

新米エンジニアのための

初歩の

# インターネット技術

《第8回 OSPFの話 その2》

浅羽 登志也

asaba@ij.ad.jp

株式会社インターネットイニシアティブ

最適経路を発見するためのプロトコルであるOSPF (Open Shortest Path First) では、ネットワーク上のすべてのルーターは同じ内容のトポロジー情報を持っていなければなりません。今回は、常に変化し続けるトポロジー情報がどのようなしくみで更新されるのかを解説してみたいと思います。

はじめに

今回はOSPFで最適経路の計算に用いられるSPF (Shortest Path First) アルゴリズムについて解説した。SPF アルゴリズムで最適経路の計算を行う際に重要になるのは、すべてのルーターが同じトポロジーの情報を持っていなければならないという点であった。

これはたとえば、いくつかの電車を乗り継いで旅行をするような場合に、旅行の計画を立てたときに見ていた電車の路線図と、途中の各乗り継ぎ駅に掲示されている路線図が違っているような場合を想像すると分かりやすいかもしれない。途中の駅まで行ってみたらそこで乗り継ぐはずの路線がなかったり、最悪の場合には同じ駅と駅の間を何度も行ったり来たりするはめになり、いつまで経っても目的地にたどり着けなかったりする場合だってあるかもしれない。

ネットワーク上でこのようなことが起こると、せっかく各ルーターが一所懸命そのルーターから各目的地への最適経路を計算していたにしても、それぞれが参照していたトポロジーの情報が異なっていたがために、ちっとも最適経路をたどることができなかったり、最悪の場合にはルーティンググループが起り、永遠に目的地にパケットが届かなかったりするわけである。

OSPF では、各ルーターは、ネットワークのトポロジーの情報をデータベースとして持っており、それに基づいて各々独立にSPFの計算を行う。したがってすべてのルーターのデータベースの情報をいかにup-to-dateに保つかが重要となるのである。

## OSPFでの隣組関係

ネットワークを作った人間の管理者様ならば、おそらく、最初にトポロジーの絵ぐらいいは書いているだろうから、全体のトポロジーを知っているのは当たり前だろう。

では、その管理者様に買いとられて遠くの町から何も分からずに連れて来られ、「おめえはここだ」とファンの音だけが聞こえる寒くて暗い部屋に押し込められた哀れなルーターたちはどうやって世間の様子を知ることができるのだろうか？

管理者様はなにやらごそごそというんなケーブルをつないで、「あんじょう頼むでえ」とだけ言い残して去って行ってしまった。ああ、このままこの暗い部屋で一生一人ぼっちで過ごすのだろうか？ ふとそんな不安な気持ちになる。

おや、しかし待てよ、管理者様のつないでいったケーブルの向こうからなにやら声が聞こえる。「HELLO」とか言っているぞ。

このケーブルの向こうには誰がいるのか？ じゃあこっちも「HELLO」と言ってみよう。なんだ、向こうにも同じような仲間がいるんじゃないか！

つまり、OSPFではまずHELLOメッセージを自分の持つすべてのリンクに対して送り出し、そのリンクの向こうにもルーターがいるのかどうかを調べる。もし双方向にHELLOでご挨拶ができれば、これでお互いはお友達ということになる(図1)。

OSPFでは、このお友達関係のことをneighborという。HELLOは定期的に繰り返し送り出され、リンクの向こうのお友達が元気であるかどうかを確認し合うのである。

複数のリンクを持っている場合には、それぞれのリンクの向こうにお友達がいる場合もある。リンクの向こうにお友達が1人もいない場合もある。1つのリンクの向こうにたくさんお友達がいる場合もある。1つのリンクの向こうにたくさんお友達がいる場

合には、どうやらその中にボスがいるらしい。ボスがいる場合は、他の者同士はあまり仲良くしていないようだ。みんながボスに報告して、他のやつらの情報はボスから聞いているようだ。とすると、とりあえずボスにはご挨拶しておかないといけなかな……。

専用回線などのリンクの場合は、リンクの向こうには1台しかルーターがない。このようなリンクは、point-to-pointリンクと呼ばれ、対向のルーターがneighborとなる。

一方、イーサネットなどのようなマルチアクセスの媒体につながっているリンクの向こうには、たくさんルーターがある場合がある。このような場合には、このたくさんのルーターを代表するボスルーターがいる。このボスルーターのことをOSPFの用語ではdesignated router (DR)という。このようなリンクはtransit networkにつながっているといい、ボスルーター(DR)がneighborとなる。またボスに万が一の事故があった場合のための“ No.2 ”もいる。この“ No.2 ”のことをOSPF用語ではback-up designated router (BDR)という。

また、イーサネットにつながっていても、その向こうにはルーターがない場合もある。このようなリンクはstub network (stub area というのもあるので、混同しな

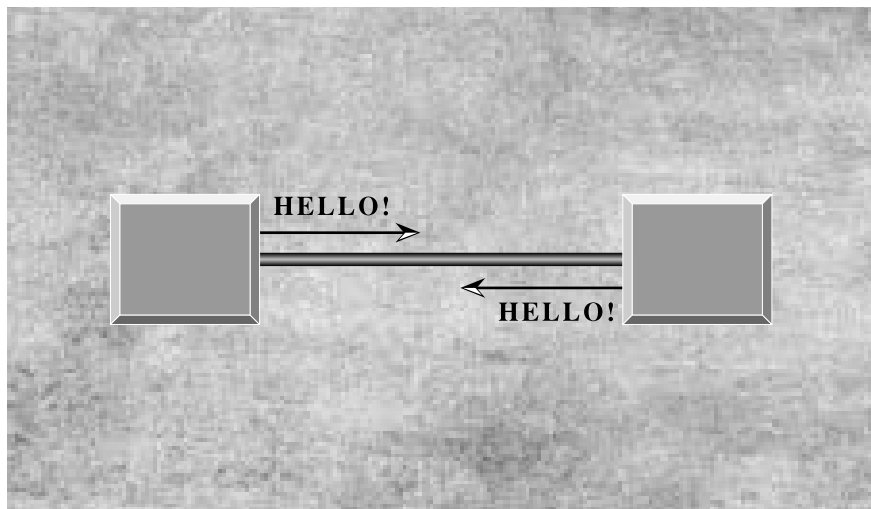


図1 HELLOメッセージ

いように)につながっているといれ、このよ  
うなリンクにはneighborはいない(図2)。

### 最初の情報交換

お友達がリンクの向こうにいることが分  
かると、少し心強くなる。たぶんあいつの  
ほうが先にここに連れて来られたのだろう  
から、ここがどういう所なのかをいろいろ  
知っているかもしれないなあ。などと思っ  
ているうちに、なんか情報を送ってきたぞ!

というわけで、それぞれのリンクにneighbor  
が見つかると、そのneighborとの間で  
ネットワークトポロジーの情報を交換する。  
OSPFでは、ネットワークトポロジーの情報  
は、それぞれのルーターが持っているリンク  
の状態によって表現されている。

リンクの状態の情報には、それをアナウ  
ンスしているルーターのIDや、リンクのコ  
スト、リンクにつけられたネットワークアド  
レスやそのネットマスクなどが含まれてい  
る。各ルーターは、自分の持つすべてのリ  
ンクの状態の情報を他のルーターに知らせ  
なければならない。このリンクの状態の情  
報をLink State (LS) と呼び、LSを他の  
ルーターに知らせることをLink State Adver  
tisement (LSA) という。

ルーターがネットワークにつながると、最  
初に自分のneighborルーターとの間で各々  
が持つトポロジーのデータベースの内容を  
交換する。これを Database Exchange  
Process と呼び、これを行うためのメッセ  
ージをDatabase Description Message と呼  
ぶ。すべての隣組の間でデータベースの内

容を一致させるというプロセスを経ること  
によって、すべてのルーターで同じ内容の  
トポロジーの情報を持つことができるので  
ある。Database Description Message中  
には各々がデータベース中に持っている  
LSAを表す情報が含まれており、特定のLS  
を自分が持っていない、もしくは、古い情  
報しか持っていない場合には、Link State  
Requestを相手に送る。Link State  
Requestを受け取ったルーターはLink State  
Updateメッセージを送り返し、要求された  
LSを相手に伝える。

たとえば、図3のようなネットワークを考  
えてみよう。図でR1とR2はルーターを表  
し、L1とL2はリンクを表す。リンクの横に  
書かれた数字は、そのリンクに対するコス  
トである。またルーターによってつながれて

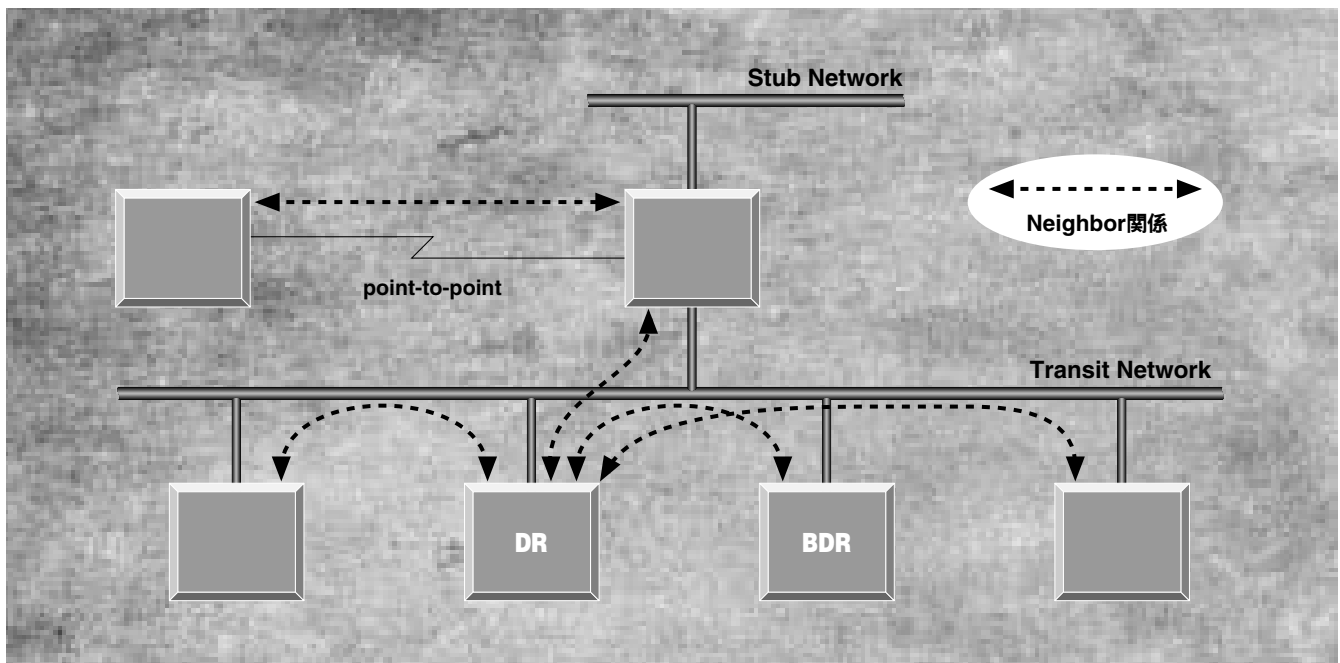


図2 リンクの種類とneighbor関係

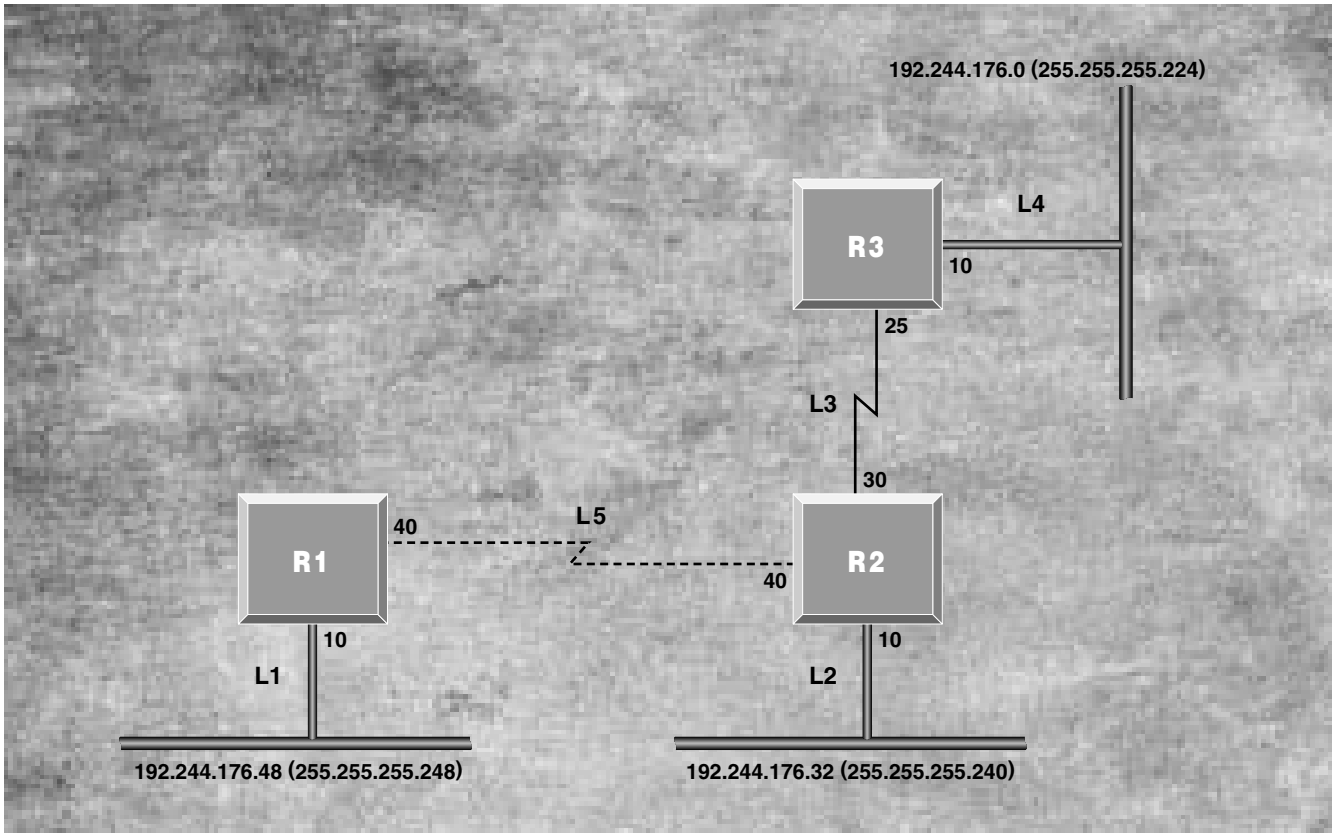


図3 ルーター間におけるトポロジー情報の交換

いるネットワークの下に書かれているアドレスは、そのネットワークのIPアドレスで、括弧の中はネットマスクである。

ここで、ルーターR1とルーターR2を新たなリンクL5を用いてつないだとする。ルーターR1は2つのリンクを持っており、接続された時点では、R1が持つトポロジーデータベースはR1のLSのみを持ち、そこには以下のような情報が含まれている。

R1のLS

- リンクL1はコスト10でstub networkにつながれている。stub networkは、ネット

ワークアドレス192.244.176.48でネットマスクが255.255.255.248である。

- リンクL5はコスト40でpoint-to-pointリンクにつながれており、neighborのルーターIDはR2である。

また、R2は3つのリンクを持っており、接続された時点では、R2とR3のLSをトポロジーデータベースに持っている。それぞれの内容は、次のとおりである。

R2のLS

- リンクL2はコスト10でstub networkにつ

ながれている。stub networkは、ネットワークアドレス192.244.176.32でネットマスクが255.255.255.240である。

- リンクL3はコスト30でpoint-to-pointリンクにつながれており、neighborのルーターIDはR3である。
- リンクL5はコスト40でpoint-to-pointリンクにつながれており、neighborのルーターIDはR1である。

R3のLS

- リンクL4はコスト10でstub networkにつながれている。stub networkは、ネット

ワークアドレス192.244.176.0でネットワークアドレスが255.255.255.224である。

リンクL3はコスト25でpoint-to-pointリンクにつながれており、neighborのルーターIDはR2である。

R1とR2のデータベースの内容が交換されると、R1とR2は、ネットワーク全体のLSをトポロジーデータベース中に持つことになる。

#### 情報の更新

さて、ではR3はどうやってR1がR2につながれたことを知るのだろうか？

R1のLSAを受け取ったR2は、受け取った情報を元に、自分の持つトポロジーデータベースを変更する。自分のデータベースを変更すると、R2はそのLSAを受け取ったインターフェイス以外のすべてのインターフェイスにLSAを転送する。R2はR1のLSAをL5から受け取っているため、L3に対してR1のLSAを転送する。

すべてのLSAには、シーケンス番号が付けられている。転送されたLSAを受け取った各ルーターでは、以前に受け取ってデータベース中に保持しているLSAのシーケンス番号と新しいLSAのシーケンス番号を比較する。比較の結果、今受け取ったLSAのほうが新しいと分かると、データベースを更新し、さらにそのLSAを、受け取ったインターフェイス以外のすべてのインターフェイスに対してさらに転送する。このようにして、新しいLSAはネットワークのすみずみまで転送されていくのである。新しい情報が洪水のようにネットワーク全体に流れていくので、このプロセスをflooding

processと呼ぶ。

floodingは、LSが更新されたときに、ただちにその更新内容をネットワーク全体に伝えるために行われる。この例のように新たなリンクが付け加えられた場合だけではなく、何らかの理由でリンクが切れた場合にも行われる。これによって、ネットワークのトポロジーの変化がいち早くネットワーク全体に伝えられるのである。

#### エリア

全体のネットワークを構成するルーターの台数が増え、その間を結ぶリンクの数も増えていくと、ルーター間で交換されるLSAの数も増え、SPFの計算量も増えていくことになる。そこで重要になるのがエリアという概念である。

すなわち、ネットワークをいくつかの部分(エリア)に分けて、エリアの内部のリンクに関してのみ完全な情報を持ち、エリアの外部に関してはサマリー情報のみでSPFの計算をするのである。エリアの外部のサマリー情報は、エリアの境界にあるルーターがエリア内部にアナウンスする。これによりルーターがデータベースに持つ情報量やSPFの計算量を減らすことができる。

OSPFでは、OSPFを用いて運用されるネットワーク全体をautonomous system(AS)と呼ぶ。ASは、必ずバックボーンエリアを持つ。ASがエリアに分割されていない場合には、全体がバックボーンエリアとなる。その他のすべてのエリアはバックボーンエリアに隣接していなければならない。エリア間の通信は必ずバックボーンエリアを経由して行われなければならない。

たとえば図4のように、ネットワークをバック

ボーン、エリア1、エリア2、エリア3の4つのエリアに分けたとする。ここで、R2、R3、R4、R5、R6のようなエリアの境界にあるルーターは、Area Border Routerと呼ばれ、そのルーターに接続されているすべてのエリアに関する完全な情報を持っている。また、R4のように、ASの外部に接続されているルーターはAutonomous System Boundary Routerと呼ばれる。また、エリア3のようにArea Border Routerを1つしか持たないエリアは、Stub Areaと呼ばれる。

このとき、たとえばエリア1内部のルーターはエリア1内部に関してはすべての情報を持つが、エリア外のN3、N4、N5、N6、N7などのネットワークに関しては、Area Border RouterであるR2、R3の各々からそれらのネットワークに到達するためのコストの情報のみを持てばよく、外部のトポロジーの情報は持たなくてよい。また、エリア3のようなStub Areaでは、外部のネットワークの情報すら持つ必要がない。エリア内部以外のネットワークに関してはすべてR6を経由して接続されているので、外部に関してはR6に向かうデフォルトの経路だけを持っていればよいからである。

#### VLSMのサポート

先の節で述べたように、OSPFのLSAには、リンクのネットワークアドレスとそのマスクの情報が含まれている。これによってVLSM(Variable Length Subnet Mask)を用いて構築されたネットワーク上でのルーティングをうまく行うことができる。実際、このVLSMのサポートが、ルーティングプロトコルとして、かくも複雑でルーターのCPUを使いまくるOSPFを採用する大きな

理由の1つになっている。最近のCIDR (Classless Inter Domain Routing) の環境では、ルーティングプロトコルがVLSMをサポートしていないと非常に不便な場合が多いのである。

OSPF 以外にVLSM をサポートしているプロトコルとして、RIP2がある。これはRIPを拡張して、ネットマスクの情報も運べるようにしたものである。プロトコルの簡単さとしてはOSPFの比ではない。

つい先日、とあるダイアルアップ系のルーターメーカーの人とお話したのだが、そのときも私の「うちはVLSMの環境だから、OSPFをサポートしてないと使えないよう～」という言葉に対して、「でもOSPFを実装しようとする、数か月かかりますぜ。多分年内には終わらないでしょうなあ……。

それよりも旦那、RIP2なら1週間で出せませ。どないでっしゃろ？」という返事が帰ってきた。私としてはVLSMが使えればなんでもよいのだが、いかんせんバックボーンのルーターがRIP2をサポートしていない。なぜ巷の人々はRIP2をあまりサポートしたがるのだろうか？

**おわりに**

OSPFは、従来用いられていたRIPに比べると、はるかに「面白い」プロトコルである。かく言う私も、「こんなところで大学のORの授業で学んだ<sup>ダイクストラ</sup>dijkstraアルゴリズムが出てくるなんて！」と思ったものである。アルゴリズムとして見るかぎり、RIPなどは比べ物にならないほど面白い。情報工学

科卒が企業の人事部からは理系とは認められないなどというとてもお寒い昨今、嬉しいかぎりではないか！（ちなみに私は情報工学科卒ではない :-P）

しかし、正直言ってOSPFなんて書くんじゃなかった。書き始めたらいくら誌面があっても足りない。ふと、「この雑誌には技術のための技術解説なんていらんだよ」という言葉を残して私の担当から去っていった、某誌副編集長の言葉を思い出す（中島さんごめんさい！）。

さて、後置きが長くなってもなんなので、この辺で止めるとしよう。次号以降では、もうちょっと広い視点で、プロバイダー間のルーティングの現状をお話できればいいなあなどと、またあとで後悔しそうなことを言い残して今回は終わりにしよう。

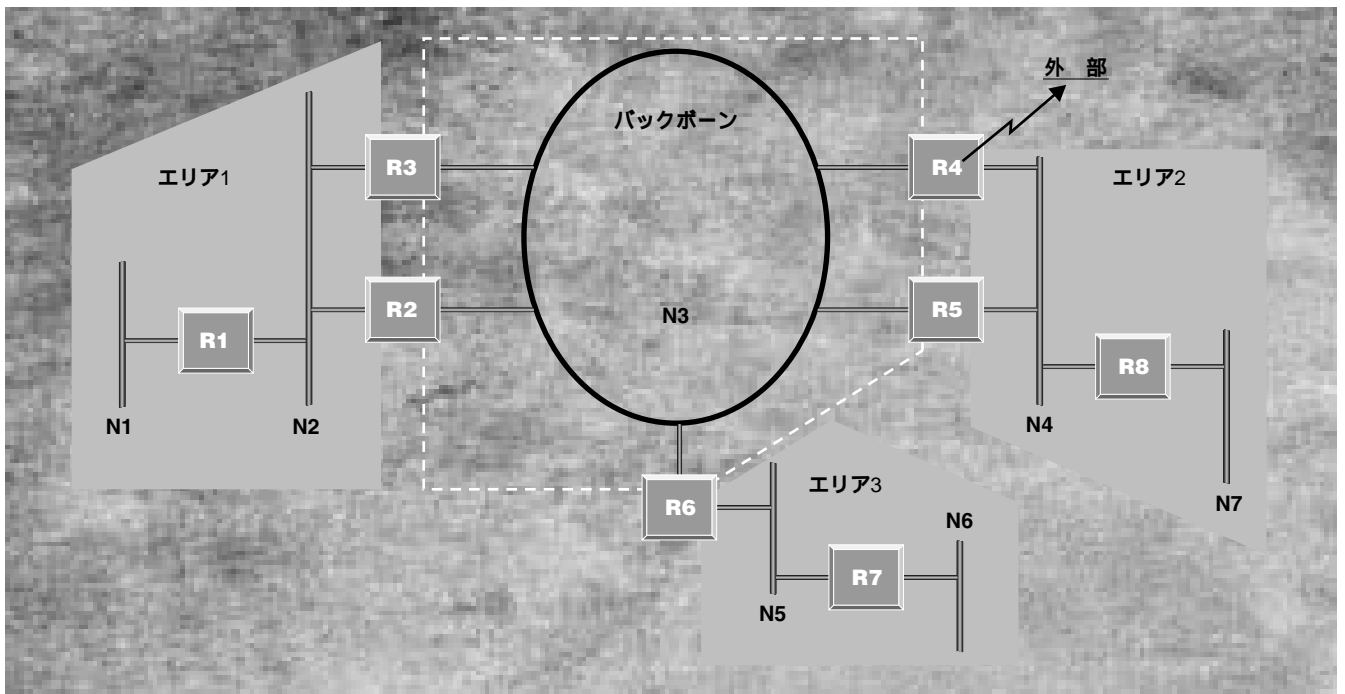


図4 ネットワークにおけるエリアの概念



## [インターネットマガジン バックナンバーアーカイブ] ご利用上の注意

このPDFファイルは、株式会社インプレスR&D(株式会社インプレスから分割)が1994年～2006年まで発行した月刊誌『インターネットマガジン』の誌面をPDF化し、「インターネットマガジン バックナンバーアーカイブ」として以下のウェブサイト「All-in-One INTERNET magazine 2.0」で公開しているものです。

<http://i.impressRD.jp/bn>

このファイルをご利用いただくにあたり、下記の注意事項を必ずお読みください。

- 記載されている内容(技術解説、URL、団体・企業名、商品名、価格、プレゼント募集、アンケートなど)は発行当時のものです。
- 収録されている内容は著作権法上の保護を受けています。著作権はそれぞれの記事の著作者(執筆者、写真の撮影者、イラストの作成者、編集部など)が保持しています。
- 著作者から許諾が得られなかった著作物は収録されていない場合があります。
- このファイルやその内容を改変したり、商用を目的として再利用することはできません。あくまで個人や企業の非商用利用での閲覧、複製、送信に限られます。
- 収録されている内容を何らかの媒体に引用としてご利用する際は、出典として媒体名および月号、該当ページ番号、発行元(株式会社インプレス R&D)、コピーライトなどの情報をご明記ください。
- オリジナルの雑誌の発行時点では、株式会社インプレス R&D(当時は株式会社インプレス)と著作権者は内容が正確なものであるように最大限に努めましたが、すべての情報が完全に正確であることは保証できません。このファイルの内容に起因する直接のおよび間接的な損害に対して、一切の責任を負いません。お客様個人の責任においてご利用ください。

このファイルに関するお問い合わせ先

**株式会社インプレスR&D**

All-in-One INTERNET magazine 編集部

[im-info@impress.co.jp](mailto:im-info@impress.co.jp)