

# PC 集群系统中 MPI 并行 FDTD 性能研究

张 玉, 李 斌, 梁昌洪

(西安电子科技大学天线与微波技术国家重点实验室, 陕西西安 710071)

**摘 要:** 本文在 MPI 编程环境下研究了不同虚拟拓扑结构方式的 3 维并行时域有限差分算法(FDTD). 将程序运行于本文组建的高性能 PC 集群系统中, 计算分析了一种新型光子带隙微带传输线结构的 S 参数特性, 并以此为研究实例, 深入地比较了不同 MPI 虚拟拓扑方式对 PC 集群系统中并行 FDTD 性能的影响, 在此研究的基础上给出了 PC 集群中 MPI 编程环境下并行 FDTD 的最佳虚拟拓扑方式.

**关键词:** 集群系统; 并行计算; 时域有限差分法; 光子带隙

**中图分类号:** TN011 **文献标识码:** A **文章编号:** 0372-2112 (2005) 09 1694-04

## Study on Performance of MPI Based Parallel FDTD on PC Clusters

ZHANG Yu, LI Bin, LIANG Chang hong

(National Key Laboratory of Antennas and Microwave Technology, Xidian Univ, Xi'an, Shaanxi 710071, China)

**Abstract:** MPI based Parallel FDTD code with different virtual topology forms is developed. The code is run on a PC cluster with high performance to analyze a new type of PBG microstrip filter structure. Taking it as an example, the influence of topology schemes on parallel performance of Parallel FDTD is studied. Finally the best MPI virtual topology for Parallel FDTD is presented.

**Key words:** PC clusters; parallel computation; FDTD; PBG

### 1 引言

FDTD(时域有限差分算法)<sup>[1]</sup>在电磁场辐射和散射、微波和毫米波电路以及电磁兼容等十分广泛的领域已经获得了非常成功的应用. 另一方面, 较长的计算时间和较大的存储空间是 FDTD 在 PC 系统上求解谐振结构、电大尺寸物体电磁场问题的瓶颈. 解决这一问题的有效方法之一是采用并行 FDTD 方法. 然而, 国外的研究者的并行 FDTD 工作<sup>[2,3]</sup>通常是在大型计算机和 workstation. 这些工作平台价格极其昂贵. 最近, 以 PC 机网络集群系统为工作平台的大型计算系统蓬勃发展, 由台式计算机构成这样的系统无需任何额外的投资, 也不会受到政策的限制, 正因为如此, 为并行计算提供了可行的、强有力的计算机硬件基础.

在网络并行 FDTD 计算方面, 尽管已经有部分文献报道<sup>[4,5]</sup>, 但仍有需进一步研究之处. 首先, 并行研究多在 PVM 并行环境下开展, 而在高性能计算系统中, 最广泛使用的编程环境是 MPI<sup>[8]</sup>, 它已经成为国际上的一种并行程序的标准, 被越来越多的硬件生产厂商、计算机用户所接受, 与 PVM 相比, MPI 本身有更多的优点<sup>[8]</sup>, 也因此 MPI 环境中并行算法的实现是国际上并行电磁计算研究的热点和主流, 所以本文的研究工作将基于 MPI 并行环境进行. 其次, 并行 FDTD 程序<sup>[4-7]</sup>未能真正运行于高性能 PC 集群中以处理大规模电磁问题, 因而无法具体考察出网络虚拟拓扑方式对并行 FDTD 性能的

影响这一关系到计算效率的重要课题. 本文将利用新组建的一个高性能 PC 集群系统, 对这个问题进行深入研究, 旨在给出最佳虚拟拓扑以获得良好的计算效率.

### 2 虚拟拓扑与场值通信

FDTD 的并行运算, 概括起来讲, 就是将整个 FDTD 计算区域划分为若干个子区域, 每个进程计算其中的一个或者多个子区域, 各个进程之间通过传递交界面上的电磁场量以确保 FDTD 的场值求解能够进行下去. 概括而言, 并行 FDTD 中, 将总的区域划分为子域有三种方法, 由此产生的数据通信方式也有三类, 下面分别给出.

#### (a) 1 维区域划分与 1 维方向通信并行 FDTD

为了清晰起见, 我们首先以 1 维并行 FDTD 为例<sup>[3]</sup>, 使用 6 个网格和 2 个节点来讲述这个问题. 如图 1(a) 所示, 当节点 2 计算  $E_{x4}$  时, 需要知道  $H_{z3.5}$  的值, 它存储于节点 1. 类似的, 当节点 1 计算  $H_{z3.5}$  时, 它需要知道存储于节点 2 的  $E_{x4}$  值. 这表明每一时间步中, 必须传递 2 个信息, 每一个沿着不同的方向.

通信的原理可以按照下面的流程来解释: 在编号为  $n$  的节点, 对每一个时间步来说, 磁场在  $t + Dt/2$  更新. 新的磁场值分别发送给节点  $n+1$  并从节点  $n-1$  接收, 所有电场值在  $t + Dt$  更新, 新的电场值发送给节点  $n-1$  并从节点  $n+1$  接收, 循环从头开始. 为了简单起见, 此处略去了激励源、外边界

条件等情况. 这个过程表明每一次迭代必须有两次通信.

对于三维 FDTD 问题, 如果计算区域沿着一个方向划分, 通信原理也是与一维 FDTD 相同. 只是传递的数据分布在一个平面, 而不再是一条线上了. 如果节点计算  $100 \times 100 \times 100$  的区域块, 那么每个传递信息量将包含  $20,000 (2 \times 100 \times 100)$  个浮点数. 计算流程如图 1(b) 所示.

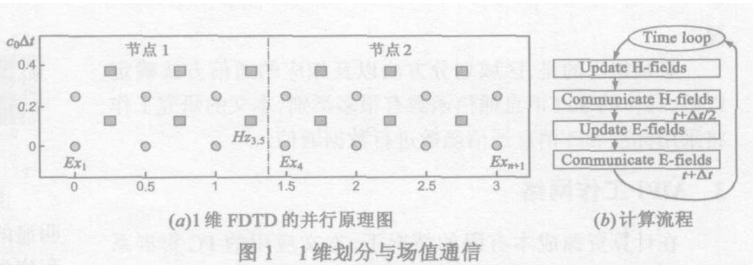


图 1 1 维划分与场值通信

结构中进程可以用笛卡尔坐标来标识, MPI 中称这类拓扑结构为笛卡尔 (Cartesian) 拓扑结构. 图 2(c) 给出了 9 个进程, 建立笛卡尔坐标的网络物理分布与  $3 \times 3$  分解的二维虚拟拓扑分布.

以中间进程 4 的场值通信为例, 如图 2(b) 所示, 在一个具有笛卡尔拓扑结构的通信器中经常在一个指定方向 (维) 上在处理器间进行数据平移 (shift), 如用 MPI\_SENDRECV 将一块数据发送给该维上后方一个进程, 同时接收从该维上前方一个进程发送来的数据. 依照图 2(c), 先利用 MPI\_CARTSHIFT 获得进程间的关系, 然后进行上、下、左、右的发送与接收就可以非常方便地实现场值通信. 可见, 尽管二维方向通信理论上是一维方向通信的外延, 但采用的是完全不同的实现方式.

(c) 3 维区域划分与 3 维方向通信并行 FDTD

这种划分方法在实现途径上与 2 维划分方法并没有本质的区别. 如图 3(a) 计算区域沿着三个方向分解为子区域, MPI 的虚拟拓扑沿着三个方向建立. 3 维方向创建的虚拟拓扑, 会出现不同方向上的交界面, 但可以分解为若干 2 维的情形来理解场值的通信. 图 3(b) 给出了一个 y 方向  $xz$  面上的通信情

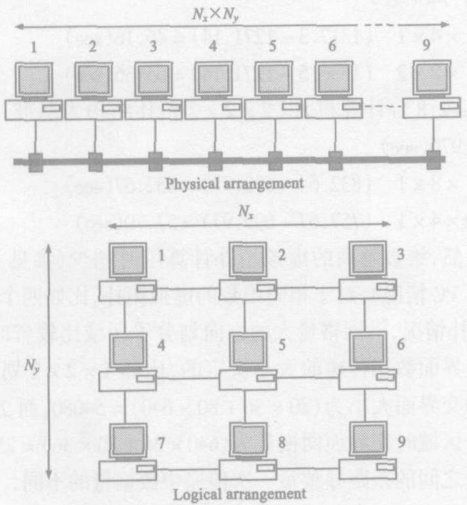
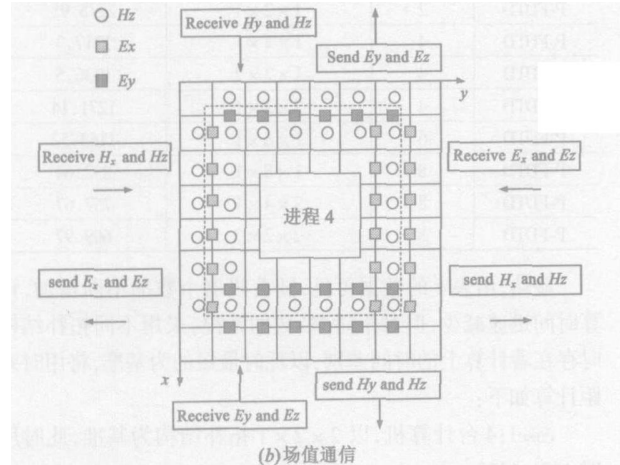
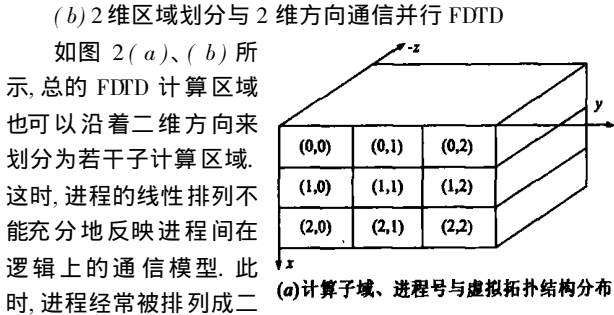


图 2 二维划分与场值通信

维或三维网格形式的拓扑模型, 而且, 通常用一个图来描述逻辑进程排列, 这种逻辑进程排列为虚拟拓扑. 应用问题中较为常见, 也是较为简单的一类进程拓扑结构具有网格形式. 这类

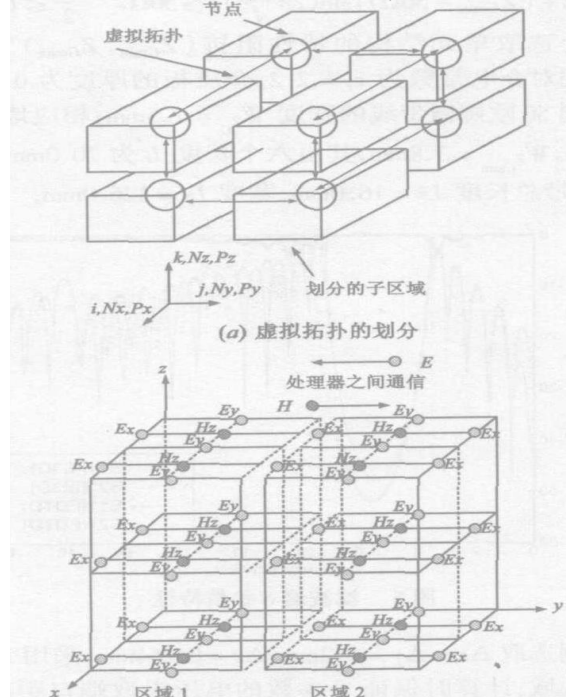


图 3 3 维划分与场值通信

况。

必须指出的是,区域划分方法以及相应的通信方式确定以后,具体的 MPI 消息通信函数有很多类别,本文的研究工作将采用标准 MPI 消息通信函数进行数据通信。

### 3 MPI 工作网络

在计算资源成本有限的情况下,本文应用的 PC 集群系统,其模型如图 4 所示,系统主要参数为:

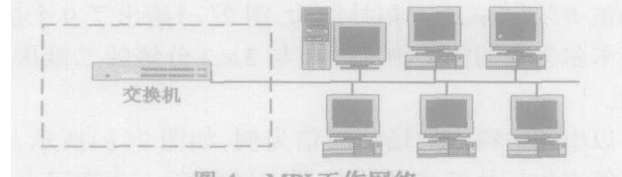


图 4 MPI 工作网络

10 节点; 每个节点 2.8GHz CPU, 1.0G Bytes 内存; 100/1000M 高速网卡; 1000M Bytes 交换机。

### 4 并行 FDTD 性能测试与应用

下面以新型微带 PBG 滤波器结构为研究对象,来考察 PC 集群中并行 FDTD 的性能。该微带 PBG 滤波器的单元结构的特性阻抗是随着微带线宽度变化而正弦变化的,它的变量仅仅是微带线的宽度。滤波器总体结构是由单元周期排列组成,而每一个单元结构的特性阻抗  $Z_c$  是沿单元的相对位置 ( $P/L$ ) 的正弦函数,其中  $L$  是单元长度,  $P$  为沿单元的位置。 $Z_c$  设置如下:

$$Z_c = 50\Omega - (50\Omega - Z_{Cmin}) \sin\left(2\pi \frac{P}{L}\right) \leq 50\Omega, \quad 0 \leq P \leq \frac{L}{2}$$

$$Z_c = 50\Omega + (Z_{Cmax} - 50\Omega) \left| \sin\left(2\pi \frac{P}{L}\right) \right| \geq 50\Omega, \quad \frac{L}{2} \leq P \leq L$$

此处选取单元结构的特性阻抗 ( $Z_{Cmin}$ ,  $Z_{Cmax}$ ) 为 (25 $\Omega$ , 100 $\Omega$ ), 相对介电常数为  $\epsilon_r = 2.2$ , 介质板的厚度为 0.508mm, 从而得到 50 欧姆微带线的宽度  $W_{Z_c} = 1.5\text{mm}$ , 相应地,  $W_{Z_{Cmin}} = 0.4\text{mm}$ ,  $W_{Z_{Cmax}} = 3.8\text{mm}$ , 其由六个长度  $L$  为 20.0mm 的单元组成, 微带总长度  $L_x = 16.0\text{mm}$ , 宽度  $L_y = 126.0\text{mm}$ 。

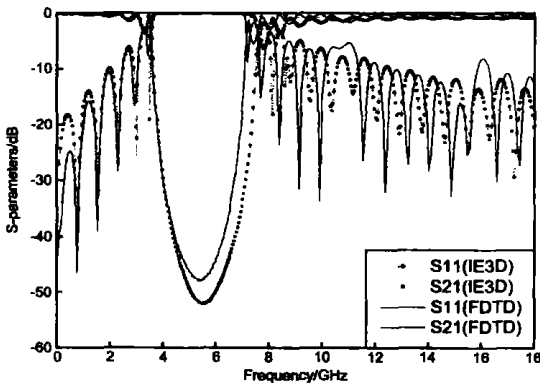


图 5 滤波器 S 参数特性

本例选取  $\Delta x = \Delta y = 0.2\text{mm}$ ,  $\Delta z = 0.254\text{mm}$ , 采用 UPML 截断计算区域, 计算时保证 S 参数的电压提取端口距离 UPML 内边界 10 个网格以上, 则最终选取的总计算区域为:  $80 \times 640 \times 36$ , FDTD 计算模型如图 6 所示。

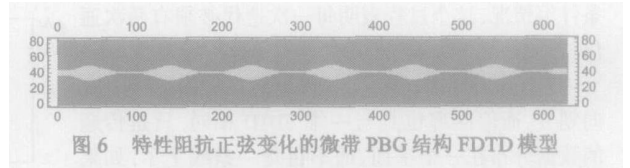


图 6 特性阻抗正弦变化的微带 PBG 结构 FDTD 模型

由图 5 可知,这种 PBG 结构在 3.5GHz~ 7.2GHz 之间具有明显的带隙,本文结果与 IE3D 软件仿真结果吻合良好,表明程序的正确性。仿真这个具有谐振特性的 PBG 结构,为了比较准确获得参数,大约要 4 万次迭代才能获得光滑的 S21 曲线,此时利用 8 台计算机大约需要 7.7 小时,单机计算则大概需要 62 小时,可见并行计算技术可以非常有效节约计算时间,扩大计算规模。

为了进一步考察并行 FDTD 方法的性能情况,我们以 1000 次迭代为例,来分析不同 PC 机数目和网络不同虚拟拓扑结构情形下的计算时间:

表 1 计算时间比较表

名称	PC 数	虚拟拓扑 ( $x \times y \times z$ )	计算耗时 (秒)
P FDTD	2	$1 \times 2 \times 1$	2275.95
P FDTD	4	$1 \times 4 \times 1$	1317.3
P FDTD	4	$1 \times 2 \times 2$	1336.5
P FDTD	4	$2 \times 2 \times 1$	1271.14
P FDTD	6	$1 \times 6 \times 1$	1163.52
P FDTD	8	$1 \times 8 \times 1$	832.64
P FDTD	8	$2 \times 4 \times 1$	757.67
P FDTD	8	$2 \times 2 \times 2$	699.97

显然,由表 1 的数据可见,随着机器个数的不断增加,计算时间迅速减少,但对相同 PC 数的情况,采用不同拓扑结构时存在着计算上的时间差别,以耗时最短的为基准,将用时差距计算如下:

case1: 4 台计算机,以  $2 \times 2 \times 1$  拓扑结构为基准,此时用时 1271.140(sec)

对比:  $1 \times 4 \times 1$  ( $1317.3 - 1271.14$ ) = 46.16(sec)

对比:  $1 \times 2 \times 2$  ( $1336.5 - 1271.14$ ) = 65.36(sec)

case2: 8 台计算机,以  $2 \times 2 \times 2$  拓扑结构为基准,此时用时 699.970(sec)

对比:  $1 \times 8 \times 1$  ( $832.64 - 699.97$ ) = 132.67(sec)

对比:  $2 \times 4 \times 1$  ( $757.67 - 699.97$ ) = 57.70(sec)

显然,维数越高的虚拟拓扑计算用时越少(参见 4 个 PC 与 8 个 PC 情况)。对于相同维数的虚拟拓扑,比如四个进程的 2 维拓扑情况,沿网格量大的方向划分子区域比较省时,这是由于交界面数据传递的大小决定的,比如  $1 \times 2 \times 2$  划分子区域产生的交界面大小为  $(80 \times 36 + 80 \times 640) = 54080$ ,而  $2 \times 2 \times 1$  划分子区域时交界面网格量为  $(640 \times 36 + 80 \times 36) = 25920$ ,显然两者之间的差距导致每一次传递中数据量的不同。

由此,对于相同 PC 数,可以得出这样的结论:

(1) 对于相同维数的虚拟拓扑,拓扑应沿着大网格量方向产生(参见 4 个 PC 情形)。

(2) 并行 FDTD 计算时,最佳的虚拟拓扑结构应该尽可能的沿着三维方向建立;如果不能沿着三维方向,则尽可能沿着

二维方向建立拓扑, 这样比一维虚拟拓扑要节省总的计算时间(参见 4 个 PC 以及 8 个 PC 情形)。

## 5 结论

本文利用并行 FDTD 计算了一种新型 PBG 结构的 S 参数特性, 以此为实例考察了 PC 集群系统中并行 FDTD 的性能, 研究表明选择恰当的 MPI 虚拟拓扑方式, 可以有效节约计算时间, 这一结论对于 PC 网络中 MPI 编程环境下并行 FDTD 的研究工作将具有现实的参考意义。

### 参考文献:

- [ 1 ] A Taflové. Computational Electrodynamics: the Finite Difference Time Domain Method, 2nd ed[M]. MA: ARTECH HOUSE, INC, 2000.
- [ 2 ] Ulf Andersson. Time Domain Methods for Maxwell Equation. Doctoral Dissertation[D]. Sweden: Royal Institute of Technology, 2001.
- [ 3 ] Gui aut C, K Mahdjoubi. A parallel FDTD algorithm using the MPI library[J]. IEEE Antennas and Propagation Magazine, 2001, 43(2): 94-103.
- [ 4 ] 闫玉波, 葛宁, 郑美艳, 葛德彪, 田春明. 网络并行 FDTD 方法分析电大目标电磁散射[J]. 电子学报, 2003, 31(6): 821-824. YAN Yu bo, Ge Ning, ZHENG Mei yan, GE De biao, TIAN Chun ming. Analysis of large Scale EM scattering problems using the parallel FDTD on PC cluster system[J]. Acta Electronica Sinica, 2003, 31(6): 821-824. (Chinese Source)
- [ 5 ] 薛正辉, 杨仕明, 高本庆, 张泽杰. FDTD 算法的网络并行运算实现[J]. 电子学报, 2003, 31(12): 1839-1843. XUE Zheng hui, YANG Shi ming, GAO Ben qing, ZHANG Ze jie. A parallel implementation strategy for the FDTD algorithm[J]. Acta Electronica Sinica, 2003, 31(12): 1839-1843. (Chinese Source)
- [ 6 ] 张玉, 宋健, 梁昌洪. 并行 FDTD 及其在基站天线分析中的应用[A]. 全国微波毫米波会议论文集[C]. 上海: 中国电子学会微波分会, 2003. 685-688. ZHANG Yu, SONG Jian, LIANG Chang hong. Parallel FDTD and its application for base station antenna[A]. Microwave and Millimeter Wave Symposium of China[C]. Shang Hai: Chinese Institute of Electronics, 2003. 685-688. (Chinese Source)

- [ 7 ] 张玉, 宋健, 梁昌洪. 并行共形 FDTD 算法及其在 PBG 结构仿真中的应用[J]. 电子学报, 2003, 31(12A): 2142-2144. ZHANG Yu, SONG Jian, LIANG Chang hong. Study on the parallel modified locally conformal FDTD algorithm on cluster of PCs for PBG structures[J]. Acta Electronica Sinica, 2003, 31(12A): 2142-2144. (Chinese Source)
- [ 8 ] 都志辉. 高性能计算并行编程技术[M]. 北京: 清华大学出版社, 2001.
- [ 9 ] Dusan Nestic, Aleksandar Nestic. BandStop Microstrip PBG Filter with Sinusoidal Variation of the Characteristic Impedance and without Etching in the Ground Plane[J]. Microwave Opt Technol Lett, 2001, 29(6): 418-420.

### 作者简介:



张 玉 男, 1978 年 4 月出生于安徽淮北, 现工作于西安电子科技大学工程学院, 主要从事电磁场并行数值计算、电磁兼容等方面研究工作, 发表论文 40 余篇。

E-mail: yuzhang@mail.xidian.edu.cn.



李 斌 男, 1978 年 4 月出生于陕西咸阳, 现为西安电子科技大学电磁场与微波技术专业博士生, 主要从事 PBG 电磁特性仿真等方面研究工作. E-mail: bli@mail.xidian.edu.cn.

梁昌洪 男, 1943 年 12 月出生于上海, 现为西安电子科技大学教授, 博士生导师, 中国电子学会会士、IEEE Senior member, 2003 年荣获首届全国教学名师奖, 研究方向包括计算场论、微波网络理论、电磁散射与逆散射、电磁兼容等方面。