

# 21世纪计算机体系结构

## 计算机体系结构共同体白皮书\*

关键词：计算机体系结构

译者：计算机体系结构国家重点实验室(筹)

### 引言与概要

信息技术 (information and communication technology, ICT) 从医疗、教育、科技、商业、政府、国防和娱乐等各个领域改变着我们的世界。如今，人们已经差不多忘记了20年前人们找资料的第一步是设法到达图书馆，10年前社交网络还主要指面对面的交流，5年前“tweets”还来自卡通角色。

重要的是，很多迹象表明信息通信技术创新正不断将科幻小说中种种炫目的想象力搬到真实世界中<sup>1</sup>。附录A列出了一些想象场景及其属性。这些未来应用包括个性化医疗技术为每个人提供不同治疗与用药方案，对潜在恐怖威胁进行精密的社交网络分析以协助国土安全，以及利用远程呈现 (telepresence) 减少交通行驶，从而减少温室气体排放。这些未来应用对计算机系统不断提出新的需求，包括处理大规模异构数据集 (大数据<sup>2</sup>)，采用分布式设计，在有限的空间下工作，既能快速部署又可保证高效运作。

半导体技术与计算机体系结构是推动过去信息通信技术创新的两个核心技术，但它们只是幕后英雄，往往不为人所知。半导体技术创新在过去40年

持续地让单个芯片在功耗与成本几乎维持不变的情况下 (Dennard scaling, 丹纳德微缩) 容纳了越来越多的晶体管 (摩尔定律)。而计算机体系结构发明了各种创新技术来利用快速增长的晶体管资源，扩展了处理器的性能，缓解了内存系统带来的性能损失。在过去几十年，两者结合的效果让信息通信领域的创新者们可以在几乎不增加成本的情况下使性能呈指数增长。

由于大多数半导体技术与计算机体系结构革新 (刻意) 是对上层应用透明的 (invisible)，应用与软件的开发人员们无需参与底层技术革新就能体验这些技术进步带来的各种好处。一方面，更高的性能可以让那些有更多计算需求的应用成为可能 (如虚拟助理、计算机视觉等)，另一方面通过支持更高层的编程抽象 (如脚本语言、可重用模块等)，使得那些对性能要求不高的应用更容易开发。事实上，计算机系统成本效益的增加带来了价值创造 (例如，分布式网页搜索成本已低廉到能用广告链接的收入来支撑)，这是领域开创者们未曾想象到的。

计算机性能增长带来的各种好处是显而易见的。最近，丹诺威茨 (Danowitz) 等人分析发现<sup>3</sup>，

\* 本文译自 Computing Community Consortium, <http://www.cra.org/ccc/docs/init/21stcenturyarchitecturewhitepaper.pdf>.

<sup>1</sup> PCAST, "Designing a Digital Future: Federally Funded Research and Development Networking and Information Technology", Dec. 2010 (<http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd-report-2010.pdf>).

<sup>2</sup> CCC, "Challenges and Opportunities with Big Data", Feb. 2012 (<http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf>).

<sup>3</sup> Danowitz, et al., "CPU DB: Recording Microprocessor History", CACM 04/2012.

表1 计算机体系结构面临的工艺技术挑战

20世纪后期	新的现实
摩尔定律——单个芯片晶体管数目每18~24个月翻一番	晶体管数目仍然每18~24个月翻一番，但见下述
丹纳德微缩——几乎是常数的单芯片功耗	不再有效。（当芯片晶体管数目翻一番）芯片功耗翻一番不可接受
晶体管不可靠性不显著，很容易隐藏（如ECC）	晶体管可靠性将不断恶化，已不再容易隐藏
更关注计算而非数据通信	芯片间、设备间、机器间的通信能力有限（如兰特规则、3G、GigE等限制因素）；通信比计算代价更大
一次性（无法重复利用的工程）成本在增加，但可以用有大容量市场的部件来分摊	设计、验证、制造和测试的一次性成本非常昂贵，尤其是面向特定市场的平台

自1985年以来，计算机体系结构革新贡献了约80倍的性能增长，与半导体技术进步的贡献相当。随着半导体技术逐渐走向“道路尽头”（见后述），未来信息通信技术创新会需要计算机体系结构扮演越来越重要的角色。但体系结构研究人员不应该再问“如何让芯片运行得更快？”，而应转向新的问题：“在失去几近完美的工艺技术按比例缩小所带来的好处时，该如何为21世纪构建从传感器到云端的基础设施，如何从性能到隐私等多个角度去增加价值？”挑战很多，但只要投入合理，也蕴藏着诸多机遇，而机遇背后的一个共同的主题就是——未来体系结构创新将需要其他信息通信技术层次的参与和投入。

### 挑战：拐点

表1总结了为信息通信技术提供动力的半导体技术所面临的严峻挑战。第一，尽管技术人员依然能够制造出更多更小的晶体管（摩尔定律），但这些晶体管已不像过去40年那样各方面都“更好”；第二，单个晶体管的功耗已不能很好地按比例缩小（丹纳德微缩定律已终结）。由于大多数产品——传感器、移动装置、终端与数据中心——都不能容忍功耗的不断增长，我们必须考虑各种手段来缓解功耗增长；第三，纳米级特征（如栅氧化层只有几个原子厚度）的制造偏差，与较大特征尺寸相比，严重地降低了晶体管的长期可靠性；第四，随着新技术（如芯片堆叠技术）与新挑战（如数据中心）的出现，必须通过局部性原

理来管理计算部件之间的通信，从而实现合理的成本、能耗目标；第五，设计、验证、制造和测试的一次性成本不断增加，导致这些成本更难被分摊，尤其是在需要平台专门化来达到高效率的场合（如手持设备、笔记本或数据中心等）。

### 机遇：21世纪的计算机体系结构

随着CMOS工艺技术按比例缩小在推动信息通信技术创新方面的作用逐步弱化，计算机架构师必须更进一步加强自身作用。然而，21世纪的计算机

表2 计算机体系结构的过去和未来

20世纪的体系结构	21世纪的体系结构	
单芯片性能	<b>体系结构即基础架构：</b> 从传感器到云计算 · 芯片到系统 · 性能加安全、隐私、可用性、可编程性……	<b>跨层次的含义：</b> 用新的接口打破当前的层次
通过对软件透明的指令级并行性（ILP）来获取性能	<b>能量优先</b> · 并行性 · 专用化 · 跨层次设计	
试验和测试过的技术：不断进行快速但可预测的改进的CMOS、DRAM和硬盘	<b>新技术：</b> 非易失存储器、接近阈值电压的运行、3D芯片、光子学，…… <b>重新考虑</b> · 内存+存储 · 可靠性 · 通信 · ……	

体系结构需要与20世纪的先行者有所不同，以体现这种作用。如表2所示，我们看到3个根本性的差异。这些差异构成了下一章节描述的未来研究计划的基础。

**体系结构作为基础架构：从传感器到云计算** 过去的体系结构研究经常集中在芯片（微处理器）或者独立计算机上，以性能作为其主要的优化目标。今后，计算机将向前成为21世纪社会结构的关键支柱。为了适应这一转变，必须对计算机体系结构研究进行扩展，必须认识到常规的计算机已经被上下文计算（如传感器、移动终端、客户端、数据中心）所取代，同时许多计算机系统规模庞大并且在地域上是分布式部署的<sup>4</sup>。这种转变要求更强调系统（如通信成为计算的一个平起平坐的伙伴）、驱动型应用（如大数据处理）以及超越性能的以人为本的设计目标（如可编程性、隐私性、安全性、可用性、电池寿命和外形尺寸）。

**能量优先** 过去的计算机体系结构大多通过软件隐形的变化来优化性能。21世纪的体系结构主要受到功耗和能量的约束，且不再能维持软件隐形创新的奢华。我们看到并行性、专用化以及跨层次设计成为能量优先时代的关键原则，但是上述三个方面都需要应对巨大的挑战。例如，尽管未来的应用蕴藏着丰富的并行性（大数据=高并行度），但是通信的能耗将大大超过计算的能耗，并且需要重新考虑如何针对1000路的并行度做设计。与通用的计算或内存单元相比，专用化可以提高100倍以上的能效，但是当前尚未找到将专用化的优势以高性价比的方式应用到广泛类型应用中的方案。跨层设计（电路—结构—运行时系统—编译器—应用）可消除不同层次上的能效浪费，但是往往需要在这许多

层次之间付出更多协同的努力。

**新技术** 过去的计算机体系结构依赖于诸如CMOS、DRAM和磁盘等稳定技术的可预测的性能提升。现在，在许多信息通信技术专业人员职业生涯中，第一次出现新兴技术对占主导地位的“试验过和测试过的”成熟技术提出了挑战的情形，但是这迫切要求体系结构的革新，以挖掘新技术的全部潜力。例如，非易失性内存技术（如闪存和相变存储器）驱动着对内存和存储系统之间关系的重新思考。接近门限电压的运行对于降低功耗有着极大的潜力，但是要以牺牲可靠性为代价，这就驱动了一门以弹性为中心设计的新学科。光子学和3D芯片堆叠根本地改变了通信的成本，这足以影响整个系统的设计。

以上各点隐含的是一种以**跨层次为主旨**的创新方法，即需要面对其他信息通信技术层次并与之进行交互。这种变化给其他层的信息通信技术创新人员带来的影响是显著的，这与近期向多核处理器的转变类似，但是其潜在影响更大。与其他层的创新人员合作可使体系结构人员能够进行更加深入的创新并获得相应的收益，但是这无疑需要巨大的投资和强大的领导力，以提供更加丰富的层间接口，这些接口对于21世纪是必要的。

## 研究方向

本节将从两条说明开始，进而讨论计算机体系结构及相关社区的重要研究方向。第一，下文所描述的内容仅代表了对本文及其他近期补充文献<sup>5,6,7</sup>作出贡献的学者的优秀思想。这些观点既不是对整个共同体的详细列举，也不是整个共同体的一致性观点，这两方

<sup>4</sup> Luiz Andr é Barroso and Urs Hölzle, "The Datacenter as a Computer", Morgan-Claypool, 2009.

<sup>5</sup> ACAR-1, "Failure is not an Option: Popular Parallel Programming", Workshop on Advancing Computer Architecture Research, August 2010 ([http://www.cra.org/ccc/docs/ACAR\\_Report\\_Popular-Parallel-Programming.pdf](http://www.cra.org/ccc/docs/ACAR_Report_Popular-Parallel-Programming.pdf)).

<sup>6</sup> ACAR-2, "Laying a New Foundation for IT: Computer Architecture for 2025 and Beyond", Workshop on Advancing Computer Architecture Research, September 2010 (<http://www.cra.org/ccc/docs/ACAR2-Report.pdf>).

<sup>7</sup> Fuller and Millett, "The Future of Computing Performance: Game Over or Next Level?", The National Academy Press, 2011 ([http://books.nap.edu/openbook.php?record\\_id=12980&page=R1](http://books.nap.edu/openbook.php?record_id=12980&page=R1)).

面在撰写这份白皮书的过程中是无法实现的。

第二，如果我们已经确信所陈述的方向对社会有价值，就需要竞争前的研究资金来支持其开发。然而即便是已经取得巨大成功的计算机公司也缺乏做这项工作的动力，其原因在于：首先，这些技术需要几年到十年的开发周期，极少有公司能够有这样的持久力；其次，成功的工作应让许多公司受益，而不是刺激任何一家公司为此付出代价；第三，我们所提倡的许多方法都需要跨越系统栈的各个层次，超出了工业标准接口（如x86）和各独立公司的专长；最后，我们需要培养下一代的技术贡献者，这可能是学术界最重要的技术转移形式。

## 将体系结构作为基础设施：从传感器延伸到云

直到最近，计算机系统还常指放置在办公桌下或者机房里的米色箱盒，而计算机体系结构的研究也主要关注于桌面、工作站或者服务器的CPU设计。新的21世纪应用的现实需要在米色箱盒之外开展更为广阔的计算机体系结构研究。在附录A中所提及的新生应用需要一个丰富的生态系统，这个系统利用无所不在的嵌入式感知/计算设备将数据传输给“云服务器”和仓库规模设施（warehouse-scale facilities），这些服务器和设施处理并提供信息给诸如平板电脑、智能手机等边缘设备（edge devices）。上述的每一类计算设备代表了特有的计算场景，都伴随着特定的挑战和机遇。但这三类设备却有一个共同的创新驱动力：能效（每瓦特功率产生的性能）。同时，为了满足新生应用的需求，未来的体系结构研究必须超越那种优化每一个孤立设备的情况，接受跨环境协同设计带来的挑战。

**智能感知和计算** 在智能感知领域，核心的约束是在非常严格的能量、尺寸和成本的限制下完成计算。现在，过滤和处理数据（在这个过程中会

生成和收集数据）（比如区分正常和异常的生物统计信号）的重要性驱动着对更强大计算能力的需求，原因在于数据的通信通常比计算更耗能。这种场景带来了很多人激动的新机遇，比如设计利用断续电源（比如环境发电，harvested energy）的系统、极端低电压（接近阈值电压和模拟电压）设计、新的通信方式（如利用建筑照明进行广播通信）以及新一代的存储技术（如非易失性存储器）等。随着传感器在健康和领域愈加重要，其安全性和可靠性也必须得到保证（考虑到诸如最近演示的对起搏器的远程黑客攻击），包括设计正确的软/硬件部件。此外，考虑到传感器数据本身具有近似性，使得近似计算有了用武之地，从而带来显著的能量节省（和复杂性降低）。

**便携边缘设备** 随着近来智能手机销售超过个人电脑<sup>8</sup>，便携计算市场呈现出爆炸式增长态势。然而，现有便携设备所具备的能力远远低于社会的预期与期望。从近期提出的增强现实谷歌眼镜，到由Star Trek在将近50年前提议的并在近来借由X Prize挑战竞赛<sup>9</sup>又兴起的医用手持分析仪，莫不如此。由于用户界面看起来对于计算能力要求很高（比如多点触控接口、语音识别、图形、全息影像、3D场景重建等），甚至移动应用也正向数据和计算密集型转变，因此，要在这类环境下充实这些应用将需要现有设备的单位能量运算率从现在的约百亿次运算/瓦提升若干数量级。正如后面章节所讨论的那样，这样的系统同时需要并行化和专用化（最新一代iPad的核心芯片就是多核芯片并且一半的芯片面积被专用部件占据）。为了亲近用户，这些设备的开发驱使研究人员在计算机设计的时候就考虑用户因素。例如利用用户反馈来调节电压/频率以节约能量；或者仅在用户关注的地方实施计算；或者压缩图片仅突出重要的部分；或者预测并提前进行用户想要进行的操作。这种情况也驱使研究人

<sup>8</sup> 如<http://mashable.com/2012/02/03/smartphone-sales-overtake-pcs/> 及<http://www.canalys.com/newsroom/smartphones-overtake-client-pcs-2011>.

<sup>9</sup> <http://www.xprize.org/x-prize-and-qualcomm-announce-10-million-tricorder-prize>.

员还要关注那些纯粹性能以外的特性，比如安全和隐私。

**基础设施——云服务器** 很多特别令人激动的新生应用，像模拟驱动的药物发现或者大规模社交网络的交互分析等，必须在合理的响应时间内完成。这种需求只能通过将成百上千个处理器的能力合并成为一个“仓库”规模计算机才能完成。互联网搜索已经展示出这种计算范式的社会重要性。但是，当前商业运行的搜索系统需要投入大量的工程精力来创建、编程和维护，并且其搜索能力还是很肤浅的（与IBM的Watson系统的潜力比较）。为了让接受传统训练的软件工程师能够在仓库规模计算系统上编程，系统架构师必须设计编程抽象、存储系统、中间件、操作系统和虚拟化层来提供支持。

尽管很多计算机的学科（操作系统、网络以及其他）在关于数据中心的创新中依然起到作用，但是为了更好地促进高层次创新，计算系统架构师必须考虑接口设计以及相应的硬件支持。此外，在数据中心系统中面临的关键挑战是如何发现局部性和实现高效局部性，这就必须克服系统工具、中间件和体系结构之间的协作阻碍。更大的挑战是要让性能可以预测；随着很多请求在系统上得到并行处理，低概率的尾延迟将成为性能瓶颈（假设单个系统有99%的概率不会产生延迟，那么在每一个请求都必须由100个系统联合响应的情况下，则有63%的请求会由于等待处理慢的系统而得到长延迟响应<sup>10</sup>）；体系结构的创新应该保证最坏情况下的绝对延迟需求。内存和存储系统的消耗在数据中心总电量中占据的比例会越来越大，利用新的接口（JEDEC标准之外）、新型的存储技术以及处理器和存储的三维堆叠技术可缓解这一趋势。

**总而言之——生态系统体系结构** 对运行时平台和虚拟化工具提出了一个需求，就是允许程序功能被分别划分到便携平台和云端，并且这种划分能够动态地适应云端上行可靠性和能效的变化。计算应该如何被划分到节点和云基础设施中？在所有

场景下，如何有效地保证安全性？如何通过让用户对私有数据拥有更多的控制权，使系统体系结构有助于保护数据的隐私性？我们是否应该对计算引擎和内存系统进行协同设计？

上述的研究方向将推动体系结构的研究，远远超越了米色箱盒的范畴。与这些研究机遇相关的最基本的研究挑战是显著地提高系统能效以及重新考虑系统的可编程性、安全性、隐私性和弹性等新需求。随着其他领域（如人机接口、机器学习和普适计算等）的迅猛发展，现在到了利用这些新技术的时候了。由于对全社区的共同努力和重要的研究基础设施的需求，因此需要对学术研究加大资金的投入才能取得成功。尽管移动设备和数据中心相对于所有计算设备来说是新出现的，但是我们讨论的研究挑战也将很可能同样适用于那些还没有出现的计算设备。

## 能耗优先

计算机系统从串行向并行（多核）的演化使得在单个芯片计算性能提升的同时，功耗大体上保持恒定。然而即便如此，当前并行计算系统已经受限于功耗或能耗。在其中一端，超级计算机和数据中心功耗需求都在数百万瓦特；在另外一端，功能各异的传感器和便携设备通常受限于电池的容量。这类传感器和便携设备通常只在短时间内需要高性能，随后便进入长时间的空闲状态。这种双模用法在很少完全空闲、也很少需要在最高速率运行的高端服务器上并不常见。因此，针对服务器领域的功耗或能耗的解决方案极有可能与针对便携设备领域的解决方案有显著的差异。然而，人们对计算性能永无止境的需求都将使得功耗、能耗和散热成为掣肘因素，无论是对高端的服务器领域还是便携、传感器设备领域。如果没有提高功耗和能效效率方法上的革新，在封装和制冷方法上没有突破，那么构建面向新应用（附录A）下满足人们对计算、存储和通信需求的高效信息通信技术系统的目标将成为

<sup>10</sup> J. Dean. "Achieving Rapid Response Times in Large Online Services". Talk in Berkeley, CA, Mar. 2012.

海市蜃楼。因此，我们迫切需要加大资金的投入来研究如何提升系统效能。

如下面的小节所述，降低能耗必须从以下三个途径加以解决：跨层次能效优化，对“并行”的再思考和有效利用“专用计算”。

### 跨层次能效优化

电子设备在工作时消耗能量（对CMOS而言，工作状态是电源接通的状态）。计算系统的所有层次：器件工艺、体系结构、软件系统（包括编译器）和应用程序，都可以对提高能量和功耗效率所有贡献。因此，我们相信跨层次的交叉研究将显著提升整个信息通信技术系统的能效。作为一个目标，我们希望到2020年计算系统的能效提高2~3个数量级，即百亿亿次级数据中心的功耗不超过10兆瓦，千万亿次级部门服务器的功耗不超过10千瓦，万亿次级便携设备的功耗不超过10瓦，十亿次级传感器系统不超过10毫瓦。只有在所有计算层次上有所变革，这一雄心勃勃的计划才有可能得以实现。

**在电路/工艺层**，我们需要研究新的更高能效的基本器件，包括传统的CMOS和新出现的器件技术。也需要研究一些能提升某些特定功能的能效的新技术，例如用于通信的光子器件、三维集成、非阻性存储器件和高效的电压转换。一些新的电路设计也亟待研究：可在超低供电电压下工作的电路，实施高效供电分布的电路，激进的功耗管理电路以及在片上支持多电压和频率域的电路。

**在体系结构层**，我们需要提出更高效、更合理的众核体系结构。我们需要构建基于异构簇的芯片组织结构，其中既包含简单的处理器核，也包含为某些功能定制的高性能处理单元，二者协同地完成计算任务。我们也需要研究如何使数据通信量最小，原因在于在新的计算环境中大部分能量被消耗在数据移动方面。特别是针对便携设备和传感器系统，将计算本地化可以显著降低高耗能的通信负载。有鉴于此，我们还需要在同步支持、高效通信和就地计算方面进行更多的研究。

**在软件层**，我们需要研究如何最大限度减少不必要的通信。需要运行时系统的支持来管理存储的层次，协调细粒度的多任务。需要研究用于管理和增强局部性的编译系统和相关工具。在编程模型层次，我们需要编程环境既可以支持专家级的程序员对整个机器进行控制，又可以对普通程序员提供简单的局部性模型。在应用层，我们需要高能效的算法级策略来减少操作数量、访存数量和处理器间的通信量，降低精度，并充分利用异构系统的优势。在编译层，我们需要提出新的方法以有效地在功耗效率与性能之间做出权衡，同时兼顾二进制代码的可靠性。总体而言，只有通过所有层次上的共同努力，才可能取得革命性的进步。

### 开发并行性支撑未来应用

在20世纪，计算系统性能提升的驱动力主要来源于单处理器性能的提升。但是2004年伊始，更快的时钟频率和更大的晶体管规模带来的好处逐渐减少，导致的问题是严重的功耗（及相关的散热）。业界通过增加每个处理器芯核数目来持续提升累积的性能和每瓦特的性能。但是，仅通过复制芯核并不是最终解决能耗和扩展性挑战的最有效途径，也不能充分利用处于“片上”这一特征和限制条件。未来计算系统性能的提升必将来自片上的大规模并行性，即与新兴应用中蕴藏的细粒度并行性相匹配的大量简单、低功耗芯核。

然而，为绝大多数用户提供一套可用的并行体系结构与相应的编程环境的目标还远没有达到。之前的研究大都关注于基于标准处理器核的粗粒度的并行。将大量并行性置于单片处理器将为并行体系结构和与之相关联的编程技术提供新的机遇。为了将并行计算的潜力以一种更加普适的形式解放出来，我们需要在广泛协作的基础上付出至少十年不懈努力，才有可能解决新的应用在所有层次的并行度问题。人类的创新很大程度上得益于计算性能的不断提升。考虑到提升性能所面临的挑战的严峻性，为了确保持续的创新，我们必须在这一领域付诸极大的资金投入。

**为并行度重新构造计算层次** 我们建议通过一项雄心勃勃的跨学科计划为并行化彻底地革新经典的计算层次——编程语言、编译器和编程工具、运行时系统、虚拟机、操作系统和体系结构。以性能为目标的同时，必须将能耗作为首要考虑的设计限制条件，从而构建一条通向未来可扩展系统的途径。当前的计算层次根源于单处理器系统，这一历史的桎梏成为了扩展并行度道路上的障碍。当然，单一、普适的编程模型可能是不存在的，我们建议开发多种模型和体系结构。进一步说，对专家和普普通程序员显然需要不同的解决方案：专家也许需要与计算机内部进行交互；而大多数普通程序员应该使用基于特定领域语言的简单、串行模型。

**聚焦于应用的体系结构研究** 我们建议一种聚焦于应用的并行体系结构研究方案，从只有几个核的多核芯片以及潜在的成百上千核的众核芯片到分布式系统、不同规模的网络结构、并行存储系统和读写（I/O）解决方案。这里的挑战是既要考虑应用的特征，但又不能过分适合某个具体的特点；我们必须借助那些在各种并行系统上工作得很好、能效也很好的机制（包括编程语言）。在体系结构方面的本质问题包括并行的类型（如数据或任务）、处理单元的组织结构（如独立的核或协处理器）以及同步和通信。最终的方案也许是各有差异，这取决于应用内在的并行性、功耗、性能和系统的规模。

**硬件/软件的协同设计** 当下的挑战需要贯穿硬件和软件（从应用、算法到系统软件和硬件）的一体化的并行性研究。如下所述，我们需要打破现有的抽象层次，提出新的机制和策略来开发局部性和并发性，高效地支持同步、通信和调度，提供可编程、高性能和高能效的平台，开发真正易于使用的并行编程的模型、框架和系统。要解决广大范围内的这些问题，需要部署广泛而深入的研究计划。

### 为性能和能效启用“专用化”

在过去的几十年中，通用计算机驱动了计算技

术的快速发展和全社会对计算系统的普遍接受。然而，同样的灵活性在使得通用计算机适用于解决大多数问题的同时也牺牲了执行很多新应用时的能效。为单一或窄类功能定制的“专用”的硬件加速单元剥离了为灵活性和通用性设计的复杂的层次结构，能效通常可以提升几个数量级。然而目前成功的案例，从医疗设备、传感器阵列到图形图像处理单元，还局限在很窄的几类问题。业界亟待两方面的研究：（1）在开发新的体系结构以改进专用单元的性能、能效的同时，扩大其适用范围和解决问题的门类；（2）降低制约定制方案发展的软件和硬件的一次性投入（non-recurring engineering, NRE）成本。

**高层次抽象支持专用化** 对通用计算机而言，我们可用多种高层语言为其编程，然后用一套复杂的工具链将其翻译成特定的指令集，为大范围的体系结构提供功能和性能的可移植性。相对而言，专用的处理单元通常采用与硬件直接交互的底层编程语言（例如Verilog）进行编程，这使得软件的功能和性能可移植性极为有限，随之而来的是高昂的一次性投入。如后面进一步讨论的那样，我们需要开发新的逻辑层次和抽象来刻画运算的结构，以便高效设计或映射到专用的硬件，而避免将负担推卸到程序员身上。这样的系统通过减少甚至消除将应用重新映射到新硬件平台的工作量，可有效地支持快速开发加速单元。

**高能效的存储层次** 对存储层次的优化既可以提升系统性能，也可以降低存储系统的能耗，但通常是把性能作为首要目标。然而，对一个浮点“乘加”操作，从存储系统里取操作数的能耗要比用于实际计算的能耗高1~2个数量级<sup>11</sup>。此外，当前的设计常常试图牺牲最坏情况下的性能来换取最大的一般情况下的性能，或者牺牲可编程性来优化最优情况下的性能。未来的存储系统必将通过专用化来提升系统的能效（例如，通过压缩和对流数据的支持），同时简化编程（例如，在需要时通过将加速

<sup>11</sup> Steve Keckler, "Life After Dennard and How I Learned to Love the Picojoule", Keynote at Micro 2011.

单元纳入到一致性和虚拟存储的范围)。对大范围各类系统,小到不间断的智能传感器,大到处理大数据的数据中心,这些机制都有潜力降低能耗。

**开发可(重)配置逻辑结构** 随着硅制造工艺复杂度的不断提升,一次性投入的代价已经变得极为高昂,这使得全定制的加速单元只可能被少数拥有大量应用场合的计算逻辑采用。目前可重配置硬件平台(例如FPGA)的采用可在某种程度上降低这类固定成本。但这种细粒度的可配置性(例如基于查表和交叉开关的可配置逻辑)牺牲了性能和能耗。未来的加速单元将基于粗粒度、部分可编程的构建模块(减少内部低效率)。同时,使用基于包交换的互连结构,以便更高效地利用昂贵的金属互连线。其它的效率将来自新兴的三维技术,比如硅基插板。该技术可以适中的代价实现有限的定制来配置系统(如顶层互连)。这些技术结合更强大的综合工具能降低一次性投入成本,从而更快、更广泛地开发和部署基于加速单元的能效优化技术。

处理器性能的不断f提升不仅极大地促进了计算领域的演进,也驱动了社会的进步。未来性能的持续提升很大程度上取决于能效的优化潜力。具体而言,很可能就是可配置和专用化的硬件的发展程度。因此,在可配置和专用处理器上的研究将使得公司、研究人员和政府得以快速并且以普遍可接受的代价将巨大的计算能力聚焦到所需解决的关键问题上。

## 工艺技术对体系结构的影响

应用对改善性能、功耗、能效和可靠性的需求驱动着业界对工艺发展不断投入。虽然标准CMOS工艺已经逼近基本的比例缩小极限,但科研人员仍持续探索可行的替代技术。诸如亚/近阈值CMOS、量子阱场效应管(QWFETs)、隧穿场效应晶体管(TFETs)和量子电导原子开关(QCAS)等技术在密度、速度、功耗和可靠性等方面具有组合优势。非易失性存储器(如闪存)在许多信息通信技术系统中已经替代了传统的旋转式硬盘,但也带来了其固有的设计挑战(如有限的可擦写次数)。其他新兴的非易失性存储技术,如自旋转移力矩存储器(STT-

RAM)、相变存储器(PCRAM)和忆阻器(memristor),则有望打破现有设计中易失性存储和非易失长期存储间的严格界限。三维集成技术通过裸片堆叠使IC设计在新的维度上扩展,但这些仍需面对现有技术和电子设计自动化工具的挑战。此外,还可以在芯片之间甚至在芯片上探索实现光互连。

除了已有的新信息通信技术的研发(如通过美国科学基金会的材料研究科学与工程中心(MRSECs)),还需要重大的体系结构进展和相应显著的资金投入来开发应用这些工艺技术。

**对存储的重新思考** 新兴技术为解决在线存储和处理大数据应用提供了新的可行途径。新兴的非易失性存储技术可以提供高得多的存储密度和功率效率,但需要针对器件能力(如更长、非对称或可变的访存延迟,以及器件耐久性)重新设计内存和外存系统结构。

**设计自动化的挑战** 新技术和新工艺驱动了电路、功能单元、微结构和系统的新设计。这些新方法仍需要与之匹配的新电子设计自动化(electronic design automation, EDA)工具,尤其是适用于混合信号和三维设计的工具。新工具必须要满足新技术的要求,能够支持功能综合、逻辑综合、版图设计等;异构计算更是挑战着我们在寄存器传输级(register transfer level, RTL)设计之前对系统建模的能力,尤其是结构的多样性和加速部件数量的急遽增加。可见,我们需要新的验证、分析和模拟工具,即验证正确操作、系统性能、功率和能耗、可靠性(如对软错误和老化效应的敏感度)和安全(如避免功率“印迹”、对信息流追踪提供体系结构级支持)的工具。

**三维集成** 裸片堆叠技术虽然拥有降低连线延迟,提高片上带宽等多种优势,但仍需面对电子设计自动化、设计和工艺支持等多种挑战。现今的计算机系统架构师已经开始调研在核上堆叠DRAM的技术,未来不仅要堆叠非易失性存储器,还将堆叠基于非兼容工艺定制的计算部件、新型互连结构(比如光互连)、含有模拟组件的传感器、射频单元等模拟器件、能源提供装置(如太阳能电池、MEMs能量收集



部件)、冷却装置(微流体冷却)等。为了应用纳米材料和纳米结构研究人员开发的新技术,需要进一步的投资,以便计算机系统设计师能够把这类“纳技术”电路变为信息通信技术系统。

## 跨层次问题和接口

计算机渗透到日常生活中越来越多的方面,一个**更好**的计算机不只是意味着速度**更快**或者能效**更高**。计算机应用需要体系结构提供以下支持:保障数据的安全性和隐私性,容忍日益变得不可靠的晶体管带来的故障,以及增强可编程性、可验证性和可移植性。为了达到这些跨领域的设计目标——“诸可易性”,需要在差异巨大的应用需求和工艺限制条件下,从根本上重新审视可长期保持不变的接口。

### 安全性、可编程性、可靠性、可验证性及其他

当前应用对一些“诸可易性”的需求变得越来越强烈,与此同时能耗约束也使得提供这些特性的代价变得越来越大。幸运的是,新的系统结构与新的技术创造了一个宝贵的转折点,让架构师有机会为这样的设计目标提供更高效的技术支撑。

**可验证性与可靠性** 保证软硬件可靠性比以往任何时候都重要;对一些植入式医疗设备而言,尤其事关生死。与此同时,CMOS工艺线宽缩小趋势使得电路变得越来越不稳定,复杂的异构系统也带来了“验证墙”的危险。所以,未来系统架构必须具备易于软硬件验证的特征。譬如,使用协处理器来检查端到端的软件不变量。目前,高度冗余的方案能效太低,我们建议研究低开销的、采用动态(硬件)检查软件中的不变量的新方法。总的来说,我们必须设计能够持续监视系统健康状况的方案(包括软件和硬件)并且实施可能的措施。最后,对于关键任务的场景(比如医疗器械),架构师必须重新考虑设计,以便保证失效情况下的安全运行。

**安全性与隐私** 对安全提供体系结构上的支

持可以追溯到几十年前,如分页、分段和保护环。近年来,一些扩展技术可有助于防止缓冲区溢出攻击,加速密码操作和隔离虚拟机。但是,现在到了自底向上重新考虑安全和隐私问题,并且定义系统结构级界面,以使硬件作为可信源更高效地支持安全服务的时候了。此类服务包括信息流追踪(减少边信道攻击)和高效执行更丰富的信息访问规则(增强隐私)。支持防干扰内存(tamper-proof memory)以及复制保护也是同等重要的议题。最后,既然安全和隐私问题与可靠/正确性本质上是息息相关的,那么这些方面的研究可以很好地结合起来。

**改善可编程性** 可编程性是指研制的软件应满足性能、能耗、可靠性以及安全性等指标,需要合理的设计以及维护投入。过去的几十年里,人们关注的是软件工程技术(比如模块性与信息隐藏),以性能和能耗为代价提高编程者的生产率<sup>12</sup>。当能效以及其他目标变得更加重要时,我们需要新的技术来削减编程抽象层次,从而消除不必要的低效。

已有的改善可编程性的技术(包括领域专用语言、动态脚本语言(如Python和Javascript)及其他)只是部分的解决方案。Facebook的HipHop可以动态地编译脚本语言程序,展现出即使存在这种抽象层的情况下如何保持高效的能力。除了跨越抽象层级的软件技术之外,我们还推荐采用**硬件**支持来改善可编程性的交叉研究方法。事务内存(transactional memory)就是一个最近的例子,它寻求大幅度地简化多线程代码中的并行化与同步操作。事务内存几乎横跨系统的所有层次,现在也逐渐成为商用主流。此外,还有些像软件调试、性能瓶颈分析、资源管理以及程序剖析(profiling)、通信管理等的议题需要投入研究。

管理应用之间的相互作用也带来了挑战。例如,这些应用如何表达服务质量(QoS)目标并且让底层的硬件、操作系统以及虚拟层共同工作来保障它们?对不断增加的虚拟化和自检(introspec-

<sup>12</sup> James Larus, "Spending Moore's Dividend", Communications of the ACM, May 2009 5(52).

tion)的支持要求协调整个软硬件系统层的资源管理,包括计算资源、互连和存储带宽。

### 跨层的交叉接口

当前的计算机体系结构定义了一系列接口,它们在这几十年里演变得很缓慢。这些接口(比如指令集架构(instruction set architecture, ISA)和虚拟内存)都是在存储作为首要考虑的因素,能量充足,软件基础架构相对简单,且几乎不考虑安全性的情况下定义的。拥有稳定的接口能够激发持续几十年的革命性系统结构创新。然而,我们目前处在技术的岔路口上,这些稳定的接口对文中谈到的很多创新而言已经成了障碍。

**更好的高层信息接口** 当前的指令集不能提供有效的方式来捕获软件意图或将关键的高层信息传递给硬件。例如,当一个程序对能效、可靠的安全性或服务等级提出需求时,指令集无法详述清楚。相反,现在的硬件必须自己来花大量能耗搜集这些信息,如指令集并行度或重复出现的分支结果序列。新的高层接口需要封装并将编程者和编译器的知识传递给硬件,从而获得更高的效率与有价值的新功能。

**更好的并行化接口** 开发并支持并行代码是个艰难的任务。编程人员一直被其中微妙的同步、死锁、任意的副作用、负载不均衡以及不可测的通信、不必要的非确定性、令人困惑的存储模型以及性能不透明性所折磨。我们需要可以使程序员能在更高层次表达并行度、局部性、计算依赖和副作用以及关键的共享和通信模式的接口。这种接口可使硬件更简单和有效,并具有高效的通信能力且能最大程度减少数据移动的同步原语。

**更好的异构性抽象接口** 支持异构并行需要新的接口。从软件角度来看,应用必须为多种不同的并行方式和存储模型进行编程,必须能够在不同的异构硬件组合之间可移植。从硬件的角度来看,我们要设计专门的计算与存储子系统。因此,需要新的硬件接口来对这些潜在的硬件所可能提供的关

键计算和通信部件进行抽象。这种接口层次要足够高,以支持软件的可移植的目标,又要足够靠近底层以高效地转换成一系列的硬件创新。

**更好的通信管理接口** 传统的计算机与编程模型主要关注计算的管理,但是目前数据通信越来越需要管理和优化。我们需要能够清晰地识别长期数据及程序依赖关系,以便软硬件调度器能够动态地识别代码中的关键路径。没有分析、管理与优化通信的能力,则没有办法保证性能、能耗与服务质量的目標。当考虑大数据应用场景时,由于涉及众多庞大系统之间的数据协作,数据管理会变得更加复杂。现有的系统缺乏合适的软硬件抽象层来描述通信关系。

**更好的安全性与可靠性接口** 已有的保护与可靠性模型不能满足当前的应用需求。我们要求接口能够在单一应用中为不同模块之间指明细粒度的保护边界,能够把安全当成第一级属性,并且能确定应用的弹性需求与期望。应用的某些部分也许要能容忍硬件故障,或者愿意为此冒险,以便运行起来能效更高。所有的这些接口都可以从合适的硬件机制中获益,比如信息流的追踪、不变量的生成和检查、事务恢复块(transactional recovery blocks)、可重构以及近似数据类型。这些都能显著提高效率。

## 结语

这份白皮书调研了计算机体系结构中的挑战和需要加强资助的有前途的研究方向,以开发更好的计算机系统,为信息通信技术革新创造条件,并改变我们的世界。

## 关于这篇文档

本文档在2012年4月和5月期间通过一个分布式的过程创建,协同写作则是在一个分布式编辑

<sup>13</sup> <http://www.cra.org/ccc/>.

器的支持下完成的。我们感谢计算社区联合体 (CCC)<sup>13</sup>，包括埃文 (Erwin Gianchandani) 和艾德 (Ed Lazowska) 给予的指导以及吉姆·雷诺 (Jim Larus) 和周以真 (Jeannette Wing) 提供的有价值的反馈。为整理成文作出贡献的研究人员用“\*”标记，总协调人用“\*\*”标记。

- Sarita Adve, University of Illinois at Urbana-Champaign \*
- David H. Albonese, Cornell University
- David Brooks, Harvard
- Luis Ceze, University of Washington \*
- Sandhya Dwarkadas, University of Rochester
- Joel Emer, Intel/MIT
- Babak Falsafi, EPFL
- Antonio Gonzalez, Intel and UPC
- Mark D. Hill, University of Wisconsin-Madison \*,\*\*
- Mary Jane Irwin, Penn State University \*
- David Kaeli, Northeastern University \*
- Stephen W. Keckler, NVIDIA and The University of Texas at Austin
- Christos
- Kozyrakis, Stanford University
- Alvin Lebeck, Duke University
- Milo Martin, University of Pennsylvania
- José F. Martínez, Cornell University
- Margaret Martonosi, Princeton University \*
- Kunle Olukotun, Stanford University
- Mark Oskin, University of Washington
- Li-Shuan Peh, M.I.T.
- Milos Prvulovic, Georgia Institute of Technology
- Steven K. Reinhardt, AMD Research
- Michael Schulte, AMD Research and University of Wisconsin-Madison
- Simha Sethumadhavan, Columbia University
- Guri Sohi, University of Wisconsin-Madison
- Daniel Sorin, Duke University
- Josep Torrellas, University of Illinois at Urbana Champaign \*
- Thomas F. Wenisch, University of Michigan \*
- David Wood, University of Wisconsin-Madison \*
- Katherine Yelick, UC Berkeley, Lawrence Berkeley National Laboratory \*

## 附录A 新兴应用的属性

有大量证据表明信息通信技术正日益加快地将许多诱人的设想从科幻变为现实。附录 A 列出了若干这样的设想，包括个性化医疗技术（为每个人提供不同治疗与用药方案）、复杂的社交网络分析（识别潜在的恐怖危险从而保障国土安全）以及远程呈现技术（减少人们在交通中产生的温室气体）。此外，可能还有很多重要的应用正在逐步兴起。试想一下，我们有多少人能在几年前预测到如今已是无处不在的社会网络？

尽管预测到和未预测到的这些未来应用将会有各种需求，但他们的共同特征都与早期的应用有所不同。当前的革新在传感器、智能手机 / 平板电脑和连接他们的数据中心，而不是在台式机上。新兴的应用有一些共同的极具挑战的属性，多数都是因为其产生数据的速度超过了当前处理数据的能力（性能、能量、可靠性及其混合）。表 A2 列出了一些属性，包括处理庞大的数据集，使用分布式设计，在尺寸受限的情况下工作，以及协调快速部署和高效操作的关系。

表A.1 新兴应用的举例

**以数据为中心的个人保健** 未来的医护系统应用植入、可穿戴和遍布全身的智能传感器，全天候地监察我们的健康（每周7天，每天24小时）。传感器内就地分析可以改进功能并通过减少通信节省设备功耗，而远程（即基于云计算的）系统可以跨越时间和患者群体进行资料聚合。这种系统使我们能够查询自己的健康数据，同时还能使医疗服务提供者不断监测病人并制定个性化的治疗方法。设计规划能够同时满足性能、功耗和能源约束的计算组织，合理地划分设备和云服务中各自实现的功能，以及在云环境中保护此类分布式的医疗信息，都是今后要面对的新挑战。

**计算驱动的科学发现** 今天先进的计算和可视化工具不断促使科学家和工程师开展模拟驱动的实验和挖掘庞大的数据集，以此作为科学发现的主要驱动力。生物模拟、蛋白质组学、纳米材料和高能物理等关键领域已经开始利用这些先进技术。随着科学研究组织开始利用真实的数据，与安全性和可重复性等相关的问题成为严峻的挑战。

**人类网络分析** 鉴于互联网和个人通信技术的发展，人们使用着十年前都难以想象的新方式进行交流。通过这些技术，人类交互产生了巨大的数据量，使我们能够分辨个体之间的行为和新的关系分类。网络科学和机器学习进展已经大大超出了计算平台有效分析这些大数据集的能力。高效的人类网络分析可以对一系列关键应用领域包括国土安全、金融市场和全球健康等产生重大的影响。

**更多** 除了这三个例子，在个性化的学习、远程呈现技术、交通、城市基础设施、机器感知/推理和增强的虚拟现实技术等领域的众多问题都在把今天的计算设施推向极限。这些问题领域的计算需求跨越了范围广泛的尺寸和体系结构，包括嵌入式传感器、手持设备和整个数据中心。

表A.2 新兴应用的属性

表A.1中的三大示例应用共享了至少3个关键属性，其展现出来的困难需要通过进一步的研究克服。

**大数据** 嵌入式传感器产生的连续数据流和由新技术带来的大量多媒体内容无处不在，我们面临数字化信息爆炸式增长。出于健康、商业和其他目的处理此类数据，我们需要高效平衡计算、通信和存储。为传感器提供足够的能力以便在生成/收集数据的地方筛选和处理它们（例如，区别正常和异常的生物信号），以达到最优能效，因为通信所需的能量大大超过计算所需。很多流源产生数据如此之快以至存储它们在成本上不可行，因而数据必须被立即处理。在其他情况下，环境的限制和聚合不同来源数据的需求影响了我们在哪儿执行这些任务。此类丰富的权衡激励出对混合架构的需求，它可以在节约能耗的同时有效地降低数据传输。

**一直在线** 为了保护我们的边界、我们的环境和我们自己，计算资源必须随时可用，并随时准备高效地提供服务。这一级别的可用性需要底层的硬件和软件提供可靠性、安全性和自我管理功能，而这些特征是日前大多数系统都不能提供的。虽然当前大型机和医疗设备都在努力实现5个9或99.999%的可用性（每年仅5分钟的不可用时间），但实现这一目标仍然要花费数百万美元。未来的解决方案需要在很多层面上实现这样的可用性，某些场合成本只有几块钱。

**安全和隐私** 随着我们对数据的依赖不断增加，我们会变得更加依赖平衡计算的性能，以确保可用、私有和安全的交易。信息安全是国家的优先考虑，而目前我们的计算系统极易受攻击。网络信息战不再是假象，通过互联网，对政府网站的攻击早已司空见惯。我们需要新的硬件系统、体系结构、固件和操作系统，以便为上面提及的应用中所需分布式信息的安全传递提供保障。

**译者：**

**包云岗：**CCF会员。中国科学院计算技术研究所副研究员。主要研究方向为计算机体系结构等。baoyg@ict.ac.cn

**隋秀峰：**CCF会员。中国科学院计算技术研究所助理研究员。主要研究方向为计算机体系结构等。suixiufeng@ict.ac.cn

**霍玮：**CCF会员。中国科学院计算技术研究所助理研究员。主要研究方向为软件可靠性、编译技术等。huowei@ict.ac.cn

**鄢贵海：**CCF会员。中国科学院计算技术研究所助理研究员。主要研究方向为计算机体系结构，容错计算，低功耗设计。yan\_guihai@ict.ac.cn

**王达：**CCF会员。中国科学院计算技术研究所助理研究员。主要研究方向为计算机体系结构微结构设计等。wangda@ict.ac.cn

**张磊：**CCF会员。中国科学院计算技术研究所副研究员。主要研究方向为可重构计算。zlei@ict.ac.cn

**谭光明：**CCF会员。中国科学院计算技术研究所副研究员。主要研究方向为高性能计算等。tgm@ict.ac.cn

**审校：**

**唐志敏：**CCF理事、CCF计算机术语审定工作委员会主任。中国科学院计算技术研究所研究员。主要研究方向为计算机体系结构和并行处理。tang@ict.ac.cn