

# 针对路径故障与局部拥塞的 NoC 容错路由算法

欧阳一鸣<sup>1</sup>, 何鑫城<sup>1</sup>, 梁华国<sup>2</sup>, 易茂祥<sup>2</sup>, 杜高明<sup>2</sup>, 安鑫<sup>1</sup>

(1. 合肥工业大学计算机与信息学院, 安徽合肥 230009; 2. 合肥工业大学电子科学与应用物理学院, 安徽合肥 230009)

**摘要:** 片上网络作为一种新型片上互连架构, 克服了片上系统在发展中遭遇的瓶颈问题. 然而, 片上网络中的路由器故障以及路由器之间的链路故障都会造成网络性能损失. 对此, 文章提出一种针对路径故障与局部拥塞的 NoC 容错路由算法. 首先, 设计了一种相隔节点间路径故障模型, 该模型下的路由器以较小的开销为代价, 动态感知两跳以内的路径故障状态. 其次, 提出了一种新颖的更能准确反映局部网络拥塞状态的拥塞模型来均衡网络流量. 最后, 当网络无故障时, 算法保证走最优路径; 有故障时, 算法不仅可以实现容错还能保证网络具有良好的性能. 实验表明, 在无故障的情况下, 本文方案相较于对比对象延迟降低了 10% ~ 20%, 吞吐率提高了 25% 左右. 在有故障的情况下, 本文方案较对比对象的优势更加明显.

**关键词:** 片上网络; 故障模型; 拥塞模型; 容错路由算法

**中图分类号:** TP302      **文献标识码:** A      **文章编号:** 0372-2112 (2016)04-0920-06

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2016.04.024

## A Fault-Tolerant Routing Algorithm Aiming at a Path Fault and Local Congestion in NoC

OUYANG Yi-ming<sup>1</sup>, HE Xin-cheng<sup>1</sup>, LIANG Hua-guo<sup>2</sup>, YI Mao-xiang<sup>2</sup>, DU Gao-ming<sup>2</sup>, AN Xin<sup>1</sup>

(1. School of Computer and Information, Hefei University of Technology, Hefei, Anhui 230009, China;

2. School of Electronic Science & Applied Physics, Hefei University of Technology, Hefei, Anhui 230009, China)

**Abstract:** As a new type of on-chip interconnection architecture, network-on-chip overcomes the bottleneck problem of the system-on-chip during the development. However, a failure arising in a router or a link between routers in network-on-chip will cause the reduction of network performance. To avoid this phenomenon, this paper puts forward a fault-tolerant routing algorithm aiming at a path fault and local congestion in network-on-chip. Firstly, the algorithm designs a fault model that reflects the fault status of the path within two hops. As a result, this novel fault model makes the router achieve a dynamic perception of path state within two hops with less cost. Secondly, a novel congestion model has been proposed for reflecting the state of the local network more accurately, contributing to balance network traffic. Finally, when a fault occurs, the algorithm not only is fault-tolerant but also makes sure the network has a good performance. What's more, the algorithm chooses the optimal path under the condition of fault-free. Experimental results show that the proposed algorithm has 10% ~ 20% lower latency in average and 25% higher throughput rate than the contrast case when the network is fault-free. In the case of defective in the network, the advantage of the present scheme has a bigger superiority.

**Key words:** network-on-chip; fault model; congestion model; fault-tolerant routing algorithm

## 1 引言

随着单个芯片上集成的核越来越多, 多核以及众核系统中同时有超过一个任务在运行的可能性也越来越大. 这就使得单任务执行的片上系统 (System-on-Chip, SoC) 在发展过程中遭遇瓶颈. 鉴于此, 有研究者提

出通过借鉴计算机网络和并行计算技术设计另一种新颖的片上互连架构——片上网络 (Network-on-Chip, NoC), 该架构达到了传统 SoC 不能实现的高带宽、低延时和可扩展性强等优点<sup>[1-5]</sup>.

因为资源共享和并行性是 NoC 的优势, 所以网络中会出现一项任务的执行引起其他任务执行效率下降

的现象.为弥补这种现象带来的性能损失,有学者提出利用路由算法<sup>[6-8]</sup>来隔离多任务,保证资源分配无冲突<sup>[9]</sup>.然而,目前的大多数算法更多的是实现 NoC 容错或考虑网络状态中的一种.文献[10]提出一种基于奇偶转向模型的容错路由算法,该算法将故障路由器划分在水平方向不相邻的矩形故障区域内,但不考虑网络中最左边两列上的路由器出现故障的情况.文献[11]设计了一种基于转弯模型来避免死锁的容错路由算法,其不足之处在于算法只能容忍一个路由器故障.文献[12]是一种分布式的容错路由算法,通过迭代寻找任意两个节点之间的通路,并将通路信息记录在路由表中.这种方案适用于数据注入率低的网络,随着网络中数据包的增多,出现死锁的可能性会增大.文献[13]提出了一种拥塞感知的容错路由算法,该算法将与当前节点相连的链路故障和下游节点输入 Buffer 故障统一为链路故障,算法采用的拥塞参数为下游 Buffer 的占有率.然而,Buffer 的占有率只能反映输入端口 Buffer 的占用情况,并不能表示下游节点的数据传输状态.文献[14]提出一种新颖的选择策略,感知相邻节点的除当前节点外剩余三个邻居节点的状态.该策略与奇偶路由算法结合,选择一条尽量避免拥塞的路径传输数据包,但是该方案仅考虑距离当前节点两跳外的节点状态.文献[15]基于 mad-y 方法提出了一种拥塞感知学习模型,该模型在每个路由器内部添加一张 Q-Table,用于存储数据在路径上的延迟.文献[15]提出的方案选择局部最优的路径进行路由,可以有效地避免拥塞、均衡网络流量,但该方案没有考虑网络中出现故障的情况.

为解决以上问题,本文提出一种针对路径故障与局部拥塞的 NoC 容错路由算法,主要有以下贡献:(1)采用 12-bit 故障向量表示节点两跳以内的路径故障状态,节省了一定的硬件开销;(2)利用下游节点各输入端口请求交叉开关(Switch)无响应的个数作为拥塞参数,使当前节点更为准确地感知下游节点的拥塞程度;(3)网络无故障时,算法保证走最优路径;(4)网络有故障时,算法不仅能容错而且能够均衡网络流量.

## 2 故障与拥塞模型

本文算法的实现基于五端口的虚通道路由器<sup>[3]</sup>.该路由器主要由东(E)、南(S)、西(W)、北(N)和本地(Local)五端口、输入缓冲 Buffer、Switch、控制模块以及相应的连接线路组成.传统的无虚通道路由器在数据传输过程中很容易引起死锁(Deadlock)<sup>[16]</sup>.本文采用了虚通道策略避免了死锁的发生<sup>[17]</sup>.

### 2.1 相隔节点间路径故障模型

容错路由算法是针对 NoC 中出现的故障而提出,

为了更好地体现算法的性能,定义一个好的故障模型显得尤为重要.针对网络中出现的永久性故障,本文提出一种新颖的故障模型,即相隔节点间路径故障模型,该模型定义相隔节点为当前节点经过两跳到达的节点.如图 1 所示,Router 0 为当前节点,Router 1 为下游节点,Router 2 与 Router 0 互为相隔节点.相隔节点间路径故障状态表示为 L-P1P2 ( $P1, P2 \in \{E, W, S, N\}$ ).如 L-EE 表示,数据从 Router 0 的 E 端口输出,到达 Router 1 后从 E 端口输出至 Router 2 所经过的路径的故障状态(包括 Router 0 与 Router 1 之间的链路故障(图 1 中①),Router 1 的输入 Buffer 故障(图 1 中②),Router 1 的内部通道故障(图 1 中③),Router 1 与 Router 2 之间的链路故障(图 1 中④),Router 2 的输入 Buffer 故障(图 1 中⑤)).L-EE 为 0 时,表示 Router 0 与 Router 2 之间的路径无故障;L-EE 为 1 时,则可能由图 1 中的①、②、③、④、⑤中一处或几处出现故障所导致.

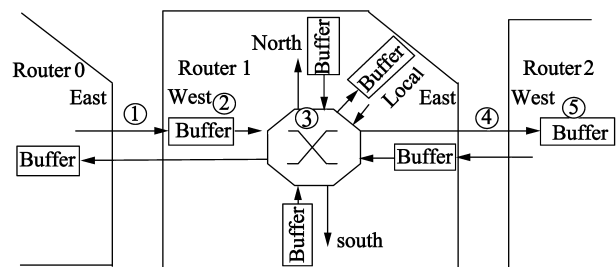


图1 相隔节点间路径故障模型

根据以上定义的路径故障模型,本文提出一种适用于该故障模型的感知区域.如图 2 所示,以 Current Node 为中心,相隔节点间路径故障模型的感知区域内共有 8 个相隔节点.其中,SE Node 与 ES Node、NW Node 和 WN Node、NE Node 和 EN Node 节点对由于选择的输出端口不同,其表示方式不一样,但在物理分布上是同一个节点.因此,Current Node 需要 12-bit 的故障向量表示到所有相隔节点的路径故障状态.

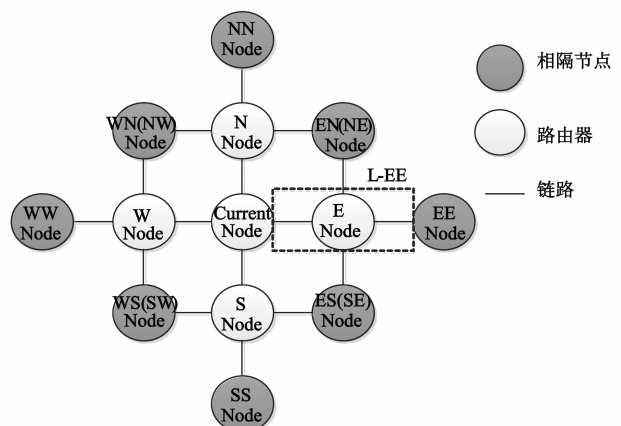


图2 节点间路径故障模型的感知区域

如表 1 所示,是 Current Node 到达感知区域内的所有相隔节点的路径故障向量.以 E 方向为例,当 L-EN、L-EE、L-ES 中至少有一个为 0 时,则表示 Current Node 与 E Node 之间路径无故障.当 L-EN、L-EE、L-ES 均为 1 时分两种情况:(1)Current Node 与 E Node 之间路径故障;(2)E Node 与 EN Node、EE Node、ES Node 三个相隔节点之间路径都出现故障.当出现以上情况,数据均不能由 E Node 继续往下传输.因此,当 L-EN、L-EE、L-ES 都为 1 时,本文规定 E Node 不可达.

表 1 Current Node 故障向量表

故障向量	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1
对应路径	L-NN	L-NW	L-NE	L-WN	L-WW	L-WS	L-SW	L-SS	L-SE	L-EN	L-ES	L-EE

## 2.2 拥塞模型

当网络中出现故障时,容错路由算法绕过故障节点或故障块到达目的节点,但在此过程中容易出现拥塞<sup>[18]</sup>.因此设计容错路由算法的过程中,需要考虑网络状态,从而在一定程度上降低拥塞的发生.基于此,本文提出一种新颖的拥塞感知模型,以动态感知下游节点的流量状况,即对局部网络流量的动态感知.该模型采用节点各端口申请输出但 Switch 未给予响应的个数作为拥塞参数,存储在拥塞寄存器 (Congestion Register, CR) 中.进行路由计算 (Routing Computation, RC) 时,CR 值会被传输至上游节点的 RC 模块,用于路由决策.为了保证 CR 值的实时性,每个时钟周期更新一次 CR 值.一旦采集到数据包申请输出端口但 Switch 未响应,CR 值加 1.当 Switch 在给定周期内未给予任何请求以应答信号,表示该交叉开关分配器故障,CR 值置为最大.模型定义的参数反映整个路由器的拥塞情况,路由数据包时选出 CR 值小的输出通道.基于这种拥塞模型下的路由算法可以均衡网络流量,降低数据包传输延迟.

## 3 容错路由算法设计

本文针对网络中出现的故障和可能出现的拥塞,提出一种针对路径故障与局部拥塞的 NoC 容错路由算法.算法设计了相隔节点间路径故障模型,从而获取路由过程中的路径故障向量;利用下游节点各端口申请输出但 Switch 未给予响应的个数作为选取路由路径的又一参数.具体思想如下,取当前节点与目的节点的 X 维和 Y 维的坐标差值,分别记作  $\Delta X$ 、 $\Delta Y$ ,目的节点与当前节点间的曼哈顿路径数记作  $K$ , $K = \frac{(|\Delta X| + |\Delta Y|)!}{(|\Delta X|)! + (|\Delta Y|)!}$ <sup>[19]</sup>.当靠近目的节点的路径无故障时,选择  $\Delta X$ 、 $\Delta Y$  中绝对值相差较大的方向作为输出,保证路径的多样性.如果检测到靠近目的节点的路径均故障,选取另外两个方向中 CR 值小的作为输出.若出现回溯,将对应的 CR 值置为最

大,再次进行路由选择.具体执行步骤:

- (1) 获取故障向量表,若寄存器值均为 0 跳至步骤 (2),否则转 (5);
- (2) 取  $\Delta X$  与  $\Delta Y$  的差.若差的绝对值大于等于 2,转 (3),否则转 (4);
- (3) 选择差值较大的且靠近目的节点的方向输出数据;
- (4) 选取 CR 值小的方向,若 CR 值相等,选择与上一跳维数不同的方向输出数据;
- (5) 当靠近目的节点的下游节点的寄存器值均为 1,转 (6),否则转 (7);
- (6) 若远离目的节点的可选路径存在,选取 CR 值小的节点输出数据;否则报错;
- (7) 选取一条 CR 值小的备选路径作为路由路径.

## 4 实验

在性能分析方面,本文采用改进后的斯坦福大学开发的仿真工具 Booksim2.0,搭建  $4 \times 4$  的拓扑,路由器设置双虚通道,每个虚通道 8flits,数据包采用 1-6flits 大小.在面积开销方面,本文采用 Synopsys 公司的 Design Compiler 进行硬件的逻辑综合,工艺库为 45nm 标准单元库.

### 4.1 网络性能

如图 3 所示,实验主要针对 XY 路由算法、文献 [14] 的 NoP 方案、文献 [15] 的 HARAQ 方案和本文方案,在均匀 (uniform) 模式下比较延迟和吞吐率.其中, NoP 方案、HARAQ 方案以及本文方案中的路由算法 (proposed 即为本文方案) 均为自适应路由算法.

从图 3 可以看出,本文方案性能优于 HARAQ 方案、NoP 方案和 XY 路由算法. XY 路由算法因为其算法的确定性,在网络流量较低的情况下和另外三种方案区别不大.随着流量的不断注入,算法的性能逐渐降低,所以在流量比较大的情况下,XY 路由算法的性能明显低于 NoP 方案、HARAQ 方案和本文方案.由于 NoP 方案仅考虑邻居以外的节点状态,忽略直接邻居节点的状态信息,即使与完全适应性路由算法结合,其性能也会受到影响.而 HARAQ 方案中每个节点均存有 Q-Table,每次路由数据包之前选择延迟最小的端口作为输出.因此, HARAQ 方案性能优于 NoP 方案.但因为 HARAQ 方案每次路由均要读取 Q-Table,使得数据包延迟增大,故其相较于 NoP 方案的优势并不是很明显.本文方案相对于 NoP 方案和 HARAQ 方案来说,在注入率达到 0.35flit/node/cycle 后优势逐渐明显.这是由于本文方案在路由数据包的过程中充分考虑网络状态,即使在注入率比较大的情况下,仍然选择拥塞度小的路径传输数据,均衡网络流量.

图 4 所示的是在 uniform 模式下,针对不同的故障比例,文献[13]的 FTCAR 方案、NoP 方案、HARAQ 方案和本文方案就网络平均延迟和吞吐率的比较.其中,由于 FTCAR 方案的容错局限性,该方案在 5% 的故障率下的性能最差. NoP 方案、HARAQ 方案在故障率达到

5% 时,性能有所衰减,当故障率达到 10% 时, NoP 方案和 HARAQ 方案性能衰减得更加明显. 当故障率达到 15% 时,本文方案性能衰减到与 10% 的故障率下的 NoP 方案、HARAQ 方案性能接近. 由此可见,本文方案的容错效果明显高于 NoP 方案、FTCAR 方案、HARAQ 方案.

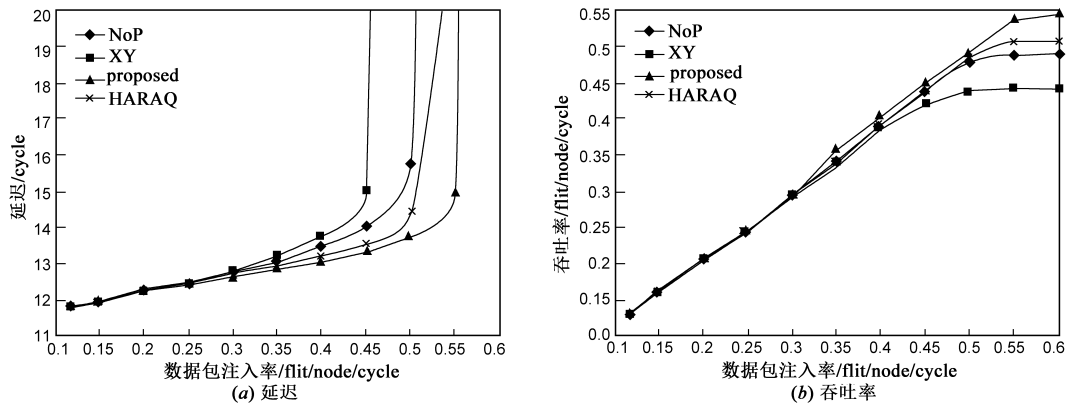


图3 不同数据包注入率下

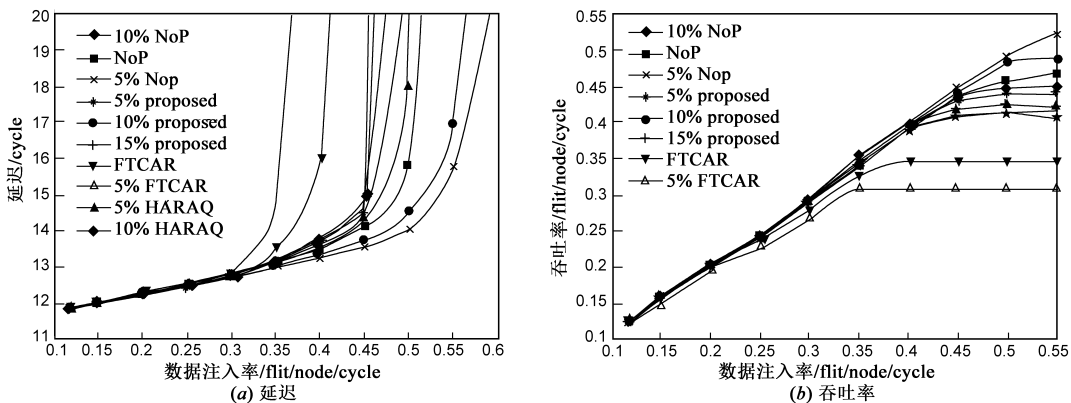


图4 uniform模式下不同故障比例下

### 4.2 面积开销和功耗分析

由于 Buffer/VC 占据路由器面积的绝大部分<sup>[20]</sup>,因此,在相同大小的 Buffer/VC 前提下,对路由器内部其他部件进行改良均不会造成面积开销过大.从图 5 可以看出,本文方案、FTCAR 方案、traditional 方案、NoP 方案以及 HARAQ 方案面积开销依次减小.主要在于前三

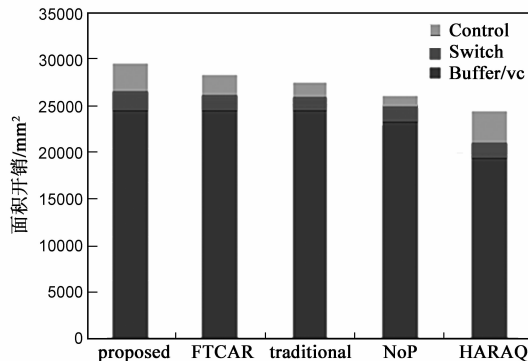


图5 面积开销

种方案均采用了双虚通道策略,增加了 Buffer 中的多路选择器、多路分配器以及 VA 控制逻辑,其次,为了实现容错和减轻拥塞,本文方案在每个路由器内部添加了故障向量和 CR,带来了可接受范围内的 Control 面积开销.

由表 2 可知,本文方案相较于 traditional 方案的功耗增加了 4.4mW. 根据上述的面积分析可知,本文方案的硬件开销和 traditional 方案相差不多,所带来的静态功耗基本接近,但是本文方案设计的路由算法会增加路由器的动态功耗,因此本文方案的功耗大于 traditional 方案. 在 HARAQ 方案中,数据每传播一跳就将该段路径上的延迟返回至上一跳节点的 Q-Table,根据 Q-Table 中原有的值与返回值计算后的延迟比较,选择其中较小的值作为该段路径下一个数据包的传输延迟,因此,该方案带来的动态功耗开销较大. 同样地, NoP 中串联的选择策略也会带来较大的动态功耗. 所以, HARAQ 方案和 NoP 方案的功耗开销均大于本文方案.

表 2 功耗

方 案	路由器功耗(mW)
traditional	25.4
FTCAR	26.6
NoP	31.3
HARAQ	33.2
proposed	29.8

## 5 总结

本文针对 NoC 路由器故障以及路由器之间的链路故障,提出了一种相隔节点间路径故障模型,利用故障向量表示当前节点两跳以内所经历的路径故障状态,再通过 CR 值统计下游节点的拥塞程度,进而设计出一种针对路径故障与局部拥塞的 NoC 容错路由算法. 算法在网络无故障的前提下走最优路径,一旦网络中出现故障,则根据故障感知模型以及 CR 值选出最优路径,不仅实现容错还可以均衡网络流量. 在设计细节方面,本文采用的路由器是虚通道路由器,可以有效地避免死锁,但又不会额外增加面积开销. 在算法实现过程中,采用最大化 CR 值的方法避免了活锁的发生. 实验结果表明,本文方案在可接受的硬件和功耗开销的前提下,大幅度提高了网络的性能,而且具有很好的容错效果.

## 参考文献

- [1] DiTomaso D, Kodi A, Louri A. QORE: A fault tolerant network-on-chip architecture with power-efficient quad-function channel (QFC) buffers [A]. IEEE 20th International Symposium on High Performance Computer Architecture (HPCA) [C]. IEEE, 2014. 320 – 331.
- [2] 王新玉, 向东, 虞志刚. TM: 一种新的片上网络拓扑结构 [J]. 计算机学报, 2014, 37(11): 2327 – 2341.  
Wang Xin-yu, Xiang Dong, Yu Zhi-gang. TM: A new topology for networks-on-chip [J]. Chinese Journal of Computers, 2014, 37(11): 2327 – 2341. (in Chinese)
- [3] 欧阳一鸣, 陈义军, 梁华国, 等. 一种故障通道隔离的低开销容错路由器设计 [J]. 电子学报, 2014, 42(11): 2142 – 2149.  
Ouyang Yi-ming, Chen Yi-jun, Liang Hua-guo. Design of a low-overhead fault channel isolated fault-tolerant router [J]. Acta Electronica Sinica, 2014, 42(11): 2142 – 2149. (in Chinese)
- [4] Liu C, Zhang L, Han Y, et al. Vertical interconnects queuing in symmetric 3D mesh network-on-chip [A]. Proceedings of the 16th Asia and South Pacific Design Automation Conference [C]. IEEE, 2011. 357 – 362.
- [5] 欧阳一鸣, 张一栋, 梁华国, 等. 三维片上网络故障及拥塞感知的容错路由器设计 [J]. 电子学报, 2013, 41(5): 912 – 917.  
Ouyang Yi-ming, Zhang yi-dong, Liang Hua-guo. A fault-tolerant design of congestion-aware router in three-dimensional network-on-chip [J]. Acta Electronica Sinica, 2013, 41(5): 912 – 917. (in Chinese)
- [6] 付斌章, 韩银和, 李华伟, 等. 面向高可靠片上网络通信的可重构路由算法 [J]. 计算机辅助设计与图形学学报, 2011, 23(3): 448 – 455.  
Fu Bin-zhang, Han Yin-he, Li Hua-wei. Building resilient NoC with a reconfigurable routing algorithm [J]. Journal of Computer-Aided Design & Computer Graphics, 2011, 23(3): 448 – 455. (in Chinese)
- [7] Feng C, Zhang M, Li J, et al. A low-overhead fault-aware deflection routing algorithm for 3D network-on-chip [A]. IEEE Computer Society Annual Symposium on VLSI (IS-VLSI) [C]. IEEE, 2011. 19 – 24.
- [8] 刘家俊, 顾华玺, 王长山. mesh 优先级容错路由 [J]. 计算机工程与应用, 2009, 45(4): 105 – 107.  
Liu Jia-jun, Gu Hua-xi, Wang Chang-shan. Priority fault-tolerant routing in mesh [J]. Computer Engineering and Applications, 2009, 45(4): 105 – 107. (in Chinese)
- [9] Ma S, Jerger N E, Wang Z. DBAR: an efficient routing algorithm to support multiple concurrent applications in networks-on-chip [A]. 38th Annual International Symposium on Computer Architecture (ISCA) [C]. IEEE, 2011. 413 – 424.
- [10] Wu J. A fault-tolerant and deadlock-free routing protocol in 2D meshes based on odd-even turn model [J]. IEEE Transactions on Computers, 2003, 52(9): 1154 – 1169.
- [11] Zhang Z, Greiner A, Taktak S. A reconfigurable routing algorithm for a fault-tolerant 2D-mesh network-on-chip [A]. 45th ACM/IEEE Design Automation Conference [C]. IEEE, 2008. 441 – 446.
- [12] Fick D, DeOrio A, Chen G, et al. A highly resilient routing algorithm for fault-tolerant NoCs [A]. Proceedings of the Conference on Design, Automation and Test in Europe [C]. European Design and Automation Association, 2009. 21 – 26.
- [13] Valinataj M, Mohammadi S, Plosila J, et al. A fault-tolerant and congestion-aware routing algorithm for networks-on-chip [A]. IEEE 13th International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS) [C]. IEEE, 2010. 139 – 144.
- [14] Ascia G, Catania V, Palesi M, et al. Implementation and analysis of a new selection strategy for adaptive routing in networks-on-chip [J]. IEEE Transactions on Computers, 2008, 57(6): 809 – 820.
- [15] Ebrahimi M, Daneshtalab M, Farahnakian F, et al. HARAQ: congestion-aware learning model for highly a-

- daptive routing algorithm in on-chip networks [ A ]. Sixth IEEE/ACM International Symposium on Networks on Chip (NoCS) [ C ]. IEEE, 2012. 19 – 26.
- [ 16 ] Ying H, Jaiswal A, Hollstein T, et al. Deadlock-free generic routing algorithms for 3-dimensional networks-on-chip with reduced vertical link density topologies [ J ]. Journal of Systems Architecture, 2013, 59(7) : 528 – 542.
- [ 17 ] Dubois F, Sheibanyrad A, Pétrot F, et al. Elevator-first: A deadlock-free distributed routing algorithm for vertically partially connected 3D-NoCs [ J ]. IEEE Transactions on Computers, 2013, 62(3) : 609 – 615.
- [ 18 ] Chen K C, Kuo C C, Hung H S, et al. Traffic-and-thermal-aware adaptive beltway routing for three dimensional network-on-chip systems [ A ]. IEEE International Symposium on Circuits and Systems (ISCAS) [ C ]. IEEE, 2013. 1660 – 1663.
- [ 19 ] Sedghi M, Koopahi E, Alaghi A, et al. An NoC test strategy based on flooding with power, test time and coverage considerations [ A ]. 21st International Conference on VLSI Design [ C ]. IEEE, 2008. 409 – 414.
- [ 20 ] DeOrio A, Fick D, Bertacco V, et al. A reliable routing architecture and algorithm for NoCs [ J ]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2012, 31(5) : 726 – 739.

#### 作者简介



欧阳一鸣 男, 1963 年生, 博士, 教授, 中国计算机学会高级会员, 研究方向: 片上网络 (NoC) 与片上系统 (SoC)、嵌入式系统的综合与测试、数字系统设计自动化。  
E-mail: oyymbox@163.com



何鑫城 女, 1990 年生, 硕士研究生, 研究方向: 片上系统以及片上网络容错方法。  
E-mail: 372521359@qq.com