

# 多核微处理器核间高速互连技术

郭广浩<sup>1</sup>, 沈绪榜<sup>2</sup>

(1. 西北工业大学 计算机学院, 陕西 西安 710072;  
2. 西安微电子技术研究所, 陕西 西安 710054)

**摘 要:**随着 VLSI 技术和半导体制造工艺的不断发展,多核处理器已经取代了单核处理器。当技术和工艺的发展使片上多处理器中核的数目增加时,各个处理器核之间的互连及其通信就成为制约处理器性能提高的瓶颈。为了能够充分发挥多核处理器的高性能,文中根据当今主流多核处理器的互连方法,通过分析各种互连方法的优势与不足,提出了针对不同的核的数目和结构采用不同的互连方法,指出将新材料、新技术、新器件与已有的成熟的多核互连方式相结合是提高多核互连效率的有效方法,并阐述了未来多核互连的研究方向和发展趋势。

**关键词:**多核微处理器;片上网络;互连通信

**中图分类号:**TP31

**文献标识码:**A

**文章编号:**1673-629X(2012)06-0030-05

## Technology of High-speed Interconnection for Multi-core Microprocessor

GUO Guang-hao<sup>1</sup>, SHEN Xu-bang<sup>2</sup>

(1. School of Computer, Northwestern Polytechnical Univ., Xi'an 710072, China;  
2. Xi'an Microelectronics Technology Institute, Xi'an 710054, China)

**Abstract:** With the development of the VLSI technology and semiconductor manufacturing process, multi-core processors have replaced single-core processor. When the development of the technology and technics make the number of core increases, interconnection and communication between each processor core will become choke point to improve the performance of processor. In order to give full play to high-performance of multi-core processors, based on the interconnection of today's mainstream method of multi-core processors, by analyzing the advantages and disadvantages of interconnection method, for different number and structure are introduced the different interconnection methods, that the new materials, new technologies, new devices combined with existing mature mode of interconnection for multi-core is the effective method to improve the efficiency of interconnection for multi-core, and the direction of research, the trends of interconnection for multi-core in future is presented.

**Key words:** multi-core microprocessor; network on chip; interconnection

## 0 引言

随着 VLSI 技术和半导体制造工艺的不断发展,单个晶体管尺寸不断减小,单个芯片上能够集成的晶体管数目也越来越多,单个处理器芯片的集成度和处理性能也越来越高,已经能够在片内集成多个处理器核以及互连接口。近来高端处理器主要利用超标量技术、流水线技术、提升主频已经很难进一步提高其性能,因为单核的功耗极限是其性能提升的瓶颈问题。采用以多核处理器为代表的先进体系结构已经逐渐成

为提高处理器性能的主要途径。多核处理器是用多个主频较低的单核通过互连,使其达到性能提高的同时还能减小功耗的目的。

多核处理器即片上多处理器(chip multi-processor, CMP),其结构主要是简化超标量结构设计,将多个相对简单的超标量处理器核集成到一个芯片上,从而避免线延迟的影响,并充分开发 TLP,提高吞吐量。当技术和工艺的发展使片上多处理器中核的数目增加时,各个处理器核之间的互连及其通信就成为制约处理器性能提高的瓶颈。为了提高多核处理器的性能,就必须对不同数目及结构的多处理器核采用不同的互连方法和技术,以达到其性能最优。文中主要介绍当今主流多核处理器的互连方法,针对不同的核数目和结构应用不同的互连方法,并简要论述了未来多核处理器互连的一些新技术、新方法及其发展方向和趋势。

收稿日期:2011-10-22;修回日期:2012-01-30

基金项目:国防微电子预研项目(41308010203)

作者简介:郭广浩(1986-),男,硕士研究生,研究方向为微电子技术与计算机系统结构;沈绪榜,研究员,中国科学院院士,研究方向为高性能计算机、微电子技术及芯片设计。

1 多核微处理器核间高速互连通信问题的提出

1996 年美国斯坦福大学首次提出片上多处理器 (CMP) 思想和首个多核结构原型,2001 年 IBM 推出第一个商用多核处理器 Power4,2005 年 Intel 和 AMD 多核处理器的大规模应用,使多核成为市场主流。在这个过程中,多核处理器从个人计算机到商用服务器,从嵌入式设备到多媒体都得到了广泛的应用。如今影响多核处理器性能的关键因素不再是其单核的运算速度,而是它们的互连通信。从单核到多核的发展必须要解决的问题就是多核核间互连的问题,不管是 IBM、Intel 或是 AMD,其双核的性能都达不到其单核性能的两倍。当单片处理器核的数目不断增加,其互连开销也会增大。不恰当的互连方法会严重影响多核处理器的性能,甚至可能比单核处理器的性能还要差。

2 多核微处理器核间的主要互连通信方法

目前多核架构已成为处理器的主流设计,而核间互连结构影响着功耗、延迟、性能等,其核间通信的效率特别是不同核间可能需要进行数据的共享和同步,是阻碍多核处理器性能提高的关键因素之一<sup>[1]</sup>,目前多核互连通信主要有总线共享结构、交叉开关互连和片上网络 (Network On Chip, NOC) 等方式。

2.1 共享总线结构 (share bus fabric, SBF)

CMP 系统中的多个处理器核、cache、输入/输出以及存储器之间的通信通过高速 SBF 连接在一起。同时 SBF 还要支持一些一致性协议。总线共享结构比较简单,但要考虑高速缓存数据一致性的问题。共享总线的局限性在于其高带宽、低延迟、大尺寸之间的相互矛盾,其对于可以连接在上面的处理器的数目有限制,到目前为止,商业计算机中连接到总线上的处理器数目最多是 36 个<sup>[2]</sup>,一般为 2~4 个核。

Intel 的 Core, Hydra, IBM 的 Power4/5 等都是总线共享结构处理器,Hydra 处理器是 1996 年美国斯坦福大学研制的一个集成了 4 个核的处理器,这在当时是一种新型的处理器结构<sup>[3]</sup>。Hydra 在一个芯片上集成了 4 个核,核间通过总线结构共享片上二级缓存、存储器端口和 I/O 访问端口,4 个核采用了通用的百万指令级 (MIPS) 处理器。Hydra 是第一个采用了高速总线核间通信方式的多核处理器设计原型。因为总线一次

只能有一个节点使用,因此对总线的使用竞争将会产生很大的延迟,这样将会对总线的速度产生很大影响。

改进的总线结构:

为了满足下一代宽带多媒体与图形应用对多核处理器高效率、低功耗的需求,IBM 与 Sony、Toshiba 合作开发了一种能够分段使用总线的多核处理器结构,它就是 Cell 处理器。如图 1 所示,Cell 处理器由一个主核和八个副核组成,通过能够分段使用的线性阵列总线连接一个存储器控制器和一个 IO 控制器<sup>[4]</sup>。

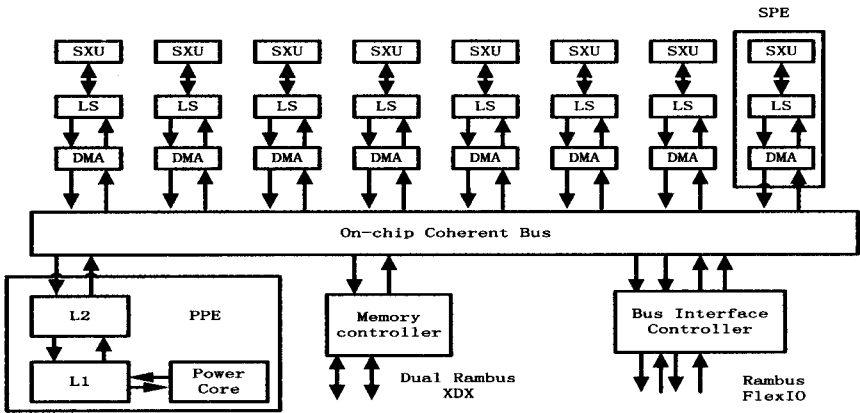


图 1 采用线性阵列总线的 Cell 处理器结构

由于 Cell 处理器使用了线性阵列的总线互连方法,它能够使处理单元和存储器、外部设备在不同区段同时分段使用总线,降低了各模块之间竞争使用总线而产生的延时,大大提高了总线的使用效率。但当连接的处理器核较多时,竞争依然会很严重。

2.2 交叉开关 (Crossbar) 互连

由于使用总线的竞争延迟较大,当多核处理器中核的数目增多时传统的共享总线将会阻碍其并行性。半导体制造工艺的迅速发展使得芯片的集成度越来越高,从而为交叉开关 (Crossbar Switch) 技术的实现创造了可能。交叉开关互连允许节点通过网络的某个路径和其他任何节点之间通信 (点对点)。使用交叉开关互连技术的多核处理器有 AMD 公司的 Athlon 双核处理器、SUN 公司的 UltraSPARC T1 处理器等。

交叉开关的结构:

交叉开关 (Crossbar) 互联系统由交叉开关 (Crossbar) 以及 Crossbar 接口逻辑构成。每个 Crossbar 由一些地址线、数据线构成,接口逻辑主要由一些加载队列构成。交叉开关能够使连接在开关结构上的节点同时进行数据交换。保证了多核处理器在空间上处理数据的并行性,极大地提高了总线的带宽。

在多核处理器中,交叉开关总线具有建立高速缓存、I/O 等与处理器核通信路径的作用。交叉开关控制的通信路径是双向的,它能为两个节点之间提供独享的数据传输通路,保障了数据传输的可靠性、高效

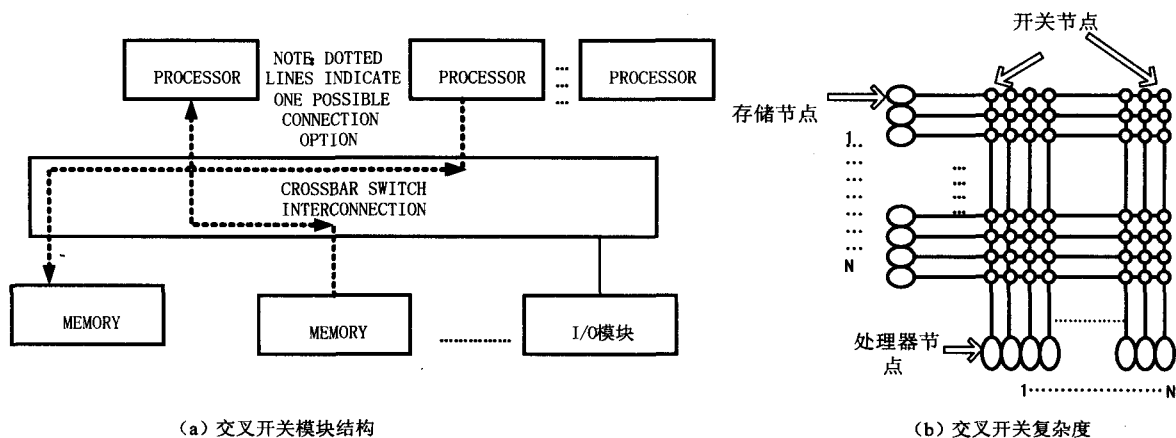


图 2 交叉开关模块结构和复杂度

性、并行性。在有些设计中,交叉开关也被做成两个模块,每个模块的数据传输是单向的,其中一个主要负责从多处理器核到缓存和 I/O 设备,另一个则是从高速缓存和 I/O 到多处理器核<sup>[5]</sup>。

交叉开关相对于总线来说其带宽会随着节点的增加而增大,而不像总线结构那样是固定不变的。并且其各个节点的连接是可以实时改变的,极大提高了多核核间通信的并行性和灵活性。交叉开关的结构如图 2(a)所示。

在理论上对于多处理器的并行计算应用,采用交叉开关进行通信,其效率是最高的<sup>[6]</sup>。然而其实现的代价较大。对于一个具有  $N$  个节点的交叉开关,其实现复杂度和代价将随着  $N^2$  增加,其结构复杂度如图 2(b)所示。与共享总线结构(SBF)一样,Crossbar 互连只适用于集成较少核的多核处理器结构,一旦集成的核数量较多,性能则会下降。

### 2.3 片上网络(Network On Chip)

片上网络是借鉴了现实中计算机互联网络的设计,将互连网络的架构用于片上系统中,有效解决片上模块之间的通信问题。与总线共享结构、交叉开关结构相比,片上网络能够连接更多的组件,其可靠性高、可扩展性强,并且具有比较低的功耗,因此片上网络被广泛用在处理器核的数目比较多的单片多核处理器中。

#### 2.3.1 片上网络结构

目前研究中的互连网络主要有环、带弦环、树形、胖树结构、星形、网格(Mesh)、环形网格(2D-torus)、立方体、带环立方体等结构<sup>[7]</sup>。片上网络的设计关键在于权衡多核互联通信效率与网络互连的开销,并保障片上网络的可重构性和可扩展性。当集成较多核时,可以使用二维 mesh 网格、3D-Torus 等片上网络结构,各个核之间利用通道相连形成网状,并在通道中发送消息实现通信<sup>[8]</sup>。

美国麻省理工学院研究的可重构 RAW 多核处理

器芯片就采用了片上二维网格结构,片上二维网格将各个处理器核需要交换的数据通过路由结构转发,通信延迟很小。每个单元能够通过片上网络和芯片上的逻辑通道,与外部存储器资源和各种 I/O 设备进行数据交换,RAW 处理器具有结构简单、可扩展性好、功耗小、在现有的工艺下易于实现等特点。

在商用多核处理器中,Tilera 公司的 Tile 系列处理器也采用了创新的二维网格片内互连架构(iMeshTM),可在单芯片内实现数十个到数百个全功能的处理器核心,同时将高性能、低功耗以及易于编程等特性集于一身。下面是 Tilera 公司 64 核处理器的 iMesh 网络互连架构,如图 3 所示。

Tilera 公司的多核处理器每个 Tile 由 Processor、Cache 和 Switch 组成,是一个具有完整功能的处理器,它的 Cache 由一级指令 Cache、数据 Cache 及二级 Cache 组成,通过 Switch 将每个 Tile 连接到网格中去,它的每个 Tile 能够独立运行一个操作系统,多个核连在一起能够运行像 SMP Linux 那样的多处理操作系统。

同样,Intel 公司也推出了 80 个核的低功耗处理器<sup>[9]</sup>。Intel 的 80 核处理器,单个处理器上集成了 80 个处理单元,80 个处理器单元按照  $10 \times 8$  的 mesh 结构连接,采用 65nm 工艺实现,一亿个晶体管,主频达到 4GHz,性能达到 TFLOPS,每个处理单元包含:两个浮点乘加器(FMAC)寄存器文件,2k 数据存储器,3k 指令存储器,五方向路由节点(Router)。

#### 2.3.2 分层网络互连方法

在分层网络中,网络的拓扑呈现为两层或多层的层次化结构,一般包括全局网和局部网两个层次,分别可以根据需要选用规则或非规则拓扑,例如全局网可以采用二维网格,局部网采用二叉树<sup>[10]</sup>。

在构造分层网络时,也可以利用片上网络互连可层次化、可扩展性好的特点先构造基本的单元组件,然后复制同样的组件构成所需要的网络拓扑结构,这种

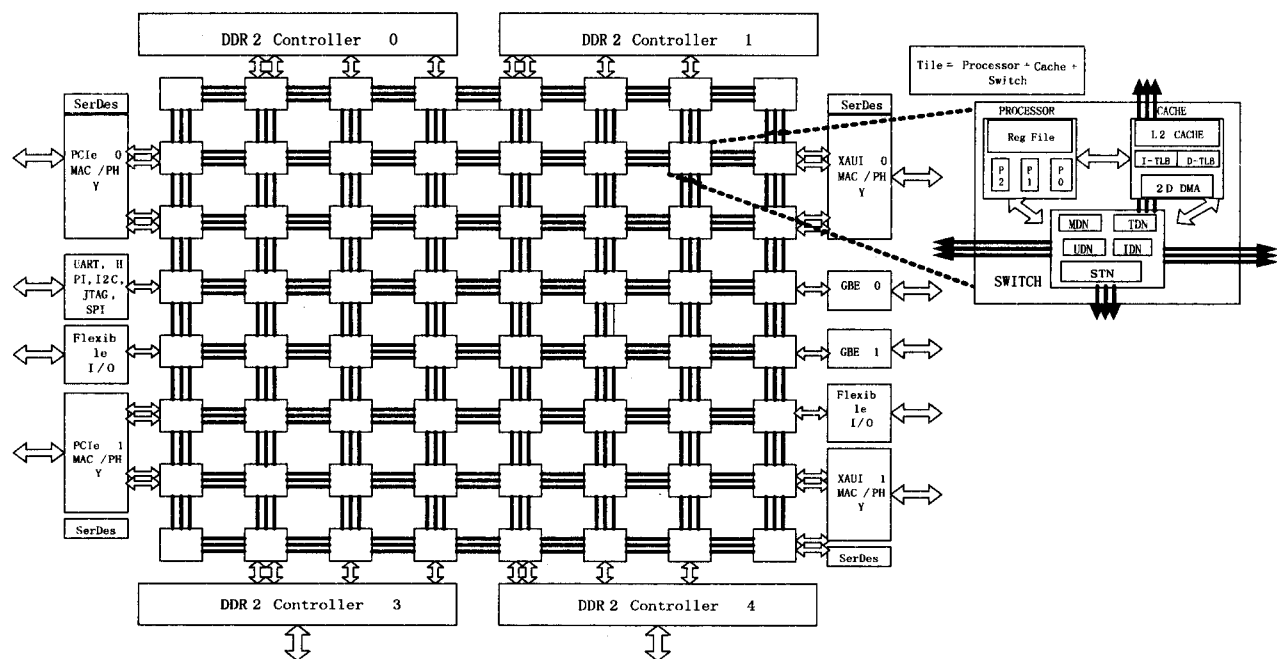


图3 iMesh片上网络架构示意图

网络拓扑结构方便构造,且形式多样,例如基三分层互连网络<sup>[11]</sup>。该结构的第0层是单个节点,如图4(a)所示,通过3条通信链路将3个节点彼此互连形成一个三角形,从而构成该结构的第1层,如图4(b)所示,1层网络是构造基三分层互连网络的基本构件<sup>[12]</sup>。如果采用递归构造方法将每个节点用一个低层次网络替代,就会得到一个高层次的网络拓扑结构,如图4所示。

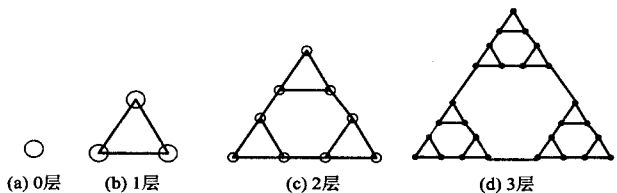


图4 基三分层互连网络拓扑结构

片上网络互连方法的改进是提高多核处理器性能的重要途径,它能在现有的工艺和制造技术的基础上尽可能地减小功耗和延迟,当多核处理器核的数目比较多时,高效的多核互连方法在功耗和延迟方面的优势将会表现的非常显著。

3 对多核互连发展趋势的展望

单核的系统架构经过几十年的发展已经基本成熟,单纯通过增加主频来提高处理器的性能已经非常困难,微处理器从单核到多核的发展已经成为不可逆转的趋势,如何提出一种行之有效的核间互连通信方式将对多核处理器的性能至关重要。虽然目前关于多核互连算法和架构的研究有很多,但大多数都无法使用到实际的产品中去,我们的研究重点应该转向新材

料和新技术上来,利用已经成熟的互连架构实现多核间的高速互连。Tilera公司能够利用二维 mesh 网格结构在传统的硅材料和现有的技术上实现一百个核的互连,新材料和新技术的使用必将使这种结构发挥更大的作用,集成成千上万个核也将会成为可能。

多核互连面临的问题主要是当核的数目比较多时互连线的功耗和延迟将会很大,如何解决这类问题将是多核处理器发展的重要课题。目前的芯片都是以硅材料为介质的,新材料、新工艺、新方法的采用将会给多核互连带来新的变化和发展。解决传统的片上电互连在带宽、时延、能耗和可靠性等问题的有效途径是采用光互连技术。它具有高带宽密度、低传输时延、低能耗和干扰小等优势,可以解决电连接所面临的通信问题。光互连技术需要一些快速、低功耗的集成光电子器件的支持,美国的 IBM、Intel、德国 Karlsruhe 大学及比利时大学等研究机构都相继研制出了速度达百 Gbit/s 的集成光电子器件<sup>[13]</sup>。新型纳米材料在集成光电子器件中的应用使得某种形态的碳纳米管可在晶体管中取代硅来控制电子流,并且碳纳米管可取代铜作为互连材料。这种碳纳米管就是石墨烯材料,以光信号驱动开关使传输速度更快,能使多核的互连实现真正意义上的高速通道。

世界最小的石墨烯光学调制器,由美国加州大学伯克利分校华裔师生共同研制诞生。这个比人的头发细 400 倍的光学调制器具备的高速信号传输能力,有望将互联网速度提高一万倍,石墨烯调制器的传输速度是 100Gbit/s,调制器用于调节光束开关,把电子信号转化成光学信号传输数字信息,如果把把这个调制器

集成到多核处理器里,使其完成多核互连的功能,将很好地解决现在多核互连在带宽、时延、能耗和可靠性等方面的问题,也必将带动多核处理器进入一个崭新的时代。

#### 4 结束语

多核处理器采用的互连方式,必须根据其实现所需核的数目和互连所带来的延迟、功耗、面积、实现工艺的复杂度等多方面进行考虑。在传统的技术和工艺框架下少于 10 核的多核处理器多采用总线的互连方式,大于 10 核的多核处理器适合采用交叉开关的互连方式,大于 36 核的结构比较适合采用片上网络结构。

随着新材料和新方法的应用及制造技术和工艺的不断进步,一些高效的互连技术已经实现,而随着片上多核处理器核的数目和结构不断的发展,还需要继续探索新型的多处理器核间互连通信技术。光互连技术的提出和各种集成光电子器件及石墨烯材料的应用为多核互连技术的发展提供了广阔的前景,一个高速、低功耗、高性能的多核处理器时代即将来临。

#### 参考文献:

- [1] Shen J P, Lipasti M. Modern Processor Design [M]. 北京:清华大学出版社,2007.
- [2] Patterson D A, Hennessy J L. 计算机组成与设计 [M]. 郑伟民,译. 北京:机械工业出版社,2007.
- [3] 黄国睿,张平,魏广博. 多核处理器的关键技术及其发展

趋势[J]. 计算机工程与设计,2009,30(10):2414-2418.

- [4] 王炜,汤志忠,乔林. 片上多处理器互连技术综述[J]. 计算机科学,2008,35(9):7-8.
- [5] 卜凡,赵忠民. 64 位多核 CPU 中交叉开关总线的设计与实现[J]. 计算机与数字工程,2008,36(11):151-154.
- [6] Zhang Y P, Jeong T. A Study of the On-chip Interconnection Network for the IBM Cyclops64 Multi-core Architecture [C]//Proceedings of 20th IEEE International Parallel Distributed Processing Symposium. [s.l.]:[s.n.],2006:1-10.
- [7] Culler D E, Singh J P, Gupta A. 并行计算机体系结构 [M]. 第 2 版. 李晓明,译. 北京:机械工业出版社,2002.
- [8] Ravankar A A, Sedukhin S G. "Mesh-of-Tori": A Novel Interconnection Network for Frontal Plane Cellular Processors [C]//Proceedings of IEEE First International Conference on Networking and Computing. [s.l.]:[s.n.],2010:281-284.
- [9] 胡晨骏,王晓蔚. 基于多核集群系统的并行编程模型的研究[J]. 计算机技术与发展,2008,18(4):70-73.
- [10] 王立炜. 片上网络架构下多核处理器系统的设计 [D]. 太原:太原理工大学,2010.
- [11] 乔保军,石峰,计卫星. 多核处理器核间互连的新型互连网络[J]. 北京理工大学学报,2007,27(6):511-516.
- [12] Haroon-Ur-Rashid Khan, Shi Feng, Jia Xinli. Performance of Triplet Based Interconnection Strategy for Multi-core On-chip Processors [C]//IEEE International Conference on High Performance Computing and Communications. [s.l.]:[s.n.],2009:163-170.
- [13] 李慧,顾华玺. 多核之间光互连技术的研究[J]. 中国集成电路,2010,19(2):50-55.

(上接第 29 页)

提出了一种通过 hashtable 保存历史数据的冗余数据过滤方法,大大减少了内存的需要。通过模拟 RFID 数据流进行仿真实验,验证了方法的有效性。文中的思想和方法对于构建 RFID 中间件(数据过滤是 RFID 中间件的重要组成部分)具有一定的参考价值。在噪音数据过滤方法中,如何更有效地提高 RFID 有效数据输出效率是下一步需要研究的问题。

#### 参考文献:

- [1] 游战清,李苏剑. 无线射频识别技术(RFID)理论与应用 [M]. 北京:电子工业出版社,2004.
- [2] Brusey J. Reasoning about Uncertainty in Location Identification with RFID [C]//RUR at IJCAI. [s.l.]:[s.n.],2003.
- [3] 郝忠孝. 主动数据库系统理论基础 [M]. 北京:科学出版社,2009.
- [4] Madden S. Continuously adaptive continuous queries over stream [C]//SIGMOD. [s.l.]:[s.n.],2002.
- [5] Oracle Sensor Edge Server [EB/OL]. 2008. [http://www.oracle.com/technology/products/iaswe/edge\\_server](http://www.oracle.com/technology/products/iaswe/edge_server).

- [6] Sybase RFID Solution [EB/OL]. 2005. <http://www.sybase.com/rfid>.
- [7] 张丰贵,程良伦. 基于 KDB 树的 RFID 事件聚合过滤算法 [J]. 计算机工程,2009,35(21):82-84.
- [8] 马岩,张延园,尹方鸣. 基于滑动窗口的 RFID 数据流多标签清洗算法 [J]. 科学技术与工程,2009,9(5):1165-1171.
- [9] 阴晓佳,鞠时光. 基于复杂事件处理机制的 RFID 数据流处理方法 [J]. 计算机应用,2009,29(10):2786-2789.
- [10] Hahn K. Adaptive workflow management to ensure transactional service composition [J]. Digital Information Management (ICDIM),2010,12(2):373-378.
- [11] Cheng Nan, Song Meina, Wang Qian. A web service process transaction framework based on compensation and proxy [J]. Pervasive Computing (JCPC),2010,24(8):369-372.
- [12] Vidyasankar K, Gottfried V. Multi-level Modeling of Web Service Compositions with Transactional Properties [J]. Journal of database management,2011,22(2):1-31.
- [13] Haddad J E. TQoS: Transactional and QoS-aware selection algorithm for automatic Web service composition [J]. IEEE Transaction on Service Computing,2010,14(8):210-218.