Computer Networking
A Top-Down Approach pp.623-656

James F. Kurose, Keith W. Roth

情報工学科 3 年 寺岡研究室
61619027 安森 涼

2019 年 3 月 7 日

1　即時対話型アプリケーションプロトコル

VoIP やビデオ会議を即時性のある対話型アプリケーションは現在，頻繁に利用される．IETF や ITU がこれらのアプリケーションに対する標準を形作っている．もし適切な標準化がされていれば個々の会社がそれぞれ相互運用できる新しい製品を開発することができる．この章では RTP と SIP の標準について述べる．

1.1　RTP

前回，VoIP アプリケーションの送信側でトランスポート層通過前に音声チャンクにヘッダフィールドを加えると述べた．ヘッダフィールドにはシーケンス番号とタイムスタンプが含まれ，音声動画ファイルとともに標準化する．RTP は RFC3550 で定義されているものが標準である．RTP は PCM,ACC,MP3 形式の音声と MPEG，H.263 形式の動画の送信に使われる．このように幅広い実装に役立っている．この節では RTP の紹介をおこなう．

RTP は UDP の上部で動く，送信側はメディアチャンクを RTP パケットにカプセル化し，UDP セグメントにカプセル化し IP に渡し，受信側が RTP パケットを UDP セグメントから取り出し，メディアチャンクを RTP パケットから取り出す．そして，チャンクをメディアプレイヤに送信し，復号化とレンダリングをおこなう．例として音声を RTP で送信することを考える．音声は 64kbps の PCM で符号化されている．チャンクは 20msec(160kbyte) ごとに送られる．送信側はそれぞれのチャンクに RTP ヘッダを加える．RTP ヘッダには音声符号化の形式，シーケンス番号とタイムスタンプを含み，基本的には 12byte で構成されている．チャンクと RTP ヘッダを合わせたものを RTP パケットと呼ぶ．これを上記のようにカプセル化して送信し，受信側で取り出していることで送受信をおこなっている．この RTP ヘッダにより，他のネットワーク接続されたマルチメディアアプリケーションにと相互運用できる．

RTP は時間通りに送信を行うことおよび，その他の QOS を保証していない．また，パケットの配達と，順序が異なる配達を防ぐことですら保証していない．事実，カプセル化はエンドシステムでのみおこなわれ，ルータは RTP を送信する IP データグラムか否かを判別できない．RTP は個別のパケットを個々のパケットのストリームに割り当てる．例えば，2 人がビデオ会議をする際に，音声と映像について，それぞれ両方向に 2 つの流れ，計 4 つのストリームを割り当てる．しかしながら MPEG1 や MPEG2 は音声と映像を一つのストリームに割り当てるので両方向計 2 つのストリームで大丈夫である．RTP ではセッションと呼ばれる参加者のグループが規定されており，参加者ごとのセッションの識別には、ネットワークアドレス、データを送信するポートの組、データを受信するポートの組が使用される．参加者は複数のセッションに同時に参加することも可能である．

1.1.1　RTP ヘッダフィールド

RTP ヘッダフィールドにはペイロードタイプ，シーケンス番号，タイムスタンプ，SSRC 識別子が含まれている．ペイロードタイプは 7bit で．音声，映像符号化の形式が含まれている．これにより受信側に通知する．続いて，シーケンス番号フィールドは 16bit 長で RTP が送られるたびにインクリメントを行い，受信者にパケットロスを通知し，パケットを回復する．86-89 のシーケンス番号を送った時に 87,88 のパケットがロスしたと受信者に通知されたら，復旧させるように試みる．タイムスタンプフィールドは 32bit 長で，先頭バイトの送信時刻を示している．前章に記述された通り，受信側がネットワークで生成されたパケットの乱れを除去する，また，同期再生を受信側に提供している．SSRC はは 32bit で RTP ストリームのソースを識別するのに利用される．RTP セッションは区別された SSRC を持っている．SSRC は送信側の IP アドレスではなく，新しいストリームが作られるたびにランダムに番号を割り振る．2 つのストリームに同じ番号が割り振られるのは稀だが，起きた時は新しい番号を振り直す．

1.2　SIP

SIP はオープンで軽いプロトコルで，IP ネットワークをもとに，発信者と着信者の間の通信でコールを確立する機構である．発信者が着信者に発信開始および終了の合図を行う．また発信者に受信者の現在の IP アドレスを通知している．また，コール管理の機構を提供している，コール中に新しいメディアストリームを加える，符号化を変更，コールに新しく参加者を加入させることとコール転送と，コール保留を行う．

1.2.1　既知の IP アドレスとのコール確立

SIP について具体的に説明する．アリスが PC でボブの PC にコール確立をしたいとする．彼らの PC にはコールの送受信の SIP に基づいたソフトウェアが装備されているとする．まずアリスがボブに招待メッセージを UDP でウェルノウンポートである 5060 番ポートで送る．招待メッセージはボブの識別子，アリスの現在の IP アドレスの指示，アリスの受信フォーマット，アリスが RTP パケットを受信したいポート番号を含んでいる．受信後，ボブは SIP の応答メッセージを送信する．この応答メッセージは 200 OK とボブの IP アドレスと受信フォーマット，ボブが RTP パケットを受信したいポート番号を含み，5060 番ポートに送信される．ボブの応答後のアリスによる SIP の Ack を送った後に，音声パケットをそれぞれ別のポートに送ることでアリスとボブが異なる符号化を利用することができる．これらの例から，SIP は out-of-band プロトコルで，つまり，SIP メッセージは，メディアファイルを送受信するソケットと異なる場所を通じて送受信されることと，SIP メッセージは ASCII で可読性があり，HTTP メッセージに似ていること，メッセージに認証が必要であるため，UDP，TCP 上で動くことがわかる．もし，ボブが符号化できない状況であったら 200 OK の代わりに 600 Not Acceptable を送信することで，アリスに招待メッセージを再送させる．

1.2.2　SIP アドレス

SIP アドレスと Email アドレスは似ている．SIP アドレスを用いて，ボブにメッセージを送るときのように IP アドレスにルーティングを行うことができる．このアドレスは Web ページに貼り付けることができる．

1.2.3　SIP 招待メッセージ

アリスがボブの SIP アドレスのみを知っており，IP アドレスを知らないときのアリスの招待メッセージに着目する．まず INVITE 行に SIP のバージョンと SIP アドレスを表示している．Via 行には通過する機器の IP アドレスを表示している．FromTo 行は Email と同じである．Call-ID 行ではコールに識別子を割り振っている．

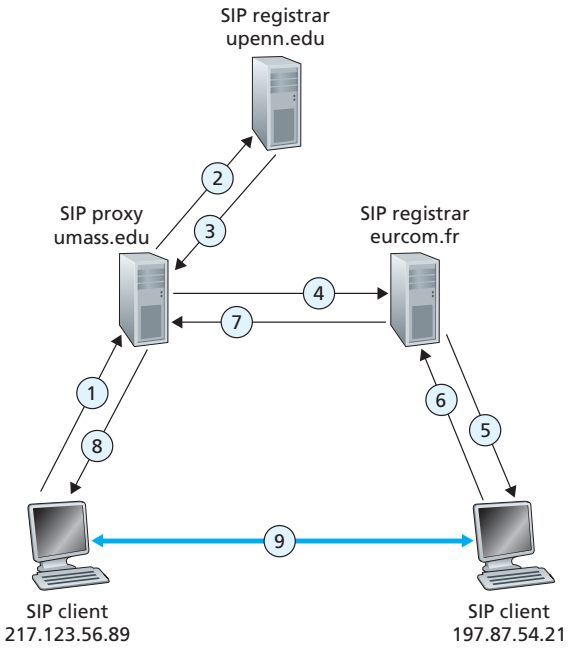

図 1　SIP プロキシとレジストラ

Content-Type 行では SIP メッセージの内容を記述するのに利用されるフォーマットが記してあり，Content-Length 行には内容の長さ (byte) を表示している．その後，メッセージの内容が記述される．

1.2.4　SIP プロキシ

IP アドレスが DHCP で動的に割り当てられることが多く，ボブは複数の端末について，IP アドレスを所持しているため，アリスの SIP 機器がボブの IP アドレスを知っているという仮定は非現実的である．アリスがボブの Email アドレスだけ知っていると仮定する. このアドレスを元に，ボブが利用中の機器の IP アドレスを特定するためにメッセージを SIP プロキシに送信する．その後プロキシはボブの Email アドレスで利用中の機器の IP アドレスを含む SIP アドレスを返信する．プロキシからの返答は発信者によって異なる．

続いて，プロキシがボブの現在の IP アドレスを決定する際に SIP の機器の一つである SIP レジストラについて述べる必要がある．全ての SIP ユーザはレジストラを保持しており，ユーザ起動した SIP アプリケーションが SIP 登録メッセージをレジストラに送信し、レジストラに現在の IP アドレスを通知する。ボブが新しい SIP アドレスに切り替えるたびに登録メッセージを送信し，同じデバイスに長期間止まっているときには，一定時間ごとに更新メッセージを送信するようにしている．多くの場合、SIP プロキシと SIP レジストラは同じホスト上で実行される．

例として，図 1 を考える．送信側を jim@umass.edu, 受信側を keith@upenn.edu とし，VoIP 接続を行う．図 1 において，まずジムが招待メッセージを SIP プロキシ umass.edu に送信する (①). その後，プロキシが upenn.edu で DNS ルックアップ (DNS を用いて、ドメイン名やホスト名から IP アドレスを、あるいはその逆を調べること) を行い，メッセージをレジストラに送信する (②). しかし受信側がもはや keith@upenn.edu に登録されていないので，レジストラはリダイレクトを行い，keith@eurecom.fr を試すよう指示する (③)．続いて, プロキシは SIP レジストラ eurecom.fr に INVITE メッセージを送信する (④)．eurecom レジストラは keith@eurecom.fr の IP アドレスを知っているため招待メッセージをホスト 197.87.54.21 に転送する (⑤). その後応答を返信し (⑥ ~ ⑧)，メディアを送信する (⑨)．このようにして SIP は、一般に通話を開始および終了するためのプロトコルとしての役割を担っている．

2　マルチメディアのためのネットワーク

ここまではアプリケーションレベルでマルチメディアを説明した．また，コンテンツ配信ネットワークと，P2P オーバーレイネットワークについても学んだ．しかし，アプリケーションレベルだけでなくネットワークもマルチメディアの送受信を支援する機構を保持している. ベストエフォートネットワーク，サービスを複数クラスへ分類，QOS 保証を説明する．

2.1　ベストエフォートネットワーク

先述のアプリケーションレベルでの機構は，パケット損失と過度のエンドツーエンド遅延がほとんど発生しない適切な規模のネットワークで正常に使用できる．また，輻輳が起こるたびに遅延，ジッタ，損失が起こる．この解決のためには単純にお金でリンク容量を拡大し，リソースの競合を回避する必要がある．これにより輻輳が完全に起こらなくなり，マルチメディアアプリケーションが完璧に機能する．ただし一定レベルのエンドツーエンド通信を実現するためにはネットワークトポロジの設計を考慮する必要がある．これは複雑な問題でこの教科書の範疇を超えている．ただし 2 つのエンドポイント間のパフォーマンスを予測し，要件を満たすことに下記の問題に対処する必要がある．

- モデルはネットワークに到達し，アプリケーションを起動するユーザレベルなどのコールレベルと生成するパケットなどのパケットレベルの両方からの指定，またワークロード (コンピュータにかかっている負荷の大きさ、実行中のソフトウェアによって処理能力が占有される度合い) が時間に応じて変化していくこと
- 会話式マルチメディアアプリケーションなどの遅延に敏感なものに対して，エンドツーエンド間の遅延が許容可能の遅延よりも短くなるようにするなどのよりよく定義された要求
- パフォーマンスを予測するための能力，および全てのユーザ要件を満たすための最小コスト帯域幅割り当てを見つける手段

技術的な観点からベストエフォート型のインターネットでマルチメディアをサポートできる．しかし，そうしない理由は経済的かつ組織的理由である．前者はユーザが ISP にお金を払うことを厭わないかを，後者はより問題で通信に必要な複数の ISP が (主に収益分配の観点から) 協力するかどうかを考えればわかることである．

2.2　サービスを複数クラスへ分類

今日のベストエフォート型インターネットにおいてトラフィックをクラス別に分類し，ISP が異なる品質のサービスを提供することが必要である．例えば HTTP や Email よりも VoIP や電話会議トラフィックに対しより高品質のサービスを提供する，コストを費やした顧客に対して高品質のサービスを提供するなどである．図 2 は，2 つのアプリケーションパケットフローが 1 つの LAN 上のホスト H1 と H2 で発生し、別の LAN 上のホスト H3 と H4 を宛先とする単純なネットワークシナリオを示す．2 つの LAN 上のルータは 1.5Mbps リンクで接続されていて，LAN の速度が 1.5Mbps を大幅に上回ると仮定し，ルータ R1 の出力キューに焦点を当てる．ここで H1 と H2 の合計送信レートが 1.5Mbps を超えるとパケット遅延と損失が発生する．さらに，1Mbps のオーディオアプリケーションと H2-H4 間の HTTPWeb ブラウジングアプリケーションと R1-R2 間の 1.5Mbps リンクを共有しているとする．この時に，分類を実現するための条件を 3 つあげる．

まず，ベストエフォート型インターネットでは，オーディオパケットと HTTP パケットは R1 の出力キューで混合され，FIFO で送信される．これにより，R1 でのバッファオーバフローによりオーディオパケットの損失や遅延が予測される．対して，HTTP の Web ブラウジングには時間的制約がないため，オーディオパケットがキューに入っていた際はこれを優先し専用リンクとし，入っていない場合は HTTP パケットを送信する必要がある．つまり，マーキング (異なるクラスのパケットを分類すること) が必要である．これが一つ目の条件である．

続いて，オーディオアプリケーションが 1.0Mbps のレートでパ

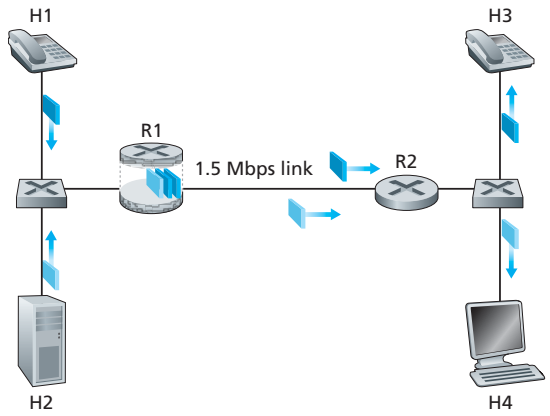

図 2　シンプルなネットワーク

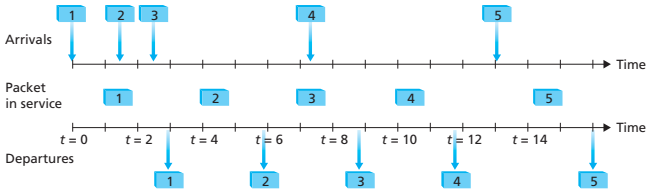

図 3　FIFO キュー

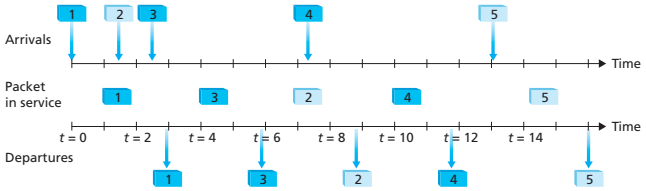

図 4　優先度付キュー

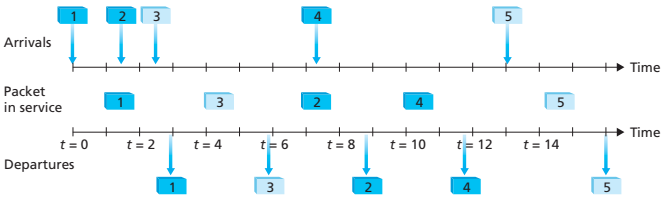

図 5　ラウンドロビン

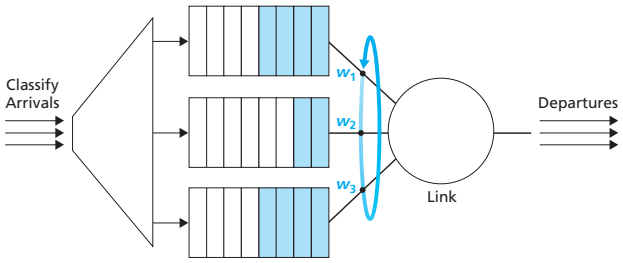

図 6　重み均等化キューイング (WFQ)

ケット送信をすれば，HTTP パケットは平均で 0.5Mbps の伝送が可能である．しかしながら，1.5Mbps 以上のレートであった時，HTTP パケットは枯渇してしまうことを考えると，リンクの帯域幅を占有し，セッションを飢餓状態にする可能性がある．つまり，あるクラスのトラフィックを他クラスのトラフィックから保護できるようにトラフィック分離を行う必要がある．

最後に，オーディオクラスに 1.0Mbps の割り当てが行われ，HTTP クラスに 1Mbps を割り当てる．ここで問題なのはオーディオパケットの生成を止めた時，HTTP パケットの送信は 0.5Mbps を超えて送信できない点である．つまり，クラスなどの分離を提供しながら，効率的にリソースを使用できるようにする必要がある．

2.3　スケジューリング

2.3.1　FIFO キューイング

リンクが現在，別のパケットの送信でビジー状態であるとき，リンクの出力キューに到着したパケットは送信を待つ．到着したパケットの保持に十分なバッファがない時，キューはそのパケットを失うか，他のパケットをキューから削除するか決定する．これをパケット廃棄と呼び，今回の説明ではパケット廃棄は無視する．FIFO では先に入ったパケットから処理を行う．図 3 にて，番号はパケットが到着した順序を表し，Packet in service では送信されている状態を表す．パケット 4 の出発後，パケット 5 が到達するまで，リンクは待機状態である．

2.3.2　優先度付キューイング

出力リンクに到着したパケットを優先クラスに分類するものである．優先クラスはパケットヘッダ上のマーキングや送信元，宛先 IP アドレス，宛先ポートなどの基準によって決定する．各優先クラスは FIFO 形式で挿入される．図 4 ではパケット 1,3,4 は高優先度クラスに分類，2,5 は低優先度クラスに分類されている．パケット 1 の送信中にパケット 2，3 が到着しそれぞれのキューに格納される．その後，優先度の高いパケット 3 が先に送信される．この方式では，一度はじまったパケットの送信は中断されない．

2.3.3　ラウンドロビン

この方式では，優先度付キューイングと同じように，クラスに分類する．しかし，クラス間でサービスを切り替える．これによりパケットがキューに入っているときはリンクが待機状態にならないようにしている．図 5 において，パケット 1，2，4 はクラス 1 に属し，パケット 3，5 は 2 番目のクラスに属す．クラス 1 からパケット 1 の送信後，サービスをクラス 2 に移し，パケット 3 を送信する．送信後，パケット 1 に写り，パケット 2 を送信する．パケット 2 の送信中にパケット 4 がキューに入った唯一のパケットであるから，次にパケット 4 を送信する．

2.3.4　重み均等化キューイング (WFQ)

続いて，図 6 にて示される，WFQ について説明する．これは各クラス i に重み w_i が割り当てられ，クラス i に $w_i/\sum w_j$ を受信することが保証され，伝送速度 R のリンクではクラス i は常に少なくとも $R \cdot w_i/\sum w_j$ のスループットが保証される．

2.4　ポリシング

これは制限レートを超えたパケットはドロップさせる，またはパケットの優先度の変更を行うことで通信速度を制限させる手法である．このポリシングにおいて，パケットレートの 3 つの側面を監視する必要がある．1 つ目は長期平均レートであり，これは時間間隔あたりのパケット数である．特に重要な問題はポリシングの期間である．なぜなら，平均レートとして 1 秒あたり 100 パケットの制限は 1 分あたり 6000 パケットの制限よりも制限が強いからである．2 つ目はピークレートであり，より短い時間にわたって送信できるパケットの最大数を表す．上記の例の後者の場合では 1 秒あたり 1000 パケットを送ることができるが，ピークレートを 800 パケットに制限することで通信速度を抑えることができる．3 つ目がバーストサイズである．これは，極端に短い時間間隔でネットワークに送信できる最大パケット数である．下記で説明するリーキーバケットは以上の 3 つの側面についてポリシングを行うものである．図 7 に示すように，リーキーバケットは最大 b 個のトークンを保持できるバケットで構成されていて，毎秒 r 個のトークンが生成され，バケットが b 個未満のトークンを保持している場合は生成されたトークンが追加され，それ以外は無視される．バケットには最大で b 個のトークンが存在する可能性があるため，最大バーストサイズは b パケットである，さらにネットワークに入ることができる最大パケット数はトークン生成率 r，時間間隔を t とすると，最大パケット数は $rt + b$ である．これにより，トークン

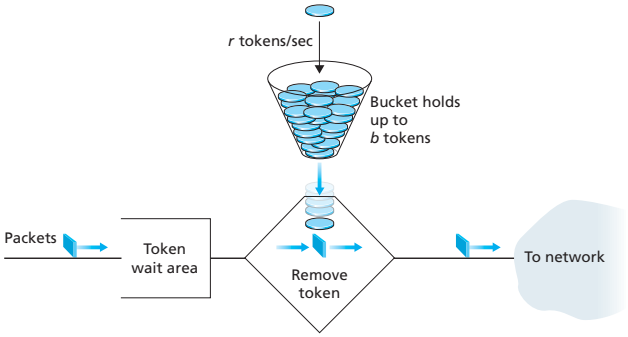

図 7　リーキーバケット

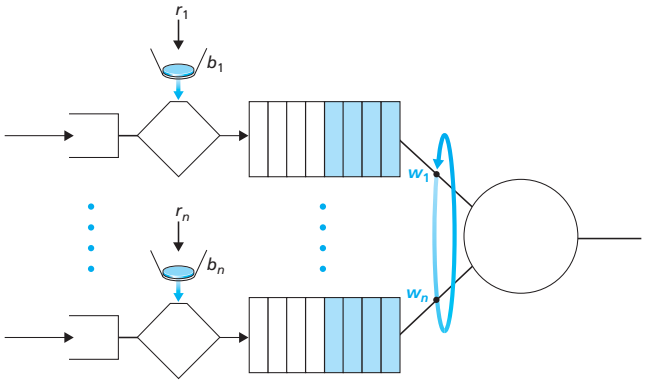

図 8　リーキーバケット + WFQ

生成率は長期平均レートを制限するように働く．2 つのリーキーバケットを利用することでフローのピークレートを制限することができるようだが，言及されていなかったので省略する．これと WFQ を組み合わせることで最大遅延の算出が可能である．これは図 8 にて示される．WFQ の仕組みより，各フロー i は少なくとも $\cdot w_i / \sum w_j$ に等しい帯域幅の割合を受け取ることが保証される．フロー 1 に着目すると $R \cdot w_1 / \sum w_j$ のレートで b_1 パケットは送信される．よって最大遅延は $\frac{b_1}{R \cdot w_1 / \sum w_j}$ である．

2.5　Diffserv

Diffserv はインターネット内の様々なトラフィッククラスを様々な方法でスケーラブルに処理する機能を提供する．何百万もの同時の送信元，宛先トラフィックがルータに存在する可能性があるという事実からスケーラビリティの必要性があげられる．図 9 に示す単純なネットワークについて，このアーキテクチャは 2 セットの要素から成る．

まず，主にパケット分類とトラフィック調整であるエッジ機能について説明する．トラフィックが最初に通るルータでパケットがマークされる．具体的には IPv4 または IPv6 パケットヘッダの Differentiated Service(DS) フィールドがある値に設定される．図 9 で H1 から H3 に送信されるパケット (以後パケット A) は R1 でマーキングされ，H2 から H4 に送信されるパケット (以後パケット B) は R2 でマーキングされる．これにより，異なるクラスのトラフィックを識別する．

続いて，転送を担うコア機能を説明する．DS フィールドにマークがついたパケットがルータに到着するとパケットのクラスに関連づけられている Per-Hop Behavior(PHB) にしたがって次のホップに転送される．これはルータによって実行されるホップごとの動作でを示す．また，アーキテクチャの原則として，パケットが属するトラフィックのクラスのみに基づいている．したがって図 9 でパケット A とパケット B が同じマーキングを受信した場合，パケットの送信元を区別せず集約して扱う．例えばパケット A,B を R3 が R4 に送信するとき，それぞれを区別しない．これによりスケーラブルになっている．

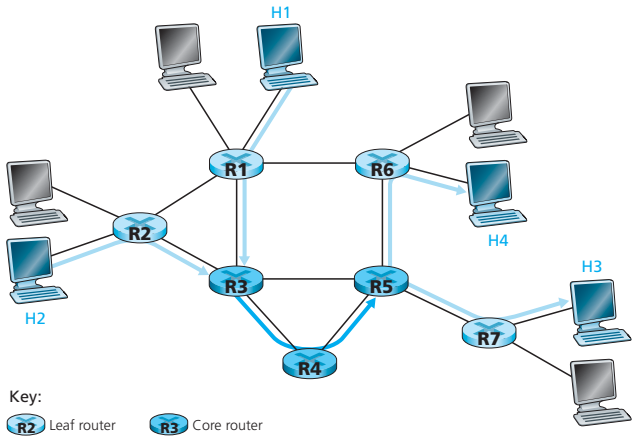

図 9　Diffserv ネットワークの例

2.6　Quality Of Service (QOS)

個々の接続に対して強いサービス保証が提供される場合に，さらに追加のネットワーク機構とプロトコルが必要になる理由をこの節では示す．2 つの 1Mbps オーディオアプリケーションが 1.5Mbps リンクを介してパケットを送信するして，2 つのフローの合計データレート (2 Mbps) がリンク容量を超えてしまう．この時 2 つのアプリケーションが帯域幅を等しく共有すると，各アプリケーションは送信パケットの 25 %を失い，これは容認できないほど低い QoS であるため、両方のオーディオアプリケーションはまったく使用できない．これを改善するため，十分なリソースが利用可能でないときは QOS を満たしているときはフローを許可し，満たしていないときはブロックしている必要がある．コール認証という機構では利用できるリソースの是非によってコールをブロックおよび実行の選択ができる．コールが目的の QoS を満たすために必要なリソース（リンク帯域幅、バッファ）を持つことを保証する唯一の方法は，それらのリソースをコールに明示的に割り当てることである．このリソースの予約により，他のすべてのコールの要求に関係なく，コールはその期間を通じてこれらのリソースにオンデマンドでアクセスできる．コールがリンク帯域幅の x Mbps の保証を予約して受信し、x より大きいレートで送信しない場合、コールは無損失および無遅延で受信ができる．コール設定シグナリングは，ルータでセッションのホップごとの QoS 要件を違反なく満たすのに十分なリソースがあるかどうかを判断する．これはプロトコルによって実現されている．