高性能通用微处理器研发现状及发展策略

胡伟武[i] 张民选[ii]

1. 高性能微处理器体系结构技术

应用需求和半导体工艺水平的提高一直是微处理器发展的最主要动力。推动微处理器性能提高的因素主要有两个：更高的主频和更先进的体系结构。处理器主频的不断提升可以直接提高性能，但随着超大规模集成电路生产工艺的发展逐渐饱和并接近极限，处理器的性能提升还必须依赖体系结构技术的不断突破。

目前高性能微处理器在体系结构技术方面主要通过开发处理器中各个层次的并行性来提高性能，包括指令级并行性（ILP[1]）、数据级并行性、以及线程/进程级并行性。其中指令级并行性包括通过流水线技术实现指令之间重叠，通过多发射或超长指令字技术实现空间重复，以及通过乱序执行技术充分发挥流水线的效率。数据级并行性主要使一条指令完成对不同数据的多个相同操作，主要包括向量处理技术以及单指令流多数据流（SIMD[2]）技术。线程/进程级并行性包括单处理器的多线程/进程技术以及多处理器的多线程/进程技术。在线程级并行方面目前比较突出的工作包括多线程/同时多线程（SMT[3]）技术以及单芯片多处理器技术（CMP[4]）。

1.1 指令级并行性的开发

指令级并行性的开发主要包括超标量/超流水技术以及超长指令字/显式并行技术。

超标量技术全部利用微处理器硬件开发指令级并行性，自从八十年代出现产品以来，技术发展已十分成熟，积累非常丰富。目前，超标量微处理器在市场上占有绝对优势。RISC[5]结构的超标量微处理器主要有：Alpha、MIPS、、SPARC、PA-RISC，而 CISC[6]结构的超标量微处理器主要由英特尔和 AMD 公司开发，如：AMD 的 Opteron 和英特尔的 Pentium 系列等。研究表明可用的 ILP 在 4~10 之间，而目前通过超标量技术开发已经达到 2~4。单纯通过硬件进一步开发 ILP 已经非常困难，实现的复杂性所带来的性能下降已越来越严重，很难获得更大的性能增益。

---

[1] Instruction Level Parallelism

[2] Single-Instruction Stream Multiple-Data Stream

[3] Simultaneous Multithreading

[4] chip multiprocessor

[5] Reduced Instruction Set Computer—精简指令集计算机

[6] Complex Instruction Set Computer—复杂指令集计算机

为提高微处理器性能，超标量技术通常与超流水技术融合，以提高芯片的频率，如英特尔 Pentium 4 系列主频已经达到 3GHz 以上。但深度多级流水线技术也使流水线清空时的代价增大，对分支预测提出了更高的要求。

进一步开发并行性的另一种方法是软、硬件协作开发，主要有超长指令字（VLIW[7]）技术和显式并行指令计算（EPIC[8]）技术。软硬件协同开发指令级并行性能够降低硬件复杂度，提高效率，便于低功耗、简单实现。VLIW 技术完全依赖于软件编译开发，硬件只提供大量的计算资源，并不负责开发指令级并行性。VLIW 技术已经广泛应用于 DSP[9]领域。

惠普和英特尔于 1994 年开始合作开发基于 EPIC 技术的 IA-64 架构。一方面，英特尔/惠普通过大力开发先进的编译技术，使编译程序在比处理器硬件指令发射窗口更大的范围内，通过程序分析和指令重新调度把原始源代码转换成指令包序列。指令包包含有编译程序识别的能够并行执行的指令信息以及各种调度信息，允许编译器为处理器的执行过程提供提示，从而优化处理器的执行效率；同时，处理器硬件提供丰富的计算资源（多条功能部件流水线和大量的寄存器）。由于编译器事先进行了大量工作，因而能够在一定程度上简化处理器硬件设计，实现高并行性和高性能。2001 年英特尔/惠普推出了 Itanium，2002 年推出了 Itanium 2，2003 年推出了改进的 Itanium 2，实践证明 EPIC 技术具有广阔的发展空间。

VLIW 和 EPIC 体系结构对软件的压力较大。处理器的开发必须与系统软件的开发和应用开发紧密结合，微处理器的效率才能得以发挥。同时，由于 VLIW 和 EPIC 处理器的指令集与传统指令集有着极大的区别，为了实现指令集兼容性还必须研究二进制兼容问题。如：Transmeta 公司[10]的 Crusoe 芯片和 Elbruk 公司[11]的 E2k 芯片，采用二进制翻译技术解决兼容性问题。

由于其良好的兼容性，沿袭传统的指令级并行方法加速单线程应用的努力将不会终止。其目标在于进一步挖掘单个应用内在的指令级并行性。目前正在开发的这类处理器采用超标量/超流水技术、前瞻推测技术、增强取指和转移预测技术、踪迹高速缓存技术等。它们的共同之处包括：增加可以同时执行的指令数，芯片内安装多级大容量缓存，显著提高内存带宽，强化转移预测功能，提高对乱序执行（Out-of-Order）的支持，增加多媒体指令和专用电路，等等。但由于超标量技术越来越向复杂化方向发展，多功能

[7] Very-Long Instruction Word

[8] Explicitly Parallel Instruction Computing

[9] Digital Signal Processing—数字信号处理

[10] 一家位于硅谷专门从事芯片设计的公司

[11] 一家位于莫斯科的高技术公司

部件的控制复杂度随部件数量平方关系增长，导致资源利用率低、电路延时大，这大大限制了更高指令级并行性和资源利用率的开发，所以提高性能的潜力很快就会达到极限。

1.2 数据级并行性的开发

向量处理技术在高性能计算机系统领域广泛应用。Cray、NEC、HITACH、Fujitsu 等公司采用自行设计的专用向量处理高性能微处理器芯片不断推出新的超级计算机系统，在 TOP500 中占有相当地位。目前向量在面向科学计算、媒体处理等特定应用中仍有极大的开发潜力，如 T0[12]、VIRAM[13]等。在高性能计算领域，向量微处理器技术有稳定的市场，值得关注。

数据级并行性开发的一个重要领域是通用处理器中的媒体加速部件，如 Pentium 4 中的 SSE2 媒体加速部件。

1.3 线程级并行性的开发

加速多线程应用是通过开发线程/进程级并行性来提高性能的方法。基于芯片容量越来越大的前提，目前高性能微处理器研究的前沿逐渐从开发单线程的指令级并行性转向开发多线程/进程级并行性（TLP[14]/PLP[15]）。多线程/进程级技术通常与超标量等技术结合，构成可以开发不同级别并行性的高性能微处理器。线程级/进程级并行性的应用主要包括两大方向：一个是同时多线程（SMT）微处理器，一个是单芯片多处理器（CMP）。

SMT 技术的核心思想是在多发射处理器中，增加对多线程硬件自动切换的支持（如每个线程一个程序计数器 PC[16]等），增加发射宽度以及相应的硬件执行资源（如寄存器堆大小和端口数、高速缓存端口、功能部件的个数），并在此基础上增加功能，允许一个时钟周期发射的指令可以取自不同的线程，以提高流水线效率。

SMT 通过复制处理器上的结构状态，让同一个处理器上的多个线程同步执行并共享处理器的执行资源，可以比单线程更有效地提高应用并行性，最大限度地实现宽发射、乱序执行的超标量处理，提高处理器运算部件的利用率，缓和由于数据相关、控制相关引起的流水线堵塞，隐藏由于高速缓存失效、通信同步带来的访问内存延时。当没有多个线程可用时，SMT 处理器几乎和传统的宽发射超标量处理器一样。SMT 只需小规模

---

[12] 由加州大学伯克利分校电机系计算机专业和该校国际计算机研究所合作开发的定点向量流媒体处理器

[13] 由加州大学伯克利分校开发的流媒体处理芯片。采用 PIM(Processor in Memory)技术，由 4 条向量处理流水线(vector lane)组成.

[14] Thread-Level-Parallelism

[15] Packet Level Parallel

[16] Program Count—程序计数器

改变处理器核心的设计，几乎不用增加额外的成本就可以显著地提升效能。这对于桌面低端系统来说无疑十分具有吸引力。

英特尔 Pentium 4 Xeon 处理器中使用了超线程（Hyper threading—HT）技术；Alpha 21464 采用 8 发射动态调度超标量流水线，可同时处理 4 个线程；IBM Power5 和 Power6、Sun 公司的 UltraSparc 5 等都设计成 4 路 SMT 微处理器。Power5 是在 Power4 的结构上进行扩展形成的 SMT 微处理器，也采用双核结构，每个核是 8 路流出的超标量结构，支持同时两个线程的执行，具有 120 个重命名寄存器和 8 个执行单元。Power5 采用 64KB 两路组相连 I-cache 和 32KB 四路组相连 D-cache。

CMP 技术在片内集成多个微处理器核，通过多核并行执行的方式开发多线程/进程级并行。CMP 最初是由美国斯坦福大学提出的，其思想是将传统的对称多处理器（SMP[17]）体系结构集成到同一芯片内，各个处理器并行执行不同的进程。当半导体工艺进入 0.18μm 以后，线延时已经超过了门延迟，这就要求微处理器在设计时要尽可能划分成许多规模更小、局部性更好的基本单元结构。由于 CMP 结构已经被划分成多个处理器核，每个核都比较简单，有利于优化设计，是一种可随工艺水平发展灵活伸缩的结构，因此很有发展前途。目前，IBM 的 Power 4 芯片和 Sun 的 MAJC5200 芯片都采用了 CMP 结构。

在 IDF2004 上，英特尔展示了 Montecito 核心的 Itanium 处理器系统。这是英特尔第一款拥有双核心的处理器。Montecito 核心的 Itanium 处理器拥有 24MB 的缓存，支持超线程技术。同时，AMD 宣布已经完成了两款双内核处理器的设计工作。其中一款为 Opteron 双内核系列芯片，是针对服务器市场而推出；而另一款 Athlon 64 双内核芯片则针对台式机而设计，预计于 2005 年下半年推出。

就像 VLIW 融合了某些超标量处理器中的机制形成 EPIC 一样，SMT 和 CMP 技术的融合已经成为微处理器研发的主要方向。在业界，英特尔采用先突破多线程，再突破多核的技术路线，而 IBM 则采用了先突破多核，再突破多线程的技术路线，殊途同归。

因此，基于 CMP 技术和 SMT 技术，在单个芯片上集成多个 SMT 处理器核以及存储、I/O、互连、高速缓存控制器，融合形成多核多线程 SOC[18]微处理器是未来微处理器的发展方向。多核多线程 SOC 微处理器能够充分开发不同层次的并行性，进一步提高微处理器芯片的性能，形成超级计算系统的高性能微处理器构建块。

1.4 其他新型结构

[17] Symmetric Multi Processor

[18] system-on-a-chip –片上系统

另外，业界和研究人员还提出了一些新的思路，如向量 IRAM[19]处理器和可重构处理器。向量 IRAM 处理器由加州大学伯克莱分校 David Patterson 研究小组提出。他们认为存储器将是未来主要的性能瓶颈，提出将可扩展多处理器嵌入到片内的大型存储器阵列中，即所谓 PIM（Processor In Memory）技术，这样可使访存延时减少为原来的 20%到 10%，存储器带宽增加 50～100 倍以上。可重构处理器由麻省理工学院计算机科学实验室提出，其基本思想是在单芯片上用几百个带有某些可重构逻辑的简单处理器来实现高度并行的体系结构，此结构允许编译器为每种应用定制相应硬件。

最近几年，斯坦福大学提出的流体系结构（stream architecture）是一种新型的体系结构。所谓流（stream）就是大量连续的、不中断的数据流。流处理就是把流作为处理对象，将应用描述成以流互连的多个核。目前研制的流处理器有硬连线流处理器 cheops[20]，可编程流处理器 Imagine、Score[21]、Raw[22]等。其中，斯坦福大学的 Imagine 于 2002 年 4 月投片成功。Imagine 用 0.18μm CMOS 工艺实现，原片面积为 16mm×16mm，单片集成 2100 万个晶体管，处理器最高频率为 312MHz。采用 1.2V 低电压工作时，性能功耗比可以超过 5GOPS/W。在关键应用上，Imagine 可达到 10GFLOPS、20GOPS 的运算能力（工作频率 250MHz），功耗为 6W。在 Imagine 中，可编程概念得到了极大扩展。几乎所有的部件都是可编程的，包括 ALU[23]间的互连开关、簇间的通信网络、寄存器组织结构等。与专用 DSP 或传统的可编程 DSP 处理器相比，Imagine 流体系结构不仅提供了更灵活的可编程能力，还达到了可与专用芯片相比的卓越性能。

1.5 微处理器结构发展趋势

在过去 60 年的发展历程中，计算机的体系结构每 20 年左右就出现一个较大突破，已经经历了一个由简单到复杂，由复杂到简单，又由简单到复杂的过程。最早期的处理器结构由于工艺技术的限制，不可能做得很复杂，一般都是串行执行；后来随着工艺技术的发展，处理器结构变得复杂，流水线技术、动态调度技术、高速缓存（C）技术、向量机技术被广泛使用，典型的代表如 IBM 360 系列的机器以及 Cray 的向量机；RISC 技术的提出使处理器结构得到一次较大的简化；但后来随着工艺技术的进一步发展以及多发射技术的实现，RISC 处理器的结构变得越来越复杂。以英特尔和惠普为代表研制的 EPIC 结构的实现并没有从根本上对处理器结构进行本质简化。在上述过程中，每一次由简单到复杂的变革都蕴涵着进一步简化的因素，例如在早期的复杂处理器 CDC

[19] Intelligent RAM—智能随机存储器

[20] 麻省理工学院(MTT)媒体实验室(Media Lab.)开发的专门处理数字流媒体的可重构处理器平台

[21] 加州理工学院(CALTECH)开发的可重构流媒体处理器

[22] 麻省理工学院计算机科学实验室(MIT Laboratory for Computer Science)开发的芯片

[23] Arithmetic-Logic Unit--算术逻辑单元

6600 以及 Cray 向量机中，已经有了只由 load 和 store 进行访存的概念，IBM 360/91 中的 Tomasulo 算法被后来的 RISC 处理器普遍使用。同样，每一次由复杂到简单的变革，也蕴涵着再次复杂的基础，例如 RISC 结构的特点使得它可以充分利用多发射以及乱序执行来提高性能，而多发射和乱序执行又会增加处理器的复杂度。

目前，包括超标量 RISC 和 EPIC 在内的指令级并行技术使得处理器核变得十分复杂，通过进一步增加处理器核的复杂度来提高性能的空间已经十分有限。现在的 RISC 或 EPIC 微处理器普遍能允许几十到上百条指令乱序执行，如 Alpha 21264 处理器的指令队列最多可以容纳 80 条指令，MIPS R10000 为 32 条，HP 8700 为 56 条，Power 4 为 200 多条，Pentium 4 为 106 条（Pentium 4 处理器虽然指令系统是 CISC，但内部的微操作则具备了很多 RISC 的特征）。而 Itianium 处理器中则包含了多达 17 个的各种功能部件。现代 RISC 处理器结构的研究重点已经从处理器核转移到外围的高速缓存结构以及支持多处理器互连。如 Alpha 21364 采用 Alpha 21264 的处理器核，在片内增加了二级高速缓存以及多处理器互连的功能。

与指令级并行技术趋于成熟相呼应的是应用的变化。在 RISC 处理器发展的鼎盛期，处理器性能的提高主要是为了满足科学和工程计算的需求，非常重视浮点运算能力；而随着网络及媒体应用的普及，对处理器性能的需求发生了变化。

随着主频的不断提高，功耗问题越来越突出。现代的通用处理器功耗峰值已经高达上百瓦。如果说传统的计算机系统结构设计追求的是每秒运行的次数（performance per second）以及每一块钱所能买到的性能（performance per dollar），那么在今天，每瓦特功耗所得到的性能（performance per watt）已经成为越来越重要的指标。尤其在移动计算领域，功耗是压倒一切的指标。降低功耗需要包括工艺技术、物理设计、体系结构设计、系统软件以及应用软件的共同作用。

集成度的提高使得单片上可用的晶体管数越来越多，为处理器的设计者提供了更多的资源来实现更高性能的芯片。但处理器复杂度的增加将大大增加设计周期和设计成本。此外，集成度的提高意味着线宽变窄，从而信号在片内传输单位距离所需的延迟也相应增大。为了解决这两个方面的问题，芯片设计越来越强调结构的层次化、功能部件的模块化和分布化，即每个功能部件都相对地简单，部件内部尽可能保持通信的局部性。

因此，对处理器的微体系结构进行突破性变革的时机已经到来。根据上述处理器的发展趋势，这场变革应该是一场由复杂到简单的变革，应该在满足科学和工程计算需求的基础上，面向网络服务和媒体的应用，考虑低功耗的要求，采用层次化、分布化的结构降低设计的复杂度。

虽然片内多处理器及多线程技术正在成为处理器体系结构设计的热点，但这些现有

的单芯片多处理器的主要思想都是把传统的板级并行处理器结构下移到片内并在处理器设计上提供某些支持，并没有解决传统的多处理器系统固有的问题，如：串行程序移植困难、对于细粒度的并行应用并行开销大等。

2. 新工艺与新材料

2.1 新工艺

半导体工艺是微处理器发展的基础和最主要动力。VLSI 的发展对微处理器的影响是深远的，微处理器的研究直接受限于半导体工艺水平参数。为了保证处理器的持续快速发展，业界正在从不同角度研发新技术。这些新技术将成为支持未来 10 年微处理器发展延续摩尔定律的新动力。

作为微处理器的构成单元，晶体管技术的发展一直引人注目。开发高密度、高性能的处理器首先要设计高速的小型晶体管。当前引人注目的一些小型晶体管技术有超微晶体管技术、多栅晶体管技术、3D 芯片技术等。

目前，超微晶体管的尺寸已迈入纳米级。2000 年 12 月，英特尔公司率先在业界开发出栅极长度为 30nm（10^{-9}m=10^{-3}μm）的单晶体管；2001 年 6 月，英特尔又将这一纪录提高到 20nm；同年 11 月，英特尔宣布已开发出栅长仅为 15nm 的新型晶体管，同时单个晶体管的实际工作频率已经达到了 2.63THz（10^{12} 赫兹）。英特尔发布的 15nm 晶体管采用“耗尽型衬底晶体管（depleted substrate transistor）”的新型结构和绝缘层上覆硅（SOI）技术及“高 k 栅电介质（high k gate dielectric)”材料，从而使制造出的芯片的晶体管密度可以达到现有微处理器的 25 倍，运行速度提高 10 倍。2002 年 12 月，IBM 宣布了当前世界上最细小的晶体管加工技术。利用该技术生产出的晶体管栅长仅为 6nm。能够以如此之小的尺寸制造出可实际动作的晶体管，意味着芯片的晶体管数量可以达到现有微处理器的 100 倍以上。专家预测，采用这种传统作法的超微晶体管技术仍然会有 10～15 年的生命期。

多栅晶体管技术是一种新型电路结构技术。传统晶体管是每个晶体管有一个栅，用来控制电流在两个结构单元之间通过或中断，进而形成计算中所需的“0”与“1”这两个二进制数字。而多栅晶体管技术是每个晶体管的栅数量增加至两个或三个，这样，晶体管控制电流的能力（也就是计算能力）也相应地得到增加，并降低了功耗，减少了电流间的相互干扰。目前，英特尔、AMD 和 IBM 公司已分别在实验室成功开发出多栅晶体管。2003 年 9 月，AMD 公布了采用全耗尽绝缘硅（Fully-depleted SOI，FDSOI）、硅锗、三栅（Tri-Gate）和金属（NiSi）栅极的栅长 20nm 的硅晶体管。IBM 则已开始致力于将双栅晶体管技术用于芯片的生产，其硅锗生产工艺等方面的同步进展会加快双栅晶体管技术的产品化。英特尔于 2003 年 6 月在实验室实现了栅长 30nm 的三栅晶体管，预

计在 2007 年前后实现三栅技术的产品化，并逐渐使三栅技术成为未来生产尺寸更小、处理性能更强的芯片的关键技术。

3D 芯片技术是 IBM 公司、Matrix 半导体公司等研发的未来芯片技术。在一块芯片的设计中，这种技术将晶体管封装成两层或三层以上。这种技术通过充分利用立体空间，在差不多同样大小的芯片里，将数量成倍的晶体管封装进去，缩短了晶体管之间金属连接导线的长度，有助于增强芯片的性能。

微处理器尺寸的缩小和晶体管数量的增加也依赖于蚀刻技术的进步。随着晶体管技术的不断发展，目前使用的可见光电路光刻技术在未来 3～4 年内将达到物理极限。作为其替代技术，一种超紫外线光刻（Extreme Ultraviolet Lithography，EUVL）技术呼之欲出。由于超紫外线的波长要短于可见光，因此，刻写的“线条”更细，可以在较小的面积上集成更多的电路。利用这种技术，半导体厂商可以在芯片上蚀刻的线宽等级低于 0.1μm，甚至可以达到 0.03μm。目前，由英特尔、摩托罗拉、AMD、Micron、德国 Infineon 等多家公司和实验室成立的联合机构已开发出基于 EUVL 的光掩模制造技术。英特尔公司预计其采用 EUVL 技术的芯片将会集成 4 亿个晶体管，是当前 Pentium 4 处理器集成的 4200 万个晶体管数量的 10 倍。有人断言，摩尔定律能否续写未来的芯片发展史，关键取决于 EUVL 技术的成败。

2.2 新材料

为了配合不断改进的晶体管技术，业界也开始考虑后 CMOS 时代的技术措施，研究在硅元素失去优势时的替代材料。

目前，AMD 提出用金属材料——镍来代替传统的硅，因为镍栅能在栅的尺寸相同的条件下通过更多电流，而且比用其他金属材料更容易降低生产成本。而英特尔和 IBM 则致力于硅锗技术。硅锗技术通过在芯片的硅原子晶格之间填入锗，利用晶体应力提高迁移率，以提高芯片处理性能和降低功耗。IBM 利用硅锗技术已加工出运行频率高达 350GHz 的晶体管，并用这种晶体管生产出频率约达 150GHz 的网络通信芯片。

另外，IBM、英特尔和惠普等公司还在加紧对另一种有前途的材料——碳纳米管的开发研究。碳纳米管（carbon nanotube）直径只有 1~2nm，只是硅晶体管尺寸的 1/500。碳纳米管因其超常的能量及半导体性能而被认为最有可能在未来取代硅，成为生产晶体管及微处理器的主要材料。此外，碳纳米管投入运行时产生的热量和功耗都比晶体管要小得多。IBM 科学家已经研制出世界上最小的计算机逻辑电路——一个由单分子碳组成的双晶体管元件。惠普公司也申请到一项关于碳纳米管技术的专利。不过，碳纳米管还处在研究试验阶段，在产品中使用碳纳米技术的时间可能需要 10 年或更长。

随着微电子技术的不断进步，世界半导体技术的发展将会加速，半导体制造厂商将

会以更先进的技术加快升级换代，以适应新的市场需求。美国半导体工业协会预测，今年将采用 0.09μm 工艺，到 2007 年将采用 0.065μm 工艺，2010 年将采用 0.045μm 工艺，2013 年 0.032μm 工艺，2016 年 0.022μm 工艺。

3. 主流高性能微处理器及主流公司现状

3.1 主流高性能微处理器现状

SUN UltraSPARC III Cu 是一款 64 位超标量微处理器，采用 SPARC V9 体系结构，VIS II 指令集，兼容 SPARC 系列，时钟频率为 900MHz、1050MHz。其体系结构采用 4 路超标量发射，6 个部件，14 级非停顿流水线，数据通路、浮点/整数运算、地址运算全部是 64 位位宽；存储系统具有 32KB 一级指令高速缓存，64KB 一级数据高速缓存，集成 8MB 二级高速缓存控制器，2 路组相联，集成了存贮管理与控制器，4MB 页、1040 项 TLB，16GB 主存、2.4GB/S（？）、改善了访存延时，支持非对齐方式取指令和数据指令预取；集成 150MHz 系统接口，支持多处理系统，具有增强的 RAS[24]性能，能够进行复杂的错误检测与恢复。

HP PA-8700 支持 64 位 PA-RISC 2.0 指令集。其体系结构采用 4 路乱序发射，具有 56 项指令重排序缓冲器，采用双套功能部件（加载/存储、整数 ALU、整数移位/归并，浮点乘一加、浮点除/平方根五个部件都是双套）；存储系统具有 1.5MB 4 路组相联数据高速缓存和 0.75MB 指令高速缓存，采用 2KB 的指令分支预测高速缓存和 32 分支目标地址高速缓存，准 LRU[25]替换策略，支持数据预取。

HP RISC-PA8800 在单芯片上集成了两个完整的 PA-8700 微处理器核，每个核有自己的 1.5MB 一级高速缓存（0.75MB L1 I 和 0.75MB L1 D），共享 32MB 片外二级高速缓存，具有新的高带宽系统总线。

MIPS R18000 支持 MIPS4 指令集。其体系结构也是 4 路乱序执行，具有两套浮点（FP[26]）执行部件；存储系统支持 52 位虚地址和 48 位物理地址，支持大型 TLB[27]（64M/256M/1G 页），三级高速缓存层次：1MB 片内二级高速缓存，片上标签（tags）支持 64MB 三级高速缓存，采用多体、无冲突操作来改进高速缓存性能；支持高带宽系

[24] Reliability Availability and Serviceability–可靠性、可用性、保持业务不中断性能

[25] Least-Recently-Used--一种算法，其基本思想是优先清除“较不常使用的数据”，并释放其空间。

[26] Floating Point—浮点

[27] Translation Lookaside Buffer—页表转换缓冲

统总线和纠错码（ECC[28]）保护。

Alpha 21264 是 500~1000MHz 时钟频率的 4 路乱序执行超标量微处理器，具有 4 个整数部件和 2 个浮点部件，支持 motion video（动态视讯）指令。21264 采用巧妙的寄存器重命名机制来达到高频率以弥补乱序执行复杂度对频率的影响：它把 4 个整数部件分成两组，每组对应独立的 80 个物理寄存器，通过硬件实现两组寄存器文件一致性的维护。21264 具有复杂的分支预测机制：局部预测、全局预测和预测选项（决定使用局部或全局预测），使得它的预测准确率达到 95%。21264 设置 64KB 数据高速缓存和 64KB 指令高速缓存，各 128 项 ITLB[29]及 DTLB[30]，高速二级缓存接口和存储控制接口，支持 48 位或 43 位虚地址，44 位物理地址，64 位数据总线，支持奇偶校验和纠错码。

Alpha 21364 在 Alpha21264 内核基础上集成 L2 高速缓存、两个存储控制器、以及多处理器路由接口等。存储控制器使 Alpha 21364 可与外部 RAM 直接相连，能够提供 12.8GB/s 的访存带宽，支持基于 64 项目录的高速缓存一致性协议；多处理器网络接口可使多个处理器直接相连，具有 4 条通信链路，每条带宽 6.4GB/s，处理器间采用异步时钟，通信延迟仅为 18ns，支持自适应路由的乱序网络通信，每个跳步都支持纠错码，还支持 3GB/s 的 I/O 带宽。

Itanium 2 起始主频 900MHz，超长指令字 6 发射，具有大量的计算资源（11 条功能部件流水线、增强的浮点能力和大容量寄存器文件）和存储资源（片内集成大容量的三级高速缓存，容量在 1.5MB 以上）；硬件与软件结合提高性能，硬件提供分支和存储器提示，编译流水优化，指明操作部件等，支持多级分支预测、数据前瞻和控制前瞻；使用基于 Infiniband 技术的总线结构，交换式系统总线，系统带宽在 2GB/s 以上。Itanium2 拥有 128 个整型寄存器，128 个 82 位浮点寄存器，64 个谓词寄存器，8 个分支寄存器。通过寄存器栈引擎机制（RSE[31]）旋转使用所有寄存器，Itanium 能够提供逻辑上无限数目的寄存器，从而大大提高处理能力。

AMD Opteron 兼容 x86-64 指令集，吸收了 RISC/CISC 的特点。其体系结构采用四级宏流水线共 32 步操作，12ns 总执行时间；具有大型高速缓存（1MB）和有效的分支预测机制；集成 128 位双通道存储控制器，每个通道支持 4 个 DDR[32] DIMM[33]s，访存带宽高，延时小，支持 PC1600，PC2100，PC2700 DDR 内存。带宽高达 5.3GB/s，在扩

[28] Error Correcting Codes

[29] Instruction TLB

[30] Data TLB

[31] Register Stack Engine

[32] double data rate--双倍速率

[33] Dual-Inline-Memory-Modules 双直列存储器模组

展到多处理系统时，可以提供的带宽和处理器数目成正比；AMD 尤其注重处理器 I/O 性能的改进，其高速 HyperTransport[34]和外设联结可能是 AMD64 系列处理器的精髓所在，Opteron 内集成了三个 HyperTransport 高速互连接口，提供了无缝扩展能力，支持高带宽的 I/O 和处理器间一致性协议，处理器间通讯口 16 位，每秒 1600MT（megatransfers/s），提供带宽 3.2GB/s；采用增强的 RAS 技术，包括：一级数据高速缓存、二级数据高速缓存的数据和 tag 的纠错码、DRAM 受纠错码保护、Chipkill 纠错码支持从整个 DRAM 芯片的错误中恢复，只读数据采用奇偶校验，机器校验体系结构 MCA[35]。

IBM Power4 处理器核支持 PowerPC 64 位体系结构，超标量乱序 8 指令发射，1350 MHz 超高频设计，支持大型物理寄存器，支持重命名，数据指令预取和指令预测执行；单芯片上集成两个处理器核，MCM[36]可达 8 个处理器，片上集成多处理器互连接口、集成 I/O 控制器、集成 L3 高速缓存控制器和集成存储控制器。

3.2 主流公司现状

影响高性能微处理器发展的因素不仅包括技术的优劣，还与产品的推广、产业环境、技术的影响度及经济政治等各方面因素有关。当前微处理器正从 32 位向 64 位过渡，各大厂商面向不同的应用大都在开发 32 位和 64 位的微处理器产品。作为一种用于装备高端计算机系统的芯片，64 位微处理器被广泛应用于一些关键应用领域。

目前通用高性能微处理器技术由三种主要的微处理器体系结构引领技术竞争，即英特尔的 IA-64、AMD 的 x86-64 以及 IBM 的 PowerPC，它们分别代表着 EPIC、超标量 CISC 以及超标量 RISC 三种体系结构的竞争。超标量/超流水技术在设计、生产和应用方面都已具有相当的积累，被目前主流的通用高性能微处理器所采用。但从技术上来讲，程序固有的最大指令级并行性限制以及复杂的乱序执行技术也使得 64 位超标量处理器的缺点逐渐暴露：乱序执行要求处理器具有较高的智能和复杂的逻辑，使得芯片的结构越来越复杂，妨碍了主频和性能的提高，设计难度和成本越来越大。从市场上来讲，各厂商的超标量体系结构自成一体，每个厂商都需要在芯片、操作系统上进行大量的投入。

英特尔和惠普联合为高端服务器和工作站市场开发出全新的、开放性的 64 位体系结构 EPIC 是微处理器发展的一个新方向，IA-64 体系结构在吸收超标量体系结构经验教训基础上另辟蹊径，从一开始就走开放性的道路，充分利用软硬件协同能力来提高指令并行度，技术上具有强劲的竞争势头。从市场发展趋势、英特尔的设计思想、制造工

[34] AMD 公司新的高速高性能母板集成电路间点对点连接，与 PCI 管座插脚数相同但速率比 PCI 高得多。

[35] Machine Check Architecture

[36] MultiChip module 多芯片组件

艺、开放性、批量生产能力、推广力度以及 OEM[37]和 ISV[38]支持程度等方面来分析，IA-64 体系结构可能发展成为高端应用的主流平台之一。

在 Mckinley（Itamium 开发代号）处理器的开发期间，英特尔开始了 IA64 处理器的商业化运作，但其他厂商和用户接纳 Itanium 2 和它的后继产品却保持了比较谨慎的态度，这可能是由于 IA-64 的研制经过了数年的失误和推迟（花费 8 年时间，投资 80 亿美元）。最终 IA-64 虽然没有像早期那些 EPIC 狂热者所预计的那样，带来一场计算机体系结构设计革命，但是因为它的积极主动和广泛的 OEM 支持，IA-64 仍然具有较强的生命力。由于 AMD 成功推出基于 x86-64 的 Opteron 和 Althon64 处理器，加之 Itanium 的市场反映并没有预期的好，英特尔也开始在桌面应用领域推出 x86 的 64 位版本处理器。处理器的开发成功与否除了技术因素之外，也取决于设计公司的市场推动能力。在微处理器领域，英特尔的地位目前还无人能敌。

AMD 公司，设计和销售一代又一代向后兼容的 x86 处理器已经成为英特尔公司在半导体行业取得霸主地位获得成功的真正法宝，而在其废弃了他们一贯的政策介入了 IA64 体系结构的时候，AMD 公司在努力探索 x86 扩展到 64 位 x86-64 的方法，虽然 AMD 公司在 x86-32 系列上也面临着较为严重的困难，但他们拾起英特尔抛弃的东西，向 64 位计算领域迈进的步伐是坚定的，目的是全面支持工作站和服务器上的各种应用。AMD 这样做的好处是可以最充分地利用现存的所有开发工具和应用软件，而且设计和生产的成本都比较低。由于核心对 32 位和 64 位应用程序有很好支持，所以从用户的观点来看，AMD 的方案从 32 位到 64 位的移植是无缝进行的。而且在技术上，x86-64 目前的优势可能更在于它卓越的 I/O 设计。在 64 位领域，x86-64 和 IA－64 的技术竞争可以说代表着高性能微处理器发展的最主要方向之一。

IBM 公司，微处理器领域另一个有力的竞争方是 IBM。IBM 拥有自己的入门级产品 PowerPC 和在高度自动化的微处理器开发工艺上采用 0.13μm 工艺实现的 Power4+，其 P 系列处理器主要在本公司的超级服务器中使用。中档服务器中则采用 IA-64。IBM 还宣布就微处理器设计技术的发展与 AMD 合作，使其可能成为 x86-64 的首个主要 OEM。这些使得 IBM 公司在微处理器的竞争中处于比较有利的位置。

Sun Microsystems 公司，目前公司面临经济下滑以及来自 DELL 和 IBM 等公司的竞争压力，收入和两年前相比降低了 32%。S 公司的市场在低端基于 x86 的服务器和高端 IBM 基于 Power4 处理器的服务器之间受到严重挤压。随着 0.13μm 工艺快速进入市场，IBM 公司很快就在单处理器性能上挑战 Sun 公司，将其 UltraSPARC 生产线置于尴

[37] Original Equipment Manufacturer--原始设备生产商

[38] Independent Software Vendor--独立软件开发商

尴的境地。

其他公司，惠普公司已经宣布原 Compaq 的所有 Alpha 处理器开发和服务都将逐步中止（至 2010 年），而原先 EV8 的开发小组已经完整地被英特尔招募了并且马上就投入了 IA64 构架的研发工作。Alpha 退出历史舞台已经是不争的事实。2001 年 9 月惠普公司在经过一笔富有争议的交易中合并了 Compaq 公司，这次交易甚至导致了惠普管理层与包括惠普家族在内的部分股东的长时间的剧烈冲突。惠普公司现在将要面对这个棘手的问题，即在扩展基于新的 64 位构架（IA64）系统生产线的同时，是否还要生产正在走下坡路的另外两个系列的产品（PA-RISC 和 Alpha）。虽然 SGI(另一款 RISC 芯片--MIPS 的开发者)和惠普公司都承诺将会推出更新更快的 RISC 产品，但是他们的处理器生产线和英特尔公司 IA64 生产线之间的性能差距巨大。

4. 国内微处理器发展状况及发展思路

4.1 目前发展状况

通过“十五”期间的发展，我国在通用处理器设计方面已经有了一定的基础，研制出多款通用微处理器和嵌入式微处理器，但还缺乏国际竞争力，要真正有所突破还任重道远。

在中国科学院知识创新工程和 863 项目的资助下，中国科学院计算技术研究所先后完成了 32 位的龙芯 1 号和 64 位的龙芯 2 号通用处理器的研制，掌握了一大批通用处理器的核心技术。其中龙芯 2 号处理器于 2003 年 8 月在上海中芯国际（SMIC）首次流片成功。到 2004 年底龙芯 2 号经过多次流片后完成了 0.18μm 的定型芯片的设计。龙芯 2 号为 4 发射超标量超流水结构，拥有 5 个强大的功能部件，片内一级指令高速缓存和数据高速缓存各 64KB，片外二级高速缓存最多可达 8MB。为了充分发挥流水线的效率，龙芯 2 号实现了先进的转移猜测、寄存器重命名、动态调度等乱序执行技术，以及非阻塞的高速缓存访问和 Load Speculation（风险加载）[39]等动态存储访问机制。在正常电压下龙芯 2 号的最高频率为 500MHz，功耗为 2~4W。用 SPEC CPU2000 进行实测，性能是龙芯 1 号的 8~10 倍，是 1GHz 的威盛处理器的 2 倍，已经达到中档 Pentium III 的水平。目前芯片样机能运行完整的 64 位中文 Linux 操作系统，全功能的 Mozilla 浏览器、多媒体播放器和 OpenOffice 办公套件，具备了桌面 PC 的基本功能。

2003 年，国防科技大学承担了 863 计划《高性能通用 CPU 芯片》重大研究课题

[39] 即在取数指令访问高速缓存后，它前面的存数指令地址还没有确定(即取数指令从高速缓存中取回的值有可能是错误的)时，先把取数指令从高速缓存中取回的值送给后续的指令用，如果后来发现它前面的存数指令和该取数指令访问的是同一个单元，再取消该取数指令后面的指令。

（YHFT64-I）。YHFT64-I 采用目前国际上最前沿 EPIC 技术，采用软硬件方法与英特尔指令集兼容，能够并发执行 8 条指令。芯片设计采用了大量先进的微体系结构技术，如：多级分支预测技术、寄存器堆栈技术、控制前瞻/数据前瞻、谓词执行技术以及低功耗技术等，能够有效开发指令集并行性，极大提高处理器性能。YHFT64-1 设计工作主频不小于 300MHz，CINT2000 和 CFP2000 的测试结果都不低于 300；支持通用操作系统，支持多处理器结构，支持数据库、WEB 等服务器应用。该芯片所有设计已经完成，采用 Chartered 0.13μm Nominal 1P8M 工艺，核心逻辑规模 5800 万晶体管，采用 HPBGA 封装，功耗 12W，面积 10×10mm^2，引腿 696 个，工作频率 300MHz ，在硬件仿真环境下成功运行 Linux 操作系统和 SPEC2000 测试程序，将在近期投片。

在科技部的支持下，中科院计算所正和国防科大一起设计主频为 1GHz 的 64 位处理器及基于该处理器的服务器和万亿次计算机系统，2005 年底完成该处理器的研制，其综合性能将达到 1.5GHz~2GHz 的 Pentium 4 的水平。

4.2 清醒认识我国微处理器技术水平

我国对于要不要做微处理器曾经有过很长一段时间的争论。反对自主研制微处理器的理由主要有三点：第一，中国的集成电路工业很落后，不仅制造工艺落后，而且设计技术也不高，人才匮乏，缺乏自主开发微处理器的基础；第二，国际上微处理器已经成为一个很成熟的产业，技术积累丰厚，专利保护严密，我们很难突破业已形成的技术壁垒；第三，微处理器的发展需要软件和操作系统的配合，我国要解决微处理器已属不易，如要同时解决操作系统和各种应用软件就更为困难，尤其是操作系统有先入为主的市场规律。因此自主开发的微处理器如走兼容的道路就无法绕开别人的知识产权，如走不兼容的道路，则面临着应用市场开发的严峻课题。在这种观念的主导下，我国在“九五”期间没有安排微处理器芯片的研究和开发，失去了几年宝贵的时间，痛失良机。

在“十五”初期，占上风的观点是我国应当开始微处理器的研发，但应当以研发专用的嵌入式处理器和 SOC 为主，高性能通用处理器的研发则处于次要地位。通过“十五”对通用微处理器和嵌入式微处理器的研究和开发，目前很多人形成了另外一种结论：认为发展自主知识产权的微处理器不是一件难事，并断言中国自己的微处理器很快可以赶上国际先进水平。这种观点的主要依据是：第一，中国的集成电路制造技术已经发展到了 0.18μm，很快可以到 0.13μm，与国际上最先进的集成电路制造技术的差距已经大大缩小。集成电路的设计技术已经有了很大的进步，获得最先进的设计工具已非难事。第二，虽然微处理器开发中有不少的知识产权壁垒，但是可以通过使用自主指令集的方法避开；尤其是开发嵌入式微处理器，可以避免知识产权的纠纷；第三，如果采用 Linux

操作系统，则可以避免没有自主操作系统的尴尬并获得比较丰富的应用软件支持。在这种观念的主导下，微处理器热已成为近年来集成电路行业的一个显著特点。无论是业界内部、还是业界外部人士都普遍持有一种乐观的情绪，认为中国自主的微处理器可以很快脱颖而出。

客观地讲，这两种观点所依赖的论据都有一定的道理，但是结论都有所偏颇。近年来，媒体不断报道我国芯片研制成功的消息，鉴定成果都是国际先进、国内领先，这种被某些媒体称为“群体突破”的表象，给人们一种感觉：似乎中国的微处理器技术已经赶上世界先进水平，长期以来人们十分担心的中国信息产业缺“芯”的局面似乎已基本改变。但实际情况是，高性能通用微处理器的研制是一个与计算机体系结构、超大规模集成电路设计、半导体工艺、验证和测试、操作系统和编译等等都密切相关的综合性技术。虽然目前我国在这几个方面都有长足的发展，与世界先进水平的差距在不断缩小，但是因此认为微处理器问题已经解决是过于乐观了。目前绝大多数国内设计的微处理器还没有真正用起来，很多甚至还只是实验室级的样品。

从技术和实现的角度看，国内与国外同行相比最大的差距，不在于上述微处理器体系结构、电路设计、工艺、验证、系统软件等单项技术上，而在于这些技术之间的结合部，在于对这些技术能够融会贯通的人才的缺乏。如国内部分研究单位在处理器的微体系结构（micro-architecture）研究方面有很强的实力，在微体系结构的核心技术如存储层次、动态调度、转移猜测等方面都有很深入的研究，但真正把这些技术在软件上和操作系统、编译等系统软件以及在物理上和硅片上晶体管、连线的行为结合在一起考虑则少有研究，技术整合实现上能力不高。设计处理器的体系结构，既要充分了解与处理器联系紧密的操作系统、编译器以及应用程序的原理和行为，以便在设计过程中在软硬件功能划分、硬件复杂度和性能之间做合理的权衡；又要充分考虑到所设计的模块和功能部件的电路结构及该电路涉及的晶体管和连线的物理行为，以便做到流水级之间的延迟平衡以及进行性能、设计复杂度、功耗等的取舍。否则，所研制的微处理器作为产品是不会有竞争力的。

从产业化和应用的角度看，国内与国外同行相比的最大差距，在于产业链的不完善。处理器芯片虽然是计算机系统的核心部件，但决不是计算机系统的全部。计算机系统还包括有关的套片、各种接口芯片、系统软件以及相应的应用等，即使是像英特尔和 IBM 这样的企业，也没有能力包揽一个系统所有软硬件的设计。越是核心技术，产业链越长，而产业链的建设需要时间和投入。国内处理器芯片研发的一个重大障碍是缺少像微软这样的系统和应用软件厂商。我们往往希望一个处理器芯片今年研发成功，明年就在市场上有很大作为，这是不现实的。即使像 IBM 和英特尔这样的芯片厂家从一个芯片研发成功到投入市场也常常需要 1~2 年时间。

造成对国内微处理器设计能力判断失误的一个重要原因是我国习惯采用鉴定会或新闻发布会方式公布成果，而不采用国际上通行的标准测试（Benchmark）。另外，国内舆论片面追求微处理器的运行频率而不是实测性能，其原因在于运行频率是非专业人员衡量微处理器性能的最直观手段。实际上，运行频率仅仅只是决定微处理器性能因素之一。如在相同性能的前提下，AMD Althon 处理器的频率就明显低于英特尔的 Pentium 4 处理器。不管是通用微处理器还是嵌入式微处理器，国际上都有公认的性能测试标准，即 SPEC CPU 2000 和 EEMBC。世界上的大芯片厂商都参加了 SPEC 和 EEMBC 组织，公布每一款芯片的性能测试结果。这些标准测试程序是在全世界范围内向实际用户征求来的（目前正在征集 SPEC CPU 2004 测试程序），经过了严格选择。比如 SPEC CPU 2000 共 14 个定点程序、12 个浮点程序，代码量达 600MB，包括科学计算、图像压缩、编译等各种应用，EEMBC 包括家电、网络、汽车、办公等几类应用。这两种标准都采取几何平均方法计算性能，这就意味着只要有一个应用程序性能很差，SPEC 或 EEMBC 分数就很低。以国内最高性能微处理器为例，SPEC 值只有 200 分左右，与国外主流微处理器也还有较大差距。

总之，我国在微处理器设计上还刚起步，在高性能通用微处理器领域要形成具有国际竞争力的产品，必须付出极其艰苦的努力，还必须团结国内外有志于高性能微处理器攻关的有生力量，集中全国力量办大事，选择突破口集中突破，而不是遍地开花，追求表面上轰轰烈烈的“群体突破”。

4.3 我国微处理器技术发展的策略

在目前的情况下，我国微处理器技术发展应该遵循什么样的策略？这是一个非常重要的问题。

（1）克服浮躁

克服浮躁的情绪和表现，是中国 IT 界的首要问题。现在没有人不认同微处理器是一个关系巨大政治经济利益的高科技产品，也没有人不希望尽快改变信息产业“无芯”的尴尬局面，甚至希望“中国芯”一步跨越世界水平；然而，被忽略的是：发展微处理器必须有一个长期的发展战略，必须打持久战。在战略上，我们要自尊自信，相信依靠中华民族的智慧可以做出高性能的微处理器；但在战术上，我们不能自高自大。关键是必须避免出现浮躁的表现、虚假的概念。要谦虚谨慎、戒骄戒躁、踏踏实实、认认真真。浮躁只会使我们的差距加大。

（2）集中力量办大事，形成优势合力

高性能通用微处理器设计是集成电路设计中最复杂，也最具有挑战性的工作。高性能微处理器的设计需要人才与经验长时间的积累。同英特尔公司 30 余年的积累相比，

国内研究单位无论在人才、资金、设计经验都是缺乏的。因此，我国在微处理器方面必须借鉴研制“两弹一星”集中力量办大事的成功经验，集中全国在微处理器方面的主要力量，形成优势合力，选择突破口集中突破。在组织实施方面，采取“总体规划，分步实施；统一设计，分工实现” 的组织实施思路，成立跨单位的强有力的技术总师组，进行总体技术方案的设计。并在统一的总体设计的基础上由各单位分别负责其中的一部分。避免使大项目成为大口袋，各单位都把自己想做的工作装进去，最后形成不了整体。

（3）优先发展高性能通用微处理器

应当优先发展高性能通用微处理器并将其列入我国微处理器发展的长期战略目标。从长远来看，我国应当在微处理器方面做出原创性的成果，这需要从基础研究层面开始，从高性能通用处理器入手，走跨越式发展的技术路线。根据国外的成功经验，CPU 的发展是“先高后低”，即先发展计算机中使用的高性能通用 CPU 芯片，突破其中的关键技术，在此基础上再发展嵌入式的应用。因此，在处理器设计上应以高性能通用 CPU 发展为龙头，通过高端通用 CPU 的发展来带动整个 CPU 产业的发展。

具体实施时，在目前达到与国外相同主频的客观条件不具备的情况下，应走通过优化处理器结构提高性能的道路，强调处理器中各个层次的并行性（包括指令级并行性、数据级并行性、以及线程级并行性）的开发以及软硬件的协同来提高性能。在指令系统的选择上，我们应该采取基于自主创新的结构，实现应用兼容的发展战略，结合自主设计与兼容性设计两方面的优点，开发自主结构的微处理器，同时在多个层次采用新型的兼容性设计，实现与主流 CPU 的应用级兼容。

中国人富于天才和智慧，只要国家决心发展微处理器，弘扬卧薪尝胆的精神，持续支持我国两到三支不怕坐冷板凳的顶尖微处理器设计队伍，一定能做出我国自己的高性能微处理器，圆我国自主知识产权微处理器之梦。

（4）从“为芯片做系统”到“为系统做芯片”

在十五期间，我们在高性能通用处理器的研发方面一直以来都是采用“为芯片做系统”的方式，即先定义芯片的性能等指标，再看这样的芯片能够用来做什么系统。现在，这种做法的弊端已经越来越显示出来。在这种思路指导下，无法开发能够包揽所有应用的通用微处理器。正确的方法是从具体面向的应用开始，“为系统做芯片”，即先定义系统的要求，再根据系统的要求设计芯片。系统的含义包括两个层次，一是应用系统的含义，如电子政务系统、Linux-PC 系统等；二是系统硬件及软件，如使用何种套片，支持何种 I/O 接口等等。军口为型号配套研制芯片是一条可以借鉴的成功经验。

（5）分阶段实施

根据目前应用的需求，我们设计的处理器芯片所面向的系统分别为高性能PC、SMP服务器以及高性能计算机系统。根据上述三类系统的要求，我们的高性能通用处理器芯片研制可以分如下三个阶段进行。

第一阶段时间跨度是从现在到“十五”末（2004年~2005年），在目前国内已经研制和正在研制的高性能通用微处理器的基础上进一步优化结构和提高主频，争取“十五”末微处理器的工作频率达到1GHz左右，SPEC CPU2000的实测性能达到和1.5GHz~2.0GHz的Pentium 4相当的水平，同时设计配套芯片组和相关系统应用软件，研制成功高性能桌面个人计算机，在“十一五”期间形成一定规模的产业。

第二阶段时间跨度从现在到“十一五”初（2004年~2006年），突破单核、单线程SOC实现技术。其中，每个SOC中嵌入一个高性能通用CPU核以及桥控制逻辑和各种接口控制逻辑，如存储器控制逻辑DDR、I/O控制逻辑PCI[40]、HT、AGP[41]等。此外还有直接支持多处理机互连的SOC，通过多处理器互连网络路由构成支持全局存储访问的多处理机系统，解决服务器用CPU问题。

第三阶段时间跨度为“十一五”初期到“十一五”末（2006年~2010年），突破多核多线程SOC技术。到“十一五”结束时设计出主频为2GHz~4GHz，片内峰值运算速度为每秒1000亿次浮点运算的多核多线程芯片。基于上述单核和多核高性能SOC芯片，在“十一五”初研制成功万亿次高性能计算机系统，在“十一五”末研制成功百万亿次高性能机系统。并下大力推广应用，着力解决国防和国民经济的核心重要应用领域依赖国外微处理器问题。

5. 结束语

发展自主知识产权的微处理器是国家的需要、时代的需要，是产业发展的需要，且我们已经初步具备了相应的条件，再没有理由怀疑我们能够做成这件事。但是也应该清醒地看到发展自主微处理器是个有相当难度的工作，需要长时间的积累。无所作为的怀疑论和盲目乐观的速胜论都是错误的。在战略上藐视它，在战术上重视它是我国发展自主知识产权微处理器时应遵循的一个基本原则。

致谢感谢国防科技大学邢座程教授对本文的大量原始贡献。

[i] **中国科学院计算技术研究所 北京 100080**

[ii] **国防科技大学计算机学院 长沙 410073**

[40] Peripheral Component Interconnect—外部设备组件互联

[41] Accelerated Graphics Port—图形加速端口