改善系统能量效率的体系结构方法:并行处理

易会战 刘永鹏
(国防科学技术大学计算机学院计算机研究所 长沙 410073)

摘 要 因为对高性能微芯片和系统设计的广泛影响,能量消耗问题受到计算机界越来越广泛的关注.多个层次的技术被用于改善系统的能量效率,并行处理是体系结构层提高能量效率的主要手段.并行处理使用性能适中的计算节点减少能量消耗,使用多个节点并行执行维持高吞吐量.文中分析了并行处理提高能量效率的基本原理,给出了并行处理的时间开销和能量开销模型.基于模型分析,对低电压并行系统、动态电压调节(Dynamic Voltage Scaling,DVS)的并行系统和多核微处理器 3 个并行处理方向进行了展望,给出了这些并行处理方向改善能量效率的空间.

关键词 并行处理;能量效率;动态电压调节;低电压设计;多核处理
中图法分类号 TP311 **DOI 号**: 10.3724/SP.J.1016.2009.02475

An Efficient Architecture Method for Improving Energy Efficiency: Parallel Processing

YI Hui-Zhan LIU Yong-Peng
(*Institute of Computer Technology, School of Computer, National University of Defense Technology, Changsha* 410073)

Abstract Energy consumption has been paid increasing attention to in the computer domain because of its deep influence on the design of high-performance chips and systems. Many techniques are proposed to improve energy efficiency of computer systems, and in the paper the author focuses on parallel processing on architecture level. Parallel processing improves energy efficiency by using some computing nodes with moderate performance, which maintain high throughput by parallel execution. In this paper, the authors present the fundamental of parallel processing improving energy efficiency, and models the time and energy overhead involved in parallel execution. Based on the models, the author investigates low-voltage parallel systems, parallel systems with dynamic voltage scaling, and multi-core microprocessors, and reveals their potential of improving energy efficiency.

Keywords parallel processing; energy efficiency; dynamic voltage scaling; low-voltage design; multi-core processing

1 引 言

当前,能量消耗问题阻碍了计算机系统性能的进一步提高.首先,大量的能量消耗提高了芯片的封装和制冷代价,为了维持稳定的运行,高性能微处理器需要采用新的封装和制冷技术.其次,环境温度和计算机的可靠性有紧密的联系,大量的能量消耗导

收稿日期:2006-04-16;最终修改稿收到日期:2009-11-11.本课题得到国家自然科学基金"软件指导的高性能计算机系统功耗和热量管理"(60903059)、国家"八六三"高技术研究发展计划项目"面向片上多处理器系统的程序设计环境"(2008AA01Z110)、国家科技重大专项(2009ZX01036-001-003-001)及高效能服务器和存储技术国家重点实验室开放基金项目(2009HSSA04)资助.**易会战**,男,1976 年生,博士,主要研究方向为低功耗编译优化、并行程序设计语言.E-mail: huizhanyi@nudt.edu.cn. **刘勇鹏**,男,1977 年生,博士研究生,主要研究方向为操作系统、功耗管理、系统容错.

致运行环境温度增加,最终降低了计算机系统的可靠性[1].最后,能量消耗的直接结果是大量的电力消耗,这增加了高性能系统的运行成本[2].

如何改善计算机系统的能量效率在多种层次上得到了广泛的研究[3-5].在计算机体系结构层,采用并行处理技术提高系统能量效率受到了最多的关注.并行处理使用性能适中的计算节点提高系统的能量效率,使用多节点的并行执行维持高的系统吞吐量.带有动态电压调节(DVS)能力的并行系统进一步提高了系统的灵活性,使得节点的性能可以根据程序的执行需求动态变化,节省了系统的能量消耗.

并行处理在提高能量效率方面的工作包括以下3个方面.首先是对低电压并行系统的尝试.Blue Gene/L[6]曾一度是运算速度最快的高性能并行系统,该系统采用了嵌入式 PowerPC 440 ASIC 芯片作为处理节点,每个节点提供了适中的性能,同时全系统的65536个节点提供了高的系统吞吐量.Blue Gene/L 的单机柜设计功耗为25kW,设计性能为5.6Tflops,节点能量效率为224Mflops/W.相比之下,同时期的 ASCI 系列系统(Red, White, Q)的能量效率只有1Mflops/W左右.

其次是对 DVS 并行系统的尝试.与低电压并行系统的设计方式相反,DVS 并行系统试图使用高性能的处理器作为系统节点,使用 DVS 动态调整系统性能和功耗.MEMO 是16个节点的机群并行系统[2],每个节点的 Intel Pentium M 处理器具有 DVS 的功能.另外的一个 DVS 系统是由 AMD Athlon-64 构成的10个节点的机群系统[5],同样每个节点具有 DVS 能力.

最后,多核处理器技术是当前最重要的研究方向之一.多核处理器采用多个性能适中的处理核心提高能量效率,使用高的数据级并行或者线程级并行提高整个处理器的性能.当前商用通用处理器一般采用了单芯片多处理技术和同时多线程技术,例如 IBM 的下一代多核处理器 IBM POWER7 计划支持8个处理核,Intel 的 Nehalem-EX 支持8个处理核16个线程.随着对高计算性能的迫切需求,图形处理器(GPU)开始应用在通用计算领域[7-9],一般作为通用处理器的加速处理器.图形处理器(GPU)具有典型的并行结构,例如 NVIDIA 的 GeForce GTX 285 包括30个核,每个核包括8个 SIMD 功能部件,构成了240个流处理器.又例如 AMD 的4870处理器包括10个核,每个核包括16个 SIMD 处理单元,每个单元包括5个乘加部件,构成800个流处理器,对于规则的矩阵乘运算,能量效率超过1Gflops/W.图形处理器简化了硬件结构,能够有效处理具有规则数据访问模式的应用,充分开发了程序中的数据级并行,能够达到很高的计算效率.

从以上分析可以看出,并行处理技术被广泛应用于改善系统的能量效率.本文通过分析的手段研究并行处理技术的节能空间.我们研究的主要问题是:

在已有的高性能系统设计下,采用低电压的并行处理技术改善系统能量效率的空间是什么?这样的高性能系统设计包括高性能并行系统和多核处理器.

在 DVS 并行系统上,通过提高并行度来改善应用的能量效率的空间是什么?这包括 DVS 并行系统和 DVS 多核处理器.

理论上并行处理能够改善系统的能量效率,但是实际上并行处理引入了一些与提高能量效率方向相反的问题.并发度的提高往往导致时间开销和能量开销增大,结果系统的能量效率反而下降.本文分析了并行处理提高系统能量效率的原理,建立了程序并行执行的时间开销和能量开销模型.基于模型分析,本文研究了并行处理对能量效率改善的作用,分析了系统扩展和应用扩展对能量节省的影响.

本文第1节给出了研究的主要问题和相关工作;第2节给出并行处理提高系统能量效率的原理;第3节分析并行处理的时间开销和能量开销,给出能量效率模型;在第4节使用能量效率模型研究并行处理对能量效率改善的作用;最后给出结论.

2 并行处理提高能量效率的原理

假定程序的执行周期数为 c,系统的运行频率为 f,那么在该系统上程序的执行时间为

$$t=c/f.$$

考虑在一个 n 个节点的并行系统上,如果程序能够完全并行执行,c 个执行周期能够均匀划分到 n 个节点上执行,每个节点只需要完成 c/n 个周期.这里要保证程序能够在 t 时间内完成,每个节点需要的执行频率为

$$f_p=(c/n)/t=(c/t)/n=f/n,$$

也就是说,每个节点的执行频率只需要原来的 $1/n$.

下面考虑两种执行的能量消耗.对于以 CMOS 工艺为基础的计算机系统,为了保证器件稳定运行,系统的电压和频率必须满足

$$f=k\cdot(V-V_T)^{\alpha}/V,$$

其中,V 为器件的供电电压,V_T 为阈值电压,α 为一个系统相关的参数.一般 V_T 远小于 V,并且 α 取为2,于是频率 f 和电压 V 近似为线性关系

$$f=V/\beta.$$

考虑 CMOS 的能量消耗公式为

$$E=C_{\text{eff}}\cdot f\cdot V^2\cdot t,$$

其中 C_{eff} 为有效切换电容.

于是串行执行时系统的能量消耗为

$$E=\beta^2\cdot C_{\text{eff}}\cdot f^3\cdot t.$$

并行执行时系统的能量消耗为

$$E_p=n\cdot C_{\text{eff}}\cdot f_p\cdot V_p^2\cdot t$$
$$=n\cdot\beta^2\cdot C_{\text{eff}}\cdot(f/n)\cdot(f/n)^2\cdot t=E/n^2.$$

这里假定单个节点的有效切换电容相同. 从上面公式可以看出,并行执行的能量消耗是原来的 $1/n^2$. 这极大地减少了系统的能量消耗,提高了能量效率.

两种情况的近似能量消耗比率为

$$r=E_p/E=1/n^2.$$

如果不采用近似公式 $f=V/\beta$,更精确的能量消耗比率为

$$r=E_p/E=V_p^2/V^2.$$

考虑在 0.18μm 工艺下的情况,阈值电压设为 0.55V,在 0.5GHz 下,运行电压设为 2V. 可以得到能量消耗比率,见图 1.

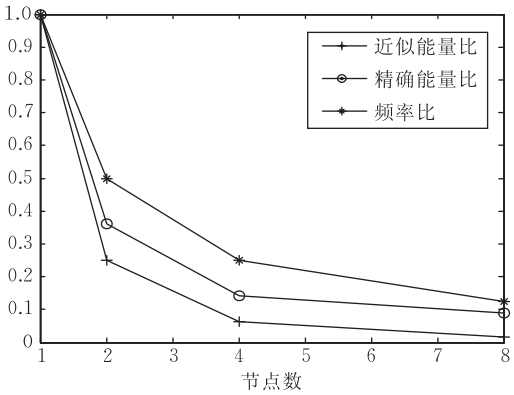

图 1　能量消耗比率

从图 1 可以看出,精确和近似方法具有相同的变化趋势,所以下面使用近似方法进行研究. 在本部分的结束,给出能量效率的明确含义. 能量效率是指单位能量消耗(或者功耗)所完成的工作量,单位为 flops/W 或者 flops/J. 在本文中,使用并行系统和串行系统的能量消耗比率来体现能量效率的高低.

3　能量效率模型

上一节的分析对应用和系统都进行了理想的假定,实际上并行处理具有时间和能量开销,这些开销最终影响了能量效率.

3.1　时间开销

在前面分析中理想的假定应用是可以完全并行的,应用可以完全划分到多个节点执行,不存在任何开销. 但是实际上应用往往不能完全并行化,存在可并行区间和串行区间,其中的串行区间不能并行化. 而且,可并行区间的并行具有相应的并行开销,例如通信和同步开销. 同时,即使在应用的可并行区间,多个节点的负载可能也不是相同的,存在负载不平衡问题.

为了讨论问题的方便,这里并不试图揭示每种并行开销. 这里的分析将并行程序划分为并行区间和串行区间,假定应用的串行区间串行执行的比率为 R_s,并行区间串行执行的比率为 $R_p=1-R_s$. 串行区间一般采用冗余执行的方式在每个节点并行执行,如果并行区间是负载平衡的,并行区间的执行是计算和通信交替的过程. 一个示意性的并行程序见图 2. 其中 S 代表串行区间,P 代表并行区间,C 为通信区间. 与串行执行相比,并行执行引入了串行区间的冗余执行和通信区间的消息交换,这些都导致系统能量消耗的增加.

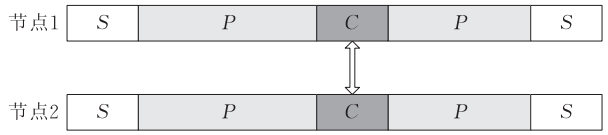

图 2　抽象的并行执行过程

下面考虑当节点数目变化时 S 区间、P 区间和 C 区间的变化. S 区间的变化最简单,当节点数增加时,每个节点上 S 区间的执行时间不变,这里仍用 R_s 表示. 假定并行节点数为 n,每个节点上 P 区间的执行时间为 R_p/n. C 区间的执行时间变化最为复杂. 假定在两个节点情况下 C 区间的执行时间比率为 $R_{2/c}$. 随着节点数的增加,C 区间执行时间的变化具有不同的规律. 文献[10]分析了不同 MPI 通信函数的时间预测函数,归结为系统大小 n 和消息长度 m 的函数. 对于点对点通信,时间预测函数为

$$\tau_1+\tau_2\cdot m.$$

对于消息启动开销较大的系统,τ_1 项占统治地位,随着节点数的增加,通信时间基本不变,这时可以简单地忽略第 2 项. 对于启动开销较小的系统,消息长度决定了程序的通信时间,这时可以简单地忽略第 1 项. 对于同一个应用,这里假定程序的消息长度随节点数的增加线性递减.

对于聚合通信,时间预测函数为

$$\tau_1+\tau_2\cdot\log_2 n,$$

聚合通信的第 2 项一般占统治地位,也就是说通信时间以 $\log_2$ 递增.

节点数目的扩展一般不会改变单个节点上通信

函数的调用次数.例如在气象模拟程序中,计算发生在规则的网格上,通信发生在网格的边界,每个节点的通信次数是不变的.因此,这里简单地假定应用的通信函数调用次数不变.

根据上面的分析,下面给出了 n 个节点的 C 区间执行时间 $R_{n/c}$ 的变化规律,分为以下 3 种情况:

$$R_{n/c}=\begin{cases}R_{2/c}, & \text{点对点通信,高起动开销(情况1)}\\ R_{2/c}\cdot 2/n, & \text{点对点通信,低起动开销(情况2)}\\ R_{2/c}\cdot \log_2 n, & \text{聚合通信(情况3)}\end{cases}$$

于是在 n 个节点上的执行时间可以表示为

$$R=R_s+R_p/n+R_{n/c}.$$

类似 Amdahl 定律,加速比函数为

$$speedup=1/R.$$

考虑在不同参数设置下的加速比,假定 $R_s=10\%$,$R_{2/c}=5\%$,则在 3 种情况下应用的加速比见图 3.从图 3 中可以看出,由于串行区间和通信区间的存在,程序的加速比受到了显著的影响.并且,对于聚合通信作为主要通信的应用,加速效果随着节点数的增加明显下降.在一个机群系统上测试了 NPB 程序[11]的加速比变化,见图 4.从图 3 和图 4 的对比可以看出:3 种应用特征的加速比曲线代表了所有 NPB 应用程序的加速比变化特性.

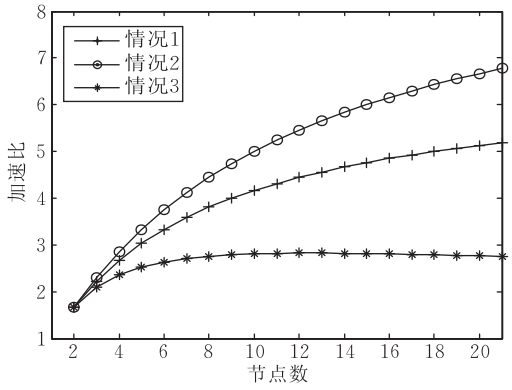

图 3 加速比曲线

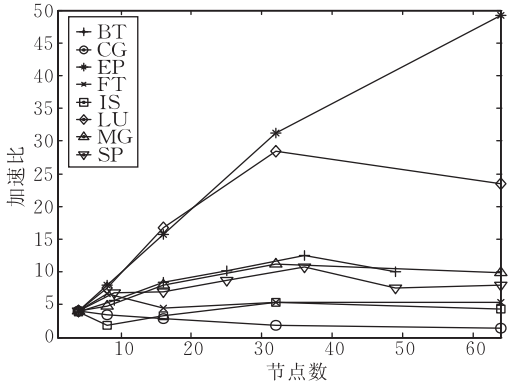

图 4 NPB 测试程序组的加速比曲线

3.2 能量开销

本文第 2 节对系统功耗进行了很多理想的假定.

(1)假定系统的性能是完全可以调节的,也就是系统的组成部分可以通过电压和频率调节,以相等的比率改变性能.实际上,对于多种部件构成的系统,系统的性能往往不是以这种方式变化的.例如,当前的 DVS 系统往往仅具有处理器的性能调节能力.

(2)没有考虑静态的能量消耗.随着工艺技术的进步,泄漏电流越来越大.

(3)系统负载不平衡下空转节点产生的能量消耗.负载不平衡下,并行应用往往产生因为同步导致的空转能量消耗.

(4)系统规模的扩展导致网络资源的能量消耗增加.

由于低频系统往往可以采用更好的泄漏控制技术,这里的研究忽略静态功耗.为了简化分析,研究忽略了负载不平衡和网络资源的能量开销.这里的分析将节点的部件分为可调节部分和不可调节部分.假定在串行执行情况下,可调节部分占总功耗的比率为 CS,不可调节部分占总功耗的比率为 $NCS=1-CS$.当进行低电压并行系统设计或者动态电压调节时,频率减少为原来的 FR,功耗成为 $NCS+CS\cdot FR^3$.

3.3 能量效率模型

根据上面对系统并行时间开销和能量开销的分析,考虑并行应用的能量消耗.假定应用的串行执行时间为 t,其 S 区间的比率为 R_s,其 P 区间的比率为 $R_p=1-R_s$.并行执行中两个节点情况下 C 区间的比率为 $R_{2/c}$.如果使用 n 个节点并行执行应用,执行时间为

$$t_p=t\cdot(R_s+R_p/n+R_{n/c}).$$

为了节省系统的能量消耗,采用降低节点的电压和频率的方法,使得应用能够在时间 t 内完成.这时频率调整的比率 FR 为

$$FR=t_p/t=R_s+R_p/n+R_{n/c}.$$

假定在串行执行情况下的系统功耗为 P,则系统可调节部分的功耗为 $CS\cdot P$,不可调节部分的功耗为 $NCS\cdot P=(1-CS)\cdot P$.

于是应用在 n 个节点上的能量消耗为

$$E_n=n\cdot(NCS+CS\cdot FR^3)\cdot P\cdot t.$$

n 个节点和单个节点的能量消耗比率为

$$\begin{aligned}E_n/E&=(n\cdot(NCS+CS\cdot FR^3)\cdot P\cdot t)/(P\cdot t)\\&=n\cdot(NCS+CS\cdot FR^3)\\&=n\cdot((1-CS)+CS\cdot(R_s+R_p/n+R_{n/c})^3).\end{aligned}$$

上式中,CS 是和系统特性紧密相关的参数,R_s

和 R_p 是和应用特性相关的参数,$R_{n/c}$ 是与系统和应用特性都相关的参数.

于是得到了并行处理的能量效率模型,下面分析并行处理在能量效率改善方面的作用.

4　并行处理的能量效率分析

首先研究系统相关参数 CS 和 $R_{2/c}$ 变化对能量效率的影响,称之为系统级能量效率可扩展性分析.然后考虑应用相关参数 R_s 和 R_p 对能量消耗的影响,称之为应用级能量效率可扩展性分析.

4.1　系统级能量效率可扩展性分析

假定 $R_s=0.1$,$R_{2/c}=0.1$,这里给出了不同 CS 的能量消耗比率,见图 5.其中包含了 3 种通信时间预测函数的曲线.从图中可以看出,系统参数 CS 对功耗的影响十分明显.当系统的全部部件都可以进行性能调节时($CS=1$),能够获得能量节省.部分部件的不可调节导致能量节省效果明显下降.例如 $CS=0.8$ 时,基本上已经很难得到能量节省.这说明在 DVS 并行系统上,除了调节主要部件的能量消耗以外,其它设备能量消耗的控制也十分重要.否则多节点累积的系统功耗很容易抵消并行带来的能量消耗节省.在当前的多机系统上,每个节点都包含了除主处理器外丰富的部件资源(存储器、通信系统和 I/O 系统),这些都是能量消耗的重要来源,因此可以预见在这样的系统上增加主处理器的 DVS 能力很难获得并行处理带来的能量节省.

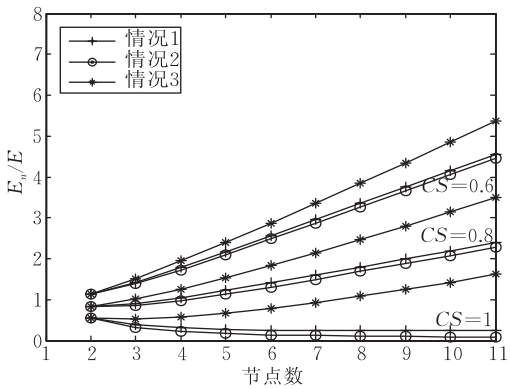

图 5　参数 CS 对能量效率的影响

全新设计的低电压并行系统可能得到成功,因为这样的系统采用了全系统的低功耗设计技术,这相当于 CS 部分的比率为 1,这样的系统是能够有效减少系统能量消耗,改善系统能量效率的.类似的 DVS 多核处理器也很有应用前景.

进一步在图 6 给出了 $CS=1$ 时的详细曲线.对于不同的通信函数,能量节省的效果不同.对于情况 3,随着节点数的增加,通信时间比率增加,很快抵消了能量节省的效果.情况 1 和 2 可以较为稳定地提高能量效率.情况 1 代表了高起动通信开销的并行系统,这对应于大规模并行处理系统.从曲线的变化情况可以看出,在原来系统的设计基础上,并行度提高 4 倍到 6 倍能够获得最佳的能量效率改善.进一步提高并行度对系统的能量效率改善没有很好的效果.

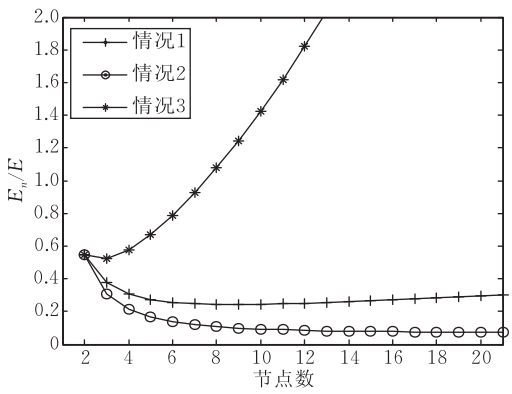

图 6　节点数目对能量效率的影响

相比之下,情况 2 的能量节省效果更为明显.回忆情况 2 代表了低通信起动开销,消息长度随系统扩展递减.情况 2 表明低延迟、高带宽的细粒度并行体系结构可能获得成功.当前多核微处理器恰恰代表了这一方向,将单个运行速度较低的处理器集成到单个芯片内,处理器之间提供低延迟、高带宽的通信链路.进一步,从低通信起动开销的短消息获得更高的能量节省可见:未来多核处理器的编程模型可能更加适宜采用中细粒度的并行方式.

下面详细分析了通信开销对能量效率的影响.从图 7 中可以看出情况 2 对通信开销是不敏感的,而情况 3 对通信开销十分敏感.情况 1 介于两者之间.从图中可以看出不管是对通信高度敏感的情况 3 还是对通信不敏感的情况 2,在超过一定节点数之后,很难观察到能量效率改善.总而言之,少量节点数并行处理提供了大量的能量节省,超过 10 个

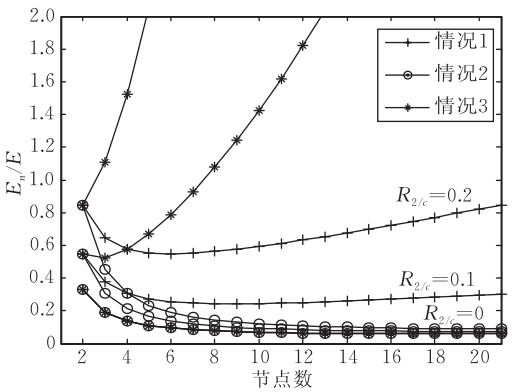

图 7　参数 $R_{2/c}$ 对能量效率的影响

节点之后，能量节省的效果已经很小，超过 16 个节点以后，几乎很少有能量节省.能量效率模型显示：最佳的能量比率变化情况为 $R_{2/c}=0$，$CS=1$，$R_s=0$，这时能量比率为

$$E_n/E=n\cdot(1/n)^3=1/n^2\rightarrow 0(n\rightarrow 0).$$

考虑 $n>16$ 的情况，这时对于任意两个节点数目 n_1 和 n_2（$n_1<n_2$），能量消耗比率变化为

$$1/n_1^2-1/n_2^2<2/n_1^2<2/16^2=0.0078125.$$

实际上工艺上的限制使得系统的运行电压必须在阈值电压之上并保持一定比例，从这个角度来说完全靠提高并行度来节省能量消耗也是不现实的.由此可以得出结论，低电压并行系统采用并行处理来提高能量效率的最高并行度为 16 左右.超过这个并行度以后，尽管可能继续改善应用的性能，但是不能提高能量效率.并且由于这是在理想情况下的结论，实际上由于系统和应用特性的因素，更多的节点可能引入很高的开销.

同样这一结论对试图采用低电压的多核结构设计的低功耗系统提出了理论上限.给定相应的制造工艺和具体的应用，设计 16 个左右的核心数目能够有效改善应用的能量效率.超过这样的数目以后，尽管可以继续提高系统的性能，但是很难改善应用的能量效率.当然也应该注意到，本文的模型假定单节点串行执行和多节点并行执行使用了相同的节点（体现在电容 C_{eff} 相同），实际上低电压多核结构设计时，多核节点因为运行在低电压和低频率，往往设计得更为简洁高效，这有利于系统能效的改善.

4.2 应用能量效率可扩展性分析

令 $R_{2/c}=0.1$，$CS=1$，下面研究了应用相关参数对系统能量消耗的影响.图 8 给出了在不同 R_s 时的功耗变化曲线.由图中可以看出，串行区间的比率对应用的能量节省有很明显的影响.对于情况 1 和情况 2，比率小于 0.2 的串行区间可以使用并行处理节省能量消耗.对于串行区间比率为 0.4 的应用，3 种情况都不能带来能量节省.

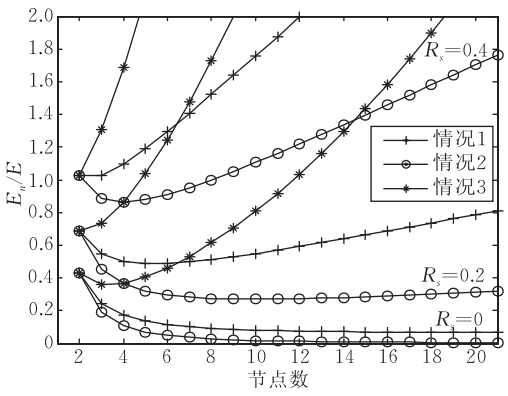

图 8 参数 R_s 对能量效率的影响

应用的规模扩展一般导致串行区间的比率下降.在图 9 中给出了应用规模扩展对能量效率的影响.从图中可以看出 R_s 的比率降低，能够得到更高的能量节省.同时，当 R_s 比率降低后，较多的节点数可能获得更高的能量节省.但是在通信开销不可扩展（情况 3）的情况下，即使在低比率的 R_s 下也很难得到能量节省.

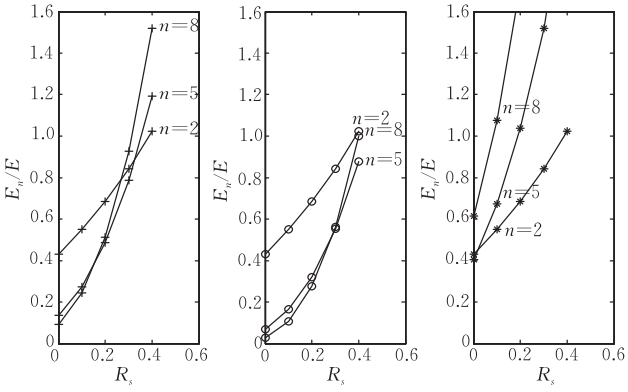

图 9 应用规模扩展对能量效率的影响

这里需要强调的是，即使在 R_s 很低的情况下，能量节省也是存在上限的，这受限于工艺的限制.

5 结 论

本文研究了并行处理改善系统能量效率的空间.主要结论包括：

在传统系统上增加 DVS 处理器，提高并行度，改善能量效率很有限的空间，受限于系统固有的不可调节部件的功耗.

低电压并行系统有很高的提高能量效率的空间，相对于传统设计，系统并行度提高 4～6 倍就利用了并行处理主要的节能空间.

单个芯片内采用提高并行度的方式提高能量效率有广泛的前景，因为单个芯片内的高带宽、低延迟特性，能量效率能够随着并行度提高稳定地改善.

应用的特性对并行处理影响很大.可以高度并行实现的应用程序能够获得更高的能量效率改善，这在 GPU 能够在处理具有规则计算模式的应用时获得很高能效得到了印证.相反，并行度低的程序很难通过并行处理提高能量效率.

参 考 文 献

[1] Hsu Chung-Hsing, Feng Wu-Chun. A power-aware run-time system for high-performance computing//Proceedings of the 2005 ACM/IEEE Conference on Supercomputing. Seattle,

Washington, USA, 2005: 1-10

[2] Ge Rong, Fen Xizhou, Cameron Kirk W. Performance-constrained distributed DVS scheduling for scientific applications on power-aware clusters//Proceedings of the 2005 ACM/IEEE Conference on Supercomputing. Seattle, Washington, USA, 2005: 34-45

[3] Rabaey J M, Chandrakasan A, Nikolic B. Digital Integrated Circuits: A Design Perspective. 2nd Edition. Beijing: Tsinghua University Press, 2004

[4] Lorch J R. Operating systems techniques for reducing processor energy consumption[Ph. D. dissertation]. University of California, Berkeley, USA, 2001

[5] Freeh Vincent W, Pan Feng, Kappiah Nandini, Lowenthal David K, Springer Rob. Exploring the energy-time tradeoff in MPI programs on a power-scalable cluster//Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium(IPDPS'05). Denver, Colorado, 2005: 41-50

[6] The BlueGene/L Team. An overview of the BlueGene/L supercomputer//Proceedings of the 2002 ACM/IEEE Conference on Supercomputing. Baltimore, USA, 2002: 1-22

[7] Fan Zhe, Qiu Feng, Kaufman Arie, Yoakum-Stover Suzanne. GPU cluster for high performance computing//Proceedings of the 2004 ACM/IEEE Conference on Supercomputing. Pittsburgh, PA, 2004. Washington, DC, USA, 2005: 47-58

[8] Luk Chi-Keung, Hong Sunpyo, Kim Hyesoon. Qilin: Exploiting parallelism on heterogeneous multiprocessors with adaptive mapping//Proceedings of the 42nd International Symposium on Microarchitecture (MICRO-42). New York, USA, 2009. Los Alamitos, CA, USA, 2002: 1-10

[9] Tomov Stanimire, Dongarra Jack, Baboulin Marc. Towards dense linear algebra for hybrid GPU accelerated many-core systems. Department of Computer Science, University of Tennessee, Knoxville, TN, USA: Technical Report UT-CS-08-632, 2008

[10] Hwang Kai, Xu Zhiwei, Arakawa Masahiro. Benchmark evaluation of the IBM SP2 for parallel signal processing. IEEE Transactions on Parallel and Distributed Systems, 1996, 7(5): 522-536

[11] Bailey Davd H, Barzcz Eric, Dagum Leonardo, Simon Horst D. NAS parallel benchmark results. IEEE Concurrency, 1993, 1(1): 43-51

YI Hui-Zhan, born in 1976, Ph.D.. His current research interests include low-power compilation optimization, parallel programming languages.

LIU Yong-Peng, born in 1976, Ph. D. candidate. His current research interests include operating system, power management, fault tolerance.

Background

Energy consumption has been paid increasing attention to in the computer domain because of its deep influence on the design of high-performance chips and systems. Many techniques are proposed to improve energy efficiency of computer systems, such as low-power techniques in device design, lower-power techniques in architecture design, low-power techniques in operating systems, and low-power techniques in compilation systems. Although low-power techniques have larger amounts of research work, the power and energy problems have not been completely solved.

Since CMOS technology has almost reached to the utmost, the main high-performance processor companies have changed the research way, from high frequency to many cores. Multi-core processors have fully utilized many transistor resources in the silicon chip, and increase the processor performance by task and data parallelism. But considering the energy and time overhead, how parallel processing improves energy efficiency has no a clear scene. In the paper, the author focuses on parallel processing on architecture level. Parallel processing improves energy efficiency by using some computing nodes with moderate performance, which maintain high throughput by parallel execution. In this paper, the authors present the fundamental of parallel processing improving energy efficiency, and models the time and energy overhead involved in parallel execution. Based on the models, the authors investigate low-voltage parallel systems, parallel systems with dynamic voltage scaling, and multi-core microprocessors, and reveals their potential of improving energy efficiency.

The project, this paper belongs to, is about power-aware and temperature-aware architecture and compilation research. For developing next-generation high-performance computer, power and thermal problem is one of the most key challenges, and power and thermal management will be one of the primary research directions for high-performance computing. In the project, the authors will investigate software-directed power and thermal management, and the planed works are listed as follows: (1) power and thermal experiment environment and analytical techniques for high-performance computing; (2) OS software directed power and thermal management techniques for high-performance computing; (3) Profile-guided power and thermal management for high-performance computing. In the project, the authors will design experiment environments and submit some excellent papers, and the research achievements will be used to the production tasks for high-performance computing. In this area, the authors have proposed the concepts of architecture energy efficiency and localizing the use of system units, and many works about low-power compilation optimization are published in Chinese journals and English conferences. This paper belongs to power and thermal experiment environment and analytical techniques for high-performance computing.