

Network Working Group
Request for Comments: 1519
Obsoletes: 1338
Category: Standards Track

V. Fuller
BARRNet
T. Li
cisco
J. Yu
MERIT
K. Varadhan
OARnet
September 1993

Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy

Status of this Memo

This RFC specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Abstract

This memo discusses strategies for address assignment of the existing IP address space with a view to conserve the address space and stem the explosive growth of routing tables in default-route-free routers.

Table of Contents

Acknowledgements	2
1. Problem, Goal, and Motivation	2
2. CIDR address allocation	3
2.1 Aggregation and its limitations	3
2.2 Distributed network number allocation	5
3. Cost-benefit analysis	6
3.1 Present allocation figures	7
3.2 Historic growth rates	8
3.3 Detailed analysis	8
3.3.1 Benefits of new addressing plan	9
3.3.2 Growth rate projections	9
4. Changes to inter-domain routing protocols and practices	11
4.1 Protocol-independent semantic changes	11
4.2 Rules for route advertisement	11
4.3 How the rules work	13
4.4 Responsibility for and configuration of aggregation	14
4.5 Intra-domain protocol considerations	15
5. Example of new allocation and routing	15

5.1	Address allocation	15
5.2	Routing advertisements	17
6.	Extending CIDR to class A addresses	18
7.	Domain Naming Service considerations	20
7.1	Procedural changes for class-C "supernets"	20
7.2	Procedural changes for class-A subnetting	21
8.	Transitioning to a long term solution	22
9.	Conclusions	22
10.	Recommendations	22
11.	References	23
12.	Security Considerations	23
13.	Authors' Addresses	24

Acknowledgements

The authors wish to express their appreciation to the members of the ROAD group with whom many of the ideas contained in this document were inspired and developed.

1. Problem, Goal, and Motivation

As the Internet has evolved and grown over in recent years, it has become evident that it is soon to face several serious scaling problems. These include:

1. Exhaustion of the class B network address space. One fundamental cause of this problem is the lack of a network class of a size which is appropriate for mid-sized organization; class C, with a maximum of 254 host addresses, is too small, while class B, which allows up to 65534 addresses, is too large for most organizations.
2. Growth of routing tables in Internet routers beyond the ability of current software, hardware, and people to effectively manage.
3. Eventual exhaustion of the 32-bit IP address space.

It has become clear that the first two of these problems are likely to become critical within the next one to three years. This memo attempts to deal with these problems by proposing a mechanism to slow the growth of the routing table and the need for allocating new IP network numbers. It does not attempt to solve the third problem, which is of a more long-term nature, but instead endeavors to ease enough of the short to mid-term difficulties to allow the Internet to continue to function efficiently while progress is made on a longer-term solution.

The proposed solution is to topologically allocate future IP address assignment, by allocating segments of the IP address space to the transit routing domains.

This plan for allocating IP addresses should be undertaken as soon as possible. We believe that this will suffice as a short term strategy, to fill the gap between now and the time when a viable long term plan can be put into place and deployed effectively. This plan should be viable for at least three (3) years, after which time, deployment of a suitable long term solution is expected to occur.

This plan is primarily directed at the first two problems listed above. We believe that the judicious use of variable-length subnetting techniques should help defer the onset of the last problem, the exhaustion of the 32-bit address space. Note also that improved tools for performing address allocation in a "supernetted" and variably-subnetted world would greatly help the user community in accepting these sometimes confusing techniques. Efforts to create some simple tools for this purpose should be encouraged by the Internet community.

Note that this plan neither requires nor assumes that already assigned addresses will be reassigned, though if doing so were possible, it would further reduce routing table sizes. It is assumed that routing technology will be capable of dealing with the current routing table size and with some reasonably small rate of growth. The emphasis of this plan is on significantly slowing the rate of this growth.

Note that this plan does not require domains to renumber if they change their attached transit routing domain. Domains are encouraged to renumber so that their individual address allocations do not need to be advertised.

This plan will not affect the deployment of any specific long term plan, and therefore, this document will not discuss any long term plans for routing and address architectures.

2. CIDR address allocation

There are two basic components of this addressing and routing plan: one, to distribute the allocation of Internet address space and two, to provide a mechanism for the aggregation of routing information.

2.1 Aggregation and its limitations

One major goal of this addressing plan is to allocate Internet address space in such a manner as to allow aggregation of routing

information along topological lines. For simple, single-homed clients, the allocation of their address space out of a transit routing domain's space will accomplish this automatically - rather than advertise a separate route for each such client, the transit domain may advertise a single aggregate route which describes all of the destinations connected to it. Unfortunately, not all sites are singly-connected to the network, so some loss of ability to aggregate is realized for the non-trivial cases.

There are two situations that cause a loss of aggregation efficiency.

- o Organizations which are multi-homed. Because multi-homed organizations must be advertised into the system by each of their service providers, it is often not feasible to aggregate their routing information into the address space any one of those providers. Note that they still may receive their address allocation out of a transit domain's address space (which has other advantages), but their routing information must still be explicitly advertised by most of their service providers (the exception being that if the site's allocation comes out of its least-preferable service provider, then that service provider need not advertise the explicit route - longest-match will insure that its aggregated route is used to get to the site on a backup basis). For this reason, the routing cost for these organizations will typically be about the same as it is today.
- o Organizations which change service provider but do not renumber. This has the effect of "punching a hole" in the aggregation of the original service provider's advertisement. This plan will handle the situation by requiring the newer service provider to advertise a specific advertisement for the new client, which is preferred by virtue of being the longest match. To maintain efficiency of aggregation, it is recommended that organizations which do change service providers plan to eventually migrate their address assignments from the old provider's space to that of the new provider. To this end, it is recommended that mechanisms to facilitate such migration, including improved protocols and procedures for dynamic host address assignment, be developed.

Note that some aggregation efficiency gain can still be had for multi-homed sites (and, in general, for any site composed of multiple, logical IP network numbers) - by allocating a contiguous power-of-two block of network numbers to the client (as opposed to multiple, independent network numbers) the client's routing information may be aggregated into a single (net, mask) pair. Also,

since the routing cost associated with assigning a multi-homed site out of a service provider's address space is no greater than the current method of a random allocation by a central authority, it makes sense to allocate all address space out of blocks assigned to service providers.

It is also worthwhile to mention that since aggregation may occur at multiple levels in the system, it may still be possible to aggregate these anomalous routes at higher levels of whatever hierarchy may be present. For example, if a site is multi-homed to two NSFNET regional networks both of whom obtain their address space from the NSFNET, then aggregation by the NSFNET of routes from the regionals will include all routes to the multi-homed site.

Finally, it should also be noted that deployment of the new addressing plan described in this document may (and should) begin almost immediately but effective use of the plan to aggregate routing information will require changes to some Inter-Domain routing protocols. Likewise, deploying classless Inter-Domain protocols without deployment of the new address plan will not allow useful aggregation to occur (in other words, the addressing plan and routing protocol changes are both required for supernetting, and its resulting reduction in table growth, to be effective.) Note, however, that during the period of time between deployment of the addressing plan and deployment of the new protocols, the size of routing tables may temporarily grow very rapidly. This must be considered when planning the deployment of the two plans.

Note: in the discussion and examples which follow, the network and mask notation is used to represent routing destinations. This is used for illustration only and does not require that routing protocols use this representation in their updates.

2.2 Distributed allocation of address space

The basic idea of the plan is to allocate one or more blocks of Class C network numbers to each network service provider. Organizations using the network service provider for Internet connectivity are allocated bitmask-oriented subsets of the provider's address space as required.

It is also worthwhile to mention that once inter-domain protocols which support classless network destinations are widely deployed, the rules described by this plan generalize to permit arbitrary super/subnetting of the remaining class A and class B address space (the assumption being that classless inter-domain protocols will either allow for non-contiguous subnets to exist in the system or that all components of a sub-allocated class A/B will be contained

within a single routing domain). This will allow this plan to continue to be used in the event that the class C space is exhausted before implementation of a long-term solution is deployed. This alternative is discussed further below in section 6.

Hierarchical sub-allocation of addresses in this manner implies that clients with addresses allocated out of a given service provider are, for routing purposes, part of that service provider and will be routed via its infrastructure. This implies that routing information about multi-homed organizations, i.e., organizations connected to more than one network service provider, will still need to be known by higher levels in the hierarchy.

The advantages of hierarchical assignment in this fashion are

- a) It is expected to be easier for a relatively small number of service providers to obtain addresses from the central authority, rather than a much larger, and monotonically increasing, number of individual clients. This is not to be considered as a loss of part of the service providers' address space.
- b) Given the current growth of the Internet, a scalable and delegatable method of future allocation of network numbers has to be achieved.

For these reasons, and in the interest of providing a consistent procedure for obtaining Internet addresses, it is recommended that most, if not all, network numbers be distributed through service providers. These issues are discussed in much greater length in [2].

3. Cost-benefit analysis

This new method of assigning address through service providers can be put into effect immediately and will, from the start, have the benefit of distributing the currently centralized process of assigning new addresses. Unfortunately, before the benefit of reducing the size of globally-known routing destinations can be achieved, it will be necessary to deploy an Inter-Domain routing protocol capable of handling arbitrary network and mask pairs. Only then will it be possible to aggregate individual class C networks into larger blocks represented by single routing table entries.

This means that upon introduction, the new addressing allocation plan will not in and of itself help solve the routing table size problem. Once the new Inter-Domain routing protocol is deployed, however, an immediate drop in the number of destinations which clients of the new protocol must carry will occur. A detailed analysis of the magnitude

of this expected drop and the permanent reduction in rate of growth is given in the next section.

It should also be noted that the present method of flat address allocations imposes a large bureaucratic cost on the central address allocation authority. For scaling reasons unrelated to address space exhaustion or routing table overflow, this should be changed. Using the mechanism proposed in this paper will have the fortunate side effect of distributing the address allocation procedure, greatly reducing the load on the central authority.

3.1 Present Allocation Figures

An informal analysis of "network-contacts.txt" (available from the DDN NIC) indicates that as of 2/25/92, 46 of 126 class A network numbers have been allocated (leaving 81) and 5467 of 16382 class B numbers have been allocated, leaving 10915. Assuming that recent trends continue, the number of allocated class B's will continue to double approximately once a year. At this rate of growth, all class B's will be exhausted within about 15 months. As of 1/13/93, 52 class A network numbers have been allocated and 7133 class B's have been allocated. We suggest that the change in the class B allocation rate is due to the initial deployment of this address allocation plan.

3.2 Historic growth rates

MM/YY	ROUTES ADVERTISED	MM/YY	ROUTES ADVERTISED
Dec-92	8561	Sep-90	1988
Nov-92	7854	Aug-90	1894
Oct-92	7354	Jul-90	1727
Sep-92	6640	Jun-90	1639
Aug-92	6385	May-90	1580
Jul-92	6031	Apr-90	1525
Jun-92	5739	Mar-90	1038
May-92	5515	Feb-90	997
Apr-92	5291	Jan-90	927
Mar-92	4976	Dec-89	897
Feb-92	4740	Nov-89	837
Jan-92	4526	Oct-89	809
Dec-91	4305	Sep-89	745
Nov-91	3751	Aug-89	650
Oct-91	3556	Jul-89	603
Sep-91	3389	Jun-89	564
Aug-91	3258	May-89	516
Jul-91	3086	Apr-89	467
Jun-91	2982	Mar-89	410
May-91	2763	Feb-89	384
Apr-91	2622	Jan-89	346
Mar-91	2501	Dec-88	334
Feb-91	2417	Nov-88	313
Jan-91	2338	Oct-88	291
Dec-90	2190	Sep-88	244
Nov-90	2125	Aug-88	217
Oct-90	2063	Jul-88	173

Table I : Growth in routing table size, total numbers
Source for the routing table size data is MERIT

3.3 Detailed Analysis

There is a small technical cost and minimal administrative cost associated with deployment of the new address assignment plan. The administrative cost is basically that of convincing the NIC, the IANA, and the network service providers to agree to this plan, which is not expected to be too difficult. In addition, administrative cost for the central numbering authorities (the NIC and the IANA) will be greatly decreased by the deployment of this plan. To take advantage of aggregation of routing information, however, it is necessary that the capability to represent routes as arbitrary network and mask fields (as opposed to the current class A/B/C

distinction) be added to the common Internet inter-domain routing protocol(s). Thus, the technical cost is in the implementation of classless interdomain routing protocols.

3.3.1 Benefits of the new addressing plan

There are two benefits to be had by deploying this plan:

- o The current problem with depletion of the available class B address space can be ameliorated by assigning more-appropriately sized blocks of class C's to mid-sized organizations (in the 200-4000 host range).
- o When the improved inter-domain routing protocol is deployed, an immediate decrease in the number routing table entries should occur, followed by a significant reduction in the rate growth of routing table size (for default-free routers).

3.3.2 Growth rate projections

As of Jan '92, a default-free routing table (for example, the routing tables maintained by the routers in the NSFNET backbone) contained approximately 4700 entries. This number reflects the current size of the NSFNET routing database. Historic data shows that this number, on average, has doubled every 10 months between 1988 and 1991. Assuming that this growth rate is going to persist in the foreseeable future (and there is no reason to assume otherwise), we expect the number of entries in a default-free routing table to grow to approximately 30000 in two years time. In the following analysis, we assume that the growth of the Internet has been, and will continue to be, exponential.

It should be stressed that these projections do not consider that the current shortage of class B network numbers may increase the number of instances where many class C's are used rather than a class B. Using an assumption that new organizations which formerly obtained class B's will now obtain somewhere between 4 and 16 class C's, the rate of routing table growth can conservatively be expected to at least double and probably quadruple. This means the number of entries in a default-free routing table may well exceed 10,000 entries within six months and 20,000 entries in less than a year.

As of Dec '92, the routing table contains 8500 routes. The original growth curves would predict over 9400 routes. At this time, it is not clear if this would indicate a significant change in the rate of growth.

Under the proposed plan, growth of the routing table in a default-

free router is greatly reduced since most new address assignment will come from one of the large blocks allocated to the service providers. For the sake of this analysis, we assume prompt implementation of this proposal and deployment of the revised routing protocols. We make the initial assumption that any initial block given to a provider is sufficient to satisfy its needs for two years.

Since under this plan, multi-homed networks must continue to be explicitly advertised throughout the system (according to Rule #1 described in section 4.2), the number multi-homed routes is expected to be the dominant factor in future growth of routing table size, once the supernetting plan is applied.

Presently, it is estimated that there are fewer than 100 multi-homed organizations connected to the Internet. Each such organization's network is comprised of one or more network numbers. In many cases (and in all future cases under this plan), the network numbers used by an organization are consecutive, meaning that aggregation of those networks during route advertisement may be possible. This means that the number of routes advertised within the Internet for multi-homed networks may be approximated as the total number of multi-homed organizations. Assuming that the number of multi-homed organization will double every year (which may be a over-estimation, given that every connection costs money), the number of routes for multi-homed networks would be expected to grow to approximately 800 in three years.

If we further assume that there are approximately 100 service providers, then each service provider will also need to advertise its block of addresses. However, due to aggregation, these advertisements will be reduced to only 100 additional routes. We assume that after the initial two years, new service providers combined with additional requests from existing providers will require an additional 50 routes per year. Thus, the total is $4700 + 800 + 150 = 5650$. This represents an annual growth rate of approximately 6%. This is in clear contrast to the current annual growth of 130%. This analysis also assumes an immediate deployment of this plan with full compliance. Note that this analysis assumes only a single level of route aggregation in the current Internet - intelligent address allocation should significantly improve this.

Clearly, this is not a very conservative assumption in the Internet environment nor can 100% adoption of this proposal be expected. Still, with only a 90% participation in this proposal by service providers, at the end of the target three years, global routing table size will be "only" $4700 + 800 + 145 + 7500 = 13145$ routes -- without any action, the routing table will grow to approximately 75000 routes during that time period.

4. Changes to inter-domain routing protocols and practices

In order to support supernetting efficiently, it is clear that some changes will need to be made to both routing protocols themselves and to the way in which routing information is interpreted. In the case of "new" inter-domain protocols, the actual protocol syntax changes should be relatively minor. This mechanism will not work with older inter-domain protocols such as EGP2; the only ways to interoperate with old systems using such protocols are either to use existing mechanisms for providing "default" routes or b) require that new routers talking to old routers "explode" supernet information into individual network numbers. Since the first of these is trivial while the latter is cumbersome (at best -- consider the memory requirements it imposes on the receiver of the exploded information), it is recommended that the first approach be used -- that older systems continue to the mechanisms they currently employ for default handling.

Note that a basic assumption of this plan is that those organizations which need to import "supernet" information into their routing systems must run IGPs (such as OSPF [1]) which support classless routes. Systems running older IGPs may still advertise and receive "supernet" information, but they will not be able to propagate such information through their routing domains.

4.1 Protocol-independent semantic changes

There are two fundamental changes which must be applied to Inter-Domain routing protocols in order for this plan to work. First, the concept of network "class" needs to be deprecated - this plan assumes that routing destinations are represented by network and mask pairs and that routing is done on a longest-match basis (i.e., for a given destination which matches multiple network+mask pairs, the match with the longest mask is used). Second, current inter-domain protocols generally do not support the concept of route aggregation, so the new semantics need to be implemented in a new set of inter-domain protocols. In particular, when doing aggregation, dealing with multi-homed sites or destinations which change service providers is difficult. Fortunately, it is possible to define several fairly simple rules for dealing with such cases.

4.2. Rules for route advertisement

1. Routing to all destinations must be done on a longest-match basis only. This implies that destinations which are multi-homed relative to a routing domain must always be explicitly announced into that routing domain - they cannot be summarized (this makes intuitive sense - if a network is multi-homed, all

of its paths into a routing domain which is "higher" in the hierarchy of networks must be known to the "higher" network).

2. A routing domain which performs summarization of multiple routes must discard packets which match the summarization but do not match any of the explicit routes which makes up the summarization. This is necessary to prevent routing loops in the presence of less-specific information (such as a default route). Implementation note - one simple way to implement this rule would be for the border router to maintain a "sink" route for each of its aggregations. By the rule of longest match, this would cause all traffic destined to components of the aggregation which are not explicitly known to be discarded.

Note that during failures, partial routing of traffic to a site which takes its address space from one service provider but which is actually reachable only through another (i.e., the case of a site which has change service providers) may occur because such traffic will be routed along the path advertised by the aggregated route. Rule #2 will prevent any real problem from occurring by forcing such traffic to be discarded by the advertiser of the aggregated route, but the output of "traceroute" and other similar tools will suggest that a problem exists within the service provider advertising the aggregate, which may be confusing to network operators (see the example in section 5.2 for details). Solutions to this problem appear to be challenging and not likely to be implementable by current Inter-Domain protocols within the time-frame suggested by this document. This decision may need to be revisited as Inter-Domain protocols evolve.

An implementation following these rules should also be generalized, so that an arbitrary network number and mask are accepted for all routing destinations. The only outstanding constraint is that the mask must be left contiguous. Note that the degenerate route 0.0.0.0 mask 0.0.0.0 is used as a default route and MUST be accepted by all implementations. Further, to protect against accidental advertisements of this route via the inter-domain protocol, this route should never be advertised unless there is specific configuration information indicating to do so.

Systems which process route announcements must also be able to verify that information which they receive is correct. Thus, implementations of this plan which filter route advertisements must also allow masks in the filter elements. To simplify administration, it would be useful if filter elements automatically allowed more specific network numbers and masks to pass in filter elements given for a more general mask. Thus, filter elements which looked like:

```
accept 128.32.0.0
accept 128.120.0.0
accept 134.139.0.0
deny 36.2.0.0
accept 36.0.0.0
```

would look something like:

```
accept 128.32.0.0 255.255.0.0
accept 128.120.0.0 255.255.0.0
accept 134.139.0.0 255.255.0.0
deny 36.2.0.0 255.255.0.0
accept 36.0.0.0 255.0.0.0
```

This is merely making explicit the network mask which was implied by the class A/B/C classification of network numbers.

4.3. How the rules work

Rule #1 guarantees that the routing algorithm used is consistent across implementations and consistent with other routing protocols, such as OSPF. Multi-homed networks are always explicitly advertised by every service provider through which they are routed even if they are a specific subset of one service provider's aggregate (if they are not, they clearly must be explicitly advertised). It may seem as if the "primary" service provider could advertise the multi-homed site implicitly as part of its aggregate, but the assumption that longest-match routing is always done causes this not to work.

Rule #2 guarantees that no routing loops form due to aggregation. Consider a mid-level network which has been allocated the 2048 class C networks starting with 192.24.0.0 (see the example in section 5 for more on this). The mid-level advertises to a "backbone" 192.24.0.0/255.248.0.0. Assume that the "backbone", in turn, has been allocated the block of networks 192.0.0.0/255.0.0.0. The backbone will then advertise this aggregate route to the mid-level. Now, if the mid-level loses internal connectivity to the network 192.24.1.0/255.255.255.0 (which is part of its aggregate), traffic from the "backbone" to the mid-level to destination 192.24.1.1 will follow the mid-level's advertised route. When that traffic gets to the mid-level, however, the mid-level *must not* follow the route 192.0.0.0/255.0.0.0 it learned from the backbone, since that would result in a routing loop. Rule #2 says that the mid-level may not follow a less-specific route for a destination which matches one of its own aggregated routes. Note that handling of the "default" route (0.0.0.0/0.0.0.0) is a special case of this rule - a network must not follow the default to destinations which are part of one of its aggregated advertisements.

4.4. Responsibility for and configuration of aggregation

The domain which has been allocated a range of addresses has the sole authority for aggregation of its address space. In the usual case, the AS will install manual configuration commands in its border routers to aggregate some portion of its address space. An domain can also delegate aggregation authority to another domain. In this case, aggregation is done in the other domain by one of its border routers.

When an inter-domain border router performs route aggregation, it needs to know the range of the block of IP addresses to be aggregated. The basic principle is that it should aggregate as much as possible but not to aggregate those routes which cannot be treated as part of a single unit due to multi-homing, policy, or other constraints.

One mechanism is to do aggregation solely based on dynamically learned routing information. This has the danger of not specifying a precise enough range since when a route is not present, it is not always possible to distinguish whether it is temporarily unreachable or that it does not belong in the aggregate. Purely dynamic routing also does not allow the flexibility of defining what to aggregate within a range. The other mechanism is to do all aggregation based on ranges of blocks of IP addresses preconfigured in the router. It is recommended that preconfiguration be used, since it more flexible and allows precise specification of the range of destinations to aggregate.

Preconfiguration does require some manually-maintained configuration information, but not excessively more so than what router administrators already maintain today. As an addition to the amount of information that must be typed in and maintained by a human, preconfiguration is just a line or two defining the range of the block of IP addresses to aggregate. In terms of gathering the information, if the advertising router is doing the aggregation, its administrator knows the information because the aggregation ranges are assigned to its domain. If the receiving domain has been granted the authority to and task of performing aggregation, the information would be known as part of the agreement to delegate aggregation. Given that it is common practice that a network administrator learns from its neighbor which routes it should be willing to accept, preconfiguration of aggregation information does not introduce additional administrative overhead.

Implementation note: aggregates which encompass the class D address space (multicast addresses) are currently not well understood. At present, it appears that the optimal strategy is to consider

aggregates to never encompass class D space, even if they do so numerically.

4.5 Intra-domain protocol considerations

While no changes need be made to internal routing protocols to support the advertisement of aggregated routing information between autonomous systems, it is often the case that external routing information is propagated within interior protocols for policy reasons or to aid in the propagation of information through a transit network. At the point when aggregated routing information starts to appear in the new exterior protocols, this practice of importing external information will have to be modified. A transit network which imports external information will have to do one of:

- a) use an interior protocol which supports aggregated routing
- b) find some other method of propagating external information which does not involve flooding it through the interior protocol (i.e., by the use of internal BGP, for example).
- c) stop the importation of external information and flood a "default" route through the internal protocol for discovery of paths to external destinations.

For case (a), the modifications necessary to a routing protocol to allow it to support aggregated information may not be simple. For protocols such as OSPF and IS-IS, which represent routing information as either a destination+mask (OSPF) or as a prefix+prefix-length (IS-IS) changes to support aggregated information are conceptually fairly simple; for protocols which are dependent on the class-A/B/C nature of networks or which support only fixed-sized subnets, the changes are of a more fundamental nature. Even in the "conceptually simple" cases of OSPF and IS-IS, an implementation may need to be modified to support supernets in the database or in the forwarding table.

5. Example of new allocation and routing

5.1 Address allocation

Consider the block of 2048 class C network numbers beginning with 192.24.0.0 (0xC0180000 and ending with 192.31.255.0 (0xC01FFF00) allocated to a single network provider, "RA". A "supernetted" route to this block of network numbers would be described as 192.24.0.0 with mask of 255.248.0.0 (0xFFFF8000).

Assume this service provider connects six clients in the following order (significant because it demonstrates how temporary "holes" may form in the service provider's address space):

- "C1" requiring fewer than 2048 addresses (8 class C networks)
- "C2" requiring fewer than 4096 addresses (16 class C networks)
- "C3" requiring fewer than 1024 addresses (4 class C networks)
- "C4" requiring fewer than 1024 addresses (4 class C networks)
- "C5" requiring fewer than 512 addresses (2 class C networks)
- "C6" requiring fewer than 512 addresses (2 class C networks)

In all cases, the number of IP addresses "required" by each client is assumed to allow for significant growth. The service provider allocates its address space as follows:

- C1: allocate 192.24.0 through 192.24.7. This block of networks is described by the "supernet" route 192.24.0.0 and mask 255.255.248.0
- C2: allocate 192.24.16 through 192.24.31. This block is described by the route 192.24.16.0, mask 255.255.240.0
- C3: allocate 192.24.8 through 192.24.11. This block is described by the route 192.24.8.0, mask 255.255.252.0
- C4: allocate 192.24.12 through 192.24.15. This block is described by the route 192.24.12.0, mask 255.255.252.0
- C5: allocate 192.24.32 and 192.24.33. This block is described by the route 192.24.32.0, mask 255.255.254.0
- C6: allocate 192.24.34 and 192.24.35. This block is described by the route 192.24.34.0, mask 255.255.254.0

Note that if the network provider uses an IGP which can support classless networks, he can (but doesn't have to) perform "supernetting" at the point where he connects to his clients and therefore only maintain six distinct routes for the 36 class C network numbers. If not, explicit routes to all 36 class C networks will have to be carried by the IGP.

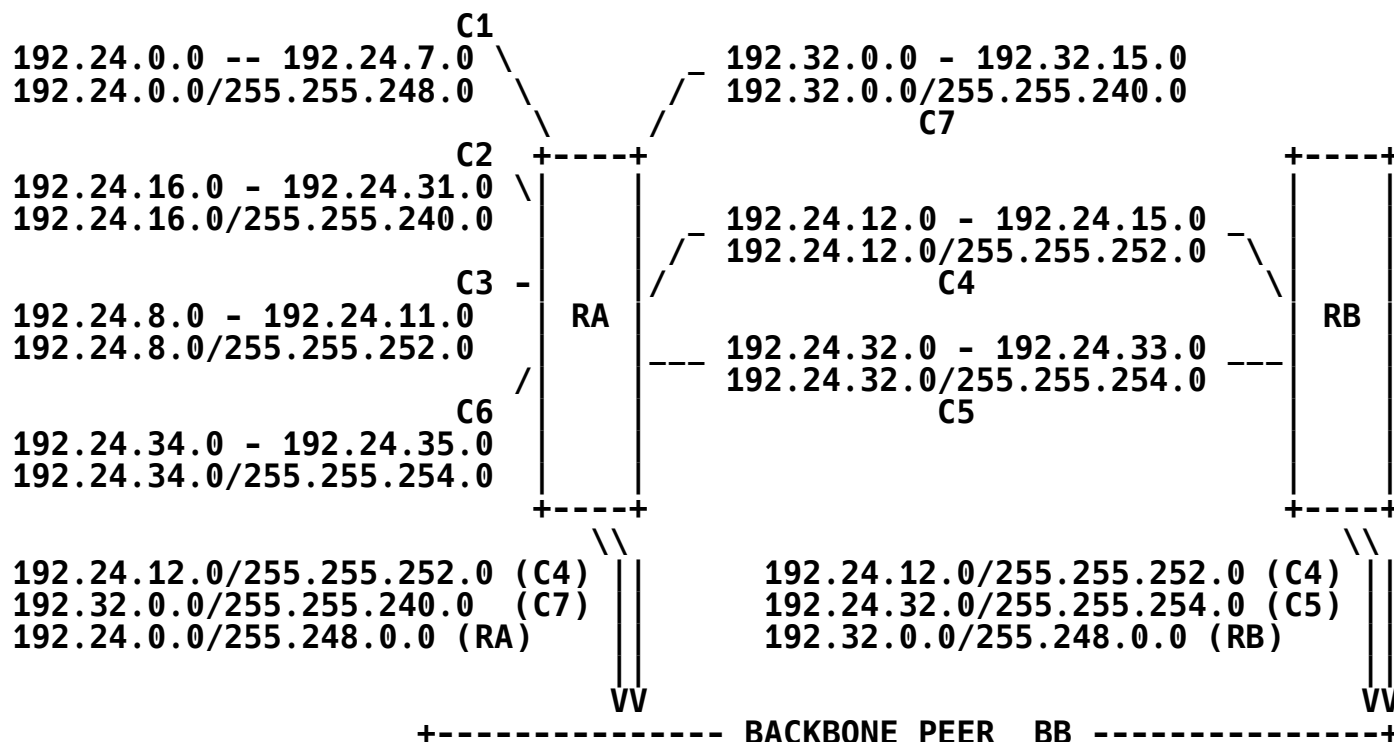
To make this example more realistic, assume that C4 and C5 are multi-homed through some other service provider, "RB". Further assume

the existence of a client "C7" which was originally connected to "RB" but has moved to "RA". For this reason, it has a block of network numbers which are allocated out "RB"'s block of (the next) 2048 class C network numbers:

C7: allocate 192.32.0 through 192.32.15. This block is described by the route 192.32.0, mask 255.255.240.0

For the multi-homed clients, we will assume that C4 is advertised as primary via "RA" and secondary via "RB"; C5 is primary via "RB" and secondary via "RA". To connect this mess together, we will assume that "RA" and "RB" are connected via some common "backbone" provider "BB".

Graphically, this simple topology looks something like this:



5.2 Routing advertisements

To follow rule #1, RA will need to advertise the block of addresses that it was given and C7. Since C4 is multi-homed and primary through RA, it must also be advertised. C5 is multi-homed and primary through RB. It need not be advertised since longest match by BB will automatically select RB as primary and the advertisement of

RA's aggregate will be used as a secondary.

Advertisements from "RA" to "BB" will be:

192.24.12.0/255.255.252.0	primary	(advertises C4)
192.32.0.0/255.255.240.0	primary	(advertises C7)
192.24.0.0/255.248.0.0	primary	(advertises remainder of RA)

For RB, the advertisements must also include C4 and C5 as well as it's block of addresses. Further, RB may advertise that C7 is unreachable.

Advertisements from "RB" to "BB" will be:

192.24.12.0/255.255.252.0	secondary	(advertises C4)
192.24.32.0/255.255.254.0	primary	(advertises C5)
192.32.0.0/255.248.0.0	primary	(advertises remainder of RB)

To illustrate the problem alluded to by the "note" in section 4.2, consider what happens if RA loses connectivity to C7 (the client which is allocated out of RB's space). In a stateful protocol, RA will announce to BB that 192.32.0.0/255.255.240.0 has become unreachable. Now, when BB flushes this information out of its routing table, any future traffic sent through it for this destination will be forwarded to RB (where it will be dropped according to Rule #2) by virtue of RB's less specific match 192.32.0.0/255.248.0.0. While this does not cause an operational problem (C7 is unreachable in any case), it does create some extra traffic across "BB" (and may also prove confusing to a network manager debugging the outage with "traceroute"). A mechanism to cache such unreachability information would help here, but is beyond the scope of this document (such a mechanism is also not implementable in the near-term).

6. Extending CIDR to class A addresses

At some point, it is expected that this plan will eventually consume all of the remaining class C address space. As of this writing, the upper half of the class A address space has already been reserved for future expansion. This section describes how the CIDR plan can be used to utilize this portion of the class A space efficiently. It is expected that this contingency would only be used if no long term solution has become apparent by the time that the class C address space is consumed.

Fundamentally, there are two differences between using a class A address and a block of class C's. First, the configuration of DNS becomes somewhat more complicated than it is without the aggregation of class A subnets. The second difference is that the routers within

the class A address would need to support and use a classless IGP.

Maintenance of DNS with a subnetted class A is somewhat painful. As part of the mechanism for providing reverse address lookups, DNS maintains a "IN-ADDR.ARPA" reverse domain. This is configured by reversing the dotted decimal network number, appending "IN-ADDR.ARPA" and using this as a type of pseudo-domain. Individual hosts then end up pointing back to a host name. Thus, for example, 131.108.1.111 has a DNS record "111.1.108.131.IN-ADDR.ARPA." Since the pseudo-domains can only be delegated on a byte boundary, this becomes painful if a stub domain receives a block of address space that does not fall on a byte boundary. The solution in this case is to enumerate all of the possible byte combinations involved. This is painful, but workable. This is discussed further below.

Routing within a class A used for CIDR is also an interesting challenge. The usual case will be that a domain will be assigned a portion of the class A address space. The domain can either use an IGP which allows variable length subnets or it can pick a single subnet mask to be used throughout the domain. In the latter case, difficulties arise because other domains have been allocated other parts of the class A address space and may be using a different subnet mask. If the domain is itself a transit, it may also need to allocate some portion of its space to a client, which might also use a different subnet mask. The client would then need routing information about the remainder of the class A.

If the client's IGP does not support variable length subnet masks, this could be done by advertising the remainder of the class A's address space in appropriately sized subnets. However, unless the client has a very large portion of the class A space, this is likely to result in a large number of subnets (for example, a mask of 255.255.255.0 would require a total of 65535 subnets, including those allocated to the client). For this reason, it may be preferable to simply use an IGP that supports variable length subnet masks within the client's domain.

Similarly, if a transit has been assigned address space from a class A network number, it is likely that it was not assigned the entire class A, and that other transit domains will get address space from this class A. In this case, the transit would also have to inject routing information about the remainder of the class A into it's IGP. This is analogous to the situation above, with the same complications. For this reason, we recommend that the use of a class A for CIDR only be attempted if IGP's with variable length subnet mask support be used throughout the class A. Note that the IGP's need not support supernetting, as discussed above.

Note that the technique here could also apply to class B addresses. However, the limited number of available class B addresses and their usage for multihomed networks suggests that this address space should only be reserved for those large single organizations that warrant this type of address. [2]

7. Domain Service considerations

One aspect of Internet services which will be notably affected by a move to either "supernetted" class-C network numbers or subdivided class-A's will be the mechanism used for address-to-name translation: the IN-ADDR.ARPA zone of the domain system. Because this zone is delegated on octet boundaries only, any address allocation plan which uses bitmask-oriented addressing will cause some degree of difficulty for those which maintain parts of the IN-ADDR.ARPA zone.

7.1 Procedural changes for class-C "supernets"

At the present time, parts of the IN-ADDR.ARPA zone are delegated only on network boundaries which happen to fall on octet boundaries. To aid in the use of blocks of class-C networks, it is recommended that this policy be relaxed and allow the delegation of arbitrary, octet-oriented pieces of the IN-ADDR.ARPA zone.

As an example of this policy change, consider a hypothetical large network provider named "BigNet" which has been allocated the 1024 class-C networks 199.0.0 through 199.3.255. Under current policies, the root domain servers would need to have 1024 entries of the form:

```
0.0.199.IN-ADDR.ARPA.  IN      NS      NS1.BIG.NET.
1.0.199.IN-ADDR.ARPA.  IN      NS      NS1.BIG.NET.
...
255.3.199.IN-ADDR.ARPA. IN      NS      NS1.BIG.NET.
```

By revising the policy as described above, this is reduced only four delegation records:

```
0.199.IN-ADDR.ARPA.    IN      NS      NS1.BIG.NET.
1.199.IN-ADDR.ARPA.    IN      NS      NS1.BIG.NET.
2.199.IN-ADDR.ARPA.    IN      NS      NS1.BIG.NET.
3.199.IN-ADDR.ARPA.    IN      NS      NS1.BIG.NET.
```

The provider would then maintain further delegations of naming authority for each individual class-C network which it assigns, rather than having each registered separately. Note that due to the way the DNS is designed, it is still possible for the root nameservers to maintain the delegation information for individual networks for which the provider is unwilling or unable to do so. This should greatly reduce the load on the domain servers for the "top" levels of the IN-ADDR.ARPA domain. The example above illustrates only the records for a single nameserver. In the normal case, there are usually several nameservers for each domain, thus the size of the examples will double or triple in the common cases.

7.2 Procedural changes for class-A subnetting

Should it be the case the class-A network numbers are subdivided into blocks allocated to transit network providers, it will be similarly necessary to relax the restriction on how IN-ADDR.ARPA naming works for them. As an example, take a provider is allocated the 19-bit portion of address space which matches 10.8.0.0 with mask 255.248.0.0. This represents all addresses which begin with the prefixes 10.8, 10.9, 10.10, 10.11, 10.12, 10.13, 10.14, an 10.15 and requires the following IN-ADDR.ARPA delegations:

8.10.IN-ADDR.ARPA.	IN	NS	NS1.MOBY.NET.
9.10.IN-ADDR.ARPA.	IN	NS	NS1.MOBY.NET.
....			
15.10.IN-ADDR.ARPA.	IN	NS	NS1.MOBY.NET.

To further illustrate how IN-ADDR.ARPA sub-delegation will work, consider a company named "F00" connected to this provider which has been allocated the 14-bit piece of address space which matches 10.10.64.0 with mask 255.255.192.0. This represents all addresses in the range 10.10.64.0 through 10.10.127.255 and will require that the provider implement the following IN-ADDR.ARPA delegations:

64.10.10.IN-ADDR.ARPA.	IN	NS	NS1.F00.COM.
65.10.10.IN-ADDR.ARPA.	IN	NS	NS1.F00.COM.
....			
127.10.10.IN-ADDR.ARPA.	IN	NS	NS1.F00.COM.

with the servers for "F00.COM" containing the individual PTR records for all of the addresses on each of these subnets.

8. Transitioning to a long term solution

This solution does not change the Internet routing and addressing architectures. Hence, transitioning to a more long term solution is not affected by the deployment of this plan.

9. Conclusions

We are all aware of the growth in routing complexity, and the rapid increase in allocation of network numbers. Given the rate at which this growth is being observed, we expect to run out in a few short years.

If the inter-domain routing protocol supports carrying network routes with associated masks, all of the major concerns demonstrated in this paper would be eliminated.

One of the influential factors which permits maximal exploitation of the advantages of this plan is the number of people who agree to use it.

If service providers start charging networks for advertising network numbers, this would be a very great incentive to share the address space, and hence the associated costs of advertising routes to service providers.

10. Recommendations

The NIC should begin to hand out large blocks of class C addresses to network service providers. Each block must fall on bit boundaries and should be large enough to serve the provider for two years. Further, the NIC should distribute very large blocks to continental and national network service organizations to allow additional levels of aggregation to take place at the major backbone networks. In addition, the NIC should modify its procedures for the IN-ADDR.ARPA domain to permit delegation along arbitrary octet boundaries.

Service providers will further allocate power-of-two blocks of class C addresses from their address space to their subscribers.

All organizations, including those which are multi-homed, should obtain address space from their provider (or one of their providers, in the case of the multi-homed). These blocks should also fall on bit boundaries to permit easy route aggregation.

To allow effective use of this new addressing plan to reduce propagated routing information, appropriate IETF WGs will specify the modifications needed to Inter-Domain routing protocols.

Implementation and deployment of these modifications should occur as quickly as possible.

11 References

- [1] Moy, J, "The OSPF Specification Version 2", RFC 1247, Proteon, Inc., January 1991.
- [2] Rekhter, Y., and T. Li, "An Architecture for IP Address Allocation with CIDR", RFC 1518, T.J. Watson Research Center, IBM Corp., cisco Systems, September 1993.

12. Security Considerations

Security issues are not discussed in this memo.

13. Authors' Addresses

Vince Fuller
BARRNet
Pine Hall 115
Stanford, CA, 94305-4122

EMail: vaf@Stanford.EDU

Tony Li
cisco Systems, Inc.
1525 O'Brien Drive
Menlo Park, CA 94025

EMail: tli@cisco.com

Jessica (Jie Yun) Yu
Merit Network, Inc.
1071 Beal Ave.
Ann Arbor, MI 48109

EMail: jyy@merit.edu

Kannan Varadhan
Internet Engineer, OARnet
1224, Kinnear Road,
Columbus, OH 43212

EMail: kannan@oar.net