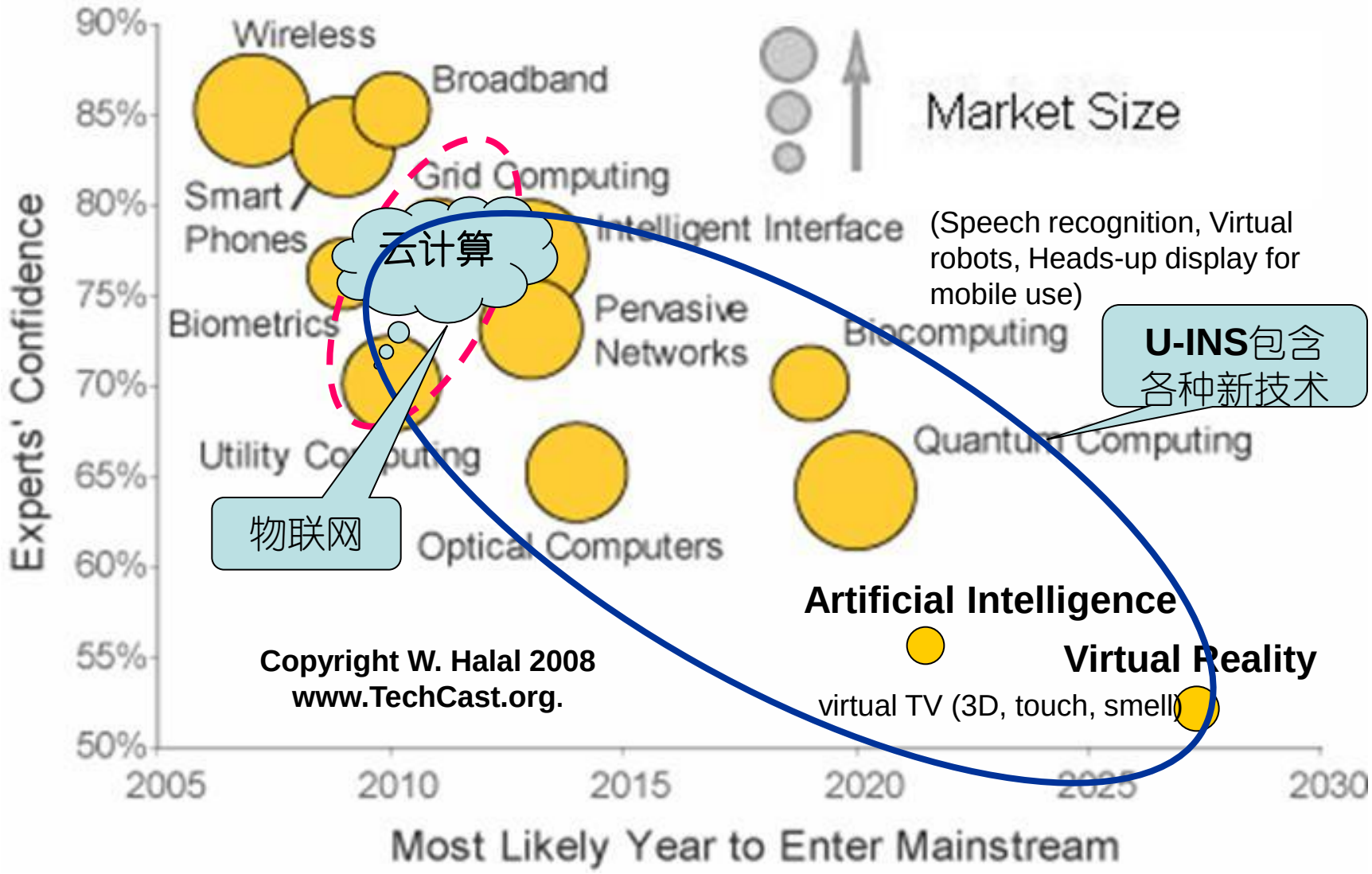

增长率%
16.2%
16.7%
17.8%
1000亿
525
610
712
839
规模(亿元)
2007年
2008年
2009年
2010年E

全国各地都在建云计算中心

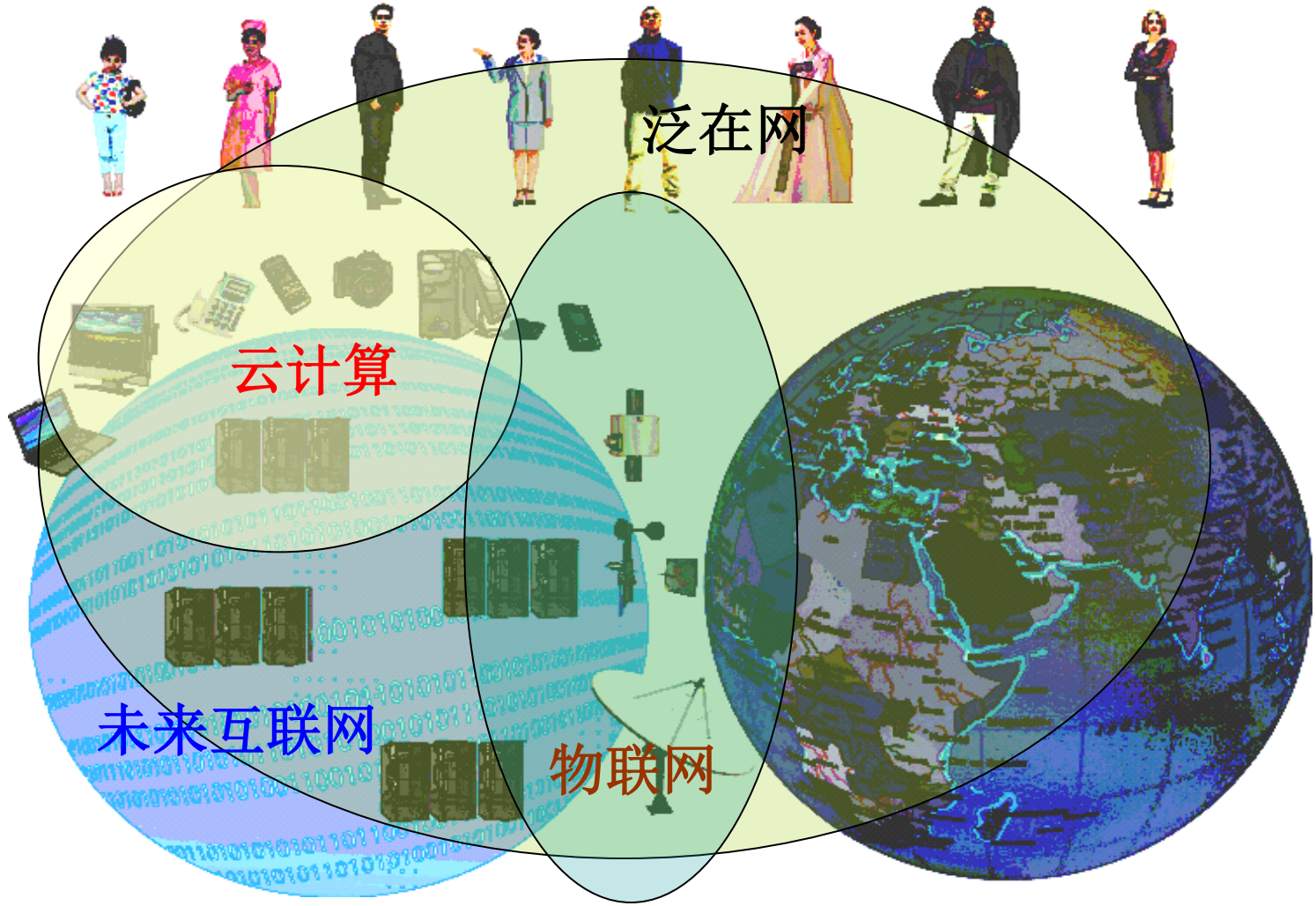
- 工信部联合发改委于10月18日联合印发《关于做好云计算服务创新发展试点示范工作的通知》，确定在北京、上海、深圳、杭州、无锡等五个城市先行开展云计算服务创新发展试点示范工作。
- 重庆计划建设国内最大的国际云计算中心，总投资４００亿元人民币，总建筑面积２０７万平方米，其中包括占地约为３平方公里的数据机房，云集上百万台服务器。
- 华为已正式发布了云计算战略及端到端的解决方案，放言“我们在云平台上要在不太长的时间里赶上、超越思科，在云业务上我们要追赶谷歌。”
- 曙光公司和中科院计算所已在成都、东莞等地实际运作云计算中心，探索符合国情的云计算发展道路。

云计算是信息技术发展的一个必经阶段

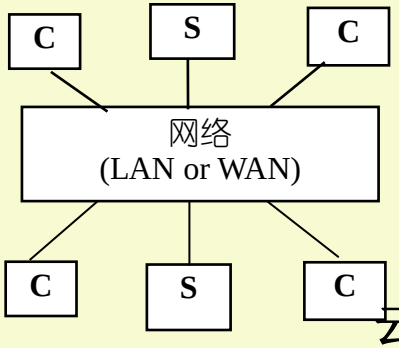
一种信息技术预测


90%
85%
80%
75%
70%
65%
60%
55%
50%
Experts' Confidence
2005
2010
2015
2020
2025
2030
Most Likely Year to Enter Mainstream
Wireless
Broadband
Smart Phones
Grid Computing
云计算
Intelligent Interface
Biometrics
Pervasive Networks
Biocomputing
Utility Computing
Optical Computers
Quantum Computing
Market Size
(Speech recognition, Virtual robots, Heads-up display for mobile use)
U-INS包含各种新技术
物联网
Artificial Intelligence
Virtual Reality
virtual TV (3D, touch, smell)
Copyright W. Halal 2008
www.TechCast.org.

未来互联网、物联网、泛在网和云计算



分（Decentralize）



如企业内部局域
网信息系统

U-INS
U社会

云计算是网络计算的新形式
“云计算是软件大型机”

云计算体现信息产业的“三国定律”：IT平台20年左右的集中-分散周期

Terminal Terminal

大型机
Mainframe

如IBM大型机

服务器	服务器
服务器	服务器

如银行业大集中
各种网站系统

天下大势，
分久必合，
合久必分。

大型机－终端 1960
客户机－服务器 1975
服务器聚集 1990
云计算 2005
普惠泛在信息网络 2020
Time

合（Consolidate）

云计算的出现具有历史必然性

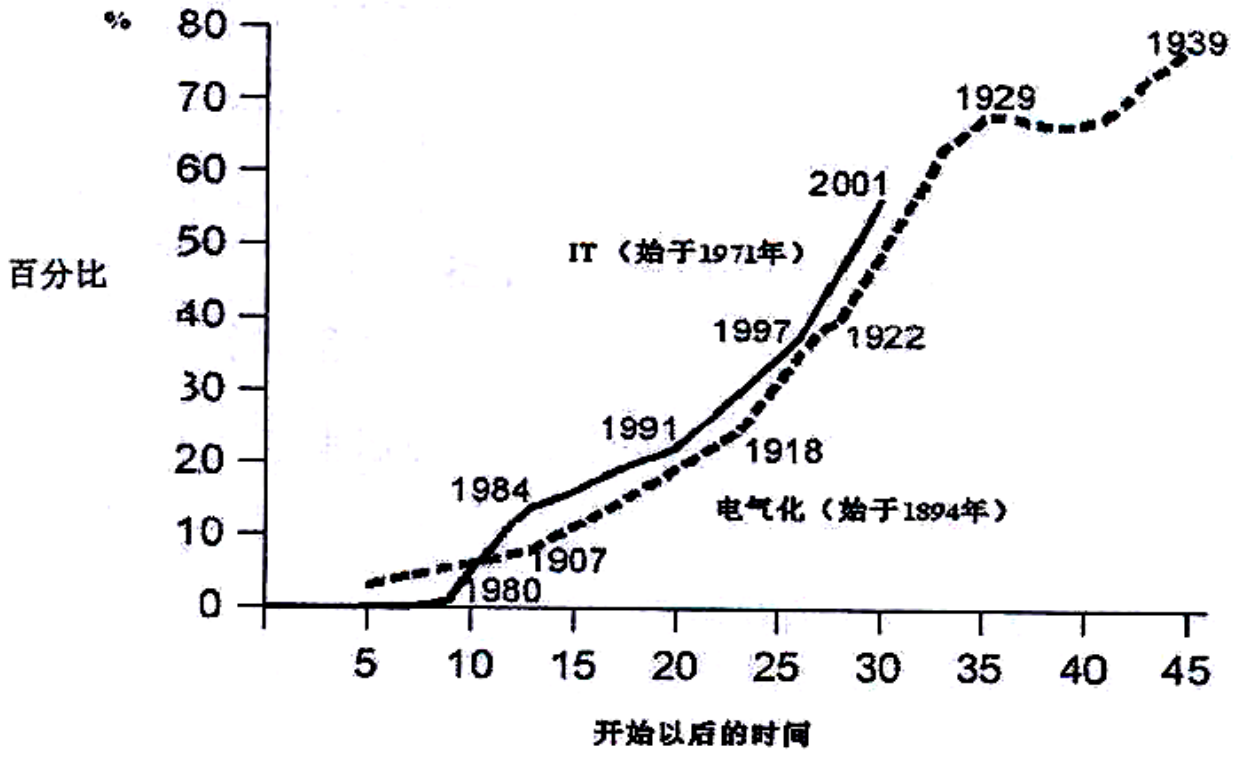
- 云计算符合“三国定律”的宏观规律，其出现有必然性。
- 云计算是信息技术发展的一个必经阶段，主要体现为资源集中和虚拟化动态优化管理，因此，有专家称“云计算是软件大型机”。
- 按螺旋形发展的“三国定律”，20年后，信息网络可能又有一轮从集中到分散的转折，强调资源的分布性和用户对网络的贡献。

- 信息技术发展已有几十年的积累，现在已进入技术发展的平台期。云计算是互联网技术发展的自然演进，体现出平台期的技术发展特点。今后10-15年互联网将有变革性的大突破（后IP时代），云计算也将会有根本性的技术突破。

一百年前电气化的主要启示

电脑普及与电气普及速度差不多

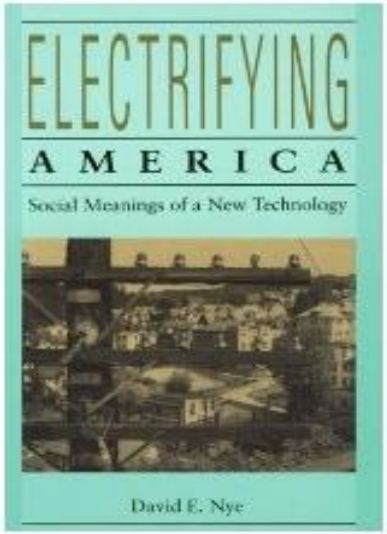
美国拥有电气和电脑的家庭比例:

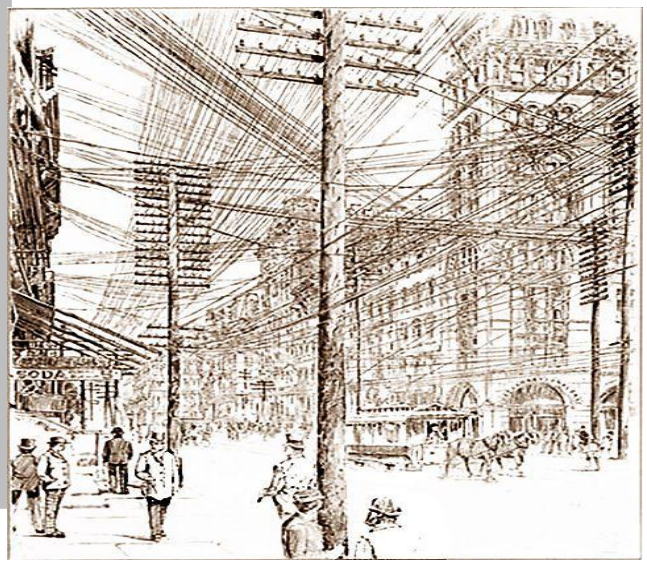


资料来源: Jovanovic and Rousseau (2003), 通用科技

美国电气化过程的启示

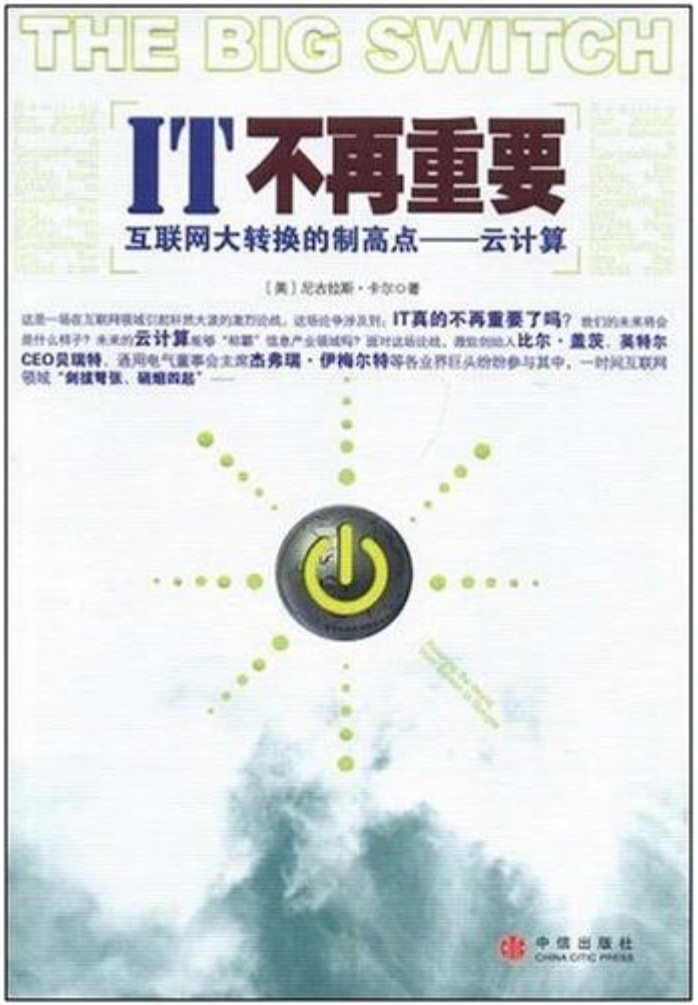
- David E. Nye，《Electrifying America: Social Meanings of a New Technology》，《美国电气化》，MIT出版社，1992年
- 在1880-1900年期间，美国和英国只有小电站，每个工厂、每条电车道都有自己的发电设备。银行和股市支持私人发展电力。
- 上世纪初，伦敦的电力有10种不同的频率、32种不同的电压、70种不同的电价。
- 为了实现电力系统的融合，美国规定地方政府可控制的地区之允许用公共电力，私人电力公司可在城市之间发展
- 近几年国外又在探讨分布式的热电联产的绿色智能电网系统。第二代能源系统成为21世纪能源工业结构调整的方向之一。





从电网普及过程看今天的互联网

- 对云计算的技术转换意义讲述最明白的书是Nicholas Carr写的“The Big Switch”,国内翻译成“IT不再重要”。
- 如同小型（直流）电站必然转变为大型（交流）电站一样，个人电脑必将让位于公共运算时代。
- 从100年前电气化中可以学到许多推广公用技术的历史经验和教训。
- 但云计算与电力网的不同，电力网只送能量不传送应用，所有的应用都是客户的责任；而信息服务应用可以通过网络传送。
- 电脑运算比发电更具模块性，数据存储、处理、传送可分拆成不同的服务。由不同公司提供，减少供应方的垄断。



上世纪初实现电气化的历史

- 在电气化的早期，各工厂为了用上电，都是自己安装发电机，恰如今天为了用上电脑，各自建立自己的信息系统一样。
- 爱迪生对经营中央电厂不感兴趣。他的商业兴趣在于通过特许经销或专利权许可方式，以卖设备和零件来赚钱。
- 将电力公用事业完成的人是另一个有前瞻眼光的英国速记员塞缪尔•英萨尔(Samuel Insull)，他最大的挑战是要让各工业企业相信，它们应该停止自己生产电力，而从中央电厂将电作为一种服务买回来。
- 公用电网在美国供电中占的比重，1907年为40%，1920年达到70%。1930年上升到80%，后来很快达到了90%以上。
- 使大型电力公司成为可能的，是发电、输电方面的一系列科学和工程上的突破，但确保其胜出的并不是技术，而是经济的发展。

从电气化历史中可学到什么？

- 实现大电站集中供电依靠几项关键技术。第1项技术是旋转变流器，它将不同电厂发出的不同电压、频率和相位的电流（包括直流、交流）组合为一种集中管理的单一系统。（在云计算中这种变流器是什么？）
- 第2项技术是负荷表。传统的电表只测量“负荷”（实际消费的千瓦数），而负荷表还能测出该客户的“负荷系数”（实际消费的瓦数在潜在使用峰值中的百分比）英萨尔在这方面的成就可与“铁路员工19世纪在管理方面做的历史性贡献相媲美”（云计算的计费系统？）
- 先落实“牵引公司”，即当时芝加哥用电大户的有轨电车公司和高架铁路公司。“牵引公司”会在早晚高峰时间消耗大量电力，与家庭用户和办公室用户互为补充。（首批用户？）
- 在市中心建立了一座“供电商店”，里面放了各种靠电马达驱动的机器作为促销展品。（“软件商店”）。

重视云计算中的软件技术

核心技术的变迁：从Inside 到Outside

- 长期以来我们认为核心技术是芯片和操作系统，用户很关注”Intel inside“，但近两年风向变了，用户买苹果公司的iPad 和iPhone, 主要关心Outside (用户体验，产业生态环境）而不是Inside(用谁的芯片）。
- 构建高效灵活的产业生态环境已成为云计算的核心技术。许多人常用“新瓶装旧酒”来嘲笑云计算，这提醒我们重视酿造“新酒”（新的Inside), 抵制忽悠和炒作。但不要忽视瓶的改变（outside创新).
- 与传统的PC模式相比，云计算主要是系统级的创新，产业环境的创新。“软件商店”（Application Store) 就是一例，可能会改变整个软件的研发与销售模式。

发挥虚拟化技术的威力

- 云计算之所以能蓬勃兴起，主要依靠虚拟化技术的成熟；而虚拟化技术能广泛推广，得益于多核技术提供充足的计算和通信能力。
- 虚拟化的重要作用是用软件模拟硬件。凡是以数字化运行的部件都可以由软件替代（被虚拟化），如录音电话等。
- 虚拟化即获得了主机时代Mainframe 的高利用率，又获得了个人电脑时代PC机的灵活性。
- 虚拟化系统是一个多租户系统，相当于按房间出租。利用率高；而过去的Client/Sever系统和IDC形式的服务外包相当于租一座楼，利用率很低。
- 3Tera公司的Applogic软件可以用拖图标的方式实现虚拟化，用户在几分中内可以生成自己的虚拟化设备系统。

云计算成功的一个要素是利益分配

- 所谓建立云计算产业链本质上是各个环节的利益分配。**Apple**公司成功的诀窍是设计了各环节都有钱可赚的利益分配机制。**IaaS**、**PaaS**、**SaaS**和软件商店的开发者的利益应合理分配。
- 联通公司的”软件商店“开通以来只收入**160**元，如何打破免费使用软件的”习惯“要认真考虑。
- 尼古拉斯**.** 卡尔的在”**IT** 不再重要“一书中指出互联网将扩大贫富差别，云计算可能产生”富豪经济“。只有少数人获得经济回报，这一趋势值得警惕。
- 近两年从国外回来一批云计算软件技术人员，他们是发展云计算的重要力量，需要充分发挥他们的作用。

重视云计算中的计算机系统技术

目前推广云计算的重点在转变商务模式

- 用户感觉到的云计算的好处主要是减少购买硬件软件的信息化开支，更好地满足动态变化的需求，降低用户端软件升级的维护成本和管理成本。这种好处主要来自**商务模式的转变**，其核心技术是**虚拟化技术**。
- 目前数据中心转向云计算平台的动力主要是**服务器的统计复用**，可以降低数据中心的运行成本，提高服务器和海量存储的**利用率**，其关键技术也是虚拟化技术。
- 虚拟化技术是一种相对门槛较低的技术，因此各大公司和各地政府都可以在较短时间内建立“云计算平台”。
- 实际上真正支持云计算的是**计算机系统技术**，这些技术用户看不见，媒体也很少宣传。
  - 与李开复的会面

云计算的挑战与机遇

	问题	机会
1	服务的可用性	选用多个云计算提供商；利用弹性来防范DDOS攻击
2	数据丢失	标准化的API；使用兼容的软硬件以进行波动计算
3	数据安全性和可审计性	采用加密技术，VLANs和防火墙；跨地域的数据存储
4	数据传输瓶颈	快递硬盘；数据备份/获取；更加低的广域网路由开销；更高带宽的LAN交换机
5	性能不可预知性	改进虚拟机支持；闪存；支持HPC应用的虚拟集群
6	可伸缩的存储	发明可伸缩的存储
7	大规模分布式系统中的错误	发明基于分布式虚拟机的调试工具
8	快速伸缩	基于机器学习的计算自动伸缩；使用快照以节约资源
9	声誉和法律危机	采用特定的服务进行保护
10	软件许可	使用即用即付许可；批量销售

—— 引自Berkeley白皮书

用户感觉不到的云计算技术

- 云计算系统的本质可以看成是：

  资源虚拟化 + 并行计算

- 云计算不等于虚拟化。虚拟服务器并不能组成一朵云，云计算的能力远远超出一般的虚拟化解决方案。
- 并行技术是藏在云计算背后的核心技术，也是Google等云计算公司具有竞争力的关键技术。
- 各个层次的互连网络在数据中心起到一个非常核心的作用，其作用可能超过服务器本身。
- 发展云计算的一个重要动机是节省能耗。电费已成为云计算中心的重要开支，目前我国的数据中心信息设备利用率低，约为8%左右，（国际先进水平是40%），能耗高PUE约为2.1~3，（国际先进水平是1.2）。

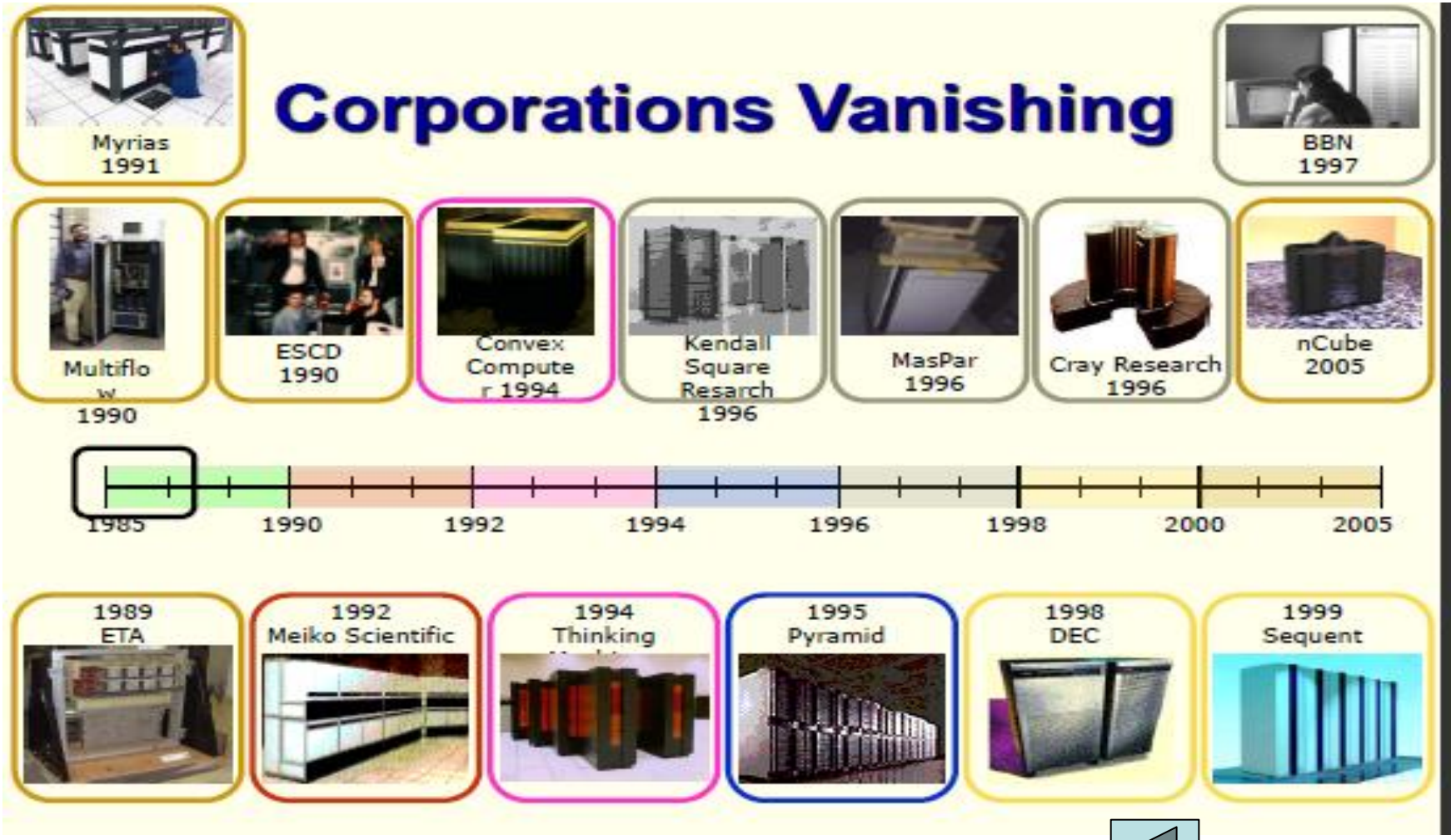
计算机系统的难点在并行处理

- 并行处理已研究了几十年，论文多如牛毛，但进展不大。
- 云计算号称用1000台服务器工作1小时的成本与用一台服务器工作1000小时相当。问题是效率怎样，如果只完成了单服务器的1/10工作，仍然不合算。
- 并行计算最关心的“如何提高计算机的性能和效率”，这个问题从来没有改变，但答案在不断变化。
- 影响并行效率的障碍一是编程（人工效率），二是通信延迟与带宽。“带宽墙”可能比“存储墙“和“功耗墙”更高。
- 历史上计算机设计的匹配规律是完成一次浮点运算需要保证一个字节的供数能力。目前主流CPU的运算速度与供数带宽之比是1：0.3-0.5，即100Gflops的芯片需要50GBps左右的内存带宽（4~8个DDR3)。GPU芯片一个字节供数要完成十次以上浮点运算，典型的“茶壶煮饺子”。

这一场“并行革命” 可能失败

- **人生三件很不愿做又不得不做的事：纳税、死亡，并行处理！**
- 2007年1月， Stanford大学校长，计算机体系结构领域的权威学者 John Hennessy在ACM杂志上指出 :“当我们谈论并行性和轻松地使用真正的并行计算机时，我们是在谈论一个计算机科学家面对的最困难的问题，如果我在计算机企业，我将感到恐慌。”
- IT产业从一个高成长的产业变成一个等待替代产品的产业，我们怎么办？如果软件不能有效地利用几十甚至上千个片内CPU核，计算机就不可能更新换代了，这是一个巨大的危机。

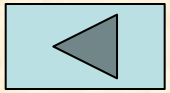
被扔进历史垃圾桶的并行计算机



“互连为王”— 数据中心现状

互连带宽和延迟是必须啃下的硬骨头

- 未来的数据中心需要与目前主流服务器不同的计算机系统来满足云计算的要求。
- 数据中心需要具有大量接口的网络交换器，其价格远远高于市面上流行的交换器，比普通交换器对分带宽bi-section bandwidth）高10倍的交换器的价格要高出100倍。
- 目前的以太网技术无法让数据的传输速率超过每秒100G，主要因为没有这么多的能量来给提供这种数据传输速度的网络系统提供电力和进行冷却。
- 3D芯片和使用硅基光电子学来制造低成本、集成、传输速率达TB级的互连是解决互连问题的希望



Google 达拉斯数据中心

- 占用了附近一个180万千瓦（长江三峡发电站的1/10）水力发电站的大部分电力输出，利用河水冷却服务器。

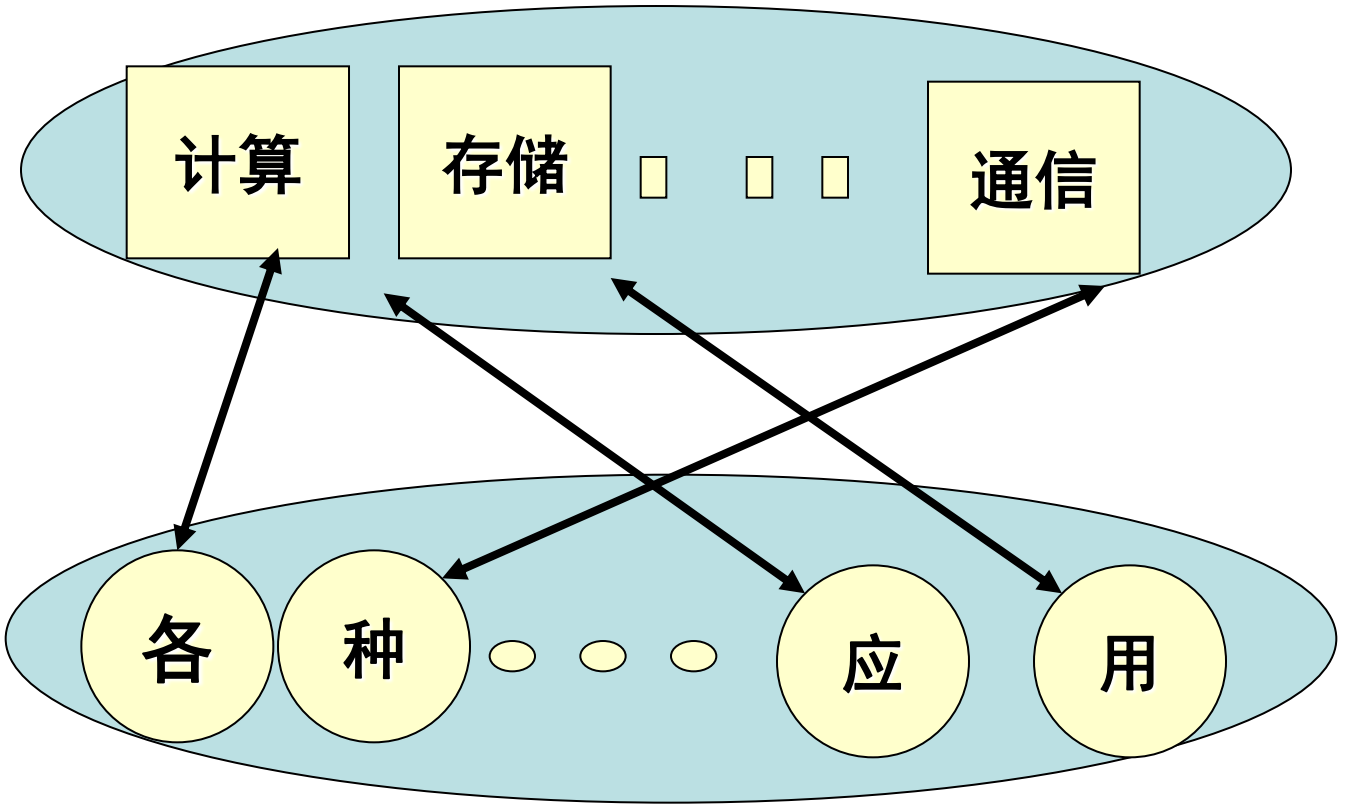
信息为什么这么“重”？
—解决计算机功耗问题的联想

- 假设要传送200TB的数据从北京到西安（1200公里），按200GB的硬盘一公斤计算，大约一吨重，按目前货运市价每吨公里）0.1元左右计算，运费大约120元，加上其他开销不会超过500元。
- 若是用租用1Gps的专线，年租费70－180万元，按天算2000－5000元，每天满打满算可传送约10TB，需要20天，租金需要2－10万元。
- 这就相当于火车运送硬盘的“原子”只有1吨重，但送过去的“BIT”按运费算超过100吨，BIT比原子“重”100倍。
- 如果把全国的网络带宽增加100倍，即从10Gb到1Tb, 可能全国发的电一半要用在网络上（现在占5%左右）

高效可靠的后台系统是云计算的关键

- 计算机系统结构（Architecture) 的概念是从上世纪60年代研制IBM360时提出来的，重要的贡献是区分了硬件与软件，定点与浮点，提出了系列计算机的概念，这些概念一直沿用至今。
- 现在和未来的云计算的workloads的特征与过去 HPC 、事务处理等应用有很大区别，从新的应用中（大量网络服务）应归纳出新的基本指令集。
- 今后决定一个云计算平台是否能存活不是光看虚拟化技术，而是看它的资源利用率，成本和可靠安全等系统因素。
- 认为建云计算中心就是买一大堆服务器，是一种幼稚的想法。把对云计算的支持偏重于中间件也是一种片面的政策。高效可靠的后台设备是云计算的关键。

计算机系统研究最基本的问题——满足应用需求的计算资源配置



云计算是MRMT系统

- 我们可以仿照Flynn的计算机分类，将计算机系统按资源自治域（Resource)－任务(Task）分成4类
  - SRST（单资源单任务系统）1对1 ,如手机
  - SRMT（单资源多任务系统）1对多如，如mainframe
  - MRST（多资源单任务系统）多对1，如某些HPC
  - MRMT（多资源单任务系统）多对多, 如云计算
- 人们常说网格是多对1的系统，云计算是1对多的系统。实际上云计算的后台有许许多多资源，很难在一个操作系统控制之下，是一个典型的多对多的系统。
- 数据中心庞大的计算、存储资源如何高效的调配是计算机系统研究的大问题。关键是资源是不是在统一的调度系统管理之下。

计算机应用的“昆虫悖论”

- 用一句话概括21世纪信息技术的发展趋势就是“为大众计算(Computing for the Masses”)。
- 数十亿用户和各行各业的应用需求一定千差万别。日本东京大学的坂村健教授曾把PC机、手机和物联网的应用种类分别比喻成哺乳类动物(2万种）、鱼类（3万种）和昆虫类（100万种）。计算机系统如何满足如此多的应用。
- 为每一种应用设计一种专用芯片和系统对供应商不经济，采用同一种通用计算机对用户而言效率不高。通用和专用是计算机系统发展中永恒的矛盾，也是最大的挑战。
- 可重构芯片可计算机，可复用的软硬件模块是计算机系统研究的追求目标。虚拟化技术也是解决此矛盾的途径之一

网络问题逐渐变成计算机系统问题

- 电信业正在进入”后电信时代。通信技术与业务正在趋向计算技术与应用；计算技术与应用正在趋向网络与服务提供，CT、IT正在真正走向融合。联通研究院将这种融合模式 称为“公众计算通信网（PCCN）“
- 在原有公众通信网的接入、交换、路由、传输要素的基础上，公众计算通信网还将实现计算处理能力、虚拟分配、调度管理以及业务开发等主要技术。
- 华为、中国移动等公司正在大量吸收懂系统结构的高端计算机人才。既懂计算机系统又懂通信协议的人才是目前最稀缺的人才。我国通信和计算机教育的分离不利于人才培养。

改变不触动“核心技术”的科研模式

超级计算中心和云服务中心

今天的数据中心与未来的HPC

- 云计算的易用性会影响传统的HPC计算模式。传统的排队批处理方式很难实现按需即时响应的科学计算， On-demand的云计算给HPC提供了更易交互的计算模式。
- 构建百万节点数量级的数据中心与今天构建petascale及今后构建 exascale的系统有许多相同的困难。Dr. Reed 认为他们是一对“双胞胎”。
- 共同的挑战包括高速互连、存储分层(包括flash, PCM等）异构多核处理器、系统可靠性和恢复能力、机柜、冷却、能耗效率、和编程等等
- 今天mega-datacenter 的经验将可用于未来的exascale 超级计算机设计。

云计算是超级计算中心的新发展

- 对高性能计算（HPC）而言，云计算并不是一个新的概念。事实上，已经发展近30年的超级计算中心也是一种早期的云计算模式：昂贵的计算资源集中部署，多个领域的用户通过互联网远程使用计算服务并依据使用量支付费用。但这种HPC服务和当前所谈论的云计算又有着一些明显的区别，如没用充分采用虚拟化技术、没有良好的用户界面等。
- 位于高端计算和桌面计算之间存在众多对高性能计算有潜在需求的用户。调研表明，阻碍这些潜在用户使用高性能计算的主要障碍包括：缺乏HPC人才、建设和运维的成本以及使用HPC应用的复杂度。而云计算正是应对这些挑战的最佳途径。
- 云计算将扩大HPC 服务的范围。 随着虚拟化技术的提高,通信延迟降低， 紧耦合的计算将在更大范围内具有吸引力。

传统HPC平台与“HPC云”的区别

	传统架构	云架构
资源管理	作业管理系统：为作业找资源，只管理处理器、应用软件	为用户、作业进行**动态地资源创建和回收，**管理处理器、内存、存储、网络和应用软件
虚拟化	不支持	服务器虚拟化、存储虚拟化、网络虚拟化
用户管理	独立的用户管理系统，用户无法独享资源	统一用户管理，用户可以独享资源
平台支持	无法修改已安装平台，无法动态修改	可以同时支持多种平台，可以动态修改
数据存储	没有备份机制，不支持异构存储	完善的备份、恢复机制，支持异构存储平台
用户使用	无资源审批流程，无法自定义资源配置	审批、拒绝、预留机制，可以自定义资源平台、软件等

云计算还不适合做尖端的超级计算

- Dan Reed：云计算绝对不是为特定目的构造的性能顶尖计算机的替代品。如果一种petascale计算需要极低的任务间通信延迟，今天的云计算肯定不适合。但是对于大多数使用较小规模设备的研究者，云计算是有吸引力的替代品.
- 目前的云模型并不支持顶尖的超级计算。动员grand challenge 应用的人做云计算就如同要说服驾驶第一方程式赛车的深受去乘公共汽车。
- HPC主要执行计算密集型的任务，CPU的利用率已经很高，虚拟化技术对提高HPC的CPU利用率作用不大。

目前的云计算做HPC效率较低

- 基于云计算理念来构建超级计算中心，除了满足传统的或现有的HPC用户需求外，更重要的是创造并吸引众多新领域的用户。
- 美国德州先进计算中心（TACC）的 Edward Walker 对Amazon EC2 上HPC应用的性能表现进行了研究，应用选择常用的基准测试程序NPB，测试结果表明：几乎相同的硬件条件下，对OpenMP版本的8个测试程序EC2性能下降7%至21%不等，MPI版本性能则下降40%至1000%不等。
- 虚拟化对计算密集型（如果数据能全部放进内存）应用的影响很小，而I/O密集型应用的性能则会有一定下降

在Amazon EC2 上运行MPI性能不高

- **Performance is below the level seen at dedicated, supercomputer centers, however, performance is comparable with low-cost cluster systems.**
- **Significant performance deficiency arises from messaging performance where latencies and bandwidths are between one and two orders of magnitude inferior to big computer center facilities.**

System	latency	uni-bw	bi-bw
LAM	81.20μs	57.85MB/s	81.98MB/s
GridMPI	83.46μs	54.60MB/s	77.07MB/s
MPICH2 nem	300μs	15.72MB/s	26.08MB/s
MPICH2 sock	85.87μs	58.49MB/s	83.42MB/s
OpenMPI	300μs	16.44MB/s	17.99MB/s
LAM/ACES	35.83μs	117.64MB/s	198.59MB/s

---MIT Constantinos Evangelinos and Chris N. Hill CCA-08 paper

标准和安全

等待标准与厂商锁定

- 尽快制定标准是发展云计算必须要做的一件重要事情，但不能等标准出来以后才考虑云计算，足够的使用经验是制定标准的前提。等待标准化可能付出机会成本，即可能因等待而失去发展机遇。
- 避免服务供应商锁定是发展云计算中需要重视的取向，要力求用户有主动选择供应商的权利。
- 很多云雾是由厂商制造的，这些厂商试图装扮他们的老产品。在这种环境中，人们很容易屈服于混乱，推迟采取行动。将这种延期美化为“避免锁定”和“确保符合标准”听起来合理和明智，但事实上可能导致被人超越。
- 在今后的五年内，云计算的特征、能力和种类将呈现出巨大的发展，停下来等待其标准化是无益的。当你处在混沌的旋涡中，坐着不动并不是明智的选择。

加快发展与云安全

- 保障“云计算”的信息安全，需要法律、行政、行业自律、技术四个层面齐头并进。
- 云的安全模式要好于内部数据中心。在客户端/服务器模式中，很多敏感信息位于客户端，打补丁是一场噩梦。在云中，集中并且保护数据更为简易，用户能够迅速锁定安全漏洞。
- 虚拟化技术重新构造现存的软件系统，将以新的方式打造更安全的系统
- 电气化推广同样碰到安全问题，大电站发出的交流电压高于小直流发展机发出的电压。交流电可能电死人（美国用交流电处理死刑犯），核电站的安全问题肯定比云计算更可怕，但不能因此不发展交流电和核电站。安全和发展是相反相成的永恒矛盾，只能在发展中解决安全问题。

请批评指正！