

文章编号:1001-9081(2009)02-0511-03

## 一个新的模块化 NoC 路由器设计

张立鹏<sup>1</sup>, 顾华玺<sup>2</sup>, 王长山<sup>1</sup>

(1. 西安电子科技大学 计算机学院, 西安 710071; 2. 西安电子科技大学 综合业务网国家重点实验室, 西安 710071)

(lipengzh1983@163.com)

**摘要:**片上网络(NoC)作为复杂片上系统的有效解决方案,已经成为研究的热点。互连网络的性能很大程度上取决于构建网络的路由器结构。基于时延、吞吐和可靠性等考虑,提出一种基于虫孔交换的模块化设计的路由器结构。该结构采用路径分组,使用更小的交叉开关,同当前设计相比,很大程度上减少了输出端口竞争。且该结构自身具有一定的容错功能。

**关键词:**片上网络;路由器;路径分组;虚信道;交叉开关

**中图分类号:** TP302 **文献标志码:** A

## Design of a novel modular Network-on-Chip router

ZHANG Li-peng<sup>1</sup>, GU Hua-xi<sup>2</sup>, WANG Chang-shan<sup>1</sup>

(1. School of Computer, Xidian University, Xi'an Shaanxi 710071, China;

2. State Key Laboratory of ISN, Xidian University Xi'an Shaanxi 710071, China)

**Abstract:** As a complex system-on-chip solution, Network-on-Chip has become a recent research focus. The performance of the interconnection network mostly depends on the fabric of network routers. A kind of modular wormhole-switched router architecture was presented in this paper considering latency, throughput and fault-tolerance and so on. Compared with the existing designs, the proposed architecture which employs path-sets and uses smaller crossbars reduces the output port contention probabilities deeply. And the architecture itself has a certain degree functionality of fault tolerance.

**Key words:** NoC; router; path-sets; virtual channel; crossbar

### 0 引言

NoC 已经逐渐代替片上总线互连,成为片上系统的解决方案。NoC 互连技术具有系统可扩展性好、低互连功耗和低延迟的特点,因此成为系统结构研究的热点。虽然互连网络在多处理器结构方面的应用已经趋于成熟,但由于深亚微米技术的发展,面积、能耗和可靠性等方面的需求,给 NoC 的互连设计带来许多新挑战。以通信为中心,基于路由器互连并且使用规则拓扑,比如 2D mesh 和 torus<sup>[1-2]</sup> 已经成为连接多个 IP 核的实际标准。

以通信为中心的结构首先在文献[3]中引入。NoC 的实现在许多文献中都有介绍, Hermes<sup>[4]</sup> 是比较典型的一种。图 1 为一个典型的 5 端口分两步执行的 NoC 路由器结构。五个端口代表东西南北和本地这五个方向。路由过程主要包括五种基本操作:路由计算(Route Computing, RC),虚信道分配(Virtual channel Allocation, VA),开关分配(Switch Allocation, SA),虚信道缓存(Virtual Channel, VC)和交叉开关传输(Crossbar Traversal, CT)。许多研究都是通过并发执行一些操作来提高网络性能。文献[5]提出了规范的虚信道路由模型,采用投机策略假设 VA 可以成功获得虚信道,VA 和 SA 并行操作,这样就大大降低了网络的时延。RC、VA 和 SA 同时执行<sup>[1]</sup>,路由的关键路径变成两个阶段,第一个阶段是 RC、VA、SA 并行,第二个阶段是 CT,这样进一步提高了网络的性能。

文献[6]提出了虚信道分组的思想,采用了提前注出策

略,并且使用不完全的交叉开关来减小开关的规模,降低开关仲裁的复杂度,得到了较好的网络性能。文献[7]比较了几种常用路由性能,由仿真结果可知,没有一种结构是最好的,只能根据具体应用采用合适的策略以达到最佳效果。

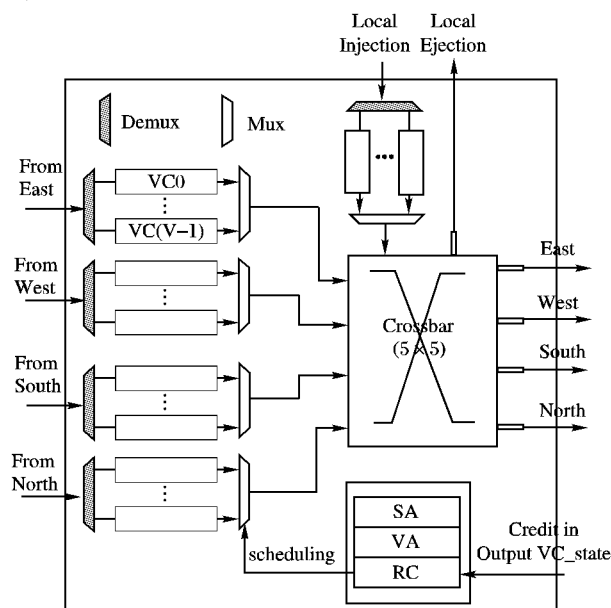


图 1 典型的虚信道路由器

本文提出了一种基于虫孔交换模块化的路由器结构。该结构兼顾时延、吞吐和容错功能,具有如下特点:支持自适应路由算法,能根据网络拥塞状况选择最优路线;使用更小的交

收稿日期:2008-09-03;修回日期:2008-10-15。 基金项目:国家自然科学基金资助项目(60803038)。

作者简介:张立鹏(1983-),男,山西运城人,硕士研究生,主要研究方向:片上网络的结构; 顾华玺(1978-),男,陕西西安人,副教授,博士,主要研究方向:片上网络、直连网络; 王长山(1950-),男,陕西西安人,副教授,主要研究方向:计算机网络。

叉开关,更简单的仲裁电路,从而节省了能耗;将路由器分成两个独立的模块,结构本身就具有一定的容错功能。

引入的路由器结构起名为 Positive-Negative 路由器,简称 Po-Ne。该路由器主要通过减少竞争率来提高性能。Po-Ne 将路由器划分为两个完全独立的模块。规定东、北方向为正方向,西、南方向为负方向,Positive Module 负责正向业务流量 (East-North), Negative Module 负责负向的业务流量 (West-South)。Po-Ne 将一般的仲裁分成路径分组和仲裁判断两部分。这样就能够使用更小、更简单的组件来减小逻辑电路的深度。每一个独立的模块需要一个简洁的  $2 \times 2$  交叉开关,相对于一般结构的  $5 \times 5$  交叉开关,其规模更小。交叉开关上的竞争通过使用一个预先的虚信道分组得到进一步减少。最后,该结构采用了 look-ahead 路由算法将到达本地节点的信息提前注出<sup>[6]</sup>。

## 1 Po-Ne 路由器

### 1.1 Po-Ne 路由器结构

图 2 是 Po-Ne 路由器结构示意图。路由器有东、西、南、北和本地注入五个端口,两组交叉开关,分别标记为 Positive-Module (East-North) 和 Negative-Module (West-South)。路由器被分成这两个独立的模块,负责各自连接处的业务传输。模块的每一个输入端口都有一组缓存 (三个虚信道组成) 用来保存从相邻路由器或者本地节点到达的数据片。这也就是路径分组的思想。该结构中数据分组提前为仲裁做了一部分前序工作,这就大大减少了输出端口处的竞争压力。提前注出模块根据提前路由得到的路由信息,将目的节点为本地端口的数据分组注出,无需经过 VA 和 CT,从而大大提高路由器性能。

路由过程如图 3 所示。如果目的节点就是本地则提前注出,否则进入相应的虚信道继续路由。

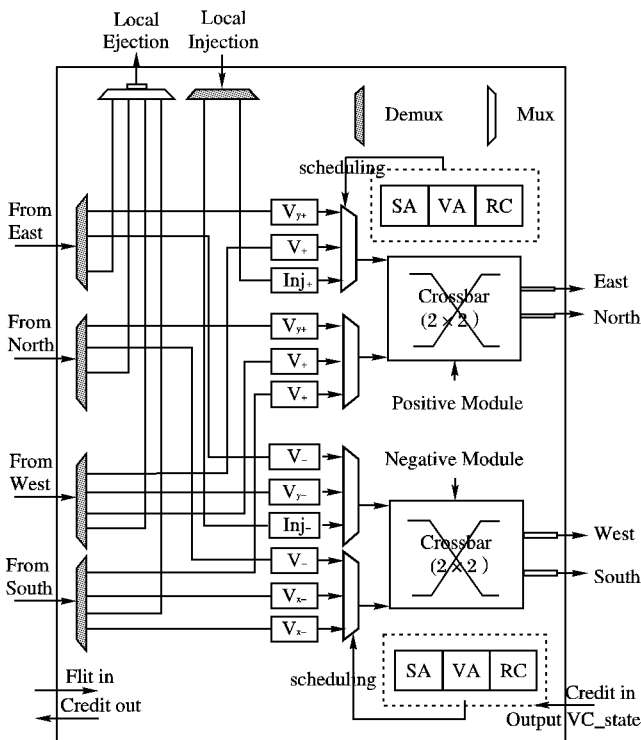


图 2 Po-Ne 路由器结构

### 1.2 虚信道设置

在图 2 中,  $V_{x+}$ ,  $V_{y+}$  用来保存到达  $x$  轴正方向和  $y$  轴正

方向的数据片。 $V_{x-}$ ,  $V_{y-}$  保存到达  $x$  轴负方向和  $y$  轴负方向传输的数据。 $V_{+}$ ,  $V_{-}$  保存到达正方向 (东、北) 和负方向 (西、南) 的数据片。比如, 一个从西到东或者到北的数据片存储在  $V_{+}$  虚信道中; 如果是从西到南的数据片就存储在  $V_{y-}$  中。本地注入的数据片存储在  $Inj_{+}$  和  $Inj_{-}$  中。采用的路由算法的不同, 虚信道缓存的数量和配置也需要跟着改变。一个死锁避免的确定性路由算法, 比如  $xy$  维序路由需要 8 个虚信道来实现 ( $2V_{+}$ ,  $2V_{-}$ ,  $1V_{y+}$ ,  $1V_{y-}$ ,  $1Inj_{+}$  和  $1Inj_{-}$ )。

死锁的避免是通过使用分离的虚信道消除网络结构中的循环依赖实现的。从不同方向到达同一个输出端口的数据片存储在不同的虚信道中, 如到达 East 端口的数据可能来自于 North、West、South 和本地端口, 且被存储在 1 个  $V_{y+}$ , 2 个  $V_{+}$  和  $1Inj_{+}$  四个不同的虚信道中, 从而实现开关内部无阻塞。在本文提出的这种结构中, 在 Po-Module 的第一组中的两个  $V_{+}$  和 Ne-Module 第一组中的两个  $V_{y-}$  用来避免直线传输方向的死锁和链式的循环依赖, 提供自适应路由。完全自适应路由需要 12 个虚信道来实现。

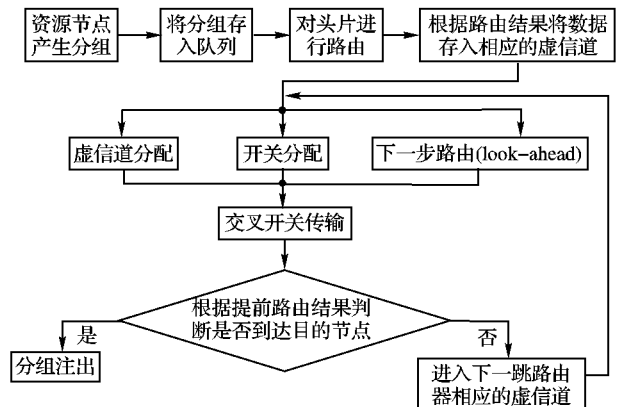


图 3 Po-Ne 路由器路由过程

### 1.3 仲裁复杂度分析

虚信道分配单元负责在所有请求相同虚信道的请求方中间做一个仲裁, 决定胜出者。图 4 比较了图 (a)、(b) 中两种结构的 VA 复杂度。第一种情况, 路由函数返回的是一个单一的虚信道, 另一种情况是路由函数返回的是一个单一的物理通道 (包含多个虚信道)。很明显, 在两种情况下 Po-Ne 路由器需要更少以及更小的仲裁器。Po-Ne 结构极大地降低了仲裁处理的复杂度。Po-Ne 路由器将虚信道分配操作划分为两个小的独立模块, 这样更进一步减小了仲裁单元的规模和复杂性。

开关分配器在所有连接到交叉开关的请求中做出仲裁, 赋予胜出者通过的权利。这样, 胜出者才能够通过交叉开关到达对应的输出链路。开关仲裁工作分两个阶段: 第一个阶段交叉开关的每个输入端口需要一个  $v$ -input 的仲裁器 (因为每个端口有  $v$  的虚信道), 也就是从每一个输出端口中选择一个虚信道里面的数据片来通过交叉开关; 第二个阶段每一个物理输出端口需要一个  $P$ -input 的仲裁器, 总共有  $P$  (表示端口数) 个仲裁器。为提高开关交换的速度, 开关分配器应该使用一种简单的算法而且要有较低的逻辑深度。为实现这个目标, Po-Ne 将开关分配模块分成两个小的模块, 每一个只负责  $2 \times 2$  的交叉开关。两种结构的 SA 复杂度对比如图 5 所示。

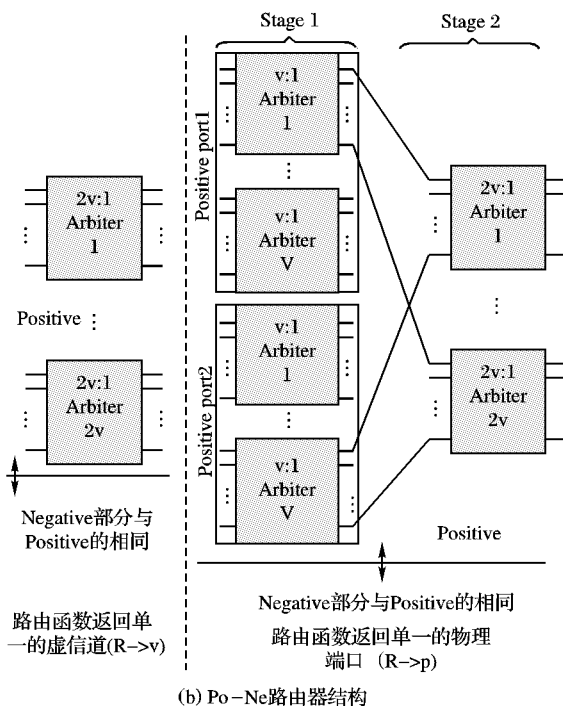
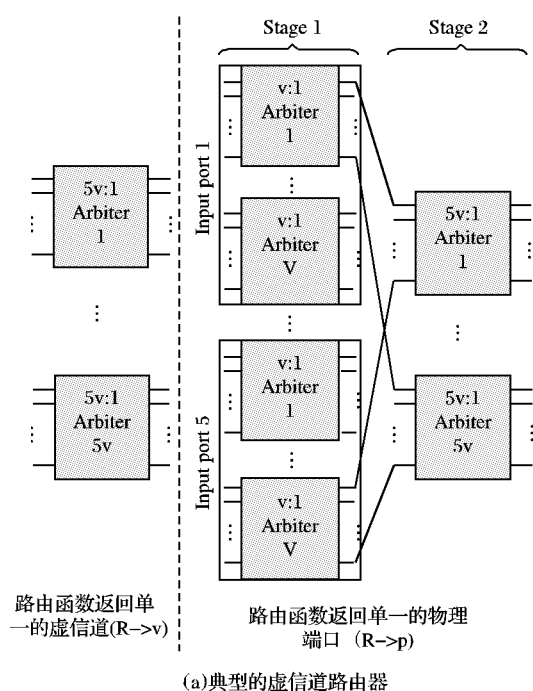


图4 VA复杂度比较

通过比较可以看到,无论是VA还是SA,Po-Ne路由器仲裁器的总个数和规模都小于常规路由器结构。而且Po-Ne路由器的仲裁判断能在两个独立的模块(每个模块都有独立的仲裁器)上并行操作,这样就大大提高了路由器的执行效率。

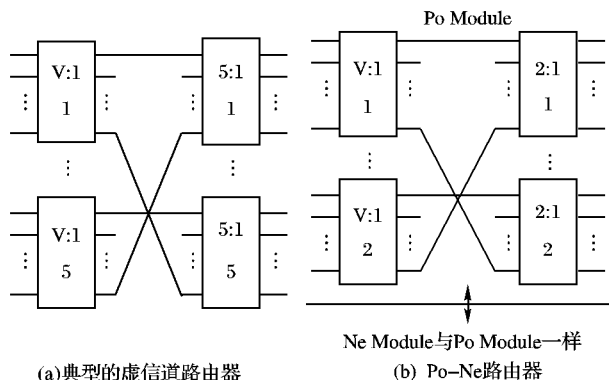


图5 SA复杂度比较

## 2 仿真

本文采用OPNET仿真软件建模,网络仿真采用 $4 \times 4$ 的2D-mesh结构,虫孔交换机制,链路速率为10 Mbps。分组长度采用确定形式——64 Byte,数据片大小为4 Byte。在均匀流量下,采用维序和奇偶转弯两种路由算法,对典型的虚信道路由器结构和Po-Ne结构的时延和吞吐进行仿真比较。典型路由器结构每个输入端口采用3个虚信道,Po-Ne总共使用12个虚信道。通过对比可以看出相同注入率的情况下Po-Ne结构的时延和吞吐性能明显优于一般的典型路由器结构。

## 3 结语

通过对现有的NoC结构研究,本文提出Po-Ne路由器结构。使用路径分组的思想将一个完整的路由器结构划分成两个几乎一样的独立模块,每个模块有自己的仲裁和交叉开关,负责各自业务流传输。这样,较大提高了每个模块的仲裁和路由效率,降低了组件的逻辑复杂度,同时很大程度上也降低了功耗。Po-Ne

结构还采用了提前路由策略,将到达本地节点的信息提前注出<sup>[6]</sup>。Po-Ne路由本身具有一定程度的容错属性。因为拥有两个独立的操作模块就意味着当其中一个模块遇到了一个错误被阻塞时,另一个模块能够继续提供服务。分析和仿真结果表明,Po-Ne路由器性能明显优于传统的典型路由器。

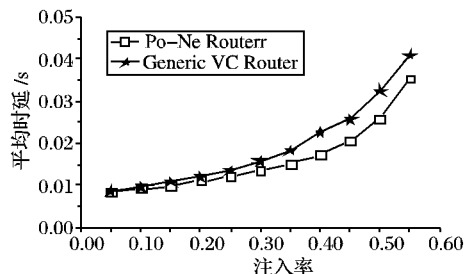


图6 维序路由的时延比较

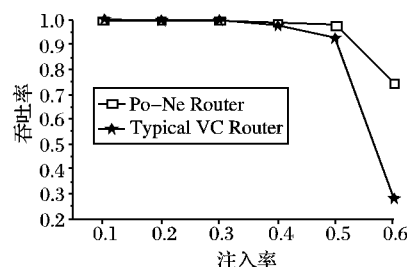


图7 维序路由的吞吐比较

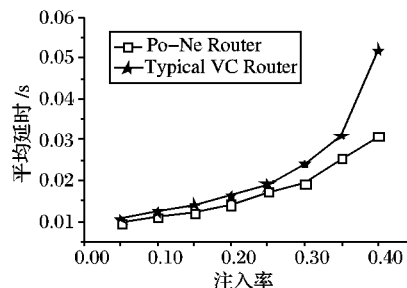


图8 奇偶转弯路由的时延比较

的所有节点的延时之和。

图2给出了本文算法与Chord在不同网络规模的环境下平均路由查找时间的对比情况。从模拟实验结果可以看出来,本文算法能够有效地降低路由查找时间,而且随着TTL的增加,路由查找时间越短,当 $TTL=1$ 时,延时减少约30%;当 $TTL=2$ 时,延时减少约35%。TTL越大,所获得的邻居信息越多,邻居信息越多,所覆盖的路由信息也就越多,路由查找总数优先考虑在邻居中进行查找,因此查找过程中节点都尽可能选择物理距离较近的节点进行路由,减少了路由过程中的网络延时。

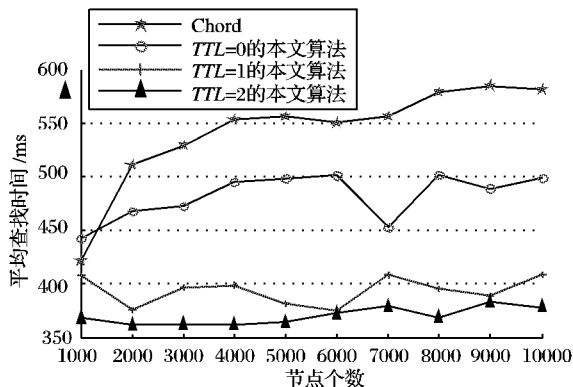


图2 本文算法与Chord的平均路由查找时间对比

图3给出了本文算法与Chord的平均查找跳数的对比。从图中可以看出,本文算法的平均路由跳数比Chord有所改进,但改善效果并不如平均查找时间那样高效,其原因在于:邻居表在一定范围内提供了逼近目的节点的信息,但不可能提供所有目的节点的信息,这种情况下,仍然需要依赖原始Chord算法进行路由。

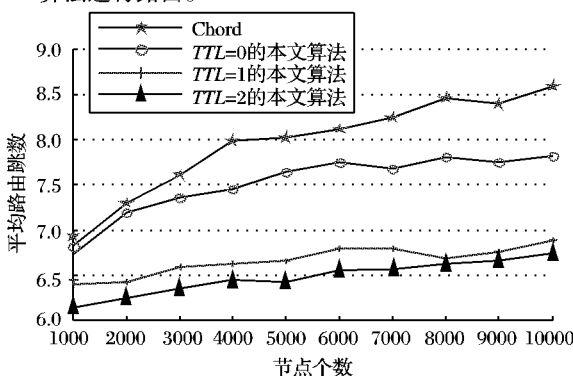


图3 本文算法与Chord的平均路由跳数对比

(上接第513页)

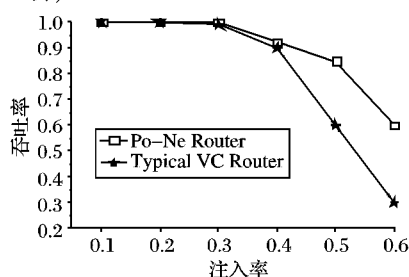


图9 奇偶转弯路由的吞吐比较

参考文献:

- [1] MULLINS R, WEST A, NOORE S. Low-latency virtual-channel routers for on-chip networks[C] // 31st Annual International Symposium on Computer Architecture. Washington: IEEE Computer Society, 2004: 188.
- [2] SANKARALINGAM K, NAGARAJAN R, LIU HAIMING, et al. Exploiting ILP, TLP, and DLP with the polymorphous TRIPS architec-

## 4 结语

本文提出一种基于有限范围组播的Chord路由改进算法,该算法充分利用IP网络层组播技术,在组播树上 $TTL \leq N(N=0,1,2,\dots)$ 范围内广播其自身的节点信息,并收集其他节点的邻居消息来建立邻居集,在路由过程中,优先考虑邻居集中更逼近目的地的邻居节点,这种路由方式在一定程度上消除了逻辑路径和物理路径之间不一致的问题。仿真实验表明,与基本Chord算法相比,本文算法能够大幅度缩短网络路由查找延时,而且TTL越大,效果越明显。下一步的工作将针对节点在Churn状态下的处理进行进一步的研究,提出更优化的邻居处理策略并进行仿真。

参考文献:

- [1] STOICA I, MORRIS R, KARGER D, et al. Chord: A scalable peer-to-peer lookup service for internet applications[C] // Proceedings of the 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM, 2001: 149-160.
- [2] ROWSTRON A, DRUSCHE P. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems[EB/OL]. [2008-06-21]. <http://www.research.microsoft.com/~antr/PAST/pastry.ps>.
- [3] RATNASAMY S, FRANCIS P, HANDLEY M, et al. A scalable content-addressable network[J]. SIGCOMM Computer Communication Review, 2001, 31(4): 161-172.
- [4] ZHAO B Y, LING HUNAG, STRIBLING J, et al. Tapestry: a resilient global-scale overlay for service deployment[J]. IEEE Journal on Selected Areas in Communications, 2004, 22(1): 41-53.
- [5] CASTRO M, DRUSCHEL P, HU Y C, et al. Proximity neighbor selection in tree-based structured peer-to-peer overlays[EB/OL]. [2008-06-21]. <http://research.microsoft.com/users/Cambridge/mcastro/publications/location-msrtr-2003-52.pdf>.
- [6] DABEK F, COX R, KAASHOEK F, et al. Vivaldi: a decentralized network coordinate system[J]. SIGCOMM Computer Communication Review, 2004, 34(4): 15-26.
- [7] Rfc2362. Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol specification[S/OL]. [2008-06-21]. <http://www.faqs.org/rfcs/rfc2362.html>.
- [8] Rfc2236. Internet Group Management Protocol, Version 2[EB/OL]. [2008-06-21]. <http://www.ietf.org/rfc/rfc2236.txt>.
- [9] PlanetSim[EB/OL]. [2008-06-21]. <http://ants.ctse.uv.es/planetsim>.
- [10] p2psim[EB/OL]. [2008-06-21]. <http://pdos.csail.mit.edu/p2psim/kingdata/>.

ture[C] // Proceedings of the 30th Annual International Symposium on Computer Architecture, New York: ACM, 2003: 422-433.

- [3] DALLY W J, TOWLES B. Route packets, not wires: On-chip interconnection networks[EB/OL]. [2008-06-05]. [http://cva.stanford.edu/publications/2001/onchip\\_dac01.pdf](http://cva.stanford.edu/publications/2001/onchip_dac01.pdf).
- [4] MORAES F G, CALAZANS N L V, De MELLO A V, et al. Hermes: an infrastructure for low area overhead packet-switching networks on chip[EB/OL]. [2008-06-05]. <http://www.inf.pucri.br/tr/tr034.pdf>.
- [5] PEH L S, DALLY W J. A delay model and speculative architecture for pipelined routers[EB/OL]. [2008-06-05]. <http://cva.stanford.edu/publications/2001/specmodel.pdf>.
- [6] KIM J M, PARK D, THEOCHARIDES T, et al. A low latency router supporting adaptivity for on-chip interconnects[EB/OL]. [2008-06-05]. <http://www.stanford.edu/group/comparch/papers/kim05.pdf>.
- [7] KELLEY K, NAOUS J, SHACHAM O. Using speculation to reduce latency in a virtual channel router[EB/OL]. [2008-06-05]. <http://cva.stanford.edu/classes/ec382c/research/speculation.pdf>.