

NoC 目录协议能耗分析与建模研究

王继禾¹⁾ 郭 兵¹⁾ 沈 艳²⁾ 伍元胜¹⁾ 刘云本^{1),2)}

¹⁾(四川大学计算机学院 成都 610065)

²⁾(成都信息工程学院控制工程学院 成都 610225)

摘 要 在 MPSoC 芯片设计中, 能耗问题是一个关键因素, NoC(Network-on-Chip 片上网络)的能耗占芯片总能耗的 50%以上. 文中首先对 NoC 应用层协议的能耗行为进行统一的数学描述, 在此基础上, 针对目前 Cache 一致性中常用的目录协议进行深入分析, 提出一种基于位置概率分布的目录协议能耗模型, 其考虑目录信息与数据块节点位置分布的无关性, 定位数据块在 NoC 中的节点位置, 计算任意节点间数据传输的概率矩阵, 并结合路由单元能耗、连线能耗以及网络拓扑结构, 准确获得目录协议中一次数据块传输的能耗. 仿真实验表明, 该文提出的模型在发包概率稳定的 Mesh(网孔)网络中可有效计算出不同类型目录协议产生的能耗, 且精度较高(误差小于 2%), 为下一步协议能耗优化研究提供可信的量化的评判依据.

关键词 MPSoC; NoC; 协议功耗; 目录协议; Cache 一致性

中图法分类号 TP306 DOI号: 10.3724/SP.J.1016.2011.01603

Analysis and Modeling of Directory Protocol Energy Consumption in NoC

WANG Ji-He¹⁾ GUO Bing¹⁾ SHEN Yan²⁾ WU Yuan-Sheng¹⁾ LIU Yun-Ben^{1),2)}

¹⁾(School of Computer Science & Engineering, Sichuan University, Chengdu 610065)

²⁾(School of Control Engineering, Chengdu University of Information Technology, Chengdu 610225)

Abstract Energy consumption has become one of the critical factors in MPSoC design. Especially, NoC (Network-on-Chip) contributes more than 50% of energy dissipation in the whole chip. After a uniform energy consumption behavior in the NoC application protocols is described, this paper analyses the processing detail of directory protocol that is a kind of widely used implementation of cache coherence protocol, and proposes an energy model based on the location probability distribution, which can locate the position of a basic data block (DB) in NoC by considering the irrelevance between the directory information and data location distribution, and calculate a probability matrix that denotes the transferring rate in any two nodes of NoC. While combining with the routing unit energy consumption, line energy consumption and network topology factors, single DB transferring energy consumption in the directory protocol can be accurately obtained. Simulation results show that this model can effectively compute the energy consumption of different kinds of directory protocols with stable transferring rate in Mesh networks, and a higher precision with less than 2% error is achieved. This model can also provide a credible quantitative criterion for the optimization of directory protocol energy consumption in the next step.

Keywords MPSoC; NoC; protocol energy consumption; directory protocol; cache coherence

收稿日期: 2011-01-05; 最终修改稿收到日期: 2011-06-13. 本课题得到国家自然科学基金(61073045)、国家“八六三”高技术研究发展计划项目基金(2008AA01Z105)、四川省杰出青年科技基金(2010JQ0011)资助. 王继禾, 男, 1986年生, 博士研究生, 主要研究方向为嵌入式实时系统. 郭兵(通信作者), 男, 1970年生, 博士后, 教授, 博士生导师, 主要研究领域为嵌入式实时系统、SoC和中间件. E-mail: guobing@scu.edu.cn. 沈艳, 女, 1973年生, 博士, 副教授, 主要研究方向为分布式测量系统、嵌入式系统开发、无线传感器网络和机器人. 武元胜, 男, 1985年生, 博士研究生, 主要研究方向为嵌入式实时系统. 刘云本, 男, 1981年生, 博士研究生, 主要研究方向为嵌入式实时系统.

1 引言

目前处理器的发展趋势倾向于采用多核集成的策略,随着芯片的集成度越来越高,设计复杂性不断上升,对计算机系统的能耗也带来严重挑战.因此,在 MPSoC(Multi-Processor SoC 多核 SoC)芯片设计中,能耗问题始终是一个关键因素^[1].

在 MPSoC 中,核间通信主要包括总线监听、交叉开关和 NoC(Network on Chip 片上网络)3 种方式.其中,总线监听一般适用于规模较小的多核网络,如 ARM Cortex-A9 处理器中采用侦听控制单元(Snoop Control Unit)控制多核在 AMBA 总线上的通信;使用交叉开关的典型例子是 AMD Athlon64 X2 处理器,利用交叉互连网络和系统请求接口对各个 CPU 核的请求进行仲裁.由于上述两种通信方式都不适合高端的多核和众核互联,所以包交换网络互连的思想被引入到 MPSoC 核间通信设计中, NoC 将成为未来 MPSoC 的主流通信方式^[2].目前 NoC 的能耗一般占芯片总能耗的 50% 以上^[3],其中静态能耗主要由器件的漏电电流造成,动态能耗由器件的 0/1 翻转和短路电流造成.由于 NoC 的动态能耗主要由协议活动产生,因此,开展 NoC 协议能耗问题的分析与建模是十分必要的.

关于 NoC 协议能耗问题的一些代表性工作包括:Benini^[4]于 2001 年首次完整描述了 NoC 系统的能耗,从能耗产生的角度将 NoC 划分成多个层次;Bhat^[5]给出了 NoC 包交换中各片上传输部件的能耗模型,并得到了各部件单次执行的能耗值;Li^[3]将 DVS(动态电压调整)用于 NoC 节点间互联,通过动态调整电压和频率降低 NoC 通信能耗;Wang^[6]采用了分组 Cache 的方法将分布式存储组织成多个自治域,从而降低了 Cache 一致性协议的能耗.但是,上述工作都没有给出明确的协议能耗模型,导致 NoC 协议能耗优化缺乏理论支撑.因此,本文将重点研究 Cache 一致性协议的能耗模型及其能耗影响因素,为能量有效的 NoC 协议设计和优化奠定基础.

尽管能耗和功耗是两个不同的概念,但是在相关研究领域中许多泛指情况下可互换使用,一些低功耗的研究工作实际上是优化系统的能量消耗.因此,在不会引起混淆的情况下,本文不详细区分能耗和功耗的概念.

本文第 2 节介绍 NoC 协议的基本组成及应用

层协议能耗的统一描述;第 3 节描述 NoC 物理层协议的能耗,分析 NoC 中能耗是如何在连线和路由单元(路由节点)中产生的;第 4 节对目录协议能耗进行建模,得出单个数据包(message)传输的能耗模型;第 5 节对提出的能耗模型进行实验验证;第 6 节对未来的工作进行展望.

2 NoC 协议栈及协议能耗的描述

2.1 NoC 协议栈

在文献[5]中,Bhat 试图将 NoC 纳入到一般网络的体系结构中,希望完全按照 ISO/OSI(Open System Interconnection 开放系统互连)的分层结构设计现有的 NoC,但是 NoC 与 OSI 之间在设计目标和工作环境方面存在一些差异:

(1) NoC 的设计目标与 OSI 不同. OSI 是满足各种应用需求的通用网络,而 NoC 是面向 Cache 一致性的专用网络,二者在低层设计上有相似之处,但在上层的实现上存在较大的差别.

(2) OSI 与 NoC 的工作环境不同. OSI 的应用环境往往不可预知,其设计十分强调 QoS(Quality of Service 服务质量),但是 NoC 一般工作在 MPSoC 中,对信息传输环境的适应性不是设计者考虑的主要问题,而令设计者头痛的是片上资源的局限性和能耗难以控制.换言之, NoC 的设计更加强调不同层次之间的统一描述和软-硬件的协同设计,而不像 OSI 追求各层协议之间相互独立.因此, NoC 更应遵循整体设计的思路.

进一步地,可将 NoC 和 OSI 的协议模型进行层次对比(如图 1 所示).

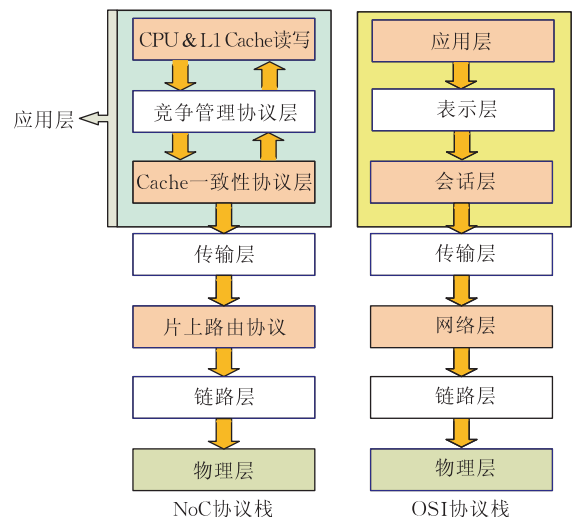


图 1 NoC 与 OSI 协议模型对比

(1) 在下层(传输层以下)协议中, NoC 与 OSI 协议栈几乎对等, NoC 依然需要物理层和链路层对信号进行编码和传输. 由于一个设计好的 NoC 拓扑结构相对简单且固定, 因此片上通信中的路由协议也相对简化, 如一种简单有效的路由协议是 X-Y 路由, 将一个数据微片(Flit, Flow Control Unit)先按照 X 轴方向传输, 到达目标列之后再按照 Y 轴方向传输, 最终到达目的节点.

(2) 在上层协议中, NoC 和 OSI 协议栈的差异较大, 功能上也不完全对等. 这是因为对于 NoC 来说, OSI 中提供服务质量保证的传输层复杂度较高, 难以利用片上较低的软-硬件成本实现. 如何在增强传输可靠性和降低设计复杂度之间取得平衡, 已成为当今 NoC 研究的一个热点问题, 如文献[7]中提出的一种增强传输可靠性的反馈控制方法, 依然难以摆脱实现成本过高的问题, 并且严重制约了传输效率.

对于 NoC 应用层协议, 可细分为 3 个子层.

(1) CPU&L1 Cache 读写层. 当 CPU 发生读-写操作时, 若目标地址所指向的内存块不在本地 L1 Cache 中, 将向 NoC 接口发出读-写请求, 并保存在接口请求队列中, 由下层协议响应这些请求, 并保证以适当的操作权限将所请求的内存块返回给 CPU 使用. 值得一提的是, 在一种具有 XBar 结构的网络中, 每个路由单元可能同时连接多个 CPU 核, 这些 CPU 拥有各自独立的 L1 Cache, 因此当这些核的读-写操作请求在本地 L1 Cache 中失效, 也将依次进入 NoC 接口的请求队列, 由下层协议无差别地依次处理^[8].

(2) 竞争管理协议层. 将 CPU 对 Cache 的读-写操作封装成独立的原子事务. 一个有效的竞争管理机制应该保证, 在任何时刻, 网络中都存在一些不能被其它读-写请求中断的原子事务, 协议确保这些事务能够完整地得到执行. 在文献[9]中, Zhang 对 Greedy 竞争管理机制与 Cache 一致性协议的关系做了较为详细的讨论.

(3) Cache 一致性协议层. 一种面向 CPU 读-写操作的同步通信机制, 应该保证, 在任何时刻, 执行事务的节点可在最短的时间内在网络上定位自己所需要的 Cache 块. 目录协议是一种主流的 Cache 一致性协议实现方式, 对共享存储器中的数据块设置目录项, 用以跟踪、记录每个数据块的状态信息, 从而判断哪些节点应对请求做出响应. 这些目录项按照固定的排列方式分布在各个路由节点中, 使得需

要这些信息的节点知道所需信息在网络中的位置^[10].

在多任务 MPSoC 环境中, 设计人员为减少片上长距离通信提出了各种映射策略, 试图降低数据远程传输对系统造成的影响. 这种映射优化虽然可以在一定程度上降低系统能耗, 但难以根本上消除并行任务对共享资源的同步、竞争这一重要特性. 在并行执行中, 任意节点上的进程在任何运行时刻都可能产生对系统中的某一共享数据的读/写请求, 这是程序员和硬件设计者无法预料的. 因此对 Cache 一致性协议的能耗进行分析和建模显得尤为重要.

2.2 应用层协议能耗的统一描述

软件能耗是指在软件运行期间, 驱动直接相关硬件(包括 CPU、存储器、I/O 控制器等)活动产生的系统能耗总和. 自从 Tiwari 提出软件能耗的概念之后, 人们越来越关注软件实现方式对系统能耗的影响程度. 协议能耗是软件能耗的一种, 表征了协议软件完成一种特定任务时所消耗的能量.

对 NoC 应用层中各层次协议的能耗, 可用下列两个定义进行统一的描述.

定义 1. 面向能耗的协议是一个由 4 个元素构成的集合 $Protocol(O, S, P, E)$, 其中, $O(o_1, o_2, \dots, o_m)$ 是协议的操作域, 即面向下层协议的接口, 为协议操作的最小粒度; $S = \{s \mid s = (o_{n_1}, o_{n_2}, \dots, o_{n_k}); o_{n_i} \in O, k < \infty\}$ 是协议的事务集, 每个 s 由一个操作序列组成, 代表协议所能够完成的功能; $P(p_1, p_2, \dots, p_n)$ 是协议的各个事务 s 发生的概率向量, 是由高层协议调用决定的; $E(e_1, e_2, \dots, e_m)$ 是操作域 O 内的各个操作 o 所对应的能耗向量.

定义 2. 协议能耗是一个 $Protocol(O, S, P, E)$ 在完成单个任务的过程中所产生的平均能耗, 可表示为 $\sum_{s_j \in S} p_j (\sum_{o_i \in s_j} e_i)$.

从定义 2 可以看出, 可将协议中各种事务的能耗视为一个基本的能耗单元, 协议能耗与协议负载无关, 只与事务发生的概率分布和单元能耗有关.

3 NoC 物理层协议的能耗分析

为了建立 NoC 目录协议能耗模型, 需要首先对物理层协议的能耗进行分析, 以明确 NoC 产生的能耗在硬件单元上的分布情况.

与通用的互联网络类似, NoC 的核心是路由单元, 处理单元通过路由单元的消息转发机制在片上

实现数据的通信,路由单元通过金属连线实现网络相互连接,这种片上通信对于能量的消耗是十分可观的,如 Alpha21364 处理器的功耗为 23 W,其中 NoC 占整个处理器功耗的 58%,而 MIT Raw 处理器的 NoC 占整个处理器功耗的 40%^[3].

NoC 产生的能耗主要由连线能耗和路由单元能耗组成.连线能耗是数据从一个路由单元发出,通过金属连线传播到相邻下一个路由单元的过程中所产生的能耗;路由单元能耗则是指数据在路由单元内部存储转发所产生的能耗. Li 等人通过对 NoC 的仿真发现,其功耗分布如图 2 所示,其中连线功耗占 82.4%,路由单元功耗占 17.6%,而在路由单元中,Buffer 功耗占 10.07%,交叉开关(Crossbar)功耗占 6.49%.

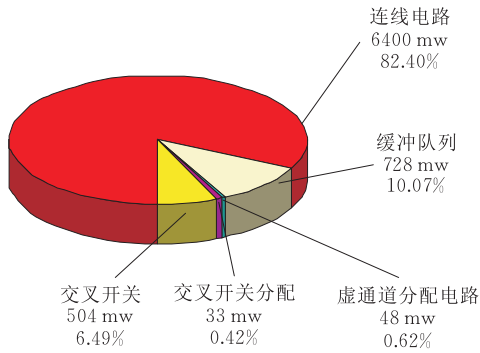


图 2 NoC 的功耗分布

下面分别对连线能耗和路由单元能耗模型进行分析介绍.

(1) 连线能耗模型. 大规模集成电路的片上通信连线几乎都采用并行通信,由于并行连线的集成度很高,所以必须考虑并行高频传输的线间电容 (interline capacitance) 效应,并且随着连线尺寸的减小,线间电容对与传输能耗的影响也逐渐增大. 在适当考虑线间电容的影响后,单个 Flit 在连线上一次传出能耗 E_{line} 可表示为下列向量形式:

$$E_{line} = d(\mathbf{V}^f)^T \mathbf{C}^t (\mathbf{V}^f - \mathbf{V}^i) \quad (1)$$

其中, d 是一个数据传输块中含有 Flit 的数量, \mathbf{V}^f 是 n 条并行连线在数据一次变化后的电平向量, $(\mathbf{V}^f)^T$ 是其转置向量, \mathbf{V}^i 是 n 条并行连线在数据一次变化前的电平向量, \mathbf{C}^t 是 $n \times n$ 的连线间电容对称矩阵,其中对角线上元素是对应位置连线到地的电容,其它位置元素是连线间的电容.

(2) 路由单元能耗模型. 目前片上路由单元的结构还没有摆脱通用路由器的基本框架. 在输入端,当数据从 E、W、N、S 和 Local 5 个方向到达路由单元后,由输入 VC (Input Virtual Channel) 完成输入

数据的接收和缓存,向路由仲裁器 (Switch Allocator) 请求并得到需要的输出端口后,数据包以 Flit 为单位通过 Crossbar 发射到下一跳 (hop) 的路由单元输入 VC 中. 如果该 Flit 已到达目的节点,那么 Crossbar 将其发送到本地 L2 Cache 接口队列 (Sink) 中. 一个 Flit 通过一个路由单元所产生的能耗 E_{Router} 可表示为

$$\begin{aligned} E_{Router} &= E_{VC} + E_{SA} + E_{Crossbar} \\ &= E_{Buffer} + E_{VCA} + E_{SA} + E_{Crossbar} \end{aligned} \quad (2)$$

其中, E_{VC} 为数据包在 VC 队列中等待的能耗, E_{SA} 和 $E_{Crossbar}$ 分别为路由仲裁器分配交叉连接的能耗和 Flit 通过 Crossbar 的能耗, E_{VC} 又可分为 VC 缓冲器能耗 E_{Buffer} 和 VC 分配电路 (VC Allocator) 能耗 E_{VCA} . 一般情况下, E_{Buffer} 和 $E_{Crossbar}$ 是整个路由单元能耗的主要来源.

在此基础上,可得出路由协议的能耗模型. 由于发送端将一个数据包分割成一组大小相等的 Flit,当数据包的第一个 Flit 通过路由到达一个新的路由节点后,由 VC 内部的分配电路为该组 Flit 分配一个数据缓冲队列,随后的 Flit 会自动地进入该缓冲队列,等待由 Crossbar 转发到下一跳. 因此,用 n 表示单个 Flit 传输所经过的跳数,一个 Flit 在传送路径上完成一次传输所产生的能耗 E_{Flit} 可表示为

$$E_{Flit} = nE_{Router} + (n-1)E_{line} \quad (3)$$

4 基于位置概率分布的目录协议能耗模型

可用一个有向图 $NoC \langle S_N, L \rangle$ 表示具有 N 个路由节点的 NoC 网络, $s_i \in S_N$ 为第 i 个路由节点, i 是节点的位置编号, $0 \leq i \leq N-1$, $l_{i,j} \in L$ 是节点 s_i 到 s_j 之间的有向边. 一个一维的连续内存 $M(n, b)$, n 为内存块的数量,将 M 平均分成 b 个区域,一个区域的大小为 $\lceil n/b \rceil$ 块. 每一个块对应一个块目录项,同一个区域中的块目录项集中存放在 NoC 的一个节点中,不同区域的目录项分布存放在不同的节点中. 假设区域的数量等于路由节点的数量,即 $b = N$,得到块地址 $block_addr$ 到路由节点位置 i 之间的映射关系为

$$i = \left\lfloor \frac{block_addr \cdot b}{n} \right\rfloor \quad (4)$$

路由节点上 PE (Processing Element 处理单元) 发出的块读-写请求会沿着 NoC 协议栈自上而下到达 Cache 一致性协议层,协议对请求的块地址

按照式(4)进行计算,得到块的目录项所在路由节点的位置 i , 目录协议根据目标节点的位置 i 完成 3-way 取数据操作(如图 3 所示). 从路由节点 s_0 请求一次取数据块的完整过程包括:

(1) 查询过程. 通过式(4)计算, 请求者 s_0 发现目标块的目录项信息静态存放在 s_{11} 的 SRAM 中, 向 s_{11} 发送请求消息;

(2) 响应过程. s_{11} 收到请求消息后, 在自己的目录项中查找发现该块的一个共享副本存储在 s_{13} 的二级缓存中, 于是向响应者 s_{13} 发送一个数据请求消息, 该消息包含了所要请求的块地址和请求节点 (s_0) 信息;

(3) 传输过程. 当响应者 s_{13} 得到请求消息后, 向 s_0 发送一份该块的副本并告知 s_{11} , s_{11} 对该块目录项的状态做出相应的更新.

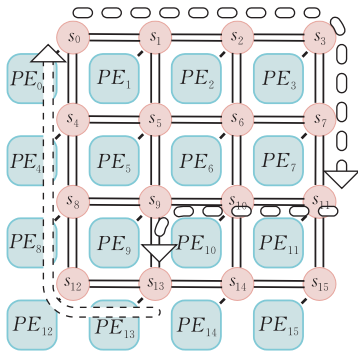


图 3 目录协议数据 3-way 传送过程

由于整个内存的块信息均匀分布存储在所有的路由节点中, 因此这种协议的网络负载较为均衡^[11].

为建立上述过程的能耗模型, 首先采用概率统计模型刻画一个 3-way 取数据的过程. 将网络中的 PE 对内存的块请求视为随机事件, 用 $q_{i,j}$ 表示 PE _{i} 的 L2 Cache 对内存中第 j 个区域中数据块请求的概率(或速率). 矩阵 Q 表示整个 NoC 中各节点对 Cache 块的请求概率, 由于 $b = N$, 该矩阵为一个方阵:

$$Q = \begin{bmatrix} q_{0,0} & q_{0,1} & \cdots & q_{0,N-1} \\ q_{1,0} & q_{1,1} & \cdots & q_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ q_{N-1,0} & q_{N-1,1} & \cdots & q_{N-1,N-1} \end{bmatrix} \quad (5)$$

矩阵对角元素 $q_{i,i}$ 不为 0, 这是因为如果一个 PE 所请求数据块的目录项恰好在本地, 不需要进行远程查询目录项的操作, 而直接向目标节点发送数据请求. 通过分析可知, $\sum_{i=0}^{N-1} q_{k,i} < 1$, 这是因为一个

PE 首先在自己的 L2 Cache 中查找所需的数据块, 如果命中, 则不向 NoC 接口发送数据请求.

一个请求信息到达目录项所在的节点后, 通过查询数据块的目录项可获得该块目前所在节点位置, 可用矩阵 P 表示响应过程的概率分布, 元素 $p_{j,k}$ 表示在节点 j 的目录项查询某一个数据块时, 该块在节点 k 的 L2 Cache 中的概率, 即在目录协议中 3-way 传输的第 2 步发生的概率分布. 响应过程的概率模型与查询过程类似, 这里不再赘述.

假设经过查询, 某一个数据块在节点 k 的 L2 Cache 中存在一个副本, 那么节点 k 将此块发送给需要它的源节点 i , 该传输过程的概率分布用方阵 T 表示, 可表示为

$$T = (Q \times P)^T \quad (6)$$

即节点 k 向节点 i 发送数据块的概率, 实际上等于节点 i 向所有节点发送请求查询概率和任何节点查询到某一个块在节点 k 中概率的乘积.

然后, 计算单个 Flit 在 NoC 中传输的平均能耗. 假设在 $c \times c$ 的 Mesh 网络中, 一个 Flit 从节点 i 产生, 目的节点为 j , 则该 Flit 经过的总跳数 $n = |j-i| \bmod(c) + \left\lfloor \frac{j}{c} \right\rfloor - \left\lfloor \frac{i}{c} \right\rfloor + 1$. 进一步地, 用 $e_{i,j}$ 表示一个 Flit 从节点 i 到 j 传输的能耗, 则根据式(3)有

$$E_{\text{Flit}} = e_{i,j} = \left[|j-i| \bmod(c) + \left\lfloor \frac{j}{c} \right\rfloor - \left\lfloor \frac{i}{c} \right\rfloor \right] \cdot (E_{\text{Router}} + E_{\text{line}}) + E_{\text{Router}} \quad (7)$$

其中, $e_{i,j}$ 包含了硬件电路、网络拓扑结构和网络层协议的所有能耗信息, 不仅是单个 Flit 对硬件资源利用程度的函数, 而且是源和目的节点之间传输跳数的函数.

引入式(2)可得到整个网络任意结点间一次数据块传输产生的能耗为

$$E_{\text{data}} = d \times \begin{bmatrix} e_{0,0} t_{0,0} & e_{0,1} t_{0,1} & \cdots & e_{0,N-1} t_{0,N-1} \\ e_{1,0} t_{1,0} & e_{1,1} t_{1,1} & \cdots & e_{1,N-1} t_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ e_{N-1,0} t_{N-1,0} & e_{N-1,1} t_{N-1,1} & \cdots & e_{N-1,N-1} t_{N-1,N-1} \end{bmatrix} \quad (8)$$

其中, d 是一个数据块中含有 Flit 的数量.

因此, 一次取数据块的过程平均能耗是任意两个节点之间传输数据能耗的期望, 可表示为

$$e_{\text{data}} = d \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} e_{i,j} t_{i,j} \quad (9)$$

以上通过对目录协议请求数据块的过程进行抽象,得到在 NoC 中数据传输的能耗模型.从式(3)可以看出,影响一个数据块传输能耗受到两方面因素影响,一方面是单个 Flit 的传输能耗,代表单个 Flit 对片上资源利用程度,对硬件传输电路和网络拓扑结构进行改进,是目前降低 $e_{i,j}$ 对 NoC 能耗影响的主要方法^[12];另一方面是面向能耗的目录协议本身的运行情况,这控制数据块的存储策略和传送过程,是对分布式存储及目录协议能耗的抽象描述,面向能耗的协议优化也是 NoC 研究领域的焦点.

5 仿真实验及分析

5.1 实验环境

仿真实验在 NS-2 平台上进行^[13],对目录协议能耗模型在不同网络发包概率的情况下进行比较和分析.相关参数设置如下:

(1) NoC 参数.在 NS-2 中建立了一个 4×4 的 Mesh 网络,相邻路由节点之间的链路传输速率为 10 Mbps,每个数据包的大小设为 512 Bytes,与一个数据块的大小相适应.在源节点上,一个 512 Bytes 的数据块被分割成 4 个大小相同的 Flit. 传送路径上的每个路由节点都只为第一个 Flit 做一次路由处理,其实质是在输入/输出端口间分配一个软连接,后继 Flit 通过此连接直接发送到输出端口.

(2) 发包参数.为了模拟不同负载环境下的 NoC 发包情况,为 NoC 中的每一个 PE 设置一个“发包向量”数据结构以控制目的地址范围,从而控制发包概率矩阵 T . PE 发包的过程分为两个相互独立的阶段:发包持续时间(Bursting Time)和空闲持续时间(Idling Time).设定任意两个 PE 之间的发包持续时间期望值和空闲持续时间期望值,以控制节点的发包概率.如果将一个轮转周期的期望设为 1 s,发包持续时间的期望 $E(BT)$ 和空闲持续时间的期望 $E(IT)$ 满足 $E(BT) + E(IT) = 1$ s,节点发包概率由 $E(BT) / [E(BT) + E(IT)] = E(BT)$ 决定.

(3) 节点距离的影响.一般来说,在一个经过优化的任务映射中,任意两节点间的通信概率会呈现如下特征:两节点 i 与 j 之间的“曼哈顿”距离越大,对应的传输概率 $t_{i,j}$ 越小.本实验通过调整矩阵 T 中的概率分布,体现不同优化深度对系统能耗的影响.具体来说, 4×4 类型模拟没有经过任务映射优化的随机发包; 3×3 类型和 2×2 类型模拟了经过一般优化和深度优化后的随机发包.

5.2 实验结果分析

实验模拟一个 4×4 Mesh 在 3 种不同发包类型情况下 60 s 的发包情况:

(1) 4×4 . 在 4×4 Mesh 网络中随机自由发包,即 T 中元素 $t_{i,j} = \frac{1}{15}$, $i \neq j$;

(2) 3×3 . 发包节点只向水平距离和垂直距离都小于 3 的节点随机自由发包;

(3) 2×2 . 发包节点只向水平距离和垂直距离都小于 2 的节点随机自由发包.

通过统计一次实验中总路由次数和产生数据包数量,得到 NoC 中单个包的平均路由次数:

$$\text{包平均路由次数} = \frac{\text{总路由次数}}{\text{数据包数量}} \quad (10)$$

同理,采用式(5)也可得到包的平均链路传输次数.根据平均路由次数和平均连路传输次数,获得 NS-2 平台与目录协议能耗模型估算的包平均能耗数据如表 1 所示.

表 1 包平均能耗和误差

类型	NS-2 平台的包平均能耗/nJ	模型估算的包平均能耗/nJ	平均误差/%	最大误差/%
4×4	0.956327	0.958100	0.185	1.767
3×3	0.783050	0.782476	0.073	0.696
2×2	0.512538	0.513286	0.146	1.350

可以看出,NS-2 平台的包平均能耗与模型估算的包平均能耗最大误差小于 2%,说明本文提出的目录协议能耗模型是有效的.

进一步地,可得到上述 3 种情况下目录协议能耗模型的包平均能耗精度与概率的变化关系(如图 4~7 所示).

从图 4 可以看出,虽然收到发包向量矩阵的影响,3 种情况下包平均路由次数有所不同,但是无论发包概率如何变化,NoC 中的包平均路由次数维持在理论值附近变化,从而保证了网络中数据包平均能耗是稳定的.

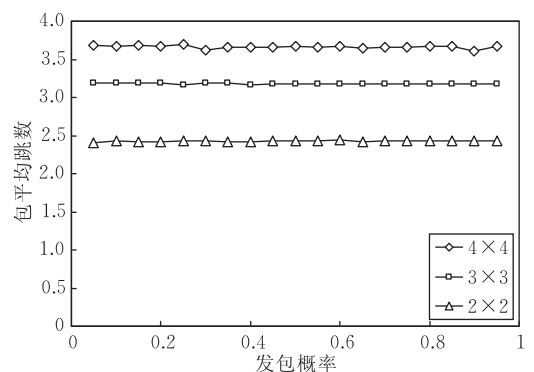


图 4 3 种情况下包平均路由次数

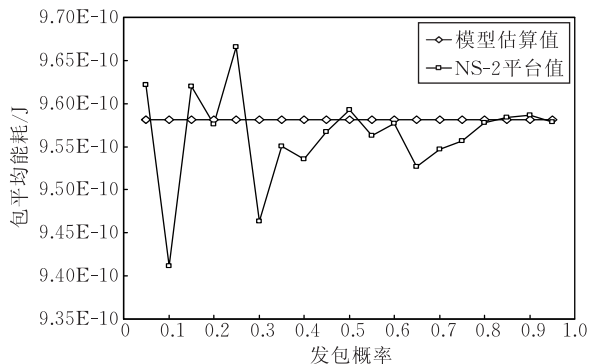


图 5 4×4 类型的包平均能耗

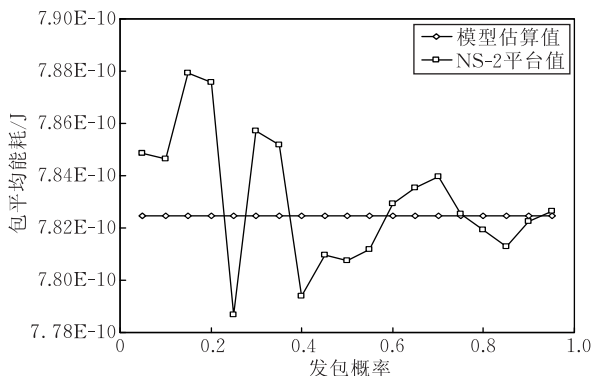


图 6 3×3 类型的包平均能耗

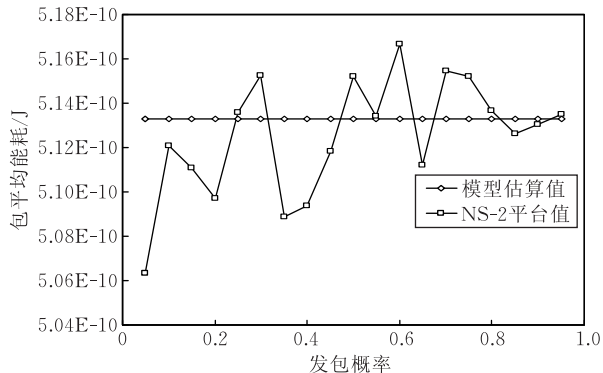


图 7 2×2 类型的包平均能耗

3 种情况下当发包概率接近 1 时, 单个包的 NS-2 平台平均能耗值接近于目录协议能耗模型估算的平均能耗值, 并且稳定性也更高(如图 5~7 所示)。随着发包概率的增大, 目录协议能耗模型估算的包平均能耗值精度也不断提高。

5.3 实验结论

从实验部分可以看出, 由式(6)得到的矩阵 T 实际上反应了 NoC 数据传输概率的集中度。这体现在网络优化中, T 的每一行发包概率应尽可能集中在与该行号的网络位置向邻近的节点上, 同时也不回避长距离传输对网络能耗的影响。利用该模型, 可对各种网络优化措施的能耗收益做定量的分析和比

较。例如在本实验中, 经过初步优化的 3×3 类型, 单个包的平均能耗为未经优化的 4×4 类型的 81.7%; 而经过深度优化的 2×2 类型, 单个包的平均能耗仅为未经优化情况的 53.6%。由此可见, 该模型为分析协议设计的能耗优化深度提供量化评判依据。

6 结束语

在 NoC 中, 能耗问题将会随着网络规模的增大而越来越受到人们的重视。本文对 NoC 协议栈中应用层协议的能耗行为进行统一的数学描述, 重点对目录协议的取数据过程所产生的能耗进行了分析与建模, 提出一种基于位置概率分布的目录协议能耗模型, 并利用发包概率稳定的 Mesh 网络验证了目录协议能耗模型的有效性。该模型不受到网络拓扑结构的影响, 同样适用于 DTorus、3-D 等其它网络拓扑结构, 通用性较强。下一步工作主要包括将目录协议的能耗模型与 Cache 一致性协议的 3 态 MSI 协议和 5 态 MOESI 协议相结合, 构建更为广泛的 NoC 能耗模型, 以及目录协议的能耗优化。

参 考 文 献

- [1] Guo Bing, Shen Yan, Shao Zi-Li. The redefinition and some discussion of green computing. Chinese Journal of Computers, 2009, 32(12): 2311-2319(in Chinese)
(郭兵, 沈艳, 邵子立. 绿色计算的重定义与若干探讨. 计算机学报, 2009, 32(12): 2311-2319)
- [2] Zhang Ge. Research on architectural layer power modeling and optimizing technology on MPSoC. Progress in Natural Science, 2009, 19(12): 1398-1409(in Chinese)
(张戈. 片上多核处理器的结构级功耗建模与优化技术研究. 自然科学进展, 2009, 19(12): 1398-1409)
- [3] Li Shang. Dynamic voltage scaling with links for power optimization of interconnection networks//Proceedings of the 9th International Symposium on High-Performance Computer Architecture (HPCA-9 2003). Anaheim, California, USA, 2003: 91-102
- [4] Benini L. Powering networks on chips//Proceedings of the 14th International Symposium on System Synthesis. Montreal, 2001: 33-38
- [5] Bhat S. Energy models for network-on-chip components. Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, 2005
- [6] Wang Zuo. Group-caching for NoC based multicore cache coherent systems//Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE'09). Nice, France, 2009: 755-760

- [7] Sharifi A. Feedback control for providing QoS in NoC based multicores//Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE 2010). Dresden, Germany, 2010: 1384-1389
- [8] Gao Xiang, Chen Yunji, Wang Huandong et al. System architecture of godson-3 Multi-Core Processors. Journal of Computer Science and Technology, 2010, 25(2): 181-191
- [9] Zhang Bo. Location-aware cache-coherence protocols for distributed transactional contention management in metric-space networks//Proceedings of the 28th IEEE International Symposium on Reliable Distributed Systems (SRDS'09). Niagara Falls, NY, 2009: 268-277
- [10] Zeng Hongbo. An interconnect-aware power efficient cache coherence protocol for CMPs//Proceedings of the IEEE International Symposium on Parallel and Distributed Processing (IPDPS 2008). Miami, FL, USA, 2008: 1-11
- [11] Zhang Michael. Latency reduction techniques in chip multi-processor cache systems [Ph. D. dissertation]. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 2006
- [12] Wu Jun-Jie. Smart multihop promotion for non-uniform Cache architecture. Chinese Journal of Computers, 2009, 32(10): 1887-1895(in Chinese)
(吴俊杰. 面向非一致 Cache 的智能多跳提升技术. 计算机学报, 2009, 32(10): 1887-1895)
- [13] Kourdy R. Using the NS-2 network simulator for evaluating multi protocol label switching in network-on-chip//Proceedings of the 2010 Second International Conference on Computer Research and Development. Kuala Lumpur, 2010: 795-799



WANG Ji-He, born in 1986, Ph. D. candidate. His research interest is embedded real-time system.

GUO Bing, born in 1970, Ph. D. professor. His current research interests include embedded real-time system, SoC

Background

This work was supported in part by the National Natural Science Foundation of China under Grant No. 61073045, the National High Technology Research and Development Program (863 Program) of China under Grant No. 2008AA01Z105, and Sichuan Science Fund for Distinguished Young Scholars under Grant No. 2010JQ0011. As a primary trend in the future embedded multi-core systems, the energy consumption of NoC (Network-on-Chip) has drawn many attentions. With the shrinking of MPSoC (Multi-Processor SoC) size and the development of new technologies, NoC contributes a larger percentage of energy consumption in MPSoC. So, it's necessary to undertake the research of NoC energy consumption. The past NoC energy consumption work focused on the communication and routing layers. They would like adding new network units into NoC to improve the transmission efficiency and reduce energy consumption, while ignored an im-

and middleware.

SHEN Yan, born in 1973, Ph. D., associate professor. Her current research interests include distributed measurement systems, embedded system development, wireless sensor networks, and robotics.

WU Yuan-Sheng, born in 1985, Ph. D. candidate. His research interest is embedded real-time system.

LIU Yun-Ben, born in 1981, Ph. D. candidate. His research interest is embedded real-time system.

portant limitation that NoC is a special network for cache coherence applications. That means the activity of protocol, or software, is the main reason of flip on components. So we try to reveal the protocol's impact on energy consumption of NoC system in this paper. In sum, this paper makes the following contributions; (1) A strict concept and definition of energy-oriented protocol and protocol energy consumption, which is a base of the following work. (2) Redivide NoC protocol stack into several layers from energy aspect. (3) A methodology to build an energy model for a NoC protocol. (4) An analysis and modeling of directory protocol energy consumption, and provides the support for energy-oriented protocol optimisation of NoC. (5) The method of using the model to compare protocols with different energy optimization depths. We hope that these works could help to evaluate novel optimized NoC in the realm of energy consumption.