ATNP/WG2/
WP/
22 April 1999

AERONAUTICAL TELECOMMUNICATIONS NETWORK PANEL

WORKING GROUP TWO

Brussels 22.4.96-277.4.96
INFORMATION PAPER

Route Aggregation - Proposed Guidance Material

Presented By Henk Hof

Prepared by Tony Whyman

SUMMARY

This is a draft of possible Guidance Material in support of Route Aggregation. The problem has been to describe a technically detailed subject in a short, concise and understandable way. This draft starts by adopting an informal "light" style before getting increasingly technical. Feedback is sought on whether this approach is appropriate, and whether the Guidance Material should be developed by continuing with the informal style or changing to a more formal approach. More generally comments are soliticted on whether this draft Guidance Material is going in the right direction to explain the subject.

DOCUMENT CONTROL LOG

SECTION	DATE	REV. NO.	REASON FOR CHANGE OR REFERENCE TO CHANGE
	22-apr	Issue 1.0	

TABLE OF CONTENTS

1. Introduction

This paper provides draft Guidance Material for Route Aggregation and Route Information Reduction. An original version was presented at the Brisbane WG2 meeting. Feedback from this meeting suggested that Guidance Material was necessary that assumed less detailed knowledge of network routing on behalf of its readers. This paper is the result.

A completely new information introduction to the subject has been written, in order to introduce the basic concepts with reference to everyday subjects, and the original contents of the paper edited to become a formal discussion of each feature that makes up Route Aggregation and its support functions. Finally, the section on implementation considerations has been updated to take into account recent work in this area and now includes an implementation model for the Phase Three Route Decision Process including Route Aggregation.

2. An Informal Introduction to Route Aggregation

2.1 What is Route Aggregation?

Route Aggregation is one of those subjects guaranteed to empty a room. Far from being the kind of subject that can be used to break the ice at parties, it seems to have as much social value as combining religion and politics in the same sentence. However, it is very relevant to the building of big Internetworks (and I mean big, not the handful of Ethernets that MIS people seem to think means big).

So, Route Aggregation is worth knowing about. But what is it? How does it help us build a big Internet, moreover, why is it relevant to the ATN?

Well, look at the signpost alongside in Figure 2-1, and imagine being confronted with it at a road junction. If you are going to one of the big cities indicated on it, then you're in luck. It points you in the right direction. But, if you are not, what do you do? Complain to the person that erected it?

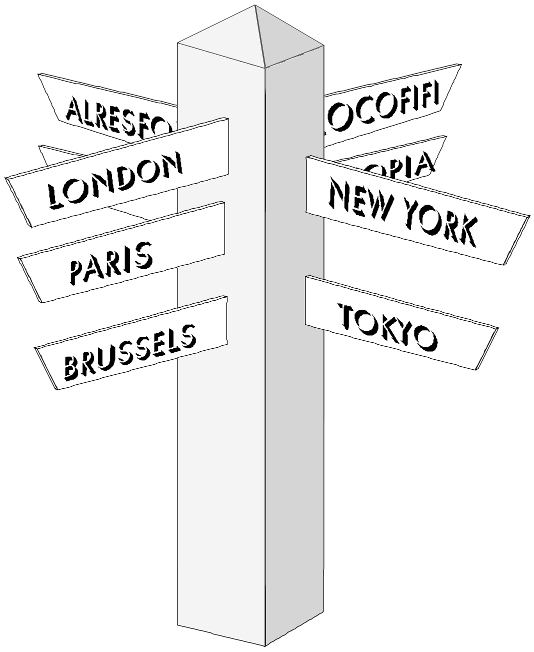

Figure 2-1 Signposting the Way

Perhaps you do. You want to go to Berlin, and you're the kind of person that complains strongly if things aren't right. The person responsible for the signpost, reacts to customer demand and adds a sign for Berlin. Off you go, a satisfied customer.

The same then happens for people wanting to go to Rome, Toulouse, Sydney, Singapore, Peking, Cape Town, Rio de Janeiro, Seattle, Moscow, Dublin, Brisbane, Winchester, Prague, Bristol, Athens, Anchorage, Stornaway, Oslo, St Petersberg, and so on, until there is no further room on the signpost to hang another sign. What does our poor signpost manager do now?

He could just erect a bigger signpost, but if he's bit cleverer, he may just realise that the problem is not one of insufficient signpost real estate, but really it's the granularity of information that is being provided. After all, London, Paris and Brussels are all in Europe (at least in a geographical sense), and hence could be replaced with a single sign indicating the direction to Europe, along with all the other cities and towns in Europe that are individually listed on the signpost.

In fact, this is a really bright idea, as it is not just the European cities that can be picked off in this way, but so can the Asian cities, the American ones, the African ones, and so on. Only those that really are local (i.e. on the same continent) need to be explicitly mentioned. What our bright signpost manager has realised is that his customers don't really need detailed information on the route for their individual destinations. There are only a few directions in which they can go anyway and, by labelling each direction with a suitable collective noun or group name, that properly and unambiguously describes what is reachable in that direction, the signpost's users will get all the information they need. After this exercise in information reduction, our signpost ended up much like that in Figure 2-2.

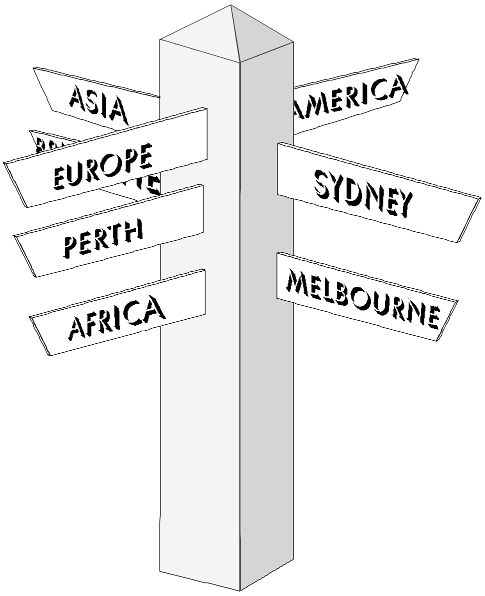

Figure 2-2 The Rationalised Signpost

This benefited the signpost's users, who didn't have to search through lots of different signs to find the one they wanted, and the signpost manager's company, as now, maintenance had been reduced to almost zero.

In fact, so much had the signpost management problem been reduced that when a new CEO was appointed shortly afterwards to the signpost management company, our signpost manager was fired in a downsizing exercise.

OK, so this is how road signs work, but is it really relevant to network routing?

Of course it is. Every router has an electronic signpost within it - its forwarding table. Each packet that it forwards, must find a sign telling it which direction to go in, otherwise it will be discarded. A Network Manager is akin to our signpost manager and must ensure that there is a suitable sign for every packet that needs to be routed.

From this you may conclude that routers adopt a principle similar to that illustrated in Figure 2-2, and minimise the amount of routing information by collecting routes together and signposting routes to appropriate group addresses. Unfortunate, you would not always be right in making such a conclusion.

For example, in the TCP/IP Internet, the routers implemented by the Internet Service Providers are much like the signpost in Figure 2-1. There's a sign for every network in the world and, when they run out of space to add new signs, the only answer is to get a bigger signpost. In fact, even this isn't true, because for most Internet Service Providers, there aren't any bigger signposts anymore.

The reason why this is so is twofold. Firstly, the network addresses used in the TCP/IP Internet are rather on the small side at only 32-bits long. Secondly, such addresses have traditionally been allocated to networks without any regard to network topology. The first problem is due to the limited horizons of the early Internet developers. No one at that time thought the Internet would grow so big and a 32-bit address was chosen for engineering reasons (i.e. efficient processing) rather than with future growth in mind. The second problem is simply due to any recognition that there needed to be a way (in network address terms) of forming the group addresses necessary to move away from the over-crowded signpost.

Network addresses are not simply names (like London or Paris) which can simply be said to also be described by a group name (e.g. Europe). Network Addresses are first of all names of systems on a network, but they are also parameters to a routing algorithm that is implemented by every router in an internetwork, and their role as parameters constrains the scope for allocating network addresses.

If our electronic signposts are to have group addresses put on them, instead of individual network addresses, then our routing algorithm must also be able to relate them to the packets that must flow in their direction in a manner that is simple and efficient.

This can be most readily achieved if similar network addresses imply that the addressed destinations are close together in the topology of the network. In this case and as illustrated in Figure 2-3, similar can be taken to imply that the addresses start off the same and only differ in the tail of the address. Indeed, how far down the address (seen as a bitstring) that the two addresses diverge, can be taken as a metric of closeness. A Group Address can then be simply the common part (or prefix) of the similar network addresses.

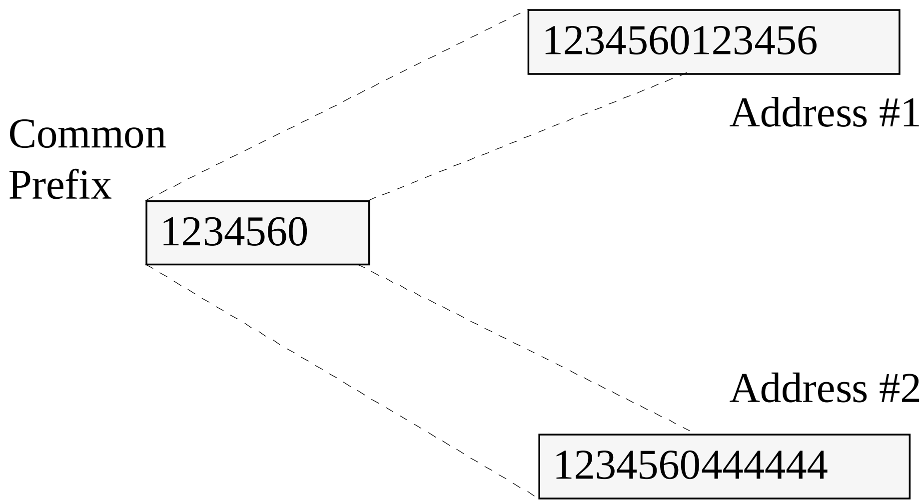

Figure 2-3 Similar Network Addresses

By allocating network addresses arbitrarily (at least on a per network basis), the early developers of the TCP/IP Internet have effectively frustrated its later growth. This is because Internet addresses are assigned on a network basis. Each network has a unique network number that is both the Group Address and prefix for every network address on that network. However, by assigning network numbers themselves arbitrarily (i.e. by not attempting to assign sequential numbers to networks that are adjacent to one another), the original TCP/IP Internet architects stopped any attempts to form Group Addresses for (e.g.) all the networks supported by a given Internet Service Provider. Hence, the electronic signposts operated by every Internet Service Provider have to have a sign for every assigned network number. The only real excuse for this is that the 32-bit address size gives only limited scope to manage the allocation of addresses in any other way.

Fortunately, for the ATN Internet, these problems were already known by the time that the ATN came to be developed and can thus be largely avoided.

The ATN specifies the use of the Connectionless Network Protocol (CLNP) instead of IP. This has the great advantage of large (variable length) addresses, and the ATN takes advantage of this to specify a 160 bit address format. Although it can be argued that such a long address is less efficient to process than a 32-bit address, 160 bits makes it much easier to ensure that similar network addresses are allocated to networks that are near each other in the ATN Internet, and can therefore be used to improve the overall routing efficiency. The electronic signposts in ATN Routers can therefore always be designed to use short Group Addresses rather than have a sign for every network in the ATN. They thus follow the signpost model of Figure 2-2.

Secondly, the ATN specifies the Inter-Domain Routing Protocol (IDRP) in support of routing information exchange. In the above analogy, IDRP would provide the information that the signpost manager uses to maintain each sign, and is essential for maintaining the Figure 2-2 signpost model. IDRP also introduces and supports the two central concepts to building large scaleable internetworks: Route Aggregation and Route Information Reduction. Formally:

- **Route Aggregation** is the process by which routes to individual destinations are brought together so that they may form the basis of a single sign on our electronic signpost. while

- **Route Information Reduction** is the process by which the network addresses that identify the individual destinations are recognised to be replaceable and replaced by a single “group address”.

These processes are directly analogous to the processes our signpost manager went through in reducing the number of signs he had to manage.

In use, Route Aggregation and Route Information Reduction are said to be policy driven. That is they do not come about spontaneously, but instead are the result of the application of specific routing policy rules laid down by a network adminstrator. To aid this process, and in particular the specification of rules for Route Information Reduction, IDRP also introduces the **Routing Domain Confederation (RDC)**. This is a containment boundary that is very useful for, amongst other things, defining the scope within which addresses appear similar.

The ATN has further extended these concepts to introduce **Route Merging**. This is a special case of Route Aggregation where routes available to the same destination, but via different directions and which are suitable for different types of traffic only, are aggregated. Special rules apply in this case.

2.2 ATN Network Addresses, Routing Domains and Routing

The “Achilles Heel” of the TCP/IP Internet is its 32-bit address size and the limited scope this gives for address management and allocation. It is clearly important that the ATN Internet learns from this and has a Network Address management and allocation strategy that permits and implements an addressing strategy that includes Group Addresses at each level of the ATN Routing Hierarchy.

The ATN Network Address (properly called an NSAP Address) is 160 bits long and, for the purposes of address administration has the syntax show in Figure 2-4, which assigns a meaning to each of the 20 octets that make up this NSAP Address. However, for the purposes of routing this figure should be ignored. ATN Routers, like all routers, do not know about the syntax of such an address, as this is for managing address allocation only. All the router sees is a bitstring that is the NSAP Address. Note that the reverse is not true. Those allocating NSAP Addresses must take account of how the routing strategy works. The ATN Addressing Plan can be implemented to give efficient routing. It can also be implemented unintelligently to give rise to the same problems found in the TCP/IP Internet.

ATN Routers are only permitted to make two assumptions about an NSAP Address. The first is that the last 8-bits are sub-addressing bits for use within the destination system, and hence only the first 152-bits are used for system addressing. The second is that the first 88-bits (the 11 octet Routing Domain Part) is used to identify a Routing Domain and the remaining 64-bits identify a system within a Routing Domain.

A **Routing Domain** is a collection of End Systems and routers under a single administration and is an important building block in a large worldwide internetwork. Essentially, a Routing Domain is permitted to be limited in size (i.e. Routing Domains do not have to be scaleable to unlimited size), and all of the systems within it operate within a common domain of trust. The routing strategy implemented within a Routing Domain is local and may be optimised for performance, whilst issues such as scaleability can be left to the inter-domain routing environment.

As the ATN Addressing Plan requires that all systems within a single Routing Domain have a common 88-bit prefix, this NSAP Address Prefix becomes an effective Group Address for all the NSAP Addresses allocated within the Routing Domain. The electronic signposts in routers outside of such a Routing Domain have only to include a single sign, pointing along the route to this Routing Domain, and may use the common NSAP Address prefix for all systems within the Routing Domain, as the Group Address on the “sign”. There is no need for these signposts to have a separate sign for each system within the Routing Domain.

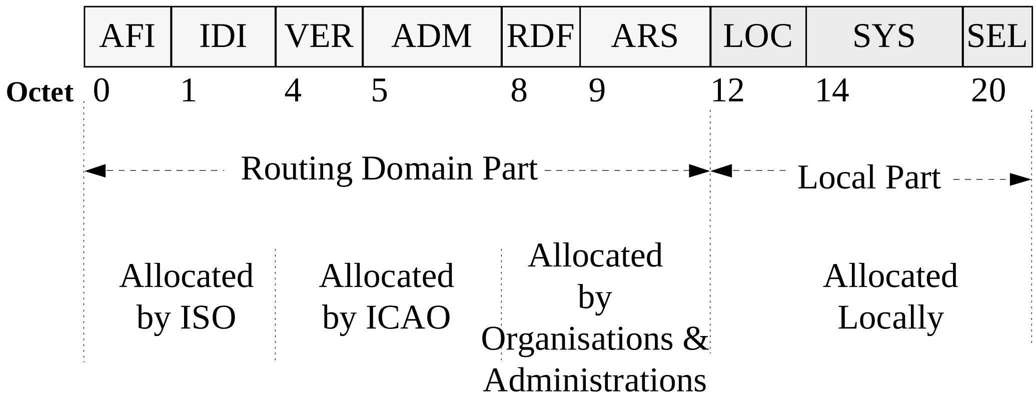

Figure 2-4 ATN NSAP Address Syntax

In fact, the Routing Domain concept allows an even greater simplification to take place. The electronic signposts that point the way to individual Routing Domains can be placed in specially designated Boundary Routers (often called Boundary Intermediate Systems, or BISs), at the edges of each Routing Domain. The routers wholly within an Routing Domain can then have a much simpler electronic signpost within them, pointing the way to each local destination (i.e. within the Routing Domain) and with one sign, labelled “All other Routes” to the nearest Boundary Router.

The problem of building big Internetworks and exchanging the routing information necessary to support routing in such internetworks, can therefore be constrained to the boundary routers. Indeed, that is how it is done in the ATN. ATN Routers which are boundary routers deal in so called “inter-domain” routes, where such a route is a path to another Routing Domain, identified by the common NSAP Address Prefix for all systems within it, and IDRP is used to support this inter-domain routing strategy.

2.3 Routes and Route Aggregation

Our signpost analogy is really only one part of the routing concept. As illustrated in Figure 2-5, signposts are just waypoints along a route between a starting point and a journey’s end and, formally, we define a route to be a combination of information that describes a path, and the NSAP Address that identifies the end point of the route. IDRP deals in such routes and allows Boundary Routers to keep each other informed about the routes that they offer.

Of course, IDRP’s routes are not to actual destination systems. They are to the Boundary Routers at the edge of the Routing Domain that contains the destination system, and the NSAP Address of the route’s end point is a Group Address - the common NSAP Address Prefix for all systems within that Routing Domain. Effectively, the Boundary Router has brought together the individual routes to each system within Routing Domain into a single route, and replaced all the individual NSAP Addresses with the appropriate single NSAP Address Prefix. We already know these two processes to be called Route Aggregation and Route Information Reduction, and these always occur implicitly, in a Boundary Router, before a route to such internal destinations is advertised to the Boundary Routers of other Routing Domains.

The question now arises as to whether there is any merit in carrying out Route Aggregation and Route Information Reduction at any other points in route distribution. The answer is a definite yes.

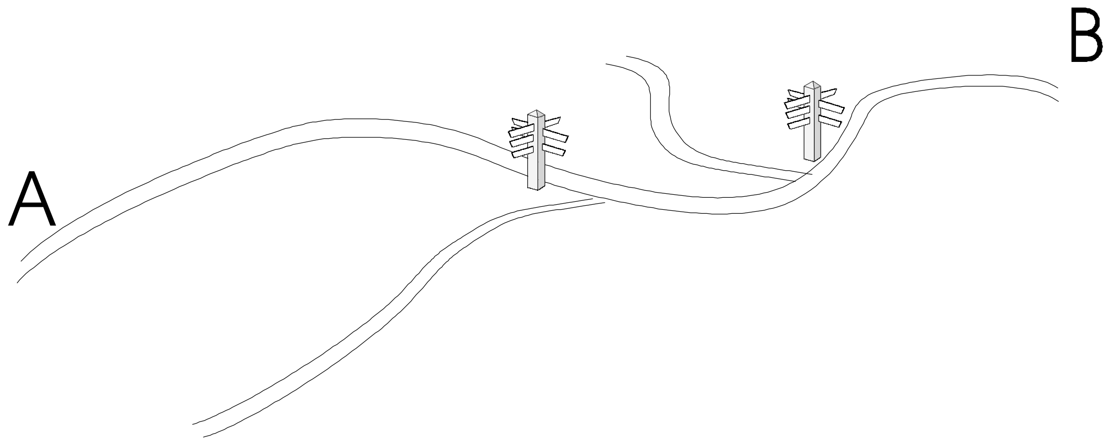

Figure 2-5 The Route - a path between A and B

Firstly, there is nothing magic about an 88-bit NSAP Address Prefix. That figure so happens to be a convenient breakpoint in the ATN Addressing Plan. In IDRP, NSAP Address Prefixes can be any number of bits in length. If routes to individual Routing Domains can be aggregated together, and their individual NSAP Address Prefixes replaced by a single shorter common prefix, then we have achieved a useful simplification not just for our local electronic signpost, but for all such signposts downstream of the point at which the routes were aggregated.

In fact, if we can achieve the general principle that the further away from a route's destination you are, the shorter the NSAP Address prefix is for the route's destination, then we have achieved the goal of a scaleable internetwork. This is because for an internetwork to be scaleable, that is to be able to grow without any serious limitation on its total size, we must never get into the situation that the TCP/IP Internet has got itself into, where there are routers which have to keep having bigger and bigger "signposts" as the internet grows. The internet then cannot grow any more, once these routers have the biggest signposts that can be purchased.

As long as the above principle is obeyed, growth can occur in the far away internet without affecting remote routers, and hence growth can continue in an almost unbounded fashion.

For example, consider the example in Figure 2-6. Here we have a service provider supporting several users, and it is assumed that the service provider has been allocated the NSAP Address Prefix "1234" for all NSAP Addresses that it allocates. It allocates the prefix "12340" to its own Routing Domain, and then allocates "12341", "12342". etc. to each of its users' Routing Domains. The systems with those Routing Domains are then allocated NSAP Addresses relative to the NSAP Address Prefixes assigned to each Routing Domain.

In each User's Routing Domain, a Boundary Router forms a route to all systems within that Routing Domain. This is a route to all systems in the Routing Domain, and the route's destination is the NSAP Address Prefix assigned to the Routing Domain. This route is then advertised using IDRP to the Service Provider's Boundary Router.

The Service Provider's Boundary Router receives a so advertised route from each user's Routing Domain and can therefore build its own electronic signpost from each of these routes, "adding a sign" for each route advertised to it. This router could just re-advertise each such route on to a Boundary Router operated by another service provider or its own users. However, because all these routes share a common NSAP Address Prefix ("1234") it is much more efficient to first aggregate the routes together, along with the route to the service provider's own Routing Domain, and then apply the Route Information Reduction procedure to end up with a single route to "1234". This is the route it then advertises on, instead of re-advertising the individual routes to each Routing Domain.

Not only is this efficient but, if for example, a new user's Routing Domain is added (and given the next NSAP Address Prefix - "12344"), then this has no impact at all on the aggregated route or the number of routes maintained by the Boundary Router in another Service Provider. The internetwork has grown locally without having a global impact, and this is what scaleability is all about.

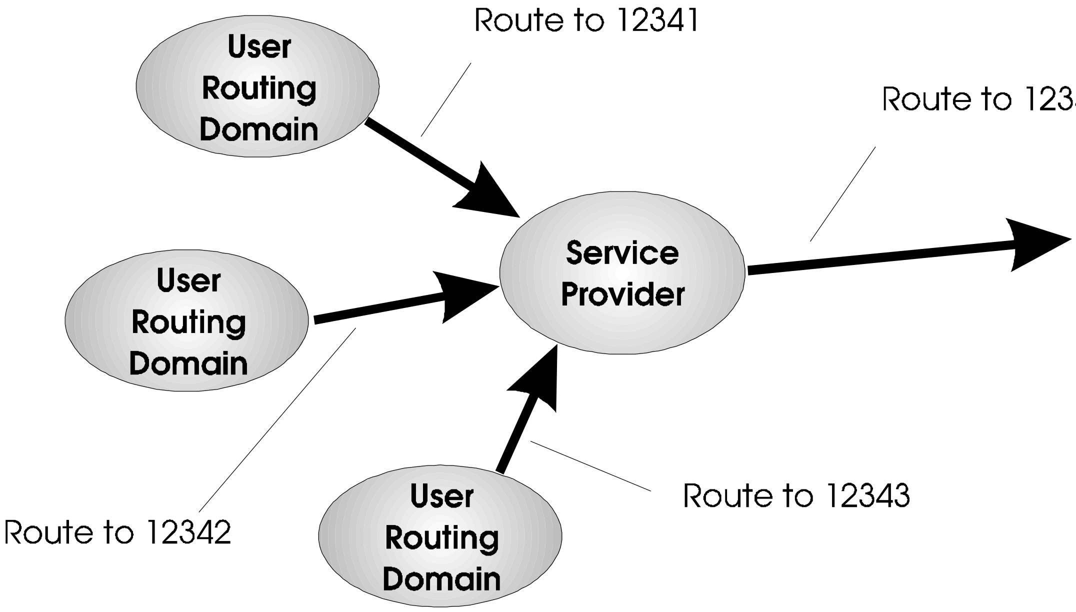

Figure 2-6 Aggregating Routes Together

This example can be readily extended. For example, if all of the Service Providers in a given country shared a common NSAP Address Prefix (e.g. "123"), then only a single route needs to be advertised internationally and which is common to all service providers. In fact, as long as the address allocation hierarchy reflects the way the network is organised, there will be many such opportunities for Route Aggregation and Route Information Reduction.

In the ATN, the addressing plan is so organised that each Administration has a single NSAP Address Prefix which will be common to all systems and Routing Domains that the maintain. Thus only a single route need be advertised between individual Administrations. Furthermore, provided that within a region, Administrations co-ordinate their addressing plans, it will be possible to form a single route to a given region keeping the overhead of inter-regional communications down to a minimum.

This principle is further exploited by the ATN Island concept. An ATN Island is essentially a regional grouping of Adminstrations with co-ordinated addressing plans. In such a situation, it is possible to form a single route to "the ATN Island", and, indeed, it is recommended that this is done prior to route advertisement to aircraft, thus keeping down the routing overhead on low bandwidth air/ground data links to a bare minimum.

2.4 Containment Boundaries and Routing Domain Confederations

Route Aggregation and Route Information Reduction generally work very well by themselves. However, to help solve the problem of when to aggregate, we need to introduce the idea of a Containment Boundary. We need some way of defining the scope of a given NSAP Address Prefix - that is to define a Containment Boundary that itself defines the limits of the domain of such an NSAP Address Prefix.

One obvious example of such a Containment Boundary is a Routing Domain. Each Routing Domain contains all systems identified by NSAP Addresses relative to the NSAP Address Prefix assigned to that Routing Domain. When routes exit a Routing Domain (i.e. at a Boundary Router), the Containment Boundary is crossed, and the router knows *a priori* that it is appropriate to aggregate the individual routes together and form a single route with its destination being the common NSAP Address Prefix for the Routing Domain.

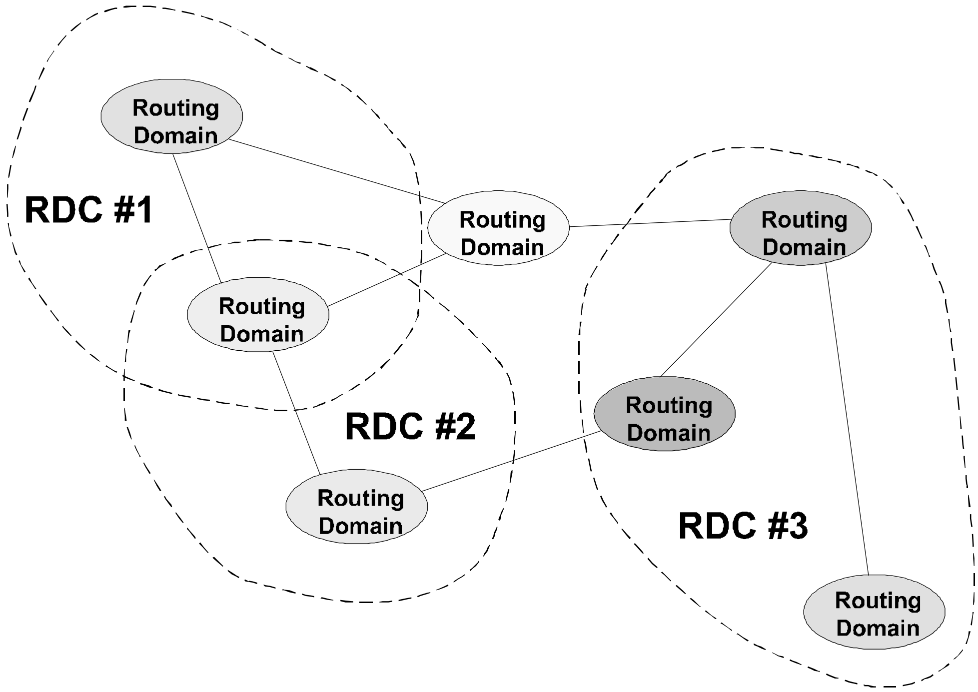

Figure 2-7 Routing Domain Confederations

In the example in 2.3 above, there is clearly some sort of Containment Boundary enclosing the Service Provider and its users. This can simply be a conventional boundary. However, IDRP does provide a means to make this more concrete in the shape of a Routing Domain Confederation (RDC).

An RDC is no more than a group of Routing Domains, as illustrated in Figure 2-7, and, at its simplest, is a means of collectively referring to a related group of Routing Domains. However, an RDC can usefully be defined to be a Containment Boundary for the domain of an NSAP Address Prefix. In the above example, we could have an RDC containing the Routing Domains of the Service Provider and its users.

With such an RDC, we can then implement a simple and effective rule for aggregating routes i.e. whenever a route that originates within the RDC is advertised across the RDC boundary, it is aggregated with all such routes to form a single route to a destination described by the common NSAP Address Prefix for all Routing Domains within the RDC. This is essentially what is happening in our example.

As happened in the example, more Routing Domains can be added to the RDC without affecting the route advertised external to the RDC. That is the internetwork has grown locally without global impact.

In the ATN, an ATN Island is an example of an RDC that contains all Routing Domains with a common NSAP Address Prefix i.e. common to all systems on the "Island". Whenever a route is advertised outside of the Island (e.g. to an aircraft) it becomes a candidate for aggregation with other such routes.

3. Route Aggregation in Detail

The preceding section has informally introduced the concepts involved in Route Aggregation and Route Information Reduction, and indicated how they apply to the ATN. It is now necessary to look at each of the main topics in detail and to consider how they are used in the ATN.

3.1 Routing Domains

In OSI routing, the most important containment boundary is formed by the Routing Domain. A Routing Domain is formally a grouping of End Systems and Intermediate Systems (i.e. routers), under a common network administrator. A Routing Domain also coincides with at least one addressing domain, but the concept is general enough to also enclose several addressing domains. The systems contained within a Routing Domain will therefore have either a single common address prefix, or at least have an address prefix chosen from a limited set. In either case, the address prefix will be unique to the Routing Domain.

Routing Domains are important because they contain systems that are subject to a common routing policy and are within a common domain of trust. Within a Routing Domain it is possible to use a connectivity based routing protocol such as that defined in ISO 10589, while between Routing Domains a policy based routing protocol, such as that defined in ISO 10747 has to be used.

In a Routing Domain that implements ISO 10589 internally, a two level approach to containment is applied within the Routing Domain, with the Routing Area forming a subsidiary containment boundary.

3.2 Inter-Domain Routing

The Inter-Domain Routing Environment may be assumed to comprise multiply interconnected Routing Domains. Some of these are Transit Routing Domains (TRDs), which will relay packets between other Routing Domains, whilst others are End Routing Domains (ERDs) which never relay between other Routing Domains, but may still have connections with many other RDs.

These RDs will be organised into many, possibly overlapping hierarchies, representing regional and organisational groupings. Service providers will operate TRDs, which will support many customers, perhaps having a single ERD, or perhaps their own grouping of TRDs and ERDs. Service Providers will themselves be interconnected, and many users will interconnect with multiple service providers.

The ISO Inter-Domain Routing Protocol (IDRP) is designed to operate in this environment. Recognising both the need to support multiple overlapping hierarchies and the lack of any common domain of trust, IDRP deals in routes, unlike the connectivity information distributed by ISO 10589. An IDRP Route comprises a set of destinations and information about the path that the routes takes through one or more RDs. The destination of a route is formally conveyed as Network Layer Reachability Information (NLRI) and is a list of NSAP Address Prefixes. A Router implementing IDRP aims to forward packets along the route which contains in its NLRI, the NSAP Address Prefix that provides the longest match with the packet’s destination NSAP Address. The destination of a route to an Routing Domain is the unique address prefix(es) that is common to all addresses within the Routing Domain.

IDRP routes are advertised from one Routing Domain to another, not just because the connectivity exists, but because organisational policy permits the route to be advertised, makes available the path offered by the route, and the resources that packets following this route will consume. Similarly, a Router is not required to always accept a route it receives - organisational policy may be to reject it, and organisational policy is also used to select between routes to the same or overlapping sets of destinations.

IDRP also gives Routing Domains a name. This is the Routing Domain Identifier (RDI), and is totally separate from the unique address prefix(es) common to all systems within the Routing Domain. An RDI is a Network Entity Title (NET) and is therefore syntactically a single address. It will typically be assigned relative to the Routing Domain’s common NSAP address prefix, but does not have to be. It is used by IDRP in the RD_Path attribute that forms part of each route’s path information, and records that the route passes through the Routing Domain. It is used for loop detection purposes and may also be referenced by routing policy.

3.3 Routing Domain Confederations

A Routing Domain Confederation (RDC) is a set of Routing Domains. It is a very general concept and RDCs may be nested within one another, and may also overlap. RDCs only have one functional role, that is to contain the expansion of trace information in a route’s path. Otherwise, they provide containment boundaries that can be readily referenced by routing policy.

Like Routing Domains, RDCs also have names, and again these are NETs, and are also called RDIs. The RDI of an RDC is also used in the IDRP RD_Path to record when a route enters an RDC and, the RDCs that the route has passed through.

The use of RDCs in controlling trace information is an important contribution to scaleability. Every time a route is advertised by a Routing Domain, the Routing Domain’s own unique Routing Domain Identifier (RDI) is added to the route’s trace information held in the RD_Path path attribute. This attribute keeps track of which RDs the route has passed through, in order to prevent routing loops. However, when a route exits an RDC, the RDIs of all RDs that are within the exited RDC are removed and replaced by the RDI of the RDC itself. This has the dual role of both reducing the overhead of the trace information, and of preventing the route from ever re-entering the RDC, as to do so would constitute a routing loop.

RDCs thus by themselves have an important role to play in reducing the overhead of distributing routes in large internets. However, by ensuring that once a route has left an RDC it cannot re-enter it, they also provide a useful basis for routing policy in scaleable internetworks, as Route Information Reduction can then be implemented without risk of ambiguous routing.

3.4 Route Aggregation

Route Aggregation is formally the process by which two or more routes are merged together to form a single route. Each IDRP path attribute has route aggregation rules associated with it, determining the mechanics of route aggregation. Route Aggregation is applied when:

a) Local Routing Policy selects a groups of routes for aggregation prior to their advertisement to another Routing Domain; or,

b) Two or more routes selected for advertisement to an adjacent Routing Domain have an identical set of NSAP Address prefixes as their destination, have the same distinguishing path attributes, including the security path attribute, but have different security information in their paths.

The former case is generally used in conjunction with Route Information Reduction in order to reduce the number of routes being advertised, while the latter is necessary to ensure proper routing when routing control procedures are in effect. This is because it is not possible to advertise two route to an adjacent Router with the same destinations and distinguishing path attributes, without one being assumed to be a replacement to the other, rather than as a different route.

The aggregation procedures are themselves fully specified in ISO 10747, or in the ATN SARPs (in the case of the security path attribute). These aggregation procedures are purely algorithmic in nature and are not affected by local policies.

It should be noted that when the NLRI is aggregated, the result is typically the set of the individual NSAP Address prefixes of each component route. Only when a shorter Address Prefix can be unambiguously used to

i) represent some or all of those NSAP Address prefixes present in the aggregated NLRI, and,

ii) only those NSAP Address prefixes,

does ISO 10747 allow substitution of some or all of the NSAP Address prefixes in the aggregated route by a single shorter common prefix. In order to reduce the size of the NLRI further requires the separate application of Route Information Reduction.

3.5 Route Information Reduction

Route Information Reduction is the policy based reduction of information in either the NLRI, or the RD_Path. Generally, reduction of the RD_Path information is carried out automatically by the RDC mechanism, and hence this is rarely needed. However, policy based reduction of the NLRI is much more useful. There are two main cases in which Route Information Reduction applies:

1) When there are known “holes” in address allocation such that once the NLRI contains a given set of NSAP Address Prefixes, they may be unambiguously replaced by a shorter common prefix. This is distinguished from the replacement of NSAP Address Prefixes during the aggregation of NLRI in that, in the case of the aggregation procedures, the replacement only occurs when it can be performed without prior knowledge of the actual state of address allocation. In this case, knowledge of which addresses are assigned and which are not and hence which will never occur in routes, can be taken into account.

2) At RDC boundaries, or some other boundary recognised by the local Routing Policy, when one or more members of a defined set of NSAP Address Prefixes are present, then they are replaced by a single shorter prefix.

In practice, Route Information Reduction is carried out after routes have been selected and aggregated. However, it may also be applied to routes which have not been aggregated. For example, there may at times be only a single route with a single NSAP Address Prefix in its NLRI that satisfies case (2) above. However, it is still important to perform the Route Information Reduction, in order to present a consistent and stable external appearance.

Route Information Reduction plays the crucial role of ensuring that the size of an Internet is not limited by the capacity of the smallest Service Provider. For example, if Route Information Reduction takes place at the boundaries of an RDC that also encloses an Addressing Domain from which RDs within the RDC are assigned NSAP Address Prefixes, then the number of RDs within a given RDC can increase without affecting Routers outside of the RDC.

3.6 Routing Policy and Route Aggregation

A Routing Domain provides a containment boundary and, regardless of which systems are switched on within the Routing Domain, there is no need to qualify the route to the Routing Domain, depending on which systems are active. The route advertised by the Routing Domain is a route to the unique NSAP Address Prefix(es) assigned it. This is always an unambiguous route.

RDCs can also be used as containment boundaries, and in a recursive fashion. They therefore enable scaleability of the inter-domain environment.

Firstly, when a group of Routing Domains share a unique NSAP Address Prefix, they may form the membership of an RDC. This RDC is then a containment boundary that coincides with the addressing domain from which those NSAP Address Prefixes were assigned.

Secondly, at the boundaries of this RDC, routing policies may be specified such that all routes that originate within the RDC are first aggregated, and then Route Information Reduction applied. The Route Information Reduction should be set up so that any NSAP Address Prefix(es) for destinations inside the RDC are replaced by the single unique NSAP Address Prefix shared by all systems within the RDC.

When the above is in place, such an RDC appears like a single Routing Domain. The internal detail is hidden from those outside, and the membership of the RDC can grow and change without affecting those outside of it. Even if whole Routing Domains are “switched off”, this fact is hidden to the outside of the RDC, given the form of Route Information Reduction applied.

As RDCs can be nested within each other, this form of information hiding can occur many times over, perhaps hiding such information on first a company, then a regional, and then on a Country basis.

What’s more, the same process can happen in reverse. At the boundaries of an RDC, routes entering the confederation can be aggregated and their destinations replaced by a single NSAP Address Prefix for the “rest of the world” - although if there is more than one entry point to the RDC, such a mechanism may require careful co-ordination if routing to outside of the RDC is not to suffer from unexpected behaviour.

3.7 Route Merging

Route Merging is specified by the ATN Internet SARPs and is a specific example of Route Aggregation. Route Merging takes place automatically (i.e. without reference to any routing policy rules), and is necessary to support routing via different air/ground subnetwork technologies on a per application basis. Without Route Merging, an intermediate router that had a choice of routes to the same aircraft via different air-ground subnetworks, would have no choice but to discard one of those routes. In consequence, any “downstream” systems would be denied access to this subnetwork.

Route Merging is also specified with simplified rules for aggregation of the RD_Path path attribute. This simplification was introduced as a short term measure to counter concerns over whether the full algorithm could be validated in time for the ATN Panel meeting.

This form of Route Aggregation corresponds to that described in 3.4 (b). There is no requirement for this aggregation to take place at RDC boundaries, neither is it likely to be used with Route Information Reduction.

3.8 Route Combination

Route Combination is the combination of two or more routes into a single UPDATE BISPDU and is not strictly part of Route Aggregation, but is instead an optimisation to reduce the number of BISPDUs exchanged between two adjacent BISs. The principle is that when a BIS has two or more routes that need to be advertised to an adjacent BIS, and when these routes have the same NLRI, they may be combined into a single UPDATE BISPDU, which encodes common path attribute values once and once only for each combined route.

However, by the same process, Route Withdrawals may also be included in the same UPDATE BISPDU as a newly advertised route. When aggregated routes are modified such that the NLRI changes, the original aggregated route has to be formally withdrawn and its replacement advertised as a new route. To prevent discontinuities in the availability of the aggregated route, it is important that the withdrawal of the older route and its replacement take place simultaneously. Route Combination, in this case combining withdrawals and updates together, is thus essential to the proper operation of Route Aggregation.

4. Route Aggregation and the ATN

4.1 In the Ground-Ground Environment

The ATN Internet SARPs do not mandate any route aggregation in support of ground-ground routing, although clearly this will become necessary as the ATN grows. However, route aggregation is mandated in support of air/ground routing.

4.2 Supporting Air/Ground Data Links

The ATN Internet SARPs defines a number of RDCs in support of air-ground routing. These are containment boundaries, and, together with specific use of Route Aggregation and Route Information Reduction, support the requirements for air-ground route information distribution.

4.2.1 The Fixed ATN RDC

The draft SARPs define the “Fixed ATN RDC” as comprising all ATN RDs other than mobile RDs and also specify an addressing plan that defines separate NSAP Address Prefixes for ground based and mobile systems. The Fixed ATN RDC is a containment boundary, containing all ground systems. Furthermore, all systems within this RDC have a common NSAP Address prefix. The Fixed ATN RDC has three main purposes:

1. To minimise the trace information sent over the air/ground data link. As this confederation is exited whenever a route is advertised from an air/ground router to an airborne router, all trace information concerning the ground environment is removed and replaced by the single RDI of the Fixed ATN RDC.

2. To ensure that erroneous routes between two air/ground routers and via an airborne router, are prevented. This is a simple consequence of the fact that routes cannot re-enter a confederation.

3. To support the recommendation in sections 3.7.1.3 and 3.7.3.3 of the ATN Internet SARPs to advertise to an airborne router an aggregated route to all ATN Destinations on other ATN Islands. It is appropriate to implement this recommendation when the Air/Ground Router has connectivity, directly or indirectly, with the Global ATN Backbone envisaged in the draft ATN SARPs.

In the last case, an aggregation rule will be required to select routes to destinations outside of the local ATN Island for aggregation into a single route. As the route is now exiting the Fixed ATN RDC it is also possible to apply a Route Information Reduction rule replacing any NSAP Address Prefix(es) for a destination outside of the ATN Island, with the common address prefix for all ground systems.

The effect of this is to offer a single route to “the rest of the ATN”, with a single address prefix in its set of destinations, and a single entry in the trace information. This enables the availability of worldwide connectivity to be efficiently advertised to an airborne router.

4.2.2 The ATN Island RDC

The ATN Internet SARPs define the ATN Island RDC as comprising all ATN RDs within the same ATN Island. This is illustrated in Figure 4-1, which also illustrates the flow of routes within an ATN Island. Like the Fixed ATN RDC, the ATN Island RDC is also a containment boundary, and it is important that there is a common NSAP Address for all systems within an ATN Island RDC. Otherwise, inefficient use of air/ground data links may result if there is no more than one or two unique NSAP Address Prefixes for all the systems within an ATN Island. The ATN Island RDC has two main purposes.

1. To minimise the trace information on routes advertised between ATN Islands.

2. To support the recommendation in sections 3.7.1.3 and 3.7.3.3 to advertise to an airborne router an aggregated route to all systems in the local ATN Island RDC.

In the last case, an aggregation rule will select routes to destinations within the ATN Island, other than those to the local Administrative Domain, and aggregate them together into a single route for advertisement to an airborne router. Provided that all systems within the ATN Island RDC do share a unique NSAP Address Prefix, then Route Information Reduction will also be appropriate here. This requires a rule that replaces any NSAP Address Prefix(es) for a destination within the ATN Island, other than those of the local Administrative Domain, with the common address prefix for all ground systems.

The effect of this will be to offer a single route to the “rest of the ATN Island”. However, the efficiency of the propagation of this route will be limited if there is no single unique address prefix for all systems within the ATN Island.

4.2.3 Administration RDCs

The ATN Internet SARPs do not require that administrations and organisations that own/operate multiple RDs, group them together into an RDC. However, such an RDC will be useful in providing a containment boundary for each administration or organisation’s systems. The current addressing plan does specify a common NSAP Address Prefix for all of an Administration or Organisation’s systems, and such an RDC would therefore be the containment boundary for systems with addresses assigned relative to this NSAP Address Prefix. However, this part of the addressing plan is not fully compatible with the goal of having a common address prefix for an ATN Island and may need to be reviewed.

4.2.4 The ATN Island Backbone RDC

The ATN Internet SARPs define the ATN Island Backbone RDC as comprising the backbone RDs within an ATN Island. It provides a containment boundary for the distribution of routes to mobiles.

Routers within a Backbone RDC are required to implement a routing policy that does not permit the advertisement of routes to individual mobile RDs, outside of the confederation, except to the preferred route to an aircraft’s Home. Instead, such routers advertise a default route to all aircraft.

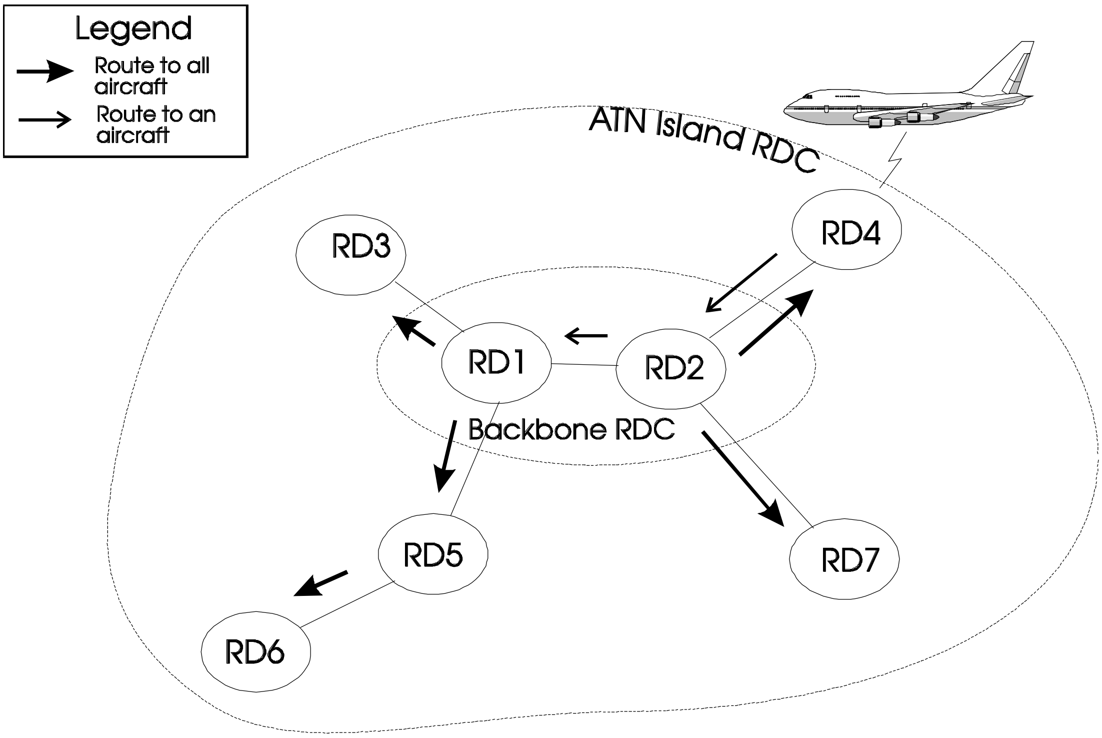

Figure 4-1 The ATN Island RDC

This RDC provides a useful boundary that may be referenced by routing policies, but no more. This is not intended to be an Addressing Domain boundary. Furthermore, the Default Route to all mobiles that is advertised from the Backbone Routing Domain to the rest of the ATN Island is not formed by Route Aggregation, but is instead generated *a priori*.

This is because, such an aggregated route would contain RDIs for (possibly all) RDs in the ATN Island. Route loop prevention functions would therefore prevent its advertisement to any other Routing Domain in the ATN Island, therefore defeating its purpose. The default route to all mobiles is therefore totally separate from individual routes to mobile systems, and is not an aggregated route. For similar reasons, the default route generated by “Home” RDs, also not formed by Route Aggregation.

5. Implementation Considerations

The preceding sections have introduced the ideas behind Route Aggregation and Route Information Reduction, and then formally described these and related concepts. This section now discusses how such policy driven concepts can be implemented in real routers.

5.1 Preliminaries

As specified in ISO 10747, each router implementing IDRP (properly known as a Boundary Intermediate System or BIS for short), maintains a Routing Information Base (RIB), a Policy Information Base (PIB) and a Forwarding Information Base (FIB). These are illustrated in Figure 5-1.

The IDRP protocol is connection mode and, for each adjacent BIS with which routing information is exchanged, a BIS-BIS connection must exist, and there is required to be a separate RIB data structure to hold all routes received from that BIS. Each such data structure is known as an *adj-RIB-in*. Similarly, for each such adjacent BIS, there is also required to be a separate RIB data structure to hold all routes currently advertised to that BIS. Each such data structure is known as an *adj-RIB-out*.

It should be observed that the contents of a given *adj-RIB-out* ought to be identical to the corresponding *adj-RIB-in* held by the adjacent BIS. To ensure this is the case, the BIS-BIS protocol provides a mechanism to "refresh" an *adj-RIB-in* from the corresponding *adj-RIB-out*. The refresh cycle is periodically performed and ensures that they are identical, thus avoiding any long term persistence of any discrepancies that might have occurred.

The routes received from another BIS are processed by a Routing Decision process, and it is the responsibility of this process to select routes from the adj-RIB-ins for local use and for propagation to other BISs. The so selected routes are recorded in a further data structure, the *loc-RIB*. This holds all currently selected routes, and it is from the loc-RIB that routing information is selected for transfer to the FIB where it is used to support the forwarding of NPDUs, and for the adj-RIB-out for transfer to other BISs.

The Routing Decision Process is formally described as a three phase process, where phases one and two are concerned with the selection of routes and their placement in the loc-RIB, whilst phase three is concerned with the processing of routes in the loc_RIB and maintaining the adj_RIB-out. The PIB may contain rules referenced by each phase of the Routing Decision Process, determining the initial selection of routes, and the aggregation and propagation of so selected routes to other BISs.

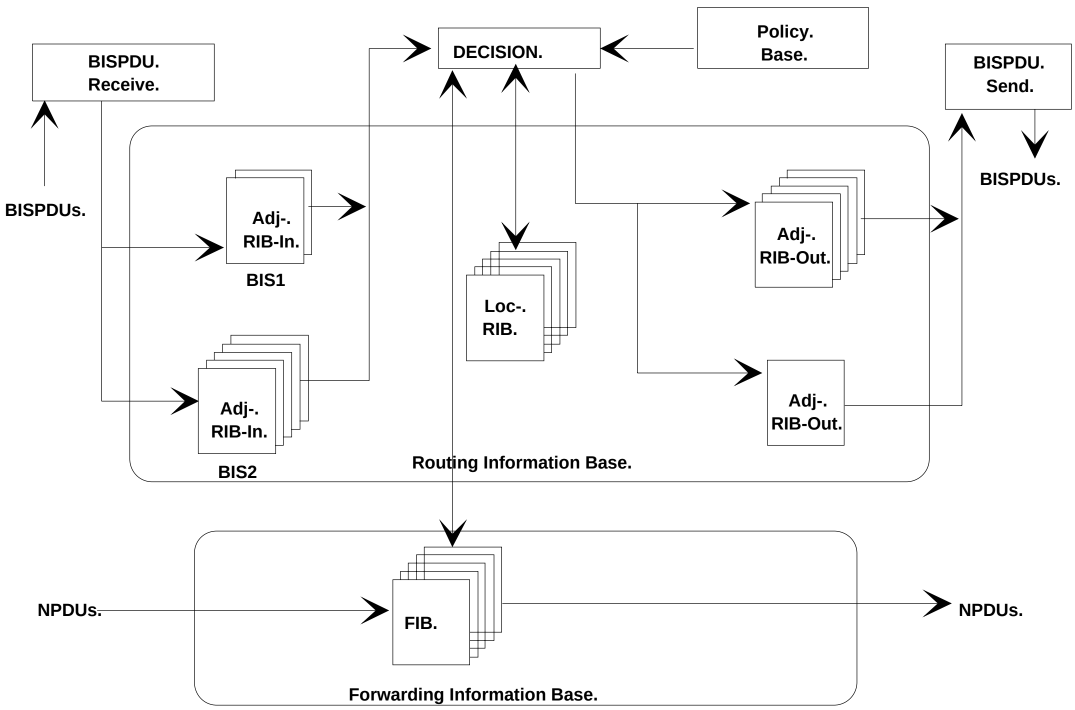

Figure 5-1 IDRP BIS Architecture

It should be noted that IDRP can handled several parallel sets of routes, where each set is distinguished by a set of “Distinguishing Path Attributes” (a RIB_Att). The purpose of this is to permit the distribution of routes according to different Quality of Service metrics and/or security information. In the ATN, two such RIB_Atts are defined: a default RIB_Att under which routes are distributed with no QoS or Security information, and a Security RIB_Att under which routes are distributed with information used to support user driven routing requirements (e.g. to select the use of specific air/ground data link types on a per application basis).

5.2 Requirements Analysis

5.2.1 Route Aggregation

The Phase Three Routing Decision Process may simply select and copy routes from the loc_RIB to an adj-RIB-out, but there is no requirement for there to be a one-to-one relationship between loc_RIB and adj-RIB-out routes because, quite simply, some of the adj-RIB-out routes may be the result of Route Aggregation.

Route Aggregation occurs as part of the Phase Three Routing Decision Process, and there are two parts to the implementation of Route Aggregation itself:

1. The selection of routes for aggregation

2. The generation of an aggregated route from the so selected routes.

5.2.1.1 Route Selection

The Selection of Routes is rule based using rules contained in the PIB. Some such rules are protocol requirements, others may be artefacts of the implementation, whilst others will be due to local policy. In the ATN, it is necessary to select routes for Route Merging, Route Aggregation and to select routes that are to be simply re-advertised on a policy driven basis.

5.2.1.1.1 Selection for Route Merging

In order to meet the ATN Route Merging requirement, when a Router's Phase Three Routing Decision Process processes the loc-RIB for the Security RIB_Att, it also selects those routes that have the same destination addresses, and which are eligible for advertisement to this particular adjacent BIS. Each such set of routes is then merged into a single route using the simplified set of merging rules.

Essentially, a filter is being repetitively applied to the routes in the loc_RIB, for each distinct set of destination addresses present. This filter has both an

- **exclusion condition** (i.e. routes are excluded if they do not contain the security path attribute, or are ineligible for advertisement to the adjacent BIS - e.g. by policy or route loop prevention), and an

- **inclusion condition** (i.e. routes are included if they have they required set of destination addresses in their NLRI).

5.2.1.1.2 Selection of Routes for Advertisement

Routes may be selected on a simple availability basis, i.e. every available route in the loc_RIB is copied to each adj-RIB_out. However, there will be many cases where a more selective policy is required. For example, local policy may permit the relaying of data between certain Routing Domains and not others. It will therefore be necessary to be selective about which routes are selected for copying to a given adj_RIB_out.

As above, a combination of exclusion and inclusion filters appears to be necessary, in order to (e.g.) include routes to or via certain RDs, while excluding those to or via others. Such filters will need to operate upon both the RD_Path and NLRI, and may be applicable to only certain RIB_Atts.

5.2.1.1.3 Selection for Route Aggregation

In the case of the ATN, it will be necessary in an air/ground router to select:

a) All routes to destinations in the local ATN Island (an RDC) except for those to the local Routing Domain; and

b) All routes to ATN destinations except for those to destinations in the local ATN Island.

Again a combination of inclusion and exclusion filters is required. This time (in addition to excluding routes which are generally not available for advertisement), both inclusion and exclusion rules operate on the RD_Path information and, in particular, operating on the originating Routing Domain or RDC.

5.2.1.1.4 The General Requirement

Filter based selection of routes for inclusion in the Adj-RIB-Out is really a general requirement of policy based routing. Selection for Route Aggregation should be part of that general requirement, and there are probably three types of route selection rule:

1) Filters which when satisfied by more than one route in the Loc-RIB require that those routes are aggregated;

2) Filters which when satisfied by more than one route in the Loc-RIB require that those routes are then grouped according to their destination, and then only those selected routes with identical NLRI are aggregated; and

3) Filters which when satisfied by more than one route in the Loc-RIB require that those routes are individually entered into the Adj_RIB_out.

The first type of filter is a general Route Aggregation selection rule, while the second is a Route Merging selection rule. The last type of filter is an individual route selection rule.

If a given route satisfies filters of both aggregation and individual selection rules then it would also appear correct to assume that the aggregation rule dominates i.e. that if a route is selected and aggregated with others, then generally it will not be also included in the Adj_RIB_out as an individual route. This appears to be a sensible approach as the whole point of aggregation is to cut down the number of routes being advertised. Special cases can always be explicitly excluded from the aggregation rule.

On the other hand, if a route is selected by more than one aggregation rule, then while it may be worthwhile issuing a warning to a network manager, it is not incorrect for it to be included in each such aggregation.

In principle, it will be necessary to specify a different set of route selection rules for each adjacent BIS. This is because different policies will apply to different adjacent Routing Domains.

5.2.1.1.5 Types of Filters

Each of the above types of selection rule will consist of a logical expression combining a number of inclusion and/or exclusion filters, selecting or excluding routes on the basis of:

a) Presence of a specific distinguishing path attribute

b) RDI of Route originator is a given value

c) Route is via an Routing Domain or RDC with a given RDI

d) Presence of an ATN Security Label tag set and the value of that tag set after the application of a logical ‘AND’ with a mask field equals a specific value.

e) Destination of route either equals or has as its prefix, a given NSAP Address Prefix.

The actual syntax of such a logical expression is implementation specific.

5.2.1.2 Aggregation Procedures

The aggregation procedures are themselves fully specified in ISO 10747, or in the ATN SARPs (in the case of the security path attribute). These aggregation procedures are purely algorithmic in nature and are not affected by local policies.

5.2.2 Route Information Reduction

As discussed earlier, there are two main cases in which Route Information Reduction of NLRI applies:

1) When there are known "holes" in address allocation such that once the NLRI contains a given set of NSAP Address Prefixes, they may be unambiguously replaced by a shorter common prefix.

2) At RDC boundaries, or some other boundary recognised by the local Routing Policy when if one or more members of a defined set of NSAP Address Prefixes are present, then they are replaced by a single shorter prefix.

In each case, this is a rule based process, and the rule can be represented as a list of NSAP Address Prefixes plus the replacement prefix.

5.2.3 Update of the Adj-RIB-out

Route Aggregation adds an additional problem to the general problem of keeping track of which routes are new routes, which have been replaced and which are withdrawn. The IDRP UPDATE BISPDU combination rules permits a route to be withdrawn and replaced by an alternative in a single BISPDU. This is very important in the case of Route Aggregation when a component route is added to or withdrawn from the aggregated route, and the NLRI changes as a result, either directly or because Route Information Reduction is then applied differently In such cases, the revised aggregated routes *replaces* the previous version of the aggregated route. This must happen in a single BISPDU if this is not to affect the other members of the aggregated route i.e. to make these routes unavailable for a period determined by the **minRouteAdvertisementInterval**.

Essentially, the routes in the Adj-RIB-out must be associated with the rule that selected them. Thus when the route that is selected by the rule changes, the proper handling can be determined. Clearly, the handling will depend upon the type of rule. Using the list numbering in 5.2.1.1.4:

- A type 1 rule results in one and only one aggregated route. If this changes, then this replaces the previous route;

- A type 2 rule results in one route for each different NSAP Address Prefix found in the NLRI of the selected routes. Such routes may be added to, replaced or withdrawn.

- A type 3 rule results in one route for each selected route. Such routes may also be added to, replaced or withdrawn.

5.3 The Phase Three Route Decision Process

The requirements analysis presented in the preceding section results in the process model for the IDRP Phase 3 Route Decision Process illustrated in Figure 5-1. Two PIB data structures are referenced: a list of "Route Selection Rules" and a list of "Reduction Rules". The former is used for grouping routes together for the purposes of Route Aggregation, while the latter is for determining when Route Information Reduction of NLRI can be performed.

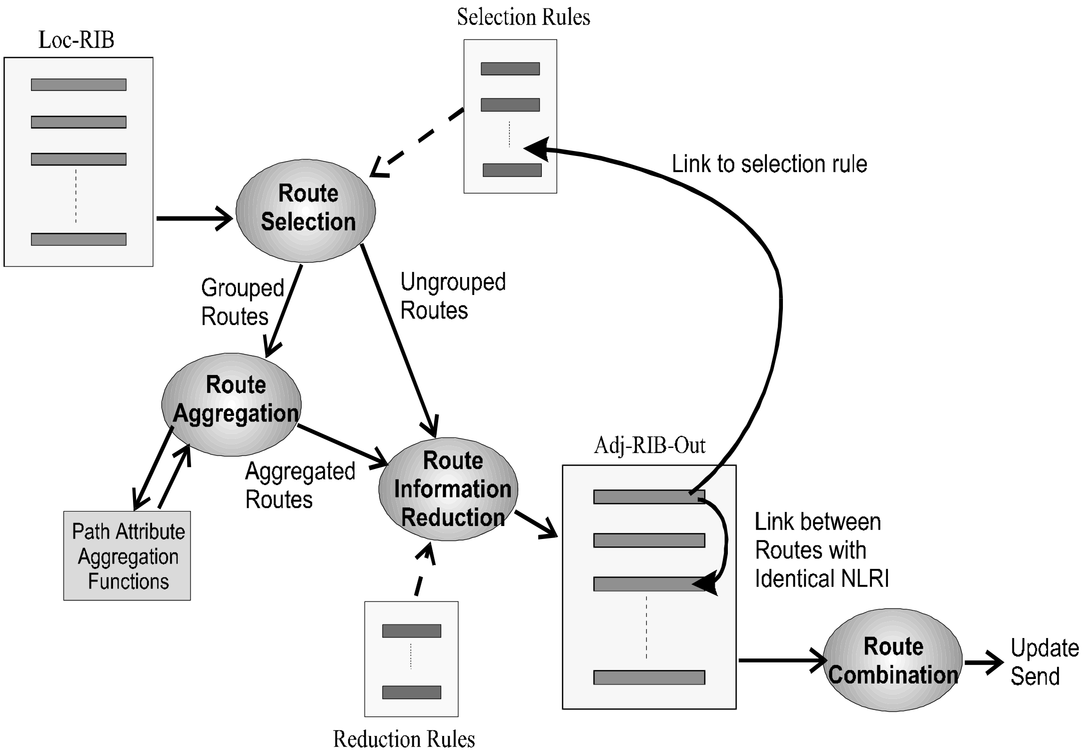

Figure 5-1 Process Model For Route Decision Phase 3

In both cases, it will be necessary to define a syntax to enable the text based definition of the rules, so that these data structures may then be created at system start up.

A "Route Selection" process is specified to pass through the Loc_RIB applying first type 1 and type 2 rules, and then applying type 3 rules to any routes in the Loc_RIB not selected by a type 1 or type 2 rule.

The routes selected by type 1 or type 2 routes are grouped routes. The routes selected by each type 1 rule form a single group. Those selected by a type 2 rule are placed in a separate group for each distinct NSAP Address Prefix in the routes' NLRI. Each group is then processed by a "Route Aggregation" process to create a single aggregated route for each such group. The aggregation process uses a library of aggregation functions to aggregate each type of path attribute.

The routes selected by a type 3 rule are ungrouped routes. Both ungrouped routes and the aggregated routes that result from the Route Aggregation process are passed to a "Route Information Reduction" process. This process inspects the NLRI of each route presented to it and applies the reduction rules to it. The application of a reduction rule will, if the rule is satisfied, result in the replacement of one or more NSAP Address Prefixes in the route's NLRI, with a single shorter prefix. The rules are applied iteratively until no further reduction can take place.

Once the reduction rules have been applied, the Route Information Reduction Process inserts the route into the Adj-RIB-out. It also links the route to the Selection Rule that originally selected it, and to any other routes in the same Adj-RIB-out that have the same NLRI. However, prior to inserting the route, the process checks the Adj-RIB-out to see if an existing route is present linked to the same Selection Rule. If this is a type 1 rule, the then new route is marked as replacing the route linked to that Selection Rule. If it is a type 2 or type 3 rule and there is an existing route in the Adj-RIB-out with the same NLRI as the new route, then again the new route is marked as replacing the existing route. Note that in both cases, if the new route is identical to the existing route in both the path attributes it contains and their values then it does not replace the existing route. The existing route may be simply viewed as refreshed.

Indeed, once the phase 3 processes complete, any routes in an Adj-RIB_out that have been neither refreshed nor replaced, must be marked as withdrawn.

Finally, when a route is passed to the Update Send process for advertisement to an adjacent BIS, a “Route Combination” process is required. This will:

a) Ensure that a route withdrawal is always advertised in the same UPDATE BISPDU as the route, if any, that replaces it; and,

b) Ensure that when a route is advertised, it is combined with any routes with the same NLRI, and which are also queued for advertisement to the adjacent BIS.

6. Conclusion

This paper has attempted to introduce the concepts behind Route Aggregation and Route Information Reduction by informally introducing a “signpost” analogy in an attempt to illustrate the processes necessary to make a scaleable internetwork. From this informal start, the various concepts have been discussed in detail, the requirements analysed and an implementation model presented. It is hoped that this will form a sufficient basis from which ATN implementation of Route Aggregation and its associated functions can proceed.

Because it is a complex subject, it is easy to try and put off the implementation of Route Aggregation. However, the lesson from the TCP/IP Internet is that if you do not design from the beginning for the functions necessary for scaleability, then they will elude you forever. Most importantly, address assignment and management have to be implemented with an understanding of how they will work in a large scale internetwork. Route Aggregation cannot therefore be overlooked. It has to be fully understood even in the early stages of ATN Deployment if mistakes are not to be made and, even in the early ATN, in such areas as the advertisement of routes over low bandwidth air/ground data links, it has an important contribution to make.