

# 一种改进的 BFT 型片上网络拓扑结构<sup>①</sup>

周文科<sup>1</sup>, 尤志强<sup>1</sup>, 邝继顺<sup>2</sup>, 刘鹏<sup>1</sup>

<sup>1</sup>(湖南大学 软件学院, 长沙 410082)

<sup>2</sup>(湖南大学 计算机与通信学院, 长沙 410082)

**摘要:** 针对片上网络(NoC)资源节点之间通信的局部性, 提出改进的蝴蝶型胖树(BFT)拓扑结构 XBFT 及相应的路由算法。该结构在 BFT 结构的基础上改变边的连接关系, 减少了路由节点数和物理连接链路数, 理论分析表明, 在 64 个 IP 核的 NoC 中, XBFT 较 BFT 路由器数目减少了 14.3%, 物理链路数减少了 10.7%, XBFT 结构比 BFT 结构拥有较小的面积开销和较低的布线难度。实验采用 gpNocsim 模拟器, 对 64 节点的网络, 基于局部负载模式和均衡负载模式进行性能比较。实验结果表明, 局部负载模式下, 当包注入速率低于 0.6 时, XBFT 结构与 BFT 性能相当, 随着包注入速率继续增加, XBFT 性能优于 BFT。均衡负载模式下, 当包注入速率低于 0.4 时, XBFT 网络性能比 BFT 稍优, 包注入速率大于 0.4, XBFT 网络性能比 BFT 稍差。

**关键词:** 片上网络; 拓扑结构; 路由算法; BFT

## Improved BFT Topology for Network on Chip

ZHOU Wen-Ke<sup>1</sup>, YOU Zhi-Qiang<sup>1</sup>, KUANG Ji-Shun<sup>2</sup>, LIU Peng<sup>1</sup>

<sup>1</sup>(Software School, Hunan University, Changsha 410082, China)

<sup>2</sup>(School of Computer and Communication, Hunan University, Changsha 410082, China)

**Abstract:** This paper presents an improved BFT topology, XBFT, and a routing algorithm to point at the local communication of network-on-chip. The new structure changes some edges link, subtracts the number of routers on the BFT topology to form XBFT. The theoretical analysis result shows that when the network size is 64 the number of routers of XBFT topology are less 14.3% than the number of routers of BFT topology and the number of physical links of XBFT topology are less than 10.7% the physical links of BFT topology. The new structure has less area cost and less layout cost than the BFT structure. Simulator analysis bases on gpNocsim using local traffic pattern and uniform traffic pattern, when the network size is 64. The simulation result shows for local traffic pattern, the average latencies of XBFT topology is less than the average latencies of BFT topology, when the injection rate of package less than 0.6, the XBFT and BFT topology had the same performance; When the injection rate greater than 0.6, the latencies of BFT is more than XBFT topology, the performance of XBFT is better than BFT topology. For uniform traffic pattern, when the injection rate is less than 0.4, the XBFT and BFT topology had the same performance; When the injection rate is greater than 0.4, the latencies of XBFT is more than BFT topology.

**Key words:** network on chip; topology; routing algorithm; BFT

由于集成电路制造工艺的进步, 芯片的集成度越来越高, 未来单个芯片上的晶体管数目将达到数十亿个, 人们可在单个芯片上集成更多的 IP (Intellectual Property)。片上系统(system on chip, SoC)上越来越多

的 IP 核之间的通信使得总线负担加剧。片上网络(NoC)的提出旨在探索 SoC 通信部分的系统级设计方法, 解决片上系统应用总线结构所带来的各种瓶颈问题。NoC 的核心思想是将计算机网络技术移植到多核

① 基金项目:国家自然科学基金(60673085,60773207);教育部留学回国人员科研启动基金

收稿时间:2011-02-27;收到修改稿时间:2011-04-06

芯片设计当中，从而解决总线结构带来的通信瓶颈问题。

现阶段关于拓扑结构方面的研究<sup>[1]</sup>,单片多处理器通常采用经典的拓扑结构,包括网格(mesh),带环网格(torus),蜘蛛网(Spidergon),胖树(SPIN)以及蝶形胖树(Butterfly fat tree BFT)等。Mesh 结构简单,可扩展性能好,物理链路利用率高,容易部署在芯片上,但是对于大规模网络,其网络直径,平均距离和功耗均较大。Torus 结构在 Mesh 结构的基础上增加行和列首尾交换结点的环形链路,缩短了结点间的平均距离,大大减小了网络直径,但是,增加的环形链路过长将导致产生额外的延迟。Spidergon 结构是一个双向环形对称的结构,每个节点和其对角节点存在一条交叉连接。SPIN 结构使得系统拥有较强的本地特性,系统的路由查找压力小,但是路由器结构复杂,增加了芯片面积。P. P Pande<sup>[2,3]</sup>提出 BFT 结构,该结构同 SPIN 结构相似,但 BFT 结构大大减少交换节点和链路数量,有效地减少了芯片面积,降低了物理设计复杂度,但是布线长度会随着树的复杂度的增加而增加。虽然树型拓扑结构有着不易实现的缺点,但随着集成 IP 核数目的增多,其优点越来越明显。考虑拓扑结构带来的好处以及在片上布线的难度,BFT 结构能给 NoC 设计提供一种较佳的拓扑结构,但是由于 NoC 中 IP 核之间的通信存在较强局部性,BFT 拓扑结构不能有效的解决局部高通信任务量带来的瓶颈问题。

本文提出在未来多核的背景下,针对片上网络资源节点之间通信的局部性特点,以降低 NoC 的面积开销和物理部署难度为目的,提高 BFT 结构局部通信的性能,提出改进的 BFT 型拓扑结构, XBFT, 并且给出其相应路由算法。

### 1 片上网络拓扑结构分析

NoC 拓扑结构体现了网络中通信节点、资源节点、链路的布局 and 互连方式。NoC 拓扑结构的选择对系统性能和芯片面积具有明显的影响,选择合适的拓扑结构能有效地缩短各种应用的设计及验证时间。衡量拓扑结构的标准通常是以理论上影响路由成本 and 性能为基础的,包括普通网络中所关心的节点数量、边的数量、节点度数、网络直径、平均距离、对分宽度和通信模式下的属性,包括网络吞吐量、传输延迟和功耗,芯片面积等。

### 1.1 BFT 与 XBFT 拓扑结构比较

在 BFT 结构中,资源节点 IP 核被放置在树的叶子节点,而路由节点被放置在拓扑结构中的非叶子节点上。图 1(a)所示为 64 个 IP 核的 BFT 拓扑结构图,每个 IP 核用坐标(0,P)表示,其中 0 表示 IP 核所在的层,P 表示 IP 核在当前层中的位置编号。最底层总共放置 N 个 IP 核,编号从 0 到(N-1),如坐标(0,n)表示放置在第 0 层的第 n 个 IP 核。路由节点用坐标(L,P)表示,含有 4 个子节点端口和 2 个父节点端口,在第一层资源节点与 N/4 个路由节点相连,每层路由器的数目取决于 IP 核的总个数。例如:如果网络中有 N 个 IP 核,则总层数为  $\log_4 N$ ,其中第 j 层中含有路由器的数目为  $N/2^{j+1}$ ,BFT 结构中的路由节点数目收敛于独立于层数的一个常数。

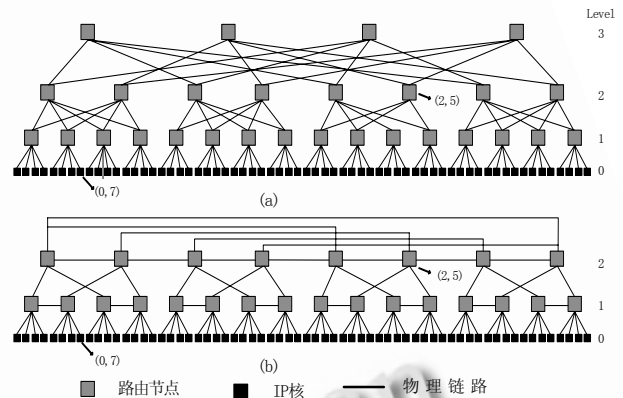


图 1 64 节点 BFT 和 XBFT 拓扑结构

同 BFT 结构相似,IP 核同样被放置在树的叶子节点,路由节点被放置在拓扑结构中的非叶子节点上。在 BFT 结构中,删除一条最底层路由交换节点与父节点的物理连接链路,增加一条与兄弟节点的物理连接链路,顶层路由节点连接成如图 2 所示的环型交叉网络,从而构成 XBFT 结构。图 1(a)所示为 64 个 IP 核的 XBFT 拓扑结构图。对于 N 个 IP 核,则总层数为  $\log_4 N - 1$ ,其中第 j 层中含有路由器的数目为  $N/2^{j+1}$ 。任意 IP 核(0,i),其父节点为(1,P<sub>1</sub>),其中  $P_1 = \lfloor i/4 \rfloor$ 。任意路由节点(L,P<sub>1</sub>),其父节点为(L+1,P<sub>2</sub>),其中 P<sub>2</sub> 由以下条件确定:

$$1. P_1 \equiv 1 \pmod{2} \begin{cases} \lfloor \frac{P_1-1}{4} \rfloor = 0, P_2 = \lfloor \frac{P_1}{2} \rfloor + 1 \\ \lfloor \frac{P_1-1}{4} \rfloor \neq 0, P_2 = \lfloor \frac{P_1}{2} \rfloor \end{cases}$$

$$2. P_1 \equiv 0 \pmod{2} \begin{cases} P_1 \equiv 0 \pmod{2}, P_2 = \lfloor \frac{P_1}{2} \rfloor \\ P_1 \equiv 1 \pmod{2}, P_2 = \lfloor \frac{P_1}{2} \rfloor - 1 \end{cases}$$

其左邻接节点为 (L, B<sub>1</sub>) 和右邻接节点为 (L, B<sub>2</sub>), B<sub>1</sub>, B<sub>2</sub> 由以下条件确定:

1. 当 L=1 时, 仅有一个邻接节点, P<sub>1</sub> 为奇数时, 无右兄弟节点, B<sub>1</sub>=P+1; P<sub>1</sub> 为偶数时, 无左兄弟节点, B<sub>2</sub>=P-1;
2. 当 L>1 时,

$$B_1 = \begin{cases} P_1+1, (0 \leq P_1 < N) \\ 0, (P_1 = N-1) \end{cases} \quad B_2 = \begin{cases} P_1-1, (0 < P_1 \leq N-1) \\ N-1, (P_1 = 0) \end{cases}$$

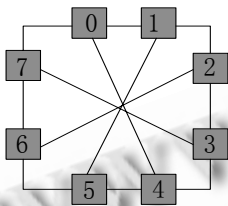


图2 XBFT 顶层互连示意图

由于 XBFT 结构中存在一条与相邻路由器节点的物理连接链路, 因此与 BFT 结构相比, 缩短了这两个相邻路由节点间相互通信的距离。表1为 BFT 与 XBFT 结构网络属性对比。在 64 个 IP 核的网络中, XBFT 结构较 BFT 结构路由器数目减少了 14.3%, 物理链路数减少了 10.7%, 对分宽度减少了 25%; 在 256 个 IP 核的网络中, XBFT 结构较 BFT 结构路由器数目减少了 20%, 物理链路数减少了 32.1%, 对分宽度减少了 56%。

表1 BFT 和 XBFT 网络属性对比

	N=16		N=64		N=256	
	BFT	XBFT	BFT	XBFT	BFT	XBFT
#Router	6	6	28	24	120	96
#Links	24	23	112	100	480	326
#Bisection Width	4	3	16	12	64	28

表2 64 节点 BFT 与 XBFT 通信节点距离对比

通信节点间距离 Hops	通信节点对数目	
	BFT	XBFT
1	192	192
2	0	256
3	1536	256
4	0	1312
5	2304	2016

## 2 XBFT的路由算法

对于基于树的拓扑结构, 路由算法可以采用最近共同祖先(least common ancestor, LCA)算法。资源节点 IP 核通信时首先要在父交换结点中查找, 如果父结点查找自己的子结点中没有目的结点, 就要向上层查询, 直到查找到源节点和目的节点的共同祖先节点时, 再往下查找目的节点。

### 2.1 XR 算法流程图

如图3所示为 XBFT 结构路由算法 XR 流程图, 其中, s 为源节点在第 0 层的位置, d 为目的节点在第 0 层的位置。

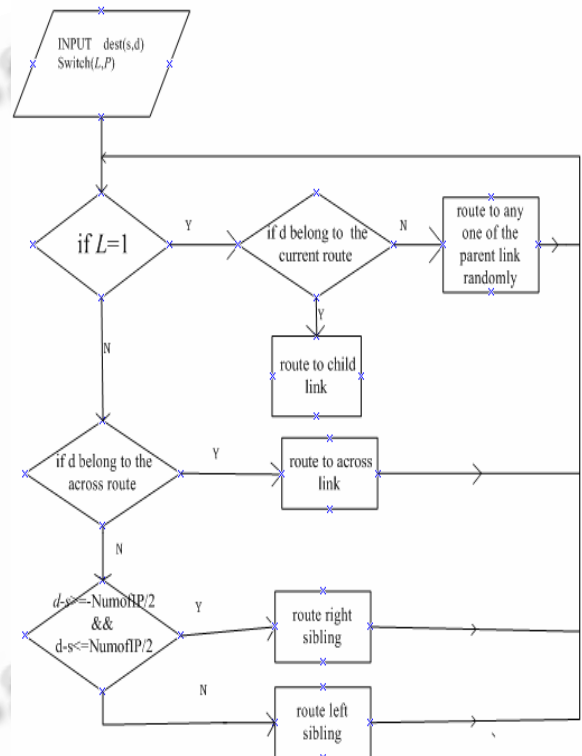


图3 XBFT 结构路由算法流程图

### 2.2 XR 算法路由描述

第一步: 输入源节点与目的节点位置(s,d)和当前路由器节点位置(L,P)。

第二步: 判断当前路由器所在层是否为第一层, 如果是则继续第三步, 否则跳到第四步。第三步: 判断目的节点是否是当前路由器的子孩子。如果是则转发到目的节点, 并退出, 否则转发到当前路由器节点的任意一个父节点, 跳转到第一步, 继续执行。

第四步: 判断目的节点是否是当前节点交叉相邻的路由节点的子孩子, 如果是则转发至交叉相连的路由节点, 转到第一步, 继续执行, 否则跳到第五步。

第五步：判断源节点与目的节点位置关系。如果  $d$  与  $s$  的差大于  $-(N-1)/2$  且小于  $(N-1)/2$  则转发到右邻接节点，并跳转到第一步，继续执行。否则，转发到左邻接节点，并跳转到第一步，继续执行。

### 3 实验结果

本文通过实验的方法对 64 个 IP 核的网络，基于局部负载模式和均衡负载模式下进行性能比较。实验工具采用了 gpNocsim<sup>[4]</sup>，主要参数如下：VC=4，Buffer=8flit，MsgLength=200 (byte)。均衡负载模式下生成数据包的源节点和目的节点都是由随机数生成。局部负载模式下，每生成一个数据包时，其目的节点以较大概率  $h$  ( $1>h>0$ ) 与源节点在同一局部范围内。

#### 3.1 局部负载模式

为了模拟两种结构在局部高通信负载环境下的网络性能，根据表 2，设置 hop 距离为 1, 2, 3, 4, 5 的  $h$  概率分别为 0.50, 0.35, 0.10, 0.04, 0.01。

表 3 为 XBFT 与 BFT 在局部负载模式下数据。当包注入速率小于 0.6 时，XBFT 与 BFT 的网络吞吐量基本一致，两种结构数据包延时基本一致。随着包注入速率的继续增加，大量的通信任务聚集在局部通信中，数据包等待处理的时间增加，致使 BFT 数据包延时的增加量明显大于 XBFT，而 XBFT 能有效的解决局部通信，故 XBFT 结构数据包延时明显小于 BFT 结构数据包延时；当包注入速率大于 0.6 时，大量的通信任务集中在局部通信中，导致 BFT 吞吐量降低，而 XBFT 能有效的适应局部频繁通信，故 XBFT 吞吐量比 BFT 稍大。

表 3 局部负载 XBFT VS BFT

注入率	吞吐量		平均延时	
	BFT	XBFT	BFT	XBFT
0.1	0.045894	0.049389	58.58905	57.52160
0.2	0.076619	0.078124	63.52598	63.7211
0.3	0.111034	0.111214	69.24117	68.8770
0.4	0.153967	0.155128	83.77574	83.42181
0.5	0.191427	0.190689	96.14745	101.4655
0.6	0.225913	0.214776	144.6760	141.7181
0.7	0.214470	0.225376	352.6815	286.9143
0.8	0.183596	0.195344	609.2719	399.3596

由以上分析可知，网络负载较低时，XBFT 与 BFT 网络吞吐量和数据包延时基本一致，网络性能相当；网络负载较高时，XBFT 比 BFT 拥有更高的网络吞吐量，较低的数据包丢失量和数据包延时，更加适应于局部通信任务量大的网络。

#### 3.2 均衡负载模式

表 4 为 XBFT 与 BFT 在均衡负载模式下数据。当包注入速率小于等于 0.4 时，XBFT 与 BFT 的网络吞吐量一致，随着包注入速率的继续增加，BFT 网络吞吐量的增加比 XBFT 的快，XBFT 网络吞吐量低于 BFT。当包注入速率小于等于 0.4 时，两种网络的包延时基本一致，随着注入速率的增加 XBFT 数据包延时大于 BFT。

表 4 均衡负载 XBFT VS BFT

注入率	吞吐量		平均延时	
	BFT	XBFT	BFT	XBFT
0.1	0.039128	0.036717	67.63914	65.82149
0.2	0.076762	0.076910	75.44086	79.5891
0.3	0.112865	0.112544	89.27287	99.73812
0.4	0.152744	0.150760	107.4476	199.8547
0.5	0.203632	0.182099	184.9830	478.6632

由以上分析可知，网络负载较低时，两种结构性性能基本一致；在网络负载比较高时，XBFT 延时比 BFT 大，因为 XBFT 的路径多样性比 BFT 少，网络负载高时，数据包在传输过程中的等待延时将有所增加。

### 4 结语

XBFT 是一种类似于 BFT 的结构，与 BFT 相比，XBFT 减少了路由器和物理连接链路数目，硬件开销明显低于 BFT。局部负载模式下，XBFT 拥有比 BFT 低的网络数据包延时，高的网络吞吐量，更加适应于任务量高的局部通信网络；均衡负载模式下，较低负载时 XBFT 的性能与 BFT 相当，较高负载时 XBFT 差于 BFT。我们目前只是在模拟器中实现并分析这个结构的性能，还需进一步考虑 XBFT 的容错性。由于 XBFT 顶层存在环形交叉网络，还需进一步调整路由算法，增加路由算法自适应性，以免不必要的拥塞降低 XBFT 的性能。

#### 参考文献

- 1 朱晓静,胡伟武,马可,章隆兵.Xmesh:一个 mesh-like 片上网络拓扑结构.软件学报,2007,18(9):2194-2204.
- 2 Pande PP, et al. Design of a switch for network on chip applications. Proc. of ISCAS, Bangkok, 2003.217-220.
- 3 Pande PP, et al. High-throughput switch-based interconnect for future SoCs. Proc. of 3rd IEEE International Workshop on System-on-chip for Real-Time Applications. Canada, 2003. 304-310.
- 4 Hossain H, Ahmed M, Al-Nayeem A, Islam TZ, Akbar M. Gpnocsim—a general purpose simulator for network on chip. International Conference on Information and Communication Technology, Dhaka, 2007.254-257.