

基于服务质量水平的带宽分层自适应*

万俊伟 卢锡城

(国防科学技术大学计算机学院 长沙 410073)

E-mail: jwwan@nudt.edu.cn

摘要 自适应服务质量技术能有效地适应随时间变化的网络环境.提出了基于服务质量水平 max-min 公平的带宽分层自适应模型.此模型建立在应用层和网络层自适应控制的机制上.网络层利用资源探测协议来实现基于服务质量水平 max-min 公平的带宽分配,而应用层实现特定于应用的自适应服务.

关键词 自适应服务质量, max-min 公平.

中图法分类号 TP393

在动态网络环境中,分布式多媒体如音频、视频等应用需要适应网络服务质量的变化,变化的因素包括网络可用带宽、报文延迟等.其中对网络可用带宽的变化以及,多媒体应用可通过改变展现给用户的时间(如帧速)和空间(如帧尺寸)上的服务质量来适应.当前,许多编码器如 MPEG-1, MPEG-2 以及 H. 261 都采用层次编码技术,使得通过有选择地加入/去除编码层次来适应网络可用带宽的变化的自适应策略成为可能.

在 ATM(asynchronous transfer mode)网络中, ABR(available bit rate)服务类用来支持 best-effort 应用.这类应用具有很强的可伸缩性,对网络可用带宽的变化具有很强的适应性.当前,这类应用的带宽分配常采用某种公平策略,如 max-min 公平策略^[1].

下一代无线移动系统,如移动 IP、移动 ATM 以及第三代蜂窝系统,要求移动用户能在任何时间、任何地点访问、操作和发布语音、视频及数据信息.但无线移动网络所特有的物理层特性和用户的移动性使报文延迟和丢失率无法保障.为了克服无线移动系统中不断变化的服务质量对应用的影响,当前一致认为自适应技术是行之有效的解决方法^[2],其中很重要的一点就是带宽的自适应技术.

当前,带宽自适应技术的目标是在充分利用网络带宽的同时,尽量减小可用带宽的变化对可见服务质量的影响.目前已提出不少传输层和网络层的自适应服务质量技术,但缺少统一的自适应模型,而且带宽的分配往往基于平均原则,并未考虑分配的带宽对应用服务质量的影响.本文针对带宽的自适应性,提出基于服务质量水平 max-min 公平的分层自适应模型.在该模型中,带宽的自适应控制分两层来实现:应用层带宽自适应控制和网络层带宽自适应控制.应用层带宽自适应控制器利用网络层自适应控制器所获得的当前应用,可用带宽信息进行特定于应用的带宽自适应控制.

本文第 1 节给出分层自适应模型,描述了带宽动态自适应过程以及带宽自适应服务 API(application programming interface).第 2 节详细描述了基于服务质量水平 max-min 公平的网络层自适应控制.第 3 节描述了特定于具体应用的应用层自适应控制.第 4 节对上述机制下的网络带宽利用率进行了讨论.最后是全文总结.

* 本文研究得到国家“九五”国防网络预研项目基金和分布与并行处理国家重点实验室基金资助.作者万俊伟,1973年生,博士,主要研究领域为自适应组播通信, QoS, 移动通信.卢锡城,1946年生,教授,博士生导师,主要研究领域为高性能计算机,高性能网络技术.

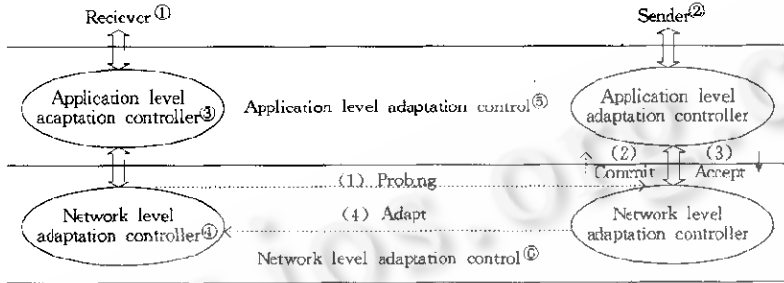
本文通讯联系人:万俊伟,长沙 410073,国防科学技术大学计算机学院

本文 1999-03-30 收到原稿,1999-08-31 收到修改稿

1 分层自适应服务模型

1.1 分层自适应控制体系

分层自适应控制体系包括网络层自适应控制和应用层自适应控制. 网络层自适应控制算法实现基于服务质量水平的 max-min 公平, 而应用层自适应控制算法支持特定于应用的一系列适应策略. 此分层自适应体系如图 1 所示.



①接收端, ②发送端, ③应用层自适应控制器, ④网络层自适应控制器, ⑤应用层自适应控制, ⑥网络层自适应控制.

Fig. 1 Split-Level adaptation control architecture
图1 分层自适应控制体系

带宽分层动态自适应控制是基于接收端周期性网络带宽资源探测过程来实现的. 在网络层, 接收端周期性地向发送端发送 probing 资源探测消息, 此消息在网络中驱动基于服务质量水平的 max-min 公平带宽分配算法, 该算法基于服务质量水平的 max-min 的公平策略, 根据当前网络中带宽使用状态为“正在进行资源探测”的会话重新分配带宽. 当探测消息到达发送端时, 其中将包含会话路径当前所能达到的速率(为当前该会话路径瓶颈链上分配的带宽), 发送端接收到该资源探测消息后, 立即将其提交(commit)给本地应用层自适应控制器, 本地应用层自适应控制器根据特定于应用的自适应策略来计算当前应用实际可用带宽, 然后通过 accept 原语将当前应用实际可用带宽信息告之网络层自适应控制器, 网络层自适应控制器将立即沿探测消息的反向路径, 向接收方返回一个 Adapt 消息, 此消息携带着当前应用实际可用带宽信息, 并用来释放会话路径上多余分配的带宽资源. 在发送端, 计算出当前应用实际可用带宽后, 应用将立即调整其当前的使用带宽, 因为接收方资源探测过程是周期性的, 根据上述自适应机制, 发送端将可以不断地根据当前网络状态对其应用的发送速率进行动态调整, 以达到带宽自适应的目的.

在资源探测过程中, 当前会话应分配的带宽资源可能小于当前会话正使用的带宽资源, 如果在探测时立即释放多余使用的带宽资源, 这部分资源可能立即就会被其他会话所占用, 而此时的发送端若仍以原速率发送数据, 就可能造成网络拥塞. 为解决该问题, 我们在资源探测过程中, 采用一种保守贪婪策略, 即在服务质量水平有所提高时, 立即分配增加的带宽资源; 在服务质量水平有所损失时, 维持原分配的带宽, 而探测过程仍真实地记录当前能获得的服务质量水平. 因此, 在探测过程中, 会话路径上的带宽只会增加, 不会减少, 从而保证了当前应用在作出自适应响应前其流的正常传输. 而探测的结果却正确地给出应用当前应调整到的使用带宽. 应用可以根据其特定的自适应策略进行响应, 如果探测的使用带宽高于当前应用的使用带宽, 应用可以考虑是否立即提高使用带宽和提高多少; 如果探测的使用带宽低于当前应用的使用带宽, 应用可以考虑是否立即降低使用带宽以及降低多少. 是否牺牲服务质量水平由具体应用来决定, 而不是由网络来决定. 根据上述保守贪婪策略, 应用调整后的速率总是小于或等于当前会话网络路径上的实际支持速率, 在探测完成后, 应用可以立即进行自适应响应.

1.2 带宽自适应服务 API

带宽自适应服务 API 用来定义自适应应用的带宽的可伸缩性和适应策略. 在我们定义的模型中, 此 API 包

含下列信息:

(1) 应用自适应策略

给出在网络可用带宽变化时,应用如何控制它们的发送速率以适应带宽的变化。

(2) 带宽-服务质量水平函数

该函数用来描述应用的自适应特性,反映应用服务质量水平与需要的可用带宽之间的关系。服务质量水平在本文中称为 SL ,它是 $[0,1]$ 区间上的数,0 代表最低的服务质量水平,其对应的带宽反映了应用最低的带宽需求; i 代表最高服务质量水平,其对应的带宽反映了应用在达到最高服务质量时所需要的带宽。

在分层自适应服务模型中,我们定义了以下 3 种类型的带宽-服务质量水平函数。

(1) 线性型,适合于 ABR 类应用,描述为 $([\min_b, \max_b], \text{Linear})$ 其中 \min_b 定义了应用最小带宽需求, \max_b 定义了应用最大带宽需求。若当前会话可用带宽为 cur_b ,令 $cur_b = \min\{cur_b, \max_b\}$,则其服务质量水平函数定义为

$$SL(cur_b) = \frac{cur_b - \min_b}{\max_b - \min_b} \quad (1)$$

(2) 层次离散型,适用于基于层次编码流的应用,描述为 $((b_1, b_2, \dots, b_k), k, \text{Discrete})$, 其中 b_1, b_2, \dots, b_k 给出各个编码层次所需要的带宽,参数 k 给出流的编码层次。若当前会话可用带宽为 cur_b ,则其服务质量水平函数定义为

$$SL(cur_b) = \frac{i-1}{k-1}, \quad \text{当 } b_i \leq cur_b < b_{i+1} \quad (2)$$

(3) 多层次连续型,适用于服务质量水平与可用带宽间是非线性连续函数关系的应用,描述为 $((b_1, b_2, \dots, b_k), k, \text{Piecewise})$, 其中 b_1 给出服务质量水平最低(为 0)时需求的带宽, b_k 给出服务质量水平最高(为 1)时需求的带宽,其他 b_i 给出原函数上按等距服务质量水平的其他关键点上的带宽信息,参数 k 给出关键点个数。若当前会话可用带宽为 cur_b ,令 $cur_b = \min\{cur_b, b_k\}$,则其服务质量水平函数定义为

$$SL(cur_b) = \frac{i + \frac{cur_b - b_i}{b_{i+1} - b_i} - 1}{k-1}, \quad \text{当 } b_i \leq cur_b < b_{i+1} \quad (3)$$

2 网络层自适应控制

网络层自适应控制是通过资源探测协议来实现的,资源探测协议驱动基于服务质量水平的 max-min 公平带宽分配算法。第 2.1 节将首先给出基于服务质量水平的 max-min 公平的定义,第 2.2 节描述资源探测协议以及实现基于服务质量水平的 max-min 公平的分布式算法。

2.1 基于服务质量水平的 max-min 公平

Max-Min 公平在网络带宽分配算法中应用是最广泛的。它的思想是,某应用流,如果提高的使用带宽不会降低其他具有相同或更低带宽的应用流的使用带宽时,才被允许提高该流的使用带宽^[3]。我们将 max-min 公平定义扩展为基于服务质量水平的 max-min 公平定义。

定义 1. 假设在当前网络中存在 n 个会话,某带宽分配方案为 $b = (b_1, b_2, \dots, b_n)$, 如果对任意会话 $i, b_i \geq \min_b[i]$ (会话 i 的最小带宽需求),而且对网络中的每条链 L , 满足条件 $\sum b_j \leq \beta^L$ (j 为任何经过链 L 的会话, β^L 为链 L 的总带宽),则定义该带宽分配方案是可行的。

定义 2. 一个带宽分配方案 $b = (b_1, b_2, \dots, b_n)$, 如果符合下列条件:

(1) 满足可行性约束;

(2) 对任意会话 i , 在可行性约束条件下,无法再增加 b_i , 除非降低某些会话 j 的带宽 b_j (其中 $SL(b_j) \leq SL(b_i)$), 则称该带宽分配方案是基于服务质量水平 max-min 公平的。

为说明基于服务质量水平的 max-min 公平带宽分配方案,我们给出图 2 的示例。在该示例中,会话 S_1D_1 , S_2D_2 , S_3D_3 的带宽-服务质量水平函数分别给定为: $S_1D_1(0, 4, \text{Linear})$; $S_2D_2(0, 6, \text{Linear})$; $S_3D_3(0, 10, \text{Linear})$;

链上用于自适应会话的带宽容量分别为 5 和 12,则在基于服务质量水平的 max-min 公平带宽分配方案下,会话 S_1D_1, S_2D_2, S_3D_3 的带宽将分别为 2、3、7,其服务质量水平分别达到 0.5、0.5 和 0.7.

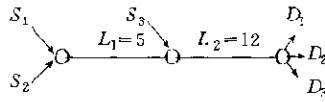


Fig. 2 Bandwidth allocation based on QoS level max-min fair
图2 基于服务质量水平的max-min公平带宽分配示例

2.2 实现基于服务质量水平 max-min 公平的分布式算法

传统的 max-min 公平实现采用集中式算法.在文献[3,4]中,集中式 max-min 公平算法改进为分布式的 max-min 公平算法,而文献[4]的算法较之文献[3]更加有效.在此,我们对文献[4]的算法进行了改进,以实现基于服务质量水平的 max-min 公平的带宽分配.下面是该算法的详细描述.

2.2.1 资源探测过程

接收端周期性地向发送端发送资源探测消息 $Probing(session_id, b_ideal, b_actual)$,其中 $session_id$ 为会话标志(网络中是唯一的); b_ideal, b_actual 为当前会话理想的和实际的需求带宽.其中 b_actual 用来跟踪可行性约束条件下的实际带宽分配,而 b_ideal 用来跟踪无可行性约束条件下的符合服务质量水平的 max-min 公平的理想带宽分配.

中间节点根据基于服务质量水平的 max-min 公平算法进行带宽分配.为了描述该算法,我们给出如下一些定义和概念.

考虑其会话 i 路径上的任一节点,令 $S_L(t)$ 为 t 时刻该节点上与会话 i 共享同一出链 L 的所有会话的集合.任一时刻, $S_L(t)$ 中的会话可以处于阻塞和满足两种状态中的一种.阻塞和满足状态的定义为:如果在最近一次带宽探测适应过程中,会话的理想带宽请求在该节点上得到满足,则会话在该链上的状态为满足状态,否则为阻塞状态.令 $S_L(t)^s$ 为 t 时刻链 L 上的所有处于满足状态的会话的集合;令 $S_L(t)^b$ 为 t 时刻链 L 上的所有处于阻塞状态的会话的集合.

定义 $alloc(i, \{i\} \cup \{\text{其他会话标志}\}, b)$ 为在可用带宽为 b 时,第 2 项参数中给定的所有会话都获得最大相同服务质量水平时,分配给会话 i 的带宽.

对会话 i ,在其经过的每条链 L 上,具有以下状态:会话的服务质量水平函数 $SL_i, r_ideal(i, L)$ 为最近一次带宽探测适应过程中,会话 i 在链 L 上按基于服务质量水平的 max-min 公平算法获得的理想带宽值; $r_actual(i, L)$ 为在最近一次带宽探测适应过程中会话 i 在链 L 上获得的实际带宽值; $status(i, L)$ 为最近一次带宽探测适应过程中会话 i 在链 L 上的状态(阻塞或满足).在每条链 L 上,还具有以下状态参数: B^L 为链 L 的带宽容量; B^s 为链 L 上所有处于满足状态的会话的 r_ideal 值的总和; B^{free} 为链 L 上当前实际剩余带宽值.

当会话 i 的资源探测消息到来时,将其看成一个新的会话请求,首先将其状态设置为空状态(既不是满足态,也不是阻塞态),然后计算会话在当前出链上可以获得的理想带宽.

$$\begin{aligned}
 b_alloc(i, L) &= alloc(i, \{i\}, B^L) && \text{如果 } S_L(t)^s \cup S_L(t)^b = \emptyset; \\
 b_alloc(i, L) &= alloc(i, \{i\} \cup \{k\}, B^L - B^s_k + r_ideal(k, L)) && \text{否则,如果 } S_L(t)^b = \emptyset; \\
 b_alloc(i, L) &= \max(alloc(i, \{i\} \cup S_L(t)^b, B^L - B^s_k), alloc(i, S_L(t), B^L)) && \text{否则,其他.}
 \end{aligned}$$

其中 $k = \arg \max \{SL_j, r_ideal(j, L)\}$; j 是任一属于 $S_L(t)^b$ 的会话.这样,基于服务质量水平的 max-min 公平带宽分配算法可描述如下:

- (1) 初始:获取资源探测消息中的参数 $session_id, b_ideal, b_actual$;
- (2) 计算 $b_alloc(i, L)$;
- (3) 如果 $b_alloc(i, L) \geq b_ideal$,认为可以满足当前需求,将会话状态设置为满足;
如果 $b_alloc(i, L) < b_ideal$,认为不能满足当前需求,将会话状态设置为阻塞;
- (4) 令 $r_ideal(i, L) = \min\{b_alloc(i, L), b_ideal\}$; //当前理想带宽
如果是新会话, $r_actual(i, L) = \min\{b_alloc(i, L), b_actual, B^{free}_L\}$,并更新 B^{free}_L 的值;

否则,如果 $r_actual(i,L) > \min\{b_alloc(i,L), b_actual, B_{free}^i\}$,

则令 $r_actual(i,L) = \min\{b_alloc(i,L), b_actual, B_{free}^i\}$,并更新 B_{free}^i 的值; //保守贪婪策略,

(5) $b_ideal - r_ideal(i,L); b_actual = \min\{b_alloc(i,L), b_actual, B_{free}^i\}$; //真实记录当前资源
//探测结果

(6) 如果 $SL_i(r_ideal(i,L)) > \text{Max}S_1$, 则

$\text{Max}S_1 Vc = i; \text{Max}S_1 = SL_i(r_ideal(i,L))$; 并将会话 i 状态设置为阻塞.

如果 $\text{Max}S_1 Vc = i$ 而且 $SL_i(r_ideal(i,L)) < \text{Max}S_1$, 则 $\text{Max}S_1 = SL_i(r_ideal(i,L))$;

//其中 $\text{Max}S_1$ 定义为当前所有会话中最大理想服务质量水平, 而 $\text{Max}S_1 Vc$ 定义为当前具有最大理想服务质量水平的会话标志.

(7) 向前发送更新后的资源探测消息.

当资源探测消息到达发送端时, b_ideal 中存放的将是路径各链上可提供给该会话的理想带宽的最小值; b_actual 中存放的是路径各链上可实际分配给该会话的带宽的最小值. 发送端将接收到的 b_ideal 和 b_actual 信息提交给应用层自适应控制器, 由其根据特定的应用层自适应控制算法计算最终的理想带宽值和实际带宽值, 发送端最终向接收端发送 adapt 消息, 该消息中带有会话当前理想带宽值和实际带宽值.

2.2.2 资源适应过程

当中间节点接收到 $adapt(i, b_ideal, b_actual)$ 消息时, 可能需要释放会话 i 多余分配的带宽, 而且可能需要改变会话的状态. 操作如下:

(1) 如果 $r_ideal(i,L) > b_ideal$, 则 $r_ideal(i,L) = b_ideal$, 并将会话状态变为满足;

(2) 如果 $r_actual(i,L) > b_actual$, 则 $r_actual(i,L) = b_actual$, 并更新 B_{free}^i 的值;

(3) 继续向前发送该 adapt 消息.

3 应用层自适应控制

当应用层自适应控制器接收到网络层自适应控制器的 $commit(session_id, b_ideal, b_actual)$ 原语时, 将根据特定于应用的应用层自适应策略计算当前理想带宽值和下一步会话实际可以使用的带宽值. 应用的自适应服务策略, 反映了不同的应用行为.

3.1 应用层自适应策略

下面, 我们给出几种典型的应用层自适应策略, 并给出其简单的实现算法:

(1) 贪婪自适应策略

该策略适合 ATM 的 ABR 服务类应用, 这类应用可以立即对当前的网络可用带宽进行适应, 其算法很简单. 下面给出该算法.

```
commit(session_id, b_ideal, b_actual)
{accept(session_id, b_ideal, b_actual);}
```

在此, 我们利用图 2 的例子来说明采用该策略的带宽动态自适应过程, 如图 3 所示. 在该示例中, 当时刻 $t=0$ 时, 会话 S_1D_1 加入, 获得的带宽为 4; $t=1$ 时, 会话 S_2D_2 加入, 获得的带宽为 1; $t=2$ 时, 会话 S_3D_3 加入, 获得的带宽为 7; 会话加入后, 其将每隔 3 个时间单位进行一次资源探测, $t=3$ 时, S_1D_1 进行资源探测, 实际带宽降为 2; $t=4$ 时, S_2D_2 进行资源探测, 实际带宽升到 3, 由此, 整个网络状态进入稳定状态. 此时, 假设 $t=7$ 时, S_2D_2 退出, 在 $t=12$ 时, 网络又进入另一个问题状态, 此时 S_1D_1 的带宽为 $24/7$, 而 S_3D_3 的带宽为 $60/7$, 如图 3 所示.

(2) 离散自适应策略

该策略适合于基于层次编码流的应用, 带宽升高和降低的幅度是以离散层次为单位的. 如果会话的带宽-服

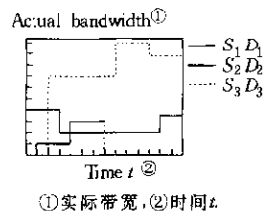


Fig. 3 Greedy adaptation scheme
图3 贪婪自适应策略

务质量水平函数为 $((b_1, b_2, \dots, b_k), k, \text{Discrete})$, 则有一种简单的离散自适应算法:

```

commit(session_id, b_ideal, b_actual)
{
    m = [1/SL_i(b_ideal)]; //[]符号在此表示下取整
    n = [1/SL_i(b_actual)];
    b_ideal = b_n; b_actual = b_n;
    accept(session_id, b_ideal, b_actual);
}

```

(3) 平滑自适应策略

该策略适合于自适应多媒体应用. 它通过限制每次带宽增长的幅度并规定相邻两次带宽变化间的间隔时间以及限制当前使用带宽下降的幅度, 以达到应用服务质量平滑改变的目的. 对平滑自适应策略, 其 API 还应包含以下3个控制参数: δ 为应用每次的最大带宽的增量; τ 为两相邻增长间的最小时间间隔; κ 为下降因子. 该自适应策略的算法如下:

```

commit(session_id, b_ideal, b_actual)
{
    if (符合  $\tau$  时间间隔约束)
    {
        if (带宽增长)
            b_actual = min{b_old -  $\delta$ , b_actual}; //b_old 为当前实际带宽值;
        else (带宽下降)
            b_actual = max{b_actual, b_old *  $\kappa$ };
        b_old = b_actual;
        accept(session_id, b_ideal, b_actual);
    }
    else accept(session_id, b_ideal, b_old);
}

```

3.2 自适应视频服务

当前的视频编码器, 如 MPEG-1, MPEG-2 和 H. 261, 多采用层次编码技术. 它们往往生成一个基本层次和多个扩展层次, 其中基本层次代表最低的视频质量, 每个扩展层次对应更高的视频质量, 每个扩展层次的加入都将提高相应的视频质量, 同时也要求更多的带宽. 传统的视频服务需要网络提供很高的常速率支持, 这在异构网络环境中往往难于满足. 利用当前视频编码的层次特性, 结合上述带宽的动态自适应策略, 我们可以实现自适应视频服务. 其大致实现如下:

在视频会话首次建立时, 如果探测到的网络理想带宽小于其基本层次的带宽, 则该会话退出. 如果探测到的网络理想带宽大于等于其基本层次的带宽, 而当前实际分配的网络带宽小于基本层次的带宽, 则会话可以选择退出或继续进行探测. 如果实际分配到的网络带宽大于等于基本层次的带宽, 则按上述离散自适应策略计算当前的发送层次. 在会话活跃过程中, 接收方周期性地激活资源探测过程, 以动态地探测当前可用网络带宽, 以便在必要时降低或提高其接收的视频质量.

因为我们的动态自适应策略采取由具体应用来决定是否升高和降低其使用带宽, 并不受网络控制, 因此, 应用可以根据其具体要求制定特定于应用的自适应策略, 如要求服务质量稳定或只希望服务质量上升等.

4 关于网络带宽利用率的讨论

在本文结束前, 我们来讨论一下上述分层带宽自适应策略对网络带宽利用率的影响. 影响大致分为两个方面: 一方面是网络中基于服务质量水平 max-min 公平带宽分配对带宽利用率的影响, 另一方面是特定于应用的自适应策略对带宽利用率的影响.

基于服务质量水平 max-min 公平的分布式带宽分配算法需要一定的收敛时间才能在网络中达到服务质量

水平的 max-min 公平。在收敛过程中,因公平性的需要,一些会话要释放一定的使用带宽,而这部分带宽需要等到其他会话探测时才可能被重新使用,从而造成在收敛过程中的某些时间网络存在闲置不用的带宽资源,因此造成一定的带宽资源浪费。比如,图3中的时刻 $t=3$ 与 $t=4$ 之间,链 L_1 上的使用带宽为3,而链 L_2 上的使用带宽为10,分别浪费了两个单位的带宽。而一旦整个网络收敛到稳定状态时,根据定义2,在服务质量水平的 max-min 公平约束下,网络的带宽利用率达到最大值。考虑到服务质量的公平对网络带宽的利用效率会有一些的负面影响,简单分析如下:假设某链的带宽为 B ,其上具有两个会话,且最大带宽分别为 \max_1, \max_2 ,并假设 \max_1 和 \max_2 都大于等于 $B/2$,按照服务质量水平公平策略,两个会话的服务质量水平都将达到 $B/(\max_1 + \max_2)$;而如果按简单的平均带宽分配原则,则两个会话的服务质量水平将分别为 $B/2\max_1$ 和 $B/2\max_2$ 。在相同的总带宽 B 的前提下,可以证明,前者的总体服务质量水平总是小于等于后者的总体服务质量水平。可见,在实现服务质量水平公平的同时,整体带宽利用效率可能会有所降低。

另外,应用的自适应策略也会对网络带宽的利用率产生影响,如平滑自适应策略,考虑到服务质量改变的平滑性,应用可能不会立即使用全部分配给它的带宽,而这部分带宽又不能被别的会话使用,从而造成一定的带宽资源的浪费。

总而言之,在实现服务质量水平 max-min 公平的同时,会损失一定的带宽利用率。

5 结束语

本文给出了基于服务质量水平 max-min 公平的带宽分层自适应控制模型,并详细介绍了该模型的具体实现机制,给出了实现基于服务质量水平 max-min 公平的分布式算法以及几种适合不同应用的应用层自适应策略及实现算法。

下一步需要进行的研究工作主要有:定量分析在该公平算法下的网络带宽利用率;研究各种应用层自适应策略的自适应行为,以求改进,使其更好地符合具体应用的需求。

参考文献

- 1 Bertsekas D, Gallager R. Data Networks. 2nd ed., Englewood Cliffs, NJ: Prentice Hall, Inc., 1992
- 2 Katz R H. Adaptation and mobility in wireless information systems. IEEE Personal Communications Magazine, 1994, 1(1): 6~17
- 3 Charny A, Ramakrishnan K K. Time scale analysis of explicit rate allocation in ATM networks. In: Onvural R O ed. Proceedings of the IEEE INFOCOM'96. Los Alamitos, CA: IEEE Computer Society Press, 1996
- 4 Kalampoukas K, Varma A, Ramakrishnan K K. An efficient rate allocation algorithm for ATM networks providing max-min fairness. In: Puigjaner R ed. Proceedings of the 6th IFIP International Conference on High Performance Networking. London: Chapman & Hall, 1995

QoS Level-Based Bandwidth Split-Level Adaptation

WAN Jun-wei LU Xi-cheng

(School of Computer National University of Defence Technology Changsha 410073)

Abstract Techniques for adaptive quality of services can efficiently respond to the network environments varying with time. In this paper, the QoS level-based max-min fair split-level bandwidth adaptation model is presented, which is built on a split-level adaptation control scheme that operates at the application and network levels. The network uses a resource probing protocol to achieve QoS level-based max-min fair bandwidth allocation, while applications realize application-specific adaptive services.

Key words Adaptive quality of service, max-min fair.