

## 基于统计时分复用技术的三维片上网络

王佳文 李丽\* 潘红兵 李伟 张荣

(南京大学微电子设计研究所 南京 210093)

(南京大学江苏省光电信息功能材料重点实验室 南京 210093)

**摘要:** 在片上网络(Network on Chip, NoC)系统中,本地子系统通常基于总线结构,而全局通信则由基于包交换的网络构成。然而,由于总线和网络之间通讯机制的差异,当本地子系统内各核访问全局资源的时候,系统整体性能将下降。在3D NoC中,由于全局网络规模的扩大,该问题将越发显著。对此,该文提出一种基于统计时分复用(Statistical Time Division Multiplex, STDM)技术的3D NoC架构。该架构首先在本地球子系统引入STDM控制单元,然后在网络接口设计中增加了计数及等待机制,并对路由节点针对STDM技术进行了优化设计,以增强对STDM的支持,减小总线、网络间的差异。同时,该文还充分利用STDM帧的特点,设计了一种新的数据包格式,以进一步降低全局通信的网络负荷。为证明新方案的高效,该文采用SystemC语言进行系统级建模,仿真结果表明:该方案在降低网络负荷、减小通信延时方面有着显著效果。最佳情况下,两者可以分别降低为传统方案的45%和20.5%。而实际应用中,尤其对于通信密集型应用而言,该方法的改善效果也同样明显。

**关键词:** 3D片上网络;统计时分复用;层次化存储架构

**中图分类号:** TP302.2

**文献标识码:** A

**文章编号:** 1009-5896(2012)10-2501-07

**DOI:** 10.3724/SP.J.1146.2012.00099

## 3D NoC Based on Statistical Time Division Multiplex

Wang Jia-wen Li Li Pan Hong-bing Li Wei Zhang Rong

(Institute of VLSI Design, Nanjing University, Nanjing 210093, China)

(Jiangsu Provincial Key Laboratory of Photonic and Electronic Materials Sciences and Technology,  
Nanjing University, Nanjing 210093, China)

**Abstract:** In the system of Network on Chip (NoC), local systems normally adopt bus-based architectures while global networks use packet-based communications. However, since the natural difference between these two architectures is unavoidable, it introduces performance degradation to the overall system when cores in local systems visit global resources. And the situation will be worse in the environment of 3D NoC due to its larger network sizes. In this paper, a 3D NoC based on Statistical Time Division Multiplexing (STDM) is proposed. Firstly, a STDM controller is introduced into the local system. Then the network interface is designed using the mechanism of counting and waiting. Finally, the router is optimized for supporting the STDM better. And novel packet formats are designed to reduce the network load and improve the system performance further. In order to demonstrate the efficiency of the novel method, a systemC model is built on system level and experimental results show that the proposed method can reduce network load and transmission delay sharply. In the best condition, the improvement can reach 45% and 20.5% compared with the conventional one. And for real applications, especially for communication-intensive ones, the proposed architecture can also improve the performance a lot.

**Key words:** 3D Network on Chip (3D NoC); Statistical Time Division Multiplex (STDM); Hierarchical memory architecture

### 1 引言

自片上网络(Network on Chip, NoC)<sup>[1]</sup>概念提

出以来,其采用的网络通信架构的确在一定程度上缓解了系统瓶颈问题。然而当核数急剧增加时,仍然无法避免全局连线过长等一系列相关问题的出现。于此同时,3D IC技术<sup>[2]</sup>日趋完善,并且由于其良好的发展前景而备受瞩目。由于其将器件分布在各层上,而层间通常采用硅通孔(Through Silicon Vias, TSV)技术垂直互连,因此其能够有效缩短器件之间的互连线长度,避免大平面跨度的长互连线,

2012-02-06 收到,2012-07-09 改回

国家自然科学基金(60876017, 61176024),江苏省产学研前瞻性联合研究项目(BY2009146),中央高校基本科研业务费专项资金(1095021031)和江苏省普通高校研究生科研创新计划(CX10B\_021Z)资助课题

\*通信作者:李丽 lili@nju.edu.cn

从而降低连线复杂度、减少互连线效应影响,以及降低拥挤度。作为上述两种技术的融合,3D NoC被提出并受到广泛关注<sup>[3,4]</sup>。考虑到3D NoC系统一般涉及较为庞大的全局网络,因此为提高系统访存能力,在设计时常采用基于层次化的设计。一般而言,层次化3D NoC系统可根据存储层次的不同分为多个等级:(1)单个处理器核,各核配置私有存储器,即三级存储单元(memory level3);(2)由若干个核构成的本地子系统,子系统内部还包含二级存储单元(memory level2),用于核间数据交互;(3)若干子系统又以某种拓扑架构互连,从而构成整个系统,并配有一级存储单元(memory level1)。

在层次化3D NoC中包含3种存储层次,因此各处理器可以分别在不同层次上同时进行访存,增强了系统对存储单元的并发访问能力,因而可以显著提高系统性能。然而,和传统NoC相比,3D NoC又具有规模较大,节点较多等特点,这也就增强了系统对全局通信能力的需求,而这正是本文研究的主要出发点。

目前,诸多研究机构针对3D NoC架构提出各种方法来改善系统性能,他们中多数从全局互连的角度出发,兼顾考虑系统性能和灵活性两个方面进行优化。文献[5]也指出,在3D NoC中,单就拓扑架构的选择这一方面就比以往更加丰富,也更具挑战性。如文献[3]利用其建立的精确仿真模型比较了多种常见的3D NoC架构在吞吐率、延迟等方面的性能差异。文献[6]利用3D IC在垂直方向互连线长度明显小于水平方向的特点,提出在网络-总线混合结构,结果表明该架构可以充分发挥3D层间通信优势。文献[7]则设计了一种网格-树架构,其可以有效提高单位面积下的系统性能。不仅拓扑架构方面,在IP核及存储布局问题上,也涌现出不少研究成果。仅就3D NoC来说,其存储,IP核布局就出现了诸如Dance-Hall架构、Sandwich架构、Per-Layer架构等<sup>[8]</sup>,该文献在统一随机交通以及本地交通两种交通情形下对上述各种架构在吞吐率、延迟等方面做出对比,进而给出其在不同情形、不同目标下各自的优劣情况。文献[9]则结合H.263编码这一具体实例,考虑3D NoC中存储、IP核的多种布局对系统性能的不同影响。

然而上述各方法多关注全局互连在整体通信中的作用,如全局拓扑选择,IP布局等,而忽略了所有传输的起点和终点均在本地子系统内部这一事实,因此也就忽略了本地子系统和全局网络交互过程的优化空间。然而,由于本地子系统通常采用总线架构而全局网络采用基于包交换的通信<sup>[10,11]</sup>,因

此这两者间的差异必然造成极大的通信开销而成为系统访存瓶颈所在。正如同文献[12]中所说,如何有效地将本地子系统内的局部通信和全局网络间的通信结合在一起已经成为迫在眉睫的问题。

本文提出一种基于统计时分复用(Statistical Time Division Multiplex, STDM)技术的层次化3D NoC。与传统做法中以时间片为基本单位不同,该技术以STDM帧为复用的基本单位,每一帧可以包含多个主设备信息。这种以帧为基本单位的传输模式和全局通信中采用的包交换传输模式一致,因此该方法可减少两者间通信的差异,有效增强各核访问网络的能力。同时,本文充分利用同一帧内数据间的相关性,改进帧格式,进一步提高信息密度,降低了网络负荷。

## 2 基于STDM技术的3D NoC架构

### 2.1 基于STDM技术的本地子系统架构

传统总线通常采用时分复用(Time Division Multiplexing, TDM)技术,即总线上的各主设备(master)以基于时间片的方式依次使用总线,每一时间片内只有一个主设备拥有总线使用权。某主设备获得总线使用权时,其他主设备均处于等待状态,如图1中基于TDM技术的总线所示。

由图1可见,当主设备1(master1)获得总线使用权后,其他所有主设备均处于等待状态,直到主设备1释放总线后,才可以申请使用。在通常情况下,某个主设备一次占用总线的时间 $t$ 较短,因此在主设备数量不多,且访问近端(挂接于本总线)从设备时,此方式不失为一种高效的复用方法。然而,当这种机制的总线被直接用于3D NoC各子系统时,则当各主设备需要访问远端(不挂接于本总线)从设备(如3D NoC中的memory level1)时,其消耗的时间 $t$ 将显著增加,对于需要从远端返回数据包的读操作而言更是如此。而在此期间,其他各主设备将只能耗费大量时间等待总线使用权,这显然会极大影响系统整体性能。

分析可知,时间开销 $t$ 可分为3段:(1)源端传输层处理时间 $t_1$ ;(2)网络层传输时间 $t_2$ ;(3)目的端传输层处理时间 $t_3$ 。其中, $t_2$ 在总时间中所占比例最大,且随着网络尺寸的增加而增加。

一般而言,在NoC系统中,在划分子系统时会采用各种映射算法<sup>[13]</sup>以使得各子系统内各核的任务之间具有较高的相关度,而各子系统之间则相关度较低。对于层次化3D NoC而言,由于其存储层次分明,这种相关度将体现得越发明显。因此对于某一子系统而言,其所需要访问的数据在存储空间上常具有相关性,具体体现为:(1)访存对象的网络地

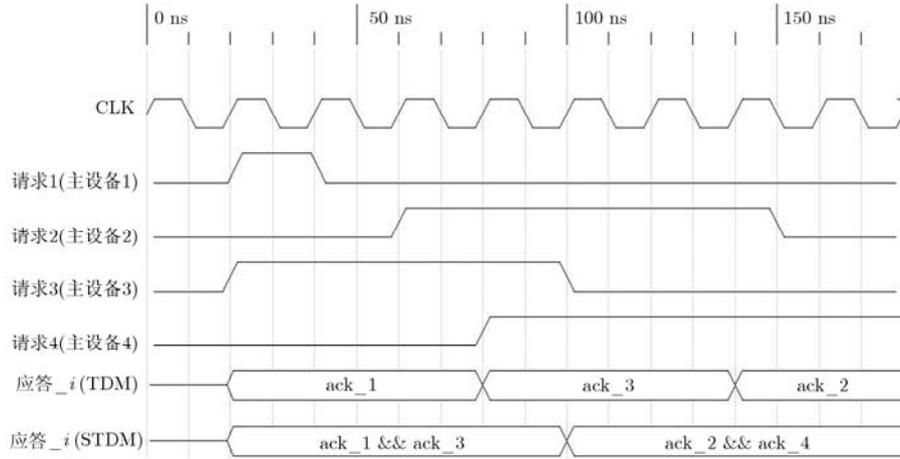


图 1 两种机制的总线访问示意图

址一致；(2)地址之间的偏移小于某设定阈值。一方面，如果充分利用这种相关性，则可以使各主设备的访存需求被打包在同一网络包内完成，以此分摊  $t_2$  的开销。另一方面，考虑到 STDM 技术中以帧为基本单位这一特点，可以将多个主设备访问相同网络地址的任务放置到同一帧内，交由 NI(Network Interface)进行统一打包、传送处理。这样，各主设备访存开销  $t$  中的  $t_2$  部分可以共用，而  $t_1$  和  $t_3$  部分则分别变为  $t_1'(t_1' > t_1)$  和  $t_3'(t_3' > t_3)$ 。然而，如前所述， $t_2$  是  $t$  中比重最大的，因此总体时间仍然可以得到有效降低。其具体访问过程如图 1 中基于 STDM 技术的总线所示。

在实际传输过程中，子系统内各核在同一时间点(在数字电路中表现为同一时间周期)发起访问总线请求的概率不高。针对这一问题，本设计将在网络接口设计中引入等待和计数机制(具体见 2.2 节)，使得即便多个主设备申请访问总线的时间稍有先后，只要其间隔小于等待时间上限，仍可视其为同时发生请求，放到同一 STDM 帧内处理，从而提高 STDM 帧形成概率，使系统实用性更强。

本地子系统最终采用的架构如图 2 所示(以 3 个主设备为例)。

由图 2 可以看出，该架构与传统总线的区别主要在于：将传统总线中从设备反馈信号里的部分逻辑分离出来，交由 STDM 控制单元处理，而其他部分则保持不变。其中，STDM 控制单元和所有涉及到 STDM 帧处理的单元进行控制交互，尤其是增加了网络接口(Network Interface, NI)和控制单元之间特殊控制信号的交互，以解决前文所述的主设备申请访问总线时间稍有先后的问题。

具体而言，当某个主设备申请使用总线时，首先由 NI 对其进行判断，如果该申请和已经记录在案

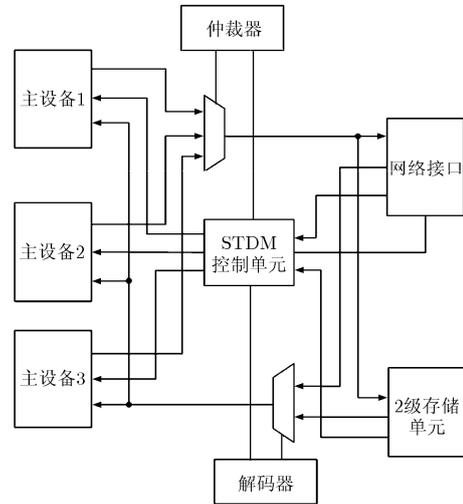


图 2 基于 STDM 技术的 3D NoC 簇内架构示意图

的申请具有相关性，则可以放在同一 STDM 帧内处理。此时，STDM 控制单元将记录下该申请号，NI 进入等待计时状态。如果等待计时  $t$  超过了允许上限  $t_{max}$ ，则处理所有已记录的主设备申请，并对总线进行锁定，其他主设备在处理过程完成之前均无法使用总线。反之，如果在等待时间内有新的主设备进行总线申请，则重复上述过程。除了等待时间有上限  $t_{max}$  以外，STDM 控制单元中所能记录下的主设备申请个数也设置了上限( $n_{max}$ )，如果记录的申请个数达到这一限制，同样会跳出等待状态，进入对申请的处理及锁定总线过程。

### 2.2 全局通信路由节点及包格式设计

由于全局网络采用基于包交换的通信技术，因此通常而言，其数据包由包头、包体和包尾 3 部分组成，其中，(1)包头：包含必要的路由和网络控制信息；(2)包体：传输的主体部分，其主要内容为 STDM 帧，通常由控制信息和有效数据两部分组成；

(3) 包尾：指示包的结束。对于任何类型的数据包，其包头和包尾均由一个数据帧(flit)构成，但包体部分则有所不同。对于普通数据包，包体部分由一个数据帧组成，而对于基于STDM技术形成的数据包，其包体部分则由无法事先确定个数的数据帧组成。因此在STDM技术引入后，网络数据包具有两种基本类型，且在STDM类型下，其包体长度为不定值。

以总线宽度为32 bit的情况为例，最终设计的数据包格式如图3所示。其中每个数据帧均为35 bit，包头中的R\_addr为目的节点路由信息，LS\_addr为初始地址。包体中各负载为实际传输内容，通常情况下为传输的有效数据，但在STDM类型的包中还包括有记录相应控制信息的负载。以与数据信息相对应的地址信息为例，在包体的第1个负载中(负载1)，在STDM类型下，为充分利用数据间相关的特性，后续数据相应地址不予以传递，而只传递其与起始数据对应地址之间的偏移关系，后续地址由目标节点根据该信息运算得出，以此降低网络负荷，从负载2开始才是相应数据信息。而普通包中则仅包含数据负载。包尾中的负载则包含本地子系统的网络标识，主要在读操作时为返回包提供相应信息。而各数据帧中均包含的H\_flag是用来区分当前数据帧为包头、包体还是包尾的，同时为了区分两种类型的数据包，包头部分的H\_flag又分为两种类型，其中，01表示普通数据包，11表示STDM数据包。各数据帧末尾的O\_flag则用来区分读/写操作。

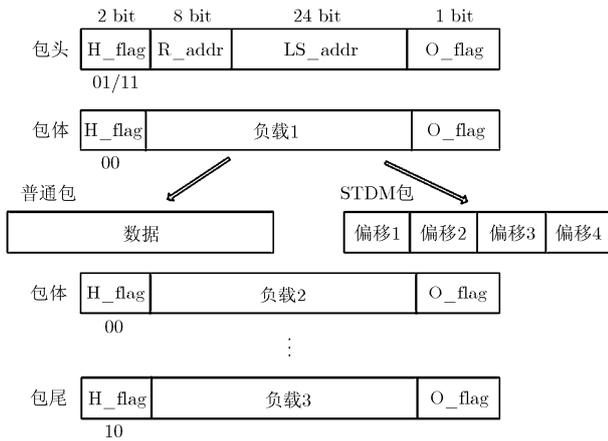


图 3 包格式示意图

为适应两种不同类型的数据包，在路由节点设计中增加了对包类型的判断以及对两种类型包的不同处理过程，其逻辑框图如图4所示。由图可见，路由节点由东、南、西、北、上、下及本地7对输入输出通道组成。每个输入通道中主要包含缓冲、路由器、输入输出握手3个模块，每个输出通道则

主要包含多路选择器、仲裁器和输出控制3个模块，模块间则由数据及控制通路相互连接。其中对包类型的区分由OFC(Output Flow Controller)模块完成，并产生相应控制信号以影响仲裁模块和数据通道的使用。

### 3 实验与分析

#### 3.1 实验环境

本设计搭建了一款基于SystemC语言的周期精确系统级仿真模型。该模型采用如图5所示的3维网格架构，其中图5(a)为整体架构示意图，图5(b)为簇内部架构示意图。其中，网络尺寸可参数化配置，本地子系统分为两类，即图2所示的基于STDM技术的运算子系统和由1级存储单元(memory level1)构成的存储子系统，且运算子系统内的主设备数目也可参数化配置。全局网络采用静态路由算法，即按照先X后Y最后Z的顺序来完成路由。同时，每个运算子系统内的处理器核数固定为6个，STDM控制单元中所能记录的主设备申请个数上限(n\_max)设置为4，等待时间上限设置为30个时钟周期。

#### 3.2 实验结果与分析

本实验主要关注新设计方案对网络负荷与执行时间的影响。

(1) 网络负荷 如图6所示，其中横坐标表示连续传输的相关数据个数，纵坐标网络负荷(单位为传输过程中所包含的数据帧个数)。由图可见，当传输单个数据(或不相关数据)时，由于本设计兼容普通数据包和STDM数据包两种类型，因此对普通数据传输没有任何不利影响。而当传输相关数据时，随着连续传输的数据个数的上升，新方案的改善效果也随着提升，并且稳定在原先值的50%-60%左右，有效地降低了网络负荷。同时，由于读操作包含发送的请求包及返回的数据包，因此其改善效果比只有发送数据包的写操作更为明显。

(2) 执行时间 除新方案自身参数配置以外，实验环境也会对执行时间的结果造成影响，如网络繁忙程度，路由算法选择，传输距离远近等，本文选取传输距离远近，即曼哈顿距离(即源节点和目的节点之间的网络距离)变化来表征新设计在传输环境变化时性能改善情况。在实验中，每个处理器核先将各自私有存储单元中的600个数连续写入到具有地址相关性的共享存储单元内，然后再将这600个数读回。上述过程保持不变，曼哈顿距离从2变到14，用传输时延比值表示其对执行时间的影响，其实验结果如图7所示。为了更明显的显示性能变化

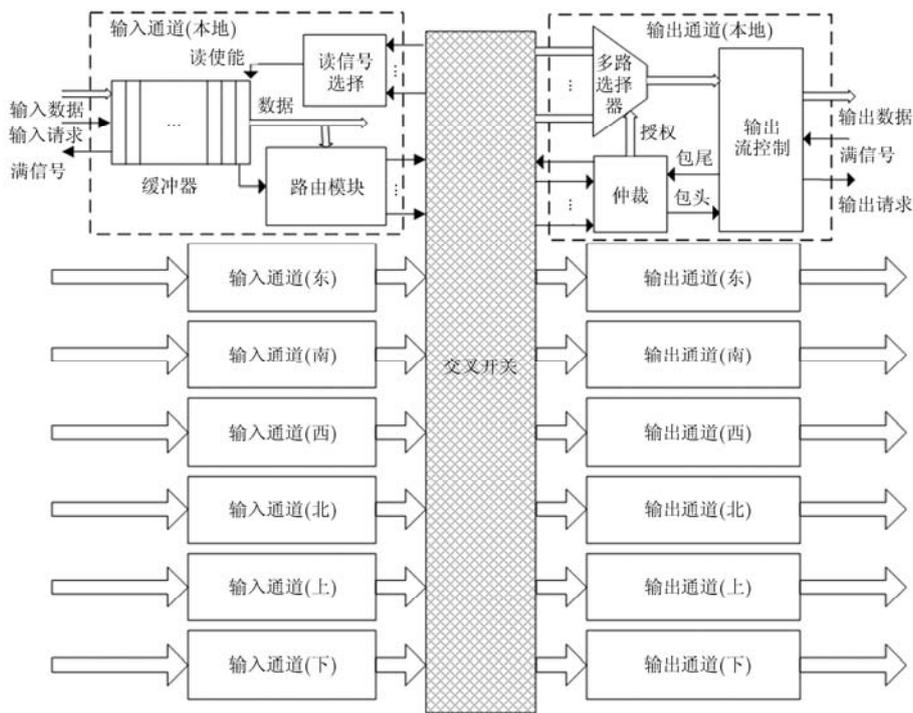


图 4 路由节点设计框图

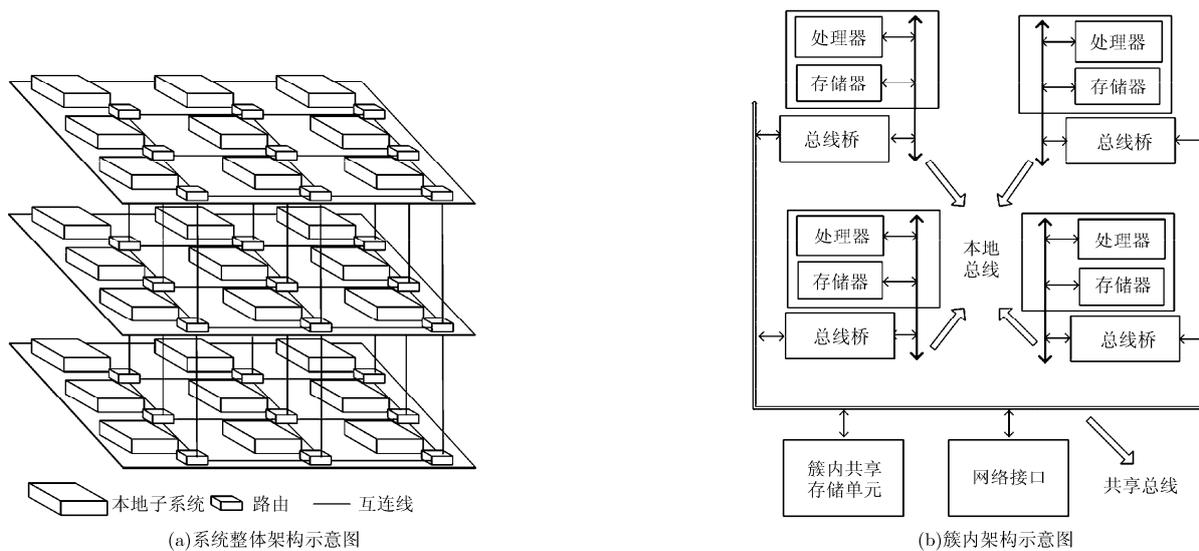


图 5 系统级模型整体架构示意图

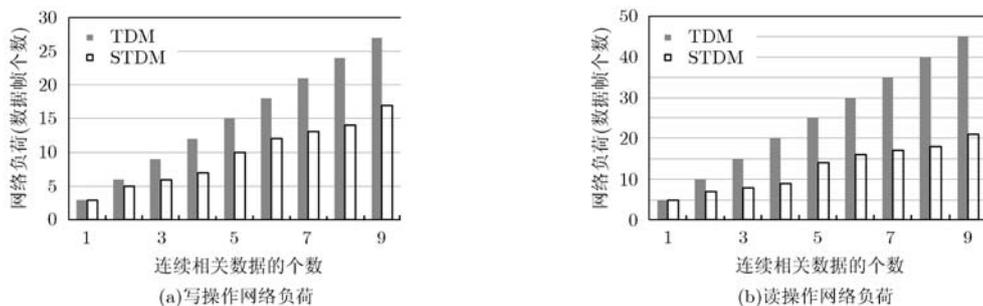


图 6 基于 STDM 技术的簇对网络负荷的影响

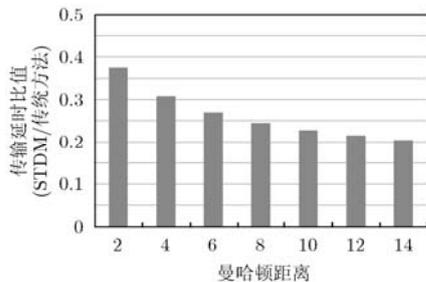


图 7 基于 STDM 技术的簇对执行时间的影响

过程, 这里将 TDM 技术下的执行时间归一化处理, 因此未在图中画出。

由图 7 可见, 新方案对执行时间有着明显的改善效果。而且随着曼哈顿距离的上升, 新方案的改善效果也随之上升。这是因为曼哈顿距离越长,  $t_2$  在  $t$  中所占比例越大, 而新方案的主要作用即多个主设备分摊这一开销, 因此当这一开销自身在整体开销中比重越大时, 新方案的改善效果也就越明显。当然, 新方案的性能受限于 NI 及 STDM 控制单元能记录的最大连续传输个数, 因此其性能的改善在达到一定值以后趋于平缓, 所以图中曲线呈现出斜率逐渐变小的状况, 在最佳状况下其执行时间可降到传统方案的 20.5% 左右。考虑到 3D NoC 相较于 2D NoC 具有网络尺寸大的特点, 正好符合本设计在较大网络距离下性能越发良好的特性, 因此本设计是非常适用于 3D NoC 平台的。

(3) 应用实例 为进一步从实用性角度验证本文所提出的方法, 以实际应用中广泛涉及的矩阵运算为例, 说明基于 STDM 技术 3D NoC 的优势。基于 STDM 技术的 3D NoC 主要针对簇内各核访问网络上 1 级存储单元(memory level1)的性能进行优化, 因此更加适用于通信密集型应用。本文在网络尺寸为  $4 \times 4 \times 4$ , 存储单元分布如引言部分所述的 3D NoC 系统中进行了实验。首先采用单核, 对两个大小为 24 阶的矩阵进行求和运算, 两个初始矩阵均存放在相距运算核曼哈顿距离为 6 的 1 级存储单元内。然后采用传统 3D NoC(簇内集成 6 核)技术对同样矩阵进行处理, 簇内各核采用循环分行法读取矩阵中相应行并完成运算, 以实现算法并行化。最后在采用基于 STDM 技术的 3D NoC(簇内仍然集成 6 核)内完成上述运算, 并行化方式和传统 3D NoC 一致。

将加速比定义为: 单核执行时间/多核执行时间, 实验最终运行结果如图 8 所示。由图 8 可知, 在传统架构下, 对于通信密集型应用, 各核只能依次通过总线访问网络, 因此效率很低, 虽然簇内集成了 6 核, 加速比却仅为 1.71, 也即多核的优势被

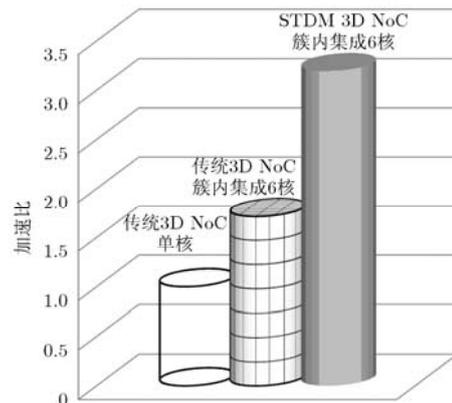


图 8 与传统 3D NoC 方案加速比对照

大量的通信开销所掩盖, 而无法充分发挥各核效率。当采用基于 STDM 技术后, 其加速比提高为 3.19, 也即各核利用率达到 53.2%, 比传统架构性能提升约 86.5%, 足见新设计对于提高算法性能有着极其显著的作用。

#### 4 结束语

本文针对 3D NoC 中总线与网络混合架构, 提出了一种基于 STDM 技术的 3D NoC 系统架构及通讯机制, 并设计了一种新的数据包格式, 以期更好地融合总线和网络两种通讯方式, 降低网络负荷, 减少通信延时, 提升系统通信能力。实验结果表明, 在最佳情况下, 本文提出的方案可以分别降低为传统方案网络负荷的 46.7% 和通信延时的 20.5%。而对于真实的通信密集型应用, 采用 STDM 技术的 3D NoC 和传统 3D NoC 相比, 加速比可以提升 86.5%。

因此基于 STDM 技术的 3D NoC 更适用于通信密集型应用, 可在网络通信、人工智能等领域得到广泛应用。

#### 参考文献

- [1] Tobias Bjerregaard, Shankar Mahadevan, et al. A survey of research and practices of Network-on-Chip[J]. *ACM Computing Surveys*, 2006, 38(1): 1-51.
- [2] Topol A W, Tulipe D C L, et al. Three-dimensional integrated circuits[J]. *IBM Journal of Research and Development*, 2006, 50(4): 491-506.
- [3] Feero B S and Pande P P. Networks-on-Chip in a three-dimensional environment: a performance evaluation[J]. *IEEE Transactions on Computers*, 2009, 58(1): 32-45.
- [4] Awet Yemane Weldezion, Matt Grange, Dinesh Pamunuwa, et al. Scalability of network-on-chip communication architecture for 3-d meshes[C]. Proceedings of the International Symposium on Networks-on-Chip, San Diego, CA, May 2009: 114-123.

- [5] Marculescu R, Ogras U Y, *et al.*. Outstanding research problems in NoC design: system, microarchitecture, and circuit perspectives[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2009, 28(1): 3-21.
- [6] Li F, *et al.*. Design and management of 3D chip multiprocessors using network-in-memory[C]. Proceedings of 33rd International Symposium on Computer Architecture, Boston, USA, 2005: 130-141.
- [7] Jeang Yuan-long, Wey Tzue-shaang, Wang Hung-yu, *et al.*. Mesh-tree architecture for network-on-chip design[C]. Proceedings of Innovative Computing, Information and Control, Kumamoto, Japan, 2007: 262-265.
- [8] Weldezion A Y, Lu Zhong-hai, Weerasekera R, *et al.*. 3-D memory organization and performance analysis for multi-processor network-on-chip architecture[C]. IEEE International Conference on 3D System Integration, San Francisco, CA, 2009: 1-7.
- [9] Xu T C, Yin A W, Liljeberg P, *et al.*. A study of 3D network-on-chip design for data parallel H.264 coding[C]. NORCHIP, Trondheim, 2009: 1-6.
- [10] Radulescu A, Dielissen J, and Pestana S G, *et al.*. An efficient on chip NI offering guaranteed services, shared memory abstraction and flexible network configuration[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2005, 24(1): 4-17.
- [11] ARTERIS. 2005. A comparison of network-on-chip and buses [OL]. <http://www.arteris.com/nocwhitepaper.pdf>, 2005.
- [12] Ahonen T and Nurmi J. Hierarchically heterogeneous network-on-chip[C]. The International Conference on "Computer as a Tool", Warsaw, Poland, 2007: 2580-2586.
- [13] Marcon C A M, Moreno E I, Calazans N L V, *et al.*. Comparison of network-on-chip mapping algorithms targeting low energy consumption[J]. *IET Computers & Digital Techniques*, 2008, 2(6): 471-482.
- 王佳文：男，1985年生，博士，研究方向为多核处理器体系架构、片上网络通信架构及通信协议。
- 李丽：女，1975年生，博士，教授，主要研究方向为研究多核处理器芯片体系结构及实现方法、SoC-IP设计方法学。