

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2007年5月3日 (03.05.2007)



PCT



(10) 国际公布号
WO 2007/048350 A1

(51) 国际专利分类号:
H04L 12/56 (2006.01)

(21) 国际申请号: PCT/CN2006/002880

(22) 国际申请日: 2006年10月27日 (27.10.2006)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
200510116784.5
2005年10月28日 (28.10.2005) CN

(71) 申请人 (对除美国外的所有指定国): 上海交通大学 (SHANGHAI JIAOTONG UNIVERSITY) [CN/CN]; 中国上海市华山路1954号, Shanghai

200030 (CN)。华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

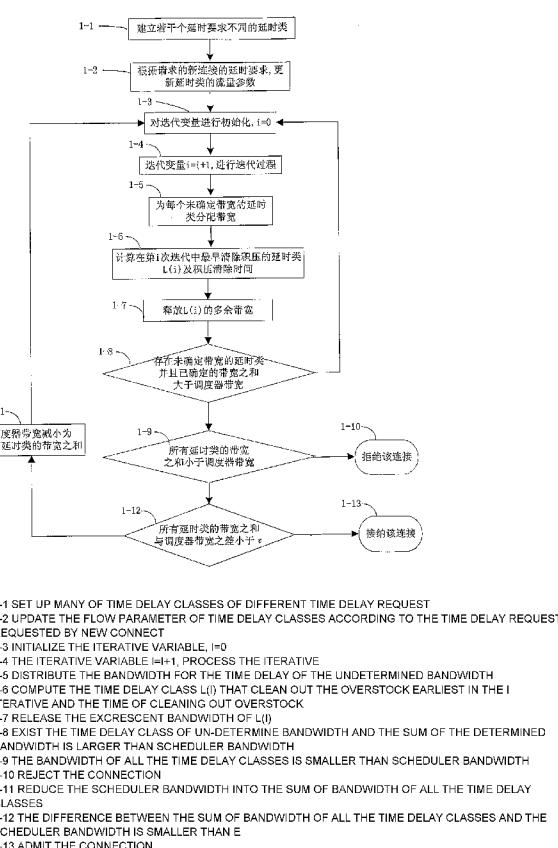
(72) 发明人; 及

(75) 发明人/申请人 (仅对美国): 赵峻 (ZHAO, Jun) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。郭勇 (GUO, Yong) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。范戈 (FAN, Ge) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。俞辉 (YU, Hui) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。惠娟娇 (HUI, Yan-jiao) [CN/CN]; 中国广东省深圳市龙岗区坂田华为

[见续页]

(54) Title: A METHOD AND EQUIPMENT FOR ADMITTING AND CONTROLLING THE INTEGRATION SERVICE MODEL

(54) 发明名称: 对集成服务模型进行接纳控制的方法及装置



(57) Abstract: A method and equipment for admitting and controlling the integration service model, includes mainly: setting up time delay classes of different time delay request of determined number in the integration service model; when a new connection requests to be set up, the determined bandwidth is distributed for the time delay classes, and according to the relation between the determined bandwidth and the scheduler bandwidth, the new connection is admitted and controlled. The integration service model is effectively admitted and controlled by utilizing the present invention, while ensures the QoS (quality of service) of connection and availability of the network.

(57) 摘要: 本发明提供了对集成服务模型进行接纳控制的方法及装置, 本发明主要包括: 在集成服务模型中建立设定数量的延时要求不同的延时类; 当有新连接请求建立时, 为所述延时类分配确定的带宽, 根据该确定的带宽和调度器带宽的关系, 对新连接进行接纳控制。利用本发明所述方法, 可以在保证连接的QoS (服务质量) 和网络的利用率的同时, 有效地对集成服务模型进行接纳控制。

WO 2007/048350 A1



总部办公楼, Guangdong 518129 (CN)。任颖(REN, Ying) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 北京凯特来知识产权代理有限公司(BEIJING CATALY IP ATTORNEY AT LAW); 中国北京市海淀区四道口路11号银辰大厦902室郑立明, Beijing 100081 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD,

SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告。

所引用双字母代码及其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

对集成服务模型进行接纳控制的方法及装置

技术领域

本发明涉及通讯领域，尤其涉及一种对集成服务模型进行接纳控制的方法及装置。

发明背景

集成服务模型是IETF（因特网工程师任务组）提出的一种网络交换节点的数据处理模型，其基本思想是在传送数据之前，根据业务的QoS（服务质量）需求进行网络资源预留，从而为该数据流提供端到端的QoS保证。

所述的集成服务模型主要由分类器、调度器和接纳控制组成。其中，所述的分类器用于将输入的数据分组按其所属的连接进行分类。所述的调度器则用于执行分组调度过程，对同一类的所有数据分组进行相同的转发处理。所述的接纳控制是在连接的建立阶段执行，用于判断节点是否具有足够的资源接纳新连接：如果是，则接纳该连接，并为其分配足够的资源；如果否，则拒绝该连接。其中，所述的接纳控制是集成服务模型的关键技术之一。

在现有技术中，提供了一种可以对集成服务模型进行接纳控制的方法，该方法具体为：基于测量的接纳控制方法，即实时测量节点的流量负载，根据测量结果判断节点是否能够接纳新连接。该方法包括两个主要部分：用于评估网络负载的测量部分和以及用于判断节点是否能够接纳新连接的决策部分。不同的接纳控制方法对这两部分可以采用不同的算法。

目前，通常采用的基于测量的接纳控制方法主要包括以下几种：

1、测量和法：通过时间窗口法来测量网络负载，以网络负载与新连接的漏桶速率之和作为决策变量。如果决策变量小于网络带宽乘以目标利用率，则判定节点能够接纳新连接，否则拒绝该新连接。

2、Hoeffding（霍夫丁）界法：以Hoeffding界测量网络的等效带宽，以等效带宽与新连接的峰值速率之和作为决策变量。如果决策变量小于网络带宽乘以目标利用率，则判定节点能够接纳新连接，否则拒绝该新连接。

3、峰值正切法：以时间点测量法测量网络负载，以 $np(1 - e^{-sp}) + e^{-sp}\hat{v}$ 作为决策变量，其中n为接纳的连接数，p为峰值速率，s为Chernoff（切尔诺夫）界的空间参数， \hat{v} 为测量的网络负载。如果决策变量小于网络总带宽，则判定节点能够接纳新连接，否则拒绝该新连接。

因此，上述基于测量的接纳控制方法包括以下缺点：

1、由于节点的流量负载具有一定的随机性，因此过去的测量数据不一定符合将来的流量情况，这就使得该方法所依据的测量数据并不准确。

2、该方法难以保证在最坏情况（例如所有业务一起以最高速率突发的情况）下的服务质量。

在现有技术中，还提供了另外一种对集成服务模型进行接纳控制的方法，具体为：基于参数的接纳控制方法。即以连接的建立或重协商阶段所约定的流量参数，而不是实时监测到的流量负载为依据，判断节点是否能够接纳新连接。这种方法一般按照最大流量情况为新连接分配带宽，因此，能够保证连接在任何时候（不违反流量参数的约定）的服务质量。该方法包括多种带宽计算方法，下面将分别进行说明。

其中，一种简单的计算方法为：为新连接分配的带宽 $r = \max\{\sigma / d, \rho\}$ ，其中 σ, ρ 为新连接的漏桶参数， d 为新连接的延时要求。如果节点的可用带宽小于 r ，则拒绝该连接。这种方法的计算量很低，但是在处理低速率低延时的业务（如语音业务）时网络带宽利用率较低。

另一种考虑了带宽的统计复用效果的带宽利用率较高的算法的主要处理过程包括如下步骤：

步骤1：对约定的流量参数变量进行初始化。

步骤2：为每个未确定带宽的连接分配带宽，并判断哪些连接的带宽可能在以后的迭代中进一步缩小（即带宽已确定），哪些连接的带宽不能进一步缩小（即带宽未确定）。

步骤3：寻找最先清除积压的连接，该连接的富裕带宽可以在以后的迭代中被其他连接使用。

步骤4：如果还存在带宽未确定的连接，并且已确定的带宽之和小于调度器带宽，转步骤2继续迭代。

步骤5：如果所有连接的带宽之和大于调度器带宽，则拒绝新连接。

步骤6：如果所有连接的带宽之和与调度器带宽之差小于一定值，则接纳新连接；否则减小调度器带宽后，转步骤1重新迭代。

可以看出，上述基于参数的接纳控制方法的缺点为：该方法虽然具有较高的带宽效率，但是计算时间正比于与连接数的平方，不具有可扩展性。也就是说，该方法是基于连接的，因此，需要为每个连接分配带宽，这就使得相应的方法在具体实现过程中，难以在计算时间和带宽效率上同时获得较好的表现。

发明内容

本发明提供了一种对集成服务模型进行接纳控制的方法及装置，从而能够在保证连接的QoS和网络的利用率的同时，有效地对集成服务模型进行接纳控制。

本发明提供了一种对集成服务模型进行接纳控制的方法，包括：

在集成服务模型中建立延时要求不同的延时类；

当有新连接请求建立时，确定该新连接所属的延时类别，为所述延时类分配确定的带宽，根据各延时类的带宽和调度器带宽的关系，对新连接进行接纳控制。

可选地，所述延时类是一组延时要求属于同一范围的连接的集合，一个延时类代表一个服务类别，属于同一个延时类的所有连接共用一个缓冲队列。

可选地，所述的确定该新连接所属的延时类别的处理具体包括：

当满足条件 $D_k \leq d < D_{k+1}$ 时，则新连接属于延时类 C_k ，其中 d 为新连接的延时要求， D_k 为延时类 C_k 的延时要求， D_{k+1} 为延时类 C_{k+1} 的延时要求；

在确定了新连接的延时类别为 C_k 后，更新 C_k 的流量参数 (σ_k, ρ_k) 为：

$\sigma_k = \sigma_k + \sigma$ ， $\rho_k = \rho_k + \rho$ ，其中 (σ, ρ) 为新连接的流量参数。

可选地，所述的为所述延时类分配确定的带宽具体包括：

当有新连接请求建立时，按照所有业务一起以最高速率突发的情况，通过迭代过程为所有延时类分配带宽，该分配的带宽包括确定的带宽和未确定的带宽；且在每次迭代时，计算首先清除积压的延时类，将该延时类释放出来的多余带宽重新分配给带宽未确定的延时类。

可选地，所述的通过迭代过程为所有延时类分配带宽具体包括：

对迭代变量进行初始化，将迭代次数变量 i 置为 0，用数组 $L(i)$ ， $i=0,1,2,\dots$ 表示第 i 个清除积压的延时类， $L(0)=0$ ，用数组 $T(i)$ ， $i=0,1,2,\dots$ 表示第 i 个清除积压的延时类的积压清除时间， $T(0)=0$ ，用数组 $S(i)$ ， $i=0,1,2,\dots$ 和 $R(i)$ ， $i=1,2,3,\dots$ 表示第 0 到第 i 个服务曲线的参数， $S(0)=0$ ， $R(1)=1$ ；

将迭代次数变量 i 加 1，按照所有业务一起以最高速率突发的情况为每个未确定带宽的延时类分配带宽 r ，如果 $T(i-1) \leq d$ ，则

$$r = \max\left\{\frac{\rho}{R(i)}, \frac{\sigma}{S(i-1) + R(i)(d - T(i-1))}\right\}, \text{ 该带宽为未确定的带宽;}$$

否则,

如果 $S(i-1)r < \sigma + \rho(T(i-1) - d)$, 则该延时类的带宽确定为:

$$r = \max\left\{\left\{\frac{\sigma + \rho(T(k) - d)}{S(k)} \mid T(k) \geq d\right\} \cup \left\{\frac{\sigma}{S(k) + R(k+1)(d - T(k))} \mid T(k) < d \leq T(k+1)\right\}\right\},$$

$$r = \frac{\rho}{R(i)}$$

否则, 该延时类的带宽确定为:

可选地, 所述的将该延时类释放出来的多余带宽重新分配给带宽未确定的延时类具体包括:

$$\frac{\sigma + \rho T(i-1) - r S(i-1)}{r R(i) - \rho}$$

在所有未清除积压的延时类中, 将使表达式 $\frac{\sigma + \rho T(i-1) - r S(i-1)}{r R(i) - \rho}$ 值最小的延时类确定为本次迭代中首先清除积压的延时类 $L(i)$, $L(i)$ 的积压清除时间为:

$$T(i) = \frac{\sigma + \rho T(i-1) - r S(i-1)}{r R(i) - \rho} + T(i-1), \text{ 其中 } \sigma, \rho, d, r \text{ 为 } L(i) \text{ 的参数;}$$

释放所述 $L(i)$ 的多余带宽, 将释放出来的带宽分配给其它未清除积压的延时类, 重

$$R(i+1) = \frac{C - \sum_{k \in P} \rho_k}{C - \sum_{k \in P} r_k},$$

新分配带宽后, 未清除积压的延时类的带宽为:

$$S(i) = S(i-1) + R(i)(T(i) - T(i-1)), \text{ 其中 } C \text{ 为调度器带宽;}$$

判断是否存在带宽未确定的延时类, 并且已确定的延时类的带宽之和小于调度器带宽, 如果是, 则重新执行所述迭代处理。

可选地, 所述的对新连接进行接纳控制具体包括:

判断是否所有已确定的延时类的带宽小于调度器带宽, 且

如果小于调度器带宽, 则判断是否所有已确定的延时类的带宽之和与调度器带宽之差小于设定值, 如果小于设定值, 则接纳所述请求建立的新连接; 如果不小于设置值, 则将调度器带宽减小为所有已确定的延时类的带宽之和, 重新执行所述迭代处理;

如果不小于调度器带宽, 则拒绝所述请求建立的新连接。

本发明还提供了一种对集成服务模型进行接纳控制的装置, 包括:

延时类建立单元, 用于在集成服务模型中建立的延时要求不同的延时类;

延时类确定处理单元, 用于在有新连接请求建立时, 确定该新连接所属的延时类别,

并触发所述延时类带宽分配单元；

延时类带宽分配单元，用于为所述延时类分配确定的带宽；

接纳控制单元，用于根据各延时类的带宽和调度器带宽的关系，对新的连接进行接纳控制。

可选地，所述的延时类确定处理单元具体包括：

延时类确定单元，用于在满足条件 $D_k \leq d < D_{k+1}$ 时，则新连接属于延时类 C_k ，

其中 d 为新连接的延时要求， D_k 为延时类 C_k 的延时要求， D_{k+1} 为延时类 C_{k+1} 的延时要求；

流量参数调整单元：用于在确定了新连接的延时类别为 C_k 后，更新 C_k 的流量参数 (σ_k, ρ_k) 为： $\sigma_k = \sigma_k + \sigma$ ， $\rho_k = \rho_k + \rho$ ，其中 (σ, ρ) 为新连接的流量参数。

可选地，所述的延时类带宽分配单元具体包括：

当有新连接请求建立时，按照所有业务一起以最高速率突发的情况，通过迭代过程为所有延时类分配带宽，该分配的带宽包括确定的带宽和未确定的带宽；且在每次迭代时，计算首先清除积压的延时类，将该延时类释放出来的多余带宽重新分配给带宽未确定的延时类。

可选地，所述的接纳控制单元具体包括：

第一判断处理单元，用于判断是否所有已确定的延时类的带宽小于调度器带宽，如果小于，则触发第二判断处理单元，否则，拒绝所述请求建立的新连接；

第二判断处理单元，用于判断是否所有已确定的延时类的带宽之和与调度器带宽之差小于设定值，如果小于，则接纳所述请求建立的新连接；否则，将调度器带宽减小为所有已确定的延时类的带宽之和，重新触发所述延时类带宽分配单元。

由上述本发明提供的技术方案可以看出，本发明通过以数量较少的延时类而不是数量很多的业务连接作为计算对象，从而可以有效地控制计算的时间复杂度，能够在保证连接的QoS和网络的利用率的同时，有效地对集成服务模型进行接纳控制。本发明能够在最坏情况（所有业务一起以最高速率突发的情况）下，保证所分配的带宽能够满足各连接的QoS需求。而且，本发明中，由于还考虑了已清除积压的延时类的带宽释放效应，因此，具有更高的带宽利用率。

附图简要说明

图1为本发明所述方法的具体实现流程示意图；

图2为集成服务模型中调度器的使用原理示意图；

图3为本发明所述的装置的具体实现结构示意图。

实施本发明的方式

本发明提供了一种对集成服务模型进行接纳控制的实现方案，本发明中，是以数量较少的延时类作为计算对象，而不再是以数量很多的业务连接作为计算对象，根据确定的延时类的带宽和调度器带宽的关系，对新连接进行接纳控制。

下面将结合附图，对本发明的具体实现方式进行详细描述。

本发明所述方法的具体实现流程如图1所示，可以包括如下步骤：

步骤1-1、建立若干个延时要求不同的延时类。

在该步骤中，首先需要建立若干个延时要求不同的延时类，所述的延时类是一组具有相似的时延要求的连接的集合；比如说，定义延时类A的时延范围为1~2秒，那么时延要求1.1秒和1.9秒的两个连接都属于延时类A；一个延时类即代表一个服务类别。

该步骤具体是在系统正式运行前进行，假设每个延时类 C_i 的延时要求为 $D_i (i=1, L, N)$ ，其原则可以为：序号较小的延时类的延时较小，即，如果 $i < j$ ，则 $D_i < D_j$ 。 $D_i (i=1, L, N)$ 的具体数值可以根据实际业务的情况来灵活设定。

在本发明中，属于同一个延时类的所有连接共用一个缓冲队列。由于数据业务的突发性，通常将来不及转发的数据存储在延时类的缓冲队列中，称作积压。过了一段时间，积压数据转发完毕，延时类的缓冲队列中没有数据，则延时类的积压清除了。从发生积压到积压清除之间的时间就是延时类的积压清除时间。

步骤1-2、根据请求的新连接的延时要求，更新延时类的流量参数。

当系统启动后，处于运行状态时，当有新连接请求建立时，首先确定该新连接的延时类别，依照步骤1-1中的假设，相应的确定方法可以是：如果 $D_k \leq d < D_{k+1}$ ，则新连接属于延时类 C_k ，其中 d 为新连接的延时要求。

在确定了新连接的延时类别为 C_k 后，然后更新 C_k 的流量参数 (σ_k, ρ_k) ，更新方法为： $\sigma_k = \sigma_k + \sigma$ ， $\rho_k = \rho_k + \rho$ ，其中 (σ, ρ) 为新连接的流量参数。

步骤1-3、对迭代变量进行初始化，i=0。

接着前面步骤，本发明需要通过迭代运算来为所有延时类分配带宽。在进行迭代运算之前，首先对迭代变量进行初始化。需要进行初始化的迭代变量包括：

i : 表示迭代次数。初始化时， $i=0$ 。

B : 表示未确定带宽、未清除积压的延时类。初始化时， $B=$ 所有延时类。

H : 表示已确定带宽、未清除积压的延时类。初始化时， $H=$ 空集。

P : 表示已确定带宽、已清除积压的延时类。初始化时， $P=$ 空集。

$L(i)$, $i=0,1,2,\dots$: $L(i)$ 表示第*i*个清除积压的延时类。例如，假设第4个清除积压的延时类是延时类3，那么 $L(4)=3$ ，初始化时， $L(0)=0$ 。

$T(i)$, $i=0,1,2,\dots$: $T(i)$ 表示第*i*个清除积压的延时类的积压清除时间。例如，假设第4个清除积压的延时类的积压清除时间为3秒，那么 $T(4)=3$ 。初始化时， $T(0)=0$ 。

$S(i)$, $i=0,1,2,\dots$ 和 $R(i)$, $i=1,2,3,\dots$: 表示服务曲线的参数。 $S(0)=0$, $R(1)=1$ 。

步骤1-4、将迭代变量加1，进行迭代过程。

将迭代变量*i*加1，即*i=i+1*，然后，进行迭代过程。

因为每次迭代可以计算出一个延时类的积压清除时间，所以迭代次数反映了各延时类的积压清除顺序。

步骤1-5、为每个未确定带宽的延时类分配带宽。

按照最坏情况（所有业务一起以最高速率突发）为每个未确定带宽的延时类分配带宽 r 。其具体计算过程为：

如果 $T(i-1) \leq d$ ，则

$$r = \max\left\{\frac{\rho}{R(i)}, \frac{\sigma}{S(i-1) + R(i)(d - T(i-1))}\right\}$$
，该延时类的带宽还没有确定，还有可能在以后的迭代中重新分配（进一步减小）。其中 $T(i-1)$ 为上次迭代计算出来的清除积压的延时类的积压清除时间；

否则，

$$r = \frac{\rho}{R(i)}$$

如果 $S(i-1)r = \sigma + \rho(T(i-1) - d)$ ，该延时类的带宽已确定。否则，将该延时类的带宽确定为：

$$r = \max\left\{\left\{\frac{\sigma + \rho(T(k) - d)}{S(k)} \mid T(k) \geq d\right\} \cup \left\{\frac{\sigma}{S(k) + R(k+1)(d - T(k))} \mid T(k) < d \leq T(k+1)\right\}\right\}$$

步骤1-6、计算在本次迭代中最早清除积压的延时类L(i)，及该延时类的积压清除时间。

计算在本次迭代（即第i次迭代）中首先清除积压的延时类L(i)。其具体方法为：在

$$\frac{\sigma + \rho T(i-1) - rS(i-1)}{rR(i) - \rho}$$

所有未清除积压的延时类中，寻找使表达式 $\frac{\sigma + \rho T(i-1) - rS(i-1)}{rR(i) - \rho}$ 最小的延时类，该延时类即为L(i)。

$$T(i) = \frac{\sigma + \rho T(i-1) - rS(i-1)}{rR(i) - \rho} + T(i-1)$$

L(i)的积压清除时间为： $\frac{\sigma + \rho T(i-1) - rS(i-1)}{rR(i) - \rho}$ ，这里 σ, ρ, d, r 为L(i)的参数。

步骤1-7、释放L(i)的多余带宽。

在集成服务模型中，调度器的使用原理图如图2所示，调度器按照一定规则，从各延时类的缓冲队列中取出数据，转发到调度器的输出端口。调度器带宽就是调度器的最大输出速率。

L(i)被清除积压后，其需要的带宽将减少，多余带宽将被释放，释放出来的带宽可以再分配给其它为未清除积压的延时类。

L(i)的多余带宽被释放后，未清除积压的延时类的实际带宽的计算方法为：

$$R(i+1) = \frac{C - \sum_{k \in P} \rho_k}{C - \sum_{k \in P} r_k}, \quad S(i) = S(i-1) + R(i)(T(i) - T(i-1))$$

其中C为调度器带宽。下面我们举例说明该计算方法。

假设调度器带宽100，4个延时类A、B、C、D分配的带宽分别为10、20、30、40，4个延时类都有积压。假设过了一段时间，延时类B首先清除积压，积压清除后只需要带宽5即可，那么未清除积压的延时类(A、C、D)的实际带宽就分别是 $10*(100-5)/(100-20)$ 、 $30*(100-5)/(100-20)$ 、 $40*(100-5)/(100-20)$ 。假设又过了一段时间，延时类C清除积压，积压清除后只需要带宽10即可，那么未清除积压的延时类(A、D)的实际带宽就分别是 $10*(100-5-10)/(100-20-30)$ 、 $40*(100-5-10)/(100-20-30)$ 。

步骤1-8、存在未确定带宽的延时类，并且已确定的带宽之和大于调度器带宽。

判断是否存在带宽未确定的延时类，并且已确定的带宽之和小于调度器带宽。如果是，执行步骤1-4，继续进行迭代过程；否则，执行步骤1-9。

步骤1-9、所有延时类的带宽之和小于调度器带宽。

判断是否所有已确定的延时类的带宽之和小于调度器带宽。如果是，执行步骤1-12；

否则执行步骤1-10。

步骤1-10、拒绝该连接。

拒绝所述请求建立的新连接。

步骤1-11、调度器带宽减小为所有延时类的带宽之和。

将调度器带宽减小为所有已确定的延时类的带宽之和，执行步骤1-3，重新开始迭代。

步骤1-12、所有延时类的带宽之和与调度器带宽之差小于 ϵ 。

判断是否所有已确定的延时类的带宽之和与调度器带宽之差小于一个预先设定的值 ϵ 。如果是，执行步骤1-13；否则，执行步骤1-11。

步骤1-13、接纳该连接。

接纳所述请求建立的新连接。

本发明还提供了一种对集成服务模型进行接纳控制的装置，其具体实现结构如图3所示，主要包括以下处理单元：

(1) 延时类建立单元，用于在集成服务模型中建立的延时要求不同的延时类，不同的延时类可以包括多个延时要求属于同一范围的连接，即所述延时类是一组延时要求属于同一范围的连接的集合，一个延时类代表一个服务类别，属于同一个延时类的所有连接共用一个缓冲队列，所述建立的所有延时类组成一个序列，按照延时要求进行排列，例如，可以为升序排列等；

(2) 延时类确定处理单元，用于在有新连接请求建立时，确定该新连接所属的延时类别，将其加入相应的延时类中，并触发所述延时类带宽分配单元；

而且，所述的延时类确定处理单元具体包括：

延时类确定单元，用于在满足条件 $D_k \leq d < D_{k+1}$ 时，则新连接属于延时类 C_k ，

其中 d 为新连接的延时要求， D_k 为延时类 C_k 的延时要求， D_{k+1} 为延时类 C_{k+1} 的延时要求；

流量参数调整单元：用于在确定了新连接的延时类别为 C_k 后，更新 C_k 的流量参数 (σ_k, ρ_k) 为： $\sigma_k = \sigma_k + \sigma$ ， $\rho_k = \rho_k + \rho$ ，其中 (σ, ρ) 为新连接的流量参数。

(3) 延时类带宽分配单元，用于为所述延时类分配确定的带宽，所述的延时类带宽分配单元具体采用的处理过程包括：

当有新连接请求建立时，按照所有业务一起以最高速率突发的情况，通过相应的迭

代过程为所有延时类分配带宽，该分配的带宽包括确定的带宽和未确定的带宽；而且，在每次迭代过程中，需要计算首先清除积压的延时类，将该延时类释放出来的多余带宽重新分配给带宽未确定的延时类；其中，具体的迭代处理过程前面已经详细描述，故在此不再重复描述。

(4) 接纳控制单元，用于根据各延时类的带宽和调度器带宽的关系，对新的连接进行接纳控制，所述的接纳控制单元具体可以包括：

第一判断处理单元，用于判断是否所有已确定的延时类的带宽小于调度器带宽，如果小于，则触发第二判断处理单元，否则，拒绝所述请求建立的新连接；

第二判断处理单元，用于判断是否所有已确定的延时类的带宽之和与调度器带宽之差小于设定值，如果小于，则接纳所述请求建立的新连接；否则，将调度器带宽减小为所有已确定的延时类的带宽之和，重新触发所述延时类带宽分配单元。

以上所述，仅为本发明较佳的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到的变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应该以权利要求的保护范围为准。

权利要求

1、一种对集成服务模型进行接纳控制的方法，其特征在于，包括：

在集成服务模型中建立延时要求不同的延时类；

当有新连接请求建立时，确定该新连接所属的延时类别，为所述延时类分配确定的带宽，根据各延时类的带宽和调度器带宽的关系，对新连接进行接纳控制。

2、根据权利要求1所述对集成服务模型进行接纳控制的方法，其特征在于，所述延时类是一组延时要求属于同一范围的连接的集合，一个延时类代表一个服务类别，属于同一个延时类的所有连接共用一个缓冲队列。

3、根据权利要求2所述对集成服务模型进行接纳控制的方法，其特征在于，所述建立的所有延时类组成一个序列，按照延时要求进行排列。

4、根据权利要求1所述对集成服务模型进行接纳控制的方法，其特征在于，所述的确定该新连接所属的延时类别的处理具体包括：

当满足条件 $D_k \leq d < D_{k+1}$ 时，则新连接属于延时类 C_k ，其中 d 为新连接的延时要求， D_k 为延时类 C_k 的延时要求， D_{k+1} 为延时类 C_{k+1} 的延时要求；

在确定了新连接的延时类别为 C_k 后，更新 C_k 的流量参数 (σ_k, ρ_k) 为：

$\sigma_k = \sigma_k + \sigma$ ， $\rho_k = \rho_k + \rho$ ，其中 (σ, ρ) 为新连接的流量参数。

5、根据权利要求1、2、3或4所述对集成服务模型进行接纳控制的方法，其特征在于，所述的为所述延时类分配确定的带宽具体包括：

当有新连接请求建立时，按照所有业务一起以最高速率突发的情况，通过迭代过程为所有延时类分配带宽，该分配的带宽包括确定的带宽和未确定的带宽；且在每次迭代时，计算首先清除积压的延时类，将该延时类释放出来的多余带宽重新分配给带宽未确定的延时类。

6、根据权利要求5所述对集成服务模型进行接纳控制的方法，其特征在于，所述的通过迭代过程为所有延时类分配带宽具体包括：

对迭代变量进行初始化，将迭代次数变量 i 置为0，用数组 $L(i)$ ， $i=0,1,2,\dots$ 表示第 i 个清除积压的延时类， $L(0)=0$ ，用数组 $T(i)$ ， $i=0,1,2,\dots$ 表示第 i 个清除积压的延时类的积压清除时间， $T(0)=0$ ，用数组 $S(i)$ ， $i=0,1,2,\dots$ 和 $R(i)$ ， $i=1,2,3,\dots$ 表示第0到第 i 个服务曲线的参数， $S(0)=0$ ， $R(1)=1$ ；

将迭代次数变量i加1，按照所有业务一起以最高速率突发的情况为每个未确定带宽的延时类分配带宽r，如果 $T(i-1) \leq d$ ，则

$$r = \max\left\{\frac{\rho}{R(i)}, \frac{\sigma}{S(i-1) + R(i)(d - T(i-1))}\right\} \text{，该带宽为未确定的带宽；}$$

否则，

如果 $S(i-1)r < \sigma + \rho(T(i-1) - d)$ ，则该延时类的带宽确定为：

$$r = \max\left\{\left\{\frac{\sigma + \rho(T(k) - d)}{S(k)} \mid T(k) \geq d\right\} \cup \left\{\frac{\sigma}{S(k) + R(k+1)(d - T(k))} \mid T(k) < d \leq T(k+1)\right\}\right\} \text{，}$$

$$r = \frac{\rho}{R(i)}$$

否则，该延时类的带宽确定为：

7、根据权利要求6所述对集成服务模型进行接纳控制的方法，其特征在于，所述的将该延时类释放出来的多余带宽重新分配给带宽未确定的延时类具体包括：

$$\frac{\sigma + \rho T(i-1) - r S(i-1)}{r R(i) - \rho}$$

在所有未清除积压的延时类中，将使表达式 $\frac{\sigma + \rho T(i-1) - r S(i-1)}{r R(i) - \rho}$ 值最小的延时类确定为本次迭代中首先清除积压的延时类 $L(i)$ ， $L(i)$ 的积压清除时间为：

$$T(i) = \frac{\sigma + \rho T(i-1) - r S(i-1)}{r R(i) - \rho} + T(i-1) \text{，其中 } \sigma, \rho, d, r \text{ 为 } L(i) \text{ 的参数；}$$

释放所述 $L(i)$ 的多余带宽，将释放出来的带宽分配给其它未清除积压的延时类，重

$$R(i+1) = \frac{C - \sum_{k \in P} \rho_k}{C - \sum_{k \in P} r_k} \text{，}$$

新分配带宽后，未清除积压的延时类的带宽为：

$$S(i) = S(i-1) + R(i)(T(i) - T(i-1)) \text{，其中 } C \text{ 为调度器带宽；}$$

判断是否存在带宽未确定的延时类，并且已确定的延时类的带宽之和小于调度器带宽，如果是，则重新执行所述迭代处理。

8、根据权利要求7所述对集成服务模型进行接纳控制的方法，其特征在于，所述的对新连接进行接纳控制具体包括：

判断是否所有已确定的延时类的带宽小于调度器带宽，且

如果小于调度器带宽，则判断是否所有已确定的延时类的带宽之和与调度器带宽之差小于设定值，如果小于设定值，则接纳所述请求建立的新连接；如果不小于设置值，则将调度器带宽减小为所有已确定的延时类的带宽之和，重新执行所述迭代处理；

如果不小于调度器带宽，则拒绝所述请求建立的新连接。

9、一种对集成服务模型进行接纳控制的装置，其特征在于，包括：

延时类建立单元，用于在集成服务模型中建立的延时要求不同的延时类；

延时类确定处理单元，用于在有新连接请求建立时，确定该新连接所属的延时类别，并触发所述延时类带宽分配单元；

延时类带宽分配单元，用于为所述延时类分配确定的带宽；

接纳控制单元，用于根据各延时类的带宽和调度器带宽的关系，对新的连接进行接纳控制。

10、根据权利要求9所述对集成服务模型进行接纳控制的装置，其特征在于，所述的延时类确定处理单元具体包括：

延时类确定单元，用于在满足条件 $D_k \leq d < D_{k+1}$ 时，则新连接属于延时类 C_k ，

其中 d 为新连接的延时要求， D_k 为延时类 C_k 的延时要求， D_{k+1} 为延时类 C_{k+1} 的延时要求；

流量参数调整单元：用于在确定了新连接的延时类别为 C_k 后，更新 C_k 的流量参数 (σ_k, ρ_k) 为： $\sigma_k = \sigma_k + \sigma$ ， $\rho_k = \rho_k + \rho$ ，其中 (σ, ρ) 为新连接的流量参数。

11、根据权利要求9或10所述对集成服务模型进行接纳控制的装置，其特征在于，所述的延时类带宽分配单元具体包括：

当有新连接请求建立时，按照所有业务一起以最高速率突发的情况，通过迭代过程为所有延时类分配带宽，该分配的带宽包括确定的带宽和未确定的带宽；且在每次迭代时，计算首先清除积压的延时类，将该延时类释放出来的多余带宽重新分配给带宽未确定的延时类。

12、根据权利要求9或10所述对集成服务模型进行接纳控制的装置，其特征在于，所述的接纳控制单元具体包括：

第一判断处理单元，用于判断是否所有已确定的延时类的带宽小于调度器带宽，如果小于，则触发第二判断处理单元，否则，拒绝所述请求建立的新连接；

第二判断处理单元，用于判断是否所有已确定的延时类的带宽之和与调度器带宽之差小于设定值，如果小于，则接纳所述请求建立的新连接；否则，将调度器带宽减小为所有已确定的延时类的带宽之和，重新触发所述延时类带宽分配单元。

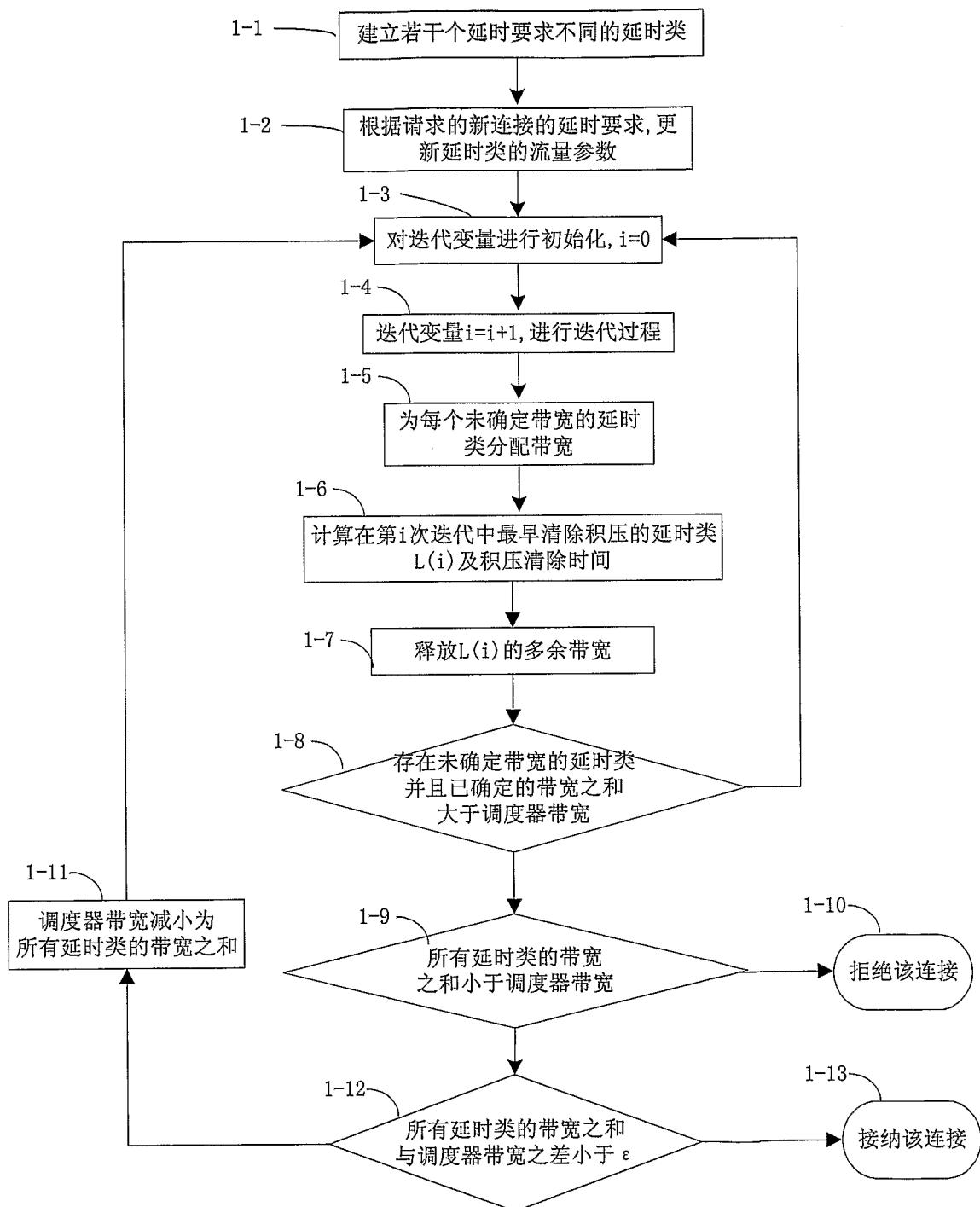


图1

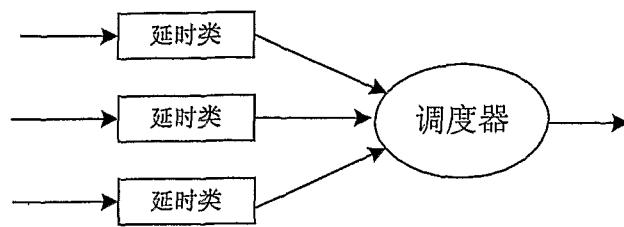


图2

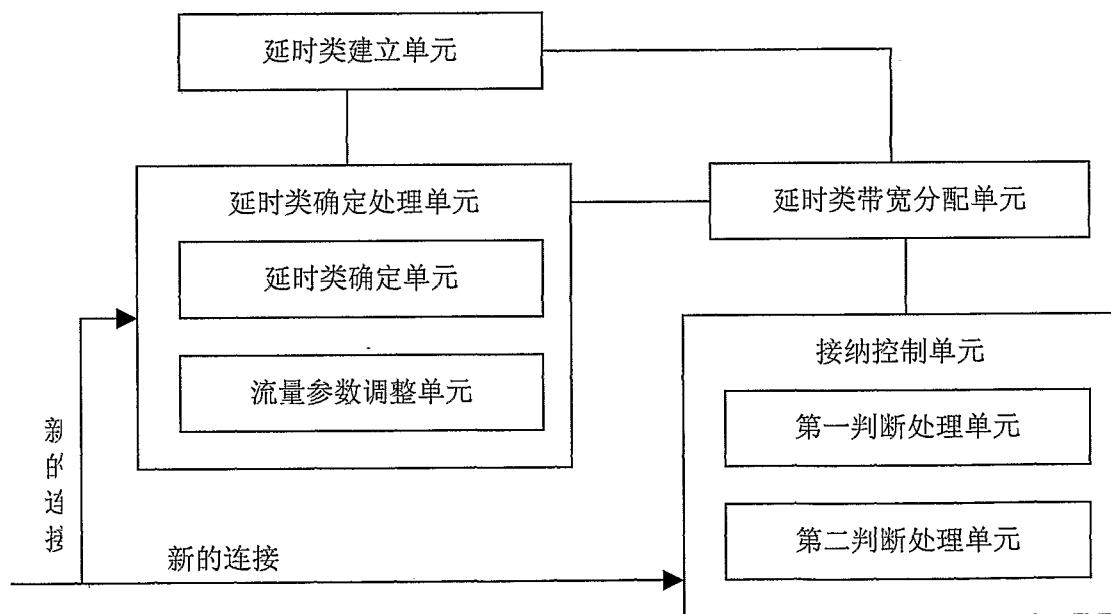


图3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2006/001898

A. CLASSIFICATION OF SUBJECT MATTER

H04L12/56 (2007.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L, H04Q, H04B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT、CNKI、WPI、EPODOC、PAJ: **dispatch+/schedul+, delay+, bandwidth, class+**

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO2004004212A2 (QUALCOMM INC) 08.Jan.2004 (08.01.2004) Claims 1、18, description paragraph 【1008】 - 【1013】	1,9
A	US2005175014A1 (PATRICK M W) 11.Aug.2005 (11.08.2005) See the whole document	1-12
A	JP2002135329A (MITSUBISHI ELECTRIC RES LABS INC et al) 10.May.2002 (10.05.2002) See the whole document	1-12
A	KR20050099241A (KT CORP) 13.Oct.2005 (13.10.2005) See the whole document	1-12
A	CN1581790A (ZTE CORP) 16.Feb.2005 (16.02.2005) See the whole document	1-12

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

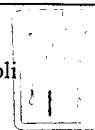
Date of the actual completion of the international search 24.Jan.2007 (24.01.2007)	Date of mailing of the international search report 08.FEB.2007 (08.02.2007)
---	--

Name and mailing address of the ISA/CN
The State Intellectual Property Office, the P.R.China
6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China
100088
Facsimile No. 86-10-62019451

Authorized officer

WANG Xiaoli

Telephone No. (86-10)62084536



INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2006/002880

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO2004004212A2	08.01.2004	US2004002341 A1 AU2003261093A KR20050016712A EP1573960 A2 TW200402239A JP2006502608T CN1714547A AU2003261093A8	01.01.2004 19.01.2004 21.02.2005 14.09.2005 01.02.2004 19.01.2006 28.12.2005 17.11.2005
US2005175014 A1	11.08.2005	NONE	
JP2002135329A	10.05.2002	EP1187404A1 US6990113B1	13.03.2002 24.01.2006
KR20050099241A	13.10.2005	NONE	
CN1581790A	16.02.2005	WO2005018154A RU2006107985A	24.02.2005 10.08.2006

国际检索报告

国际申请号

PCT/CN2006/002880

A. 主题的分类

H04L12/56 (2007.01)i

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04L, H04Q, H04B

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT、CNKI、WPI、EPODOC、PAJ:延时/时延/延迟, 调度, 带宽, 类, 连接, dispatch+/schedul+, delay+, bandwidth, class+

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	WO2004004212A2 (高通股份有限公司) 08.1月2004 (08.01.2004) 权利要求 1、18, 说明书【1008】至【1013】	1,9
A	US2005175014A1 (PATRICK M W) 11.8月2005 (11.08.2005) 参见全文	1-12
A	JP2002135329A (MITSUBISHI ELECTRIC RES LABS INC 等) 10.5月2002 (10.05.2002) 参见全文	1-12
A	KR20050099241A (KT CORP) 13.10月2005 (13.10.2005) 参见全文	1-12
A	CN1581790A (中兴通讯股份有限公司) 16.2月2005 (16.02.2005) 参见全文	1-12

 其余文件在 C 栏的续页中列出。 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

24.1月2007 (24.01.2007)

国际检索报告邮寄日期

08·2月2007 (08·02·2007)

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号: (86-10)62019451

受权官员

王晓丽



电话号码: (86-10) 62084536

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2006/002880

检索报告中引用的专利文件	公布日期	同族专利	公布日期
WO2004004212A2	08.01.2004	US2004002341 A1 AU2003261093A KR20050016712A EP1573960 A2 TW200402239A JP2006502608T CN1714547A AU2003261093A8	01.01.2004 19.01.2004 21.02.2005 14.09.2005 01.02.2004 19.01.2006 28.12.2005 17.11.2005
US2005175014 A1	11.08.2005	无	
JP2002135329A	10.05.2002	EP1187404A1 US6990113B1	13.03.2002 24.01.2006
KR20050099241A	13.10.2005	无	
CN1581790A	16.02.2005	WO2005018154A RU2006107985A	24.02.2005 10.08.2006