

Internet-Draft  
Expiration Date: May 1997

Richard Woundy  
Arun Viswanathan  
Nancy Feldman  
Rick Boivie  
IBM Corp.

November 1996

## **ARIS: Aggregate Route-Based IP Switching**

[<draft-woundy-aris-ipswitching-00.txt>](#)

### Status of This Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

To learn the current status of any Internet-Draft, please check the "1id-abstracts.txt" listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), nic.nordu.net (Europe), munnari.oz.au (Pacific Rim), ds.internic.net (US East Coast), or ftp.isi.edu (US West Coast).

### Abstract

IP based networks use a number of routing protocols, including RIP, OSPF, IS-IS, and BGP, to determine how packets ought to be routed. Among these protocols, OSPF and BGP are IETF-recommended standards that have been extensively deployed and exercised in many networks. In this memo, we describe a mechanism which uses these protocols as the basis for switching IP datagrams, by the addition of a simple

protocol ("ARIS") that establishes switched paths through a network. The ARIS protocol allows us to leverage the advantages of switching technologies in an internet network.

This memo is defined with respect to ATM. ARIS can be easily extended to other switching technologies.

## **1. Introduction**

In this paper, an Integrated Switch Router (ISR), is a standard IP router that has been augmented with ATM virtual circuit (VC) switching support. The ISR at an entry point to the ATM switching environment performs standard IP forwarding of datagrams, but the "next hop" of the IP forwarding table has been extended to include a reference to a VC (switched path). Each VC may have an endpoint at a neighboring router (comparable to today's IP next hops on conventional routers), or may traverse a series of ISRs along the best IP forwarding path, to an egress ISR endpoint. This allows datagrams to be switched at hardware speeds through an entire ISR network.

The key link between the IP network routing protocols and the ARIS VC establishment protocol is the "egress identifier". The egress identifier refers to an egress ISR that forwards traffic either to a foreign routing domain, or across an area boundary within the same network. ARIS establishes VCs towards each unique egress identifier. Since thousands of IP destinations can map to the same egress identifier, ARIS minimizes the number of VCs required in an ISR network. This allows a large network to switch all of its IP traffic, resulting in improved aggregate IP throughput.

Egress ISRs initiate the setup of VCs by sending Establishment messages to their upstream neighbors, typically within the same domain. These upstream neighbors forward the messages to their own upstream neighbors in Reverse Path Multicast style, after ensuring that the VC path is loop-free. Eventually, all ISRs establish VCs to all egress ISRs.

The VC to an egress point, in general, takes the form of a tree. A tree results because of the "merging" of VCs that occurs at a node when multiple upstream VCs for a given egress point are spliced to a single downstream VC for that egress point.

## **2. VC Conservation**

An important goal of the ARIS protocol is to minimize the number of



VCs required by ISRs to switch all IP traffic in the switching cloud. Since switches may support only a limited VPI/VCI space, ARIS restrains its VC consumption so that VCs are available as needed for its own use, as well as for ATM services and other applications, such as RSVP. Further benefits include simplification of network management, both for automated tools and for human comprehension and analysis, and VC-setup overhead.

The consumption of VCs is minimized:

- o by the use of egress routers that may map thousands of IP destinations to the same VC, and
- o by enabling the merging of VCs.

The merging of VCs enables ARIS to create VC trees, each of which connects all of the ingresses to a given egress. This results in  $n$  trees, where  $n$  is the number of egresses in a network, while still providing the benefits of full mesh connectivity (without  $O(n^2)$  VCs).

The network routing domain has the greatest performance and VC conservation when all routers in the domain are ISRs. Maximum ARIS benefits are also tied closely to an IP network routing topology with a high ratio of IP destinations to egresses, as exists in a typical IP backbone. However, ARIS is flexible enough to be highly beneficial even in networks with partial ISR deployments or arbitrary network routing topologies.

The ability of the ARIS protocol to conserve the number of VCs depends on the hardware capabilities of the ISR. Some ATM switching components can "merge" multiple inbound VCs onto one outbound VC at close to standard switching rates. These merge-capable components are able to reassemble cells from the inbound VCs into frames, and inject the frames into the outbound VC, without interleaving cells from different frames.

### **3. Loop Prevention**

The ARIS protocol guarantees that VC loops are prevented, even in the presence of transient IP routing loops. With datagram forwarding loops, each hop decrements the TTL, so traffic is eventually dropped. However, ATM switching does not have a counter similar to the TTL, so traffic persists in a VC loop as long as the VC loop exists. At best, the traffic in the VC loop steals bandwidth from other (UBR) VCs; at worst, the traffic interferes with IP routing traffic, slows down routing convergence, and lengthens the life of the VC loop.



The ARIS protocol avoids creating VC loops by the use of an "ISR ID" list, similar in function to the BGP AS\_PATH attribute. Each ISR in the VC establishment path appends its own unique ISR ID to each establishment message it forwards. In this way, an ISR is able to determine the path a message has traversed, and can ensure that no loops are formed.

Further, if an ISR modifies or deletes an egress due to an IP route change, or receives a message that modifies an existing VC to an egress, the ISR must unsplice any established upstream VC from the downstream VC. This unsplicing forces inbound traffic to be forwarded at the IP network layer, so that transient IP routing loops, potentially created by the route change, cannot produce VC loops. The ISR must then re-establish a new VC to the modified egress. Note that ARIS does not attempt to suppress transient IP routing protocol loops; it only avoids establishing VC loops with this information.

#### 4. ARIS Messages

The ARIS protocol runs directly over IP, using IP protocol TBD. The following messages are used to manage the ISR switching cloud:

##### Init

This is the first message sent by an ISR to each of its neighbors, as notification of its existence. It is periodically transmitted until a positive acknowledgment is received. The Init message may include the neighbor timeout period, acceptable VC label range, encapsulation scheme [1], and other adjacency information.

##### KeepAlive

This message is sent by an ISR to inform its neighbors of its continued existence. It is the first message that is transmitted after an adjacency has been established. In order to prevent the neighbor timeout period from expiring, ARIS messages must be periodically sent to neighbors. The VC KeepAlive will only be sent when no other ARIS messages have been transmitted within the periodic interval time. A recommended transmission interval is 1/3 of the exchanged neighbor timeout period.

##### Establishment

This message is initiated by the egress ISR, and is periodically sent to each upstream neighbor to setup or refresh a VC. It is also sent by any ISR in response to a Trigger message. Each ISR that receives an Establishment message for an egress identifier must verify that the path is correct and loop free. If the Establishment message changes a previously known VC path to the



egress identifier, the ISR unsplices the obsolete VC. The ISR creates a downstream VC for the egress identifier, and replies with an Acknowledgment message. It then creates a VC for each of its upstream neighbors, forwards the Establishment message to the upstream neighbors with its unique ISR ID appended to the ISR ID path and the VC label for the upstream neighbor to use for forwarding, and waits for an Acknowledgment message. This pattern continues until all ISRs are reached.

#### Trigger

This message is sent by an ISR when it has detected that a local IP routing change has modified its path to the egress identifier. After unsplicing the obsolete VC, the ISR sends a Trigger message to its new downstream neighbor requesting an Establishment message.

#### Teardown

This message may be sent when an ISR has lost all connectivity to an egress identifier, or when a downstream node to an egress identifier has become an upstream node due to routing changes. In the former case, the message traverses the upstream ISR paths of the VC, unsplicing each VC along the way. In the latter case, the message is sent single hop to the new upstream (previously downstream) node, unsplicing the obsolete VC.

#### Acknowledgment

This message is sent as a response to ARIS messages. When an ISR receives a positive acknowledgment to Init message, it responds with a KeepAlive message. When an ISR receives a positive acknowledgment to an Establishment message, it splices the upstream VC to the downstream VC.

## **5. ISR Information Bases**

The ISR needs three logical information bases to compute routes and forward datagrams: the routing information base, the forwarding information base, and the VC information base. The first, the routing information base (RIB), is used for the computation of best-effort routes by various IP routing protocols. The RIB for the ISR is essentially unchanged from the RIB on a standard router. In the ISR context, the RIB is also used to identify egress points and egress identifiers for the other two information bases.

The forwarding information base (FIB) of the ISR has been extended beyond the content of the FIB on a standard router to include an egress identifier in each next hop entry. The FIB tends to contain many IP destination prefix entries, which point to a small number of





next hop entries that describe the hop-by-hop forwarding operation(s). Next hop entries on the ISR consist of an outgoing interface, next hop IP address, egress identifier, and the associated established downstream VC. The association of the next hops with the egress identifiers is the responsibility of the routing protocols, while the association of the next hop/egress identifiers with the established VCs is the responsibility of the ARIS protocol.

The VC information base (VCIB), which does not exist on a standard router, maintains for each egress identifier the upstream to downstream VC mappings and related states. This mapping is controlled by the ARIS protocol.

## 6. Egress Identifiers

The ARIS protocol uses egress identifiers that balance the desire to share the same egress identifier among many IP destination prefixes, with the desire to maximize switching benefits. To provide flexibility, ARIS supports many types of egress identifiers. ISRs choose the type of egress identifier to use based on routing protocol information and local configuration.

The first type of egress identifier is the IP destination prefix. This type results in each IP destination prefix sustaining its own VC tree, and thus will not scale in large backbone and enterprise networks. However, this is the only information that some routing protocols, such as RIP, can provide. This type of identifier may work well in networks where the number of destination prefixes is limited, such as in campus environments, or even in a wide-area network of a private enterprise.

The second type of egress identifier is the egress IP address. This type is used primarily for BGP protocol updates, which carry this information in the NEXT\_HOP attribute [2]. There are certain types of OSPF routes that also use this type. See "BGP Interaction" and "OSPF Interaction" for detailed information.

The third type of egress identifier is the OSPF Router ID, which allows aggregation of traffic on behalf of multiple datagram protocols routed by OSPF. The latest version of OSPF, OSPFv3, supports the Router ID for both IP and IPv6 [3]. See "OSPF Interaction" for further information.

The fourth type of egress identifier is the multicast (source, group) pair [4], used by multicast protocols, such as DVMRP [5], MOSPF [6] and PIM ([7], [8]). The fifth is the (ingress-of-source, group), used for such multicast protocols as MOSPF and PIM. See "IP



Multicast Interaction" for IP multicast protocol details.

Other egress identifier types may be defined, including but not limited to IS-IS NSAP addresses, NLSP IPX addresses, IPv6 destination prefixes, etc.

A hierarchy amongst the egress identifiers may be introduced to allow more flexible control over egress identifier selection. This allows an ISR to autolearn or be configured with non-default egress identifiers, and to select which egress identifiers to use in various routing situations.

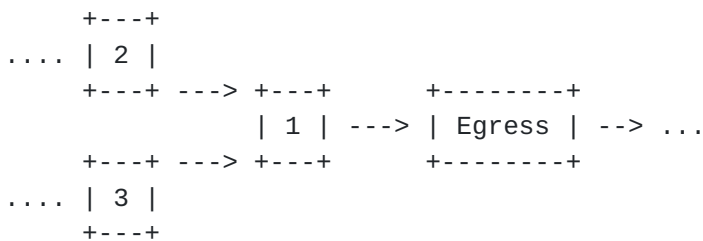
It should be noted that a network achieves a further performance optimization with ARIS when egress identifiers refer to the next hop router of the egress ISR. This allows datagrams to be switched entirely from the ingress point in the routing domain to the router past the egress ISR.

7. Egress ISRs

In the ARIS protocol, Establishment messages are originated from the egress ISR. An ISR is considered an egress ISR, with respect to a particular egress identifier, under any of the following conditions:

1. The egress identifier refers to the ISR itself (including one of its directly attached interfaces).
2. The egress identifier is reachable via a next hop router that is outside the ISR switching infrastructure.
3. The egress identifier is reachable by crossing a routing domain boundary, such as another area for OSPF summary networks, or another autonomous system for OSPF AS externals and BGP routes.

8. Establishment Initiation Example



Example: Egress initiates Establishment

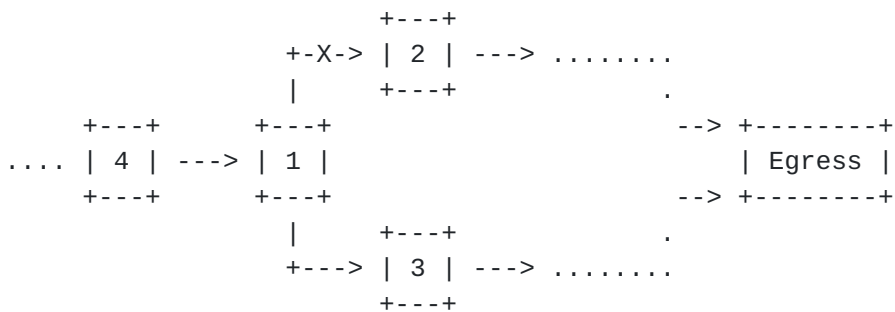


- a) The Egress ISR learns of an egress identifier that indicates the egress is itself (see "Egress ISRs"). It creates a FIB entry for its next hop and egress identifier (itself).
- b) The Egress creates a VCIB entry with an upstream VC to ISR1, and initiates a VC Establishment message with the upstream VC label to use, and itself in the ISR ID path.
- c) ISR1 verifies that the Establish message was received from the expected next hop (Egress) by matching its FIB entry, and that the ISR ID path is loop free. It then creates a VCIB entry and a downstream VC to the Egress with the given VC label, replaces the default VC in the FIB with this new value, and replies to the Egress with an Acknowledgment message.
- d) ISR1 creates an upstream VC to each of its upstream neighbors, ISR2 and ISR3, and updates the corresponding VCIB entry. It forwards the Establishment message to each upstream neighbor, with its own ISR ID appended to the ISR ID path and with the VC label to use.
- e) When ISR1 receives each acknowledgment from each upstream neighbor, it updates the VCIB and splices the corresponding upstream VC to its Egress downstream VC.

All upstream nodes recursively follow the same procedures as ISR1, until all Ingress nodes have been added to the VC path to the Egress.

The Egress ISR is responsible for periodically sending refresh VC Establishment messages, to prevent VC timeouts. If a refresh is not received in the allotted time, VCs are unspliced and discarded.

**9. Establishment Trigger Example**



Example: ISR1 changes routes from ISR2 to ISR3



- a) ISR1 learns of a new path to the Egress via ISR3 from the routing protocols. It removes the FIB and VCIB entries for the next hop ISR2/Egress. ISR1 creates a new FIB entry for the next hop ISR3/Egress with a default VC to the next hop.
- b) ISR1 sends a Trigger message to new downstream node ISR3 requesting an Establish message for the VC path to the Egress.
- c) ISR3 creates an upstream VC, updates its corresponding VCIB entry, and replies with an Establish message to ISR1, containing the full ISR ID path and the VC label.
- d) ISR1 verifies that the Establish message was received from the expected next hop (ISR3), and that the ISR ID path is loop free. It then creates a new VCIB entry and downstream VC to ISR3 with the given VC label, and replaces the default VC in the FIB with this new value.
- e) ISR1 sends an Acknowledgment message to ISR3.
- f) ISR3 receives the Acknowledgment, updates the VCIB and splices its ISR1 upstream VC to its downstream VC.
- g) ISR1 appends its ISR ID to the Establish message, and forwards the message to ISR4 with the upstream VC label.
- h) ISR4 verifies the Establish message, updates the VCIB, and unsplices the current VC to ISR1/Egress from its upstream node(s), and sends an Acknowledgment to ISR1.
- i) ISR1 receives the Acknowledgment, updates the VCIB and splices the ISR4 upstream VC to the ISR3 downstream VC.
- j) ISR4 appends its ISR ID to the path, and forwards the establishment message to its upstream neighbors with a VC label. When ISR4 receives an Acknowledgment from an upstream neighbor, it updates the VCIB and splices the upstream VC to the ISR1 downstream VC.

All upstream nodes recursively follow the same procedure as ISR4, until all ingress nodes have been updated.

## **10. TTL Decrement**

In order to comply with the requirements for IPv4 routers, the IP datagram Time-To-Live (TTL) field must be decremented on each hop it traverses [9]. Currently, switched packets within ATM cannot





decrement the TTL. However, ARIS can decrement TTLs appropriately by maintaining a hop-count per egress identifier. This hop-count is calculated by including a hop-count field in the Establish message, which is incremented at each ISR as it traverses the upstream path. Before forwarding a packet on a VC, an ingress ISR decrements the TTL by the hop-count plus one. If the decrement value is greater than or equal to the TTL of the packet, the packet is forwarded hop-by-hop.

## **11. Multipath Considerations**

Many IP routing protocols such as OSPF support the notion of equal-cost multipath routes, in which a router maintains multiple next hops for one destination prefix when two or more equal-cost paths to the prefix exist. Unfortunately, because of limitations in most ATM switching hardware, each path needs its own VC. Therefore, ingress ISRs may maintain a number of VCs to one egress ISR, each VC representing a different equal-cost path to the egress. In this case, the ingress ISR will make multipath decisions for traffic on behalf of all downstream ISRs.

Each ISR that receives multiple (legal) Establishment messages from downstream ISRs with different paths to the same egress identifier can choose one of four different approaches for sending Establishment messages upstream. One approach is to send multiple Establishment messages upstream, preserving multiple VCs to the egress ISR. Each Establishment message requires an additional numeric identifier to be able to distinguish multiple distinct VCs to the destination, so that successive Establishment messages for distinct VCs are not misinterpreted as consecutive replacements of the same VC. When multiple Establishment VCs are preserved upstream, they require distinct VPI/VCI assignments, which works against conservation of VCs.

Another approach, that conserves VCs at the cost of switching performance, is to originate one Establishment message upstream, and to forward datagrams at the IP network layer on the multipath point ISR.

A third approach is to propagate only one Establishment message from the downstream ISRs to the upstream ISRs, and ignore the content of other VC Establishment messages. This conserves VCs and maintains switching performance, but may not balance loads across downstream links as well as the first two approaches, even if VCs are selectively dropped.

The final approach is to propagate one Establishment message that carries the content of all downstream Establishment messages, so that



only one upstream VC is created to the multipath point. This requires that the ATM switching hardware on the multipath ISR be capable of correctly distributing the traffic of upstream VCs onto multiple downstream VCs. Furthermore, the Establishment message to send upstream must concatenate the ISR ID lists from downstream messages, in order to preserve the VC loop-free property. The ISR ID list concatenation is similar to using AS\_SETs for aggregation in the BGP protocol. This final approach has the benefit of both VC conservation and performance, although it requires a slightly more complex implementation.

## **12. BGP Interactions with ARIS**

The BGP implementation of the ISR uses the NEXT\_HOP attribute as the egress identifier. When the BGP border ISR injects routes into the BGP mesh, it may use its own IP address or the address of its external BGP peer as the value of the NEXT\_HOP attribute. This choice of NEXT\_HOP attribute value creates different Establishment behaviors with ARIS.

If the BGP border ISR uses its own IP address as the NEXT\_HOP attribute in its injected routes, then all of these BGP routes share the same egress identifier. This approach establishes only one tree to the BGP border ISR, and the ISR must forward traffic at the IP layer towards its external BGP neighbors.

If the BGP border ISR uses the external BGP peer as the NEXT\_HOP attribute in its injected routes, then the BGP routes from each unique external BGP neighbor share the same egress identifier. This approach establishes one VC tree per external BGP neighbor of the BGP border ISR. The BGP border ISR can switch traffic directly to its external BGP neighbors.

## **13. OSPF Interactions with ARIS**

The OSPF protocol exchanges five types of "link state advertisements" to create OSPF routing tables. All types of advertisements contain an "Advertising Router" field, which identifies the OSPF Router ID of the router that originates the advertisement. The ISR uses this OSPF Router ID as the egress identifier.

The use of the OSPF Router ID as an egress identifier allows a new level of destination prefix abstraction. In a typical network, a router may be connected to several LANs (Ethernets, Token Rings, etc.), and may communicate to remote networks outside of its routing domain via adjacent routers. The remote destination networks may be



injected into the link state routing domain via static configuration, or via other routing protocols (such as RIP or BGP). These local and remote networks may be represented in the router forwarding tables as many destination prefixes, which cannot be aggregated into shorter prefixes (even when using CIDR [10]). Router labels (OSPF Router ID) provide a compact means to represent a number of destination prefixes that exit the link state routing domain at the same egress router. The association between destination prefixes and router labels is an easy by-product of the normal SPF computation.

The one exception to using the OSPF Router ID is when ISRs receive an AS external link advertisement with a non-zero forwarding address. The OSPF protocol uses the forwarding address to allow traffic to bypass the router that originates the advertisement. Since the OSPF Router ID refers to the bypassed router, it is inadequate as an egress identifier in this case. Instead, the ARIS protocol must use the forwarding address as the egress identifier.

Using the forwarding address as the egress identifier provides significant benefits. Since the AS external forwarding address and the BGP NEXT\_HOP attribute are both external IP addresses, they are compatible types of egress identifiers, which may allow BGP and OSPF routes to share the same VC. Further, the OSPF AS boundary ISR can switch traffic directly to its external neighbors, just like BGP.

The ISR identifies itself as an OSPF egress when the ISR is an area border router or an AS boundary router, or when it is directly attached to a LAN.

#### **14. IP Multicast Interactions with ARIS**

The ARIS protocol can be used to setup VCs for IP multicast traffic, in particular for multicast protocols using Reverse Path Multicasting (RPM). The typical RPM forwarding information base maps a source IP network address and multicast group pair, (S,G), to an expected incoming interface and a set of outgoing interfaces. The ISR extends the forwarding information base to include one egress identifier per (S,G).

The current choice of egress identifier is the (S,G) pair itself. This egress identifier creates one source-based VC tree per source address and group pair. The VC tree carries traffic from the ingress ISR to all egress ISRs, using multicast switching within intermediate ISRs. Egress ISRs for multicast are similar to egress ISRs for the unicast case, except that multicast egress ISRs are determined by group membership location, instead of egress point reachability. An ISR becomes an egress for a particular (S,G) when it forwards traffic



from a source S to a group G over a non-ISR link.

Having multicast VCs set up on the basis of (S,G) works well with the IGMPv3 Group-Source messages, since these IGMP messages can create unique trees for each sender within the same group [[11](#)].

An alternative egress identifier choice is to use the "ingress" of the source address S in the (S,G) pair. This choice creates one ingress-based VC tree for all of the sources of a group that share a given ingress, instead of one for each of the sources. This permits a greater amount of VC aggregation in the ISR cloud. The ingress of a source address is calculated in a similar fashion to calculating an egress identifier for a destination prefix. Unfortunately, one cannot calculate useful ingress identifiers for DVMRP, for the same reason that one cannot calculate useful egress identifiers for RIP.

Furthermore, since some protocols permit source-specific multicast pruning, the multicast distribution tree for a particular group may differ according to source address, even if sources share the same ingress point. However, the advantages this approach offers with regards to VC conservation on those protocols capable of supporting the ingress of source may outweigh the disadvantages of wasting bandwidth by sending traffic to leaf networks where a particular source may be filtered.

Based on the topology of the multicast distribution tree, there may be multiple egress ISRs for the egress identifier (S,G). Each ISR can send one multicast Establishment message to the one upstream ISR on the path back toward the source address. The ISR ID lists of multicast downstream ISRs, with the current ISR ID, are concatenated (like BGP AS\_SETs) before sending the Establishment message to the upstream ISR.

The observant reader may note that ARIS uses a multicast scheme to build unicast VCs, and a unicast scheme to build multicast VCs.

## **[15. Virtual Path Extension](#)**

The ARIS usage of "merged" VC flows requires that ATM switching hardware have the capability of preventing cell interleaving (see "VC Conservation"). Unfortunately, much of the existing ATM switching hardware cannot support VC merging. One solution to this problem is to use virtual paths (VPs) to egress points, rather than virtual circuits (VCs). The virtual path extension merges VPs, creating trees of VPs to the egress points, instead of merging VCs. Cell interleaving is prevented by the assignment of unique VC identifiers (VCIs) within each VP.





The ISRs within a network are assigned unique VCIs to prevent VCI collisions when paths from different ISRs are merged. Each ISR requires a block of VCIs as labels to distinguish between cell paths to the same egress identifier. By assigning a unique block of VCIs to each ISR, ARIS guarantees that an ISR at a network merge point can safely merge upstream VP flows for an egress identifier to a single downstream VP without VCI collisions.

Although the virtual path extension uses VCs much less efficiently than a VC merging implementation, it reduces network latency and hardware requirements because frame reassembly and re-segmentation is not required on intermediate ISRs. In addition, although this variation uses more VC space, the work involved in establishing and maintaining switched paths is still  $O(n)$ .

An alternative approach to the VC merging problem is to use an end-to-end VC label upstream allocation. This allows the ingress node to choose the downstream VC. In this approach, ISRs acknowledge the Establishment message with a label only after they receive an Acknowledgment message from their own upstream neighbor. Thus, the Establishment message traverses fully to the ingress node before being acknowledged. Ingress ISRs immediately acknowledge the Establishment message with the VC label. These acknowledgements may be merged as they travel downstream to the egress node. This method adds latency in the VC set up, and removes the benefits of ARIS VC aggregation ( $O(n^2)$  versus  $O(n)$  VCs). However, it adds the flexibility of performing VC-switching instead of VP-switching, which also makes switching possible at the routing boundaries.

## **16. Multiprotocol Support**

ARIS is capable of multi-protocol support. Further, the egress aggregation feature of ARIS is applicable to those network layer topologies (IP, CLNP, IPX) that use a link state routing protocol. In particular, integrated IS-IS can calculate routes for CLNP, IP4, and IP6 simultaneously (with one Dijkstra calculation), and OSPFv3 (the new draft of OSPF) can calculate routes for IP4 and IP6 simultaneously. Both integrated IS-IS and OSPFv3 use a single router label to represent a single router that supports multiple network layer protocols. In this context, ARIS can minimize switching paths by using a single switching path for traffic from multiple network protocols destined to the same egress multiprotocol router.

## **17. ARIS And Frame Switching Technology**

The ARIS protocol is easily extendible to other switching



environments. Though the present document illustrates ARIS in ATM cell switching technology, it may later be extended to other switching technologies. In fact, ARIS applies well to frame switching technology.

While ARIS solves the problem of cell interleaving in the case of ATM by Virtual Path switching, it naturally and easily maps to a frame switching environment. This is due to the fact that multiple upstream flows can be merged into a single downstream flow without the problems of cell interleaving.

## **18. Quality of Service**

ARIS can be extended to support Quality of Service (QoS) parameters. This will be addressed in a future ARIS revision.

## **19. ARIS Advantages**

The ARIS approach to IP switching offers the following advantages:

- o VC aggregation
  - Use of egress identifiers allows many route (CIDR) prefixes to map to the same VC
  - Ability to create  $O(n)$  ARIS VCs (not  $O(n^2)$ ) on full mesh, where  $n$  is the number of nodes in a network.
- o Low VC setup overhead
  - Few VCs to setup due to routing-based and aggregated VCs
  - VC teardown/setup only on internal route changes
- o Switches large proportion of traffic
  - All traffic assigned to VCs
  - VCs created independent of traffic patterns
- o Guaranteed loop-free VC paths
- o Preserves VC's for RSVP flows or standard ATM connections
- o Scales to large networks

## **20. Security Consideration**

Security considerations are not addressed in this memo.



## **21. Intellectual Property Considerations**

International Business Machines Corporation may seek patent or other intellectual property protection for some or all of the aspects discussed in the forgoing document.

## **22. Acknowledgements**

The authors wish to acknowledge the following people for their input and support: Ed Bowen, Jerry Marin, Wayne Pace, Dean Skidmore, and Vijay Srinivasan.

## **23. References**

- [1] Juha Heinanen, "Multiprotocol Encapsulation over ATM Adaptation Layer 5", [RFC 1483](#), Telecom Finland, July 1993
- [2] Y. Rekhter, T. Li, "A Border Gateway Protocol 4 (BGP-4)", [RFC 1771](#), IBM Corp, Cisco Systems, March 1995
- [3] J. Moy, "OSPF Version 2", [RFC 1583](#), Proteon Inc, March 1994
- [4] S. Deering, "Host extensions for IP multicasting", [RFC 1112](#), Stanford University, August 1989
- [5] D. Waitzman, C. Partridge, S. Deering, "Distance Vector Multicast Routing Protocol", [RFC 1075](#), BBN, Stanford University, November 1988
- [6] J. Moy, "Multicast Extensions to OSPF", [RFC 1584](#), Proteon Inc, March 1994
- [7] S. Deering, D. Estrin, D. Farinacci, V. Jacobson, C. Liu, L. Wei, P. Sharma, A. Helmy, "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification", Internet Draft <[draft-ietf-idmr-pim-spec-02.txt](#)>, Xerox, Cisco Systems, USC, LBL, September 1995
- [8] D. Estrin, D. Farinacci, V. Jacobson, C. Liu, L. Wei, P. Sharma, A. Helmy, "Protocol Independent Multicast-Dense Mode (PIM-DM): Protocol Specification", Internet Draft <[draft-ietf-idmr-pim-dm-spec-01.txt](#)>, USC, Cisco Systems, LBL, January 1996
- [9] F. Baker (Editor), "Requirements for IP Version 4 Routers", [RFC 1812](#), Cisco Systems, June 1995



- [10] V. Fuller, T. Li, J. Yu, K. Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy", [RFC 1519](#), BARRNET, Cisco Systems, MERIT, OARnet, September, 1993
- [11] B. Cain, S. Deering, A. Thyagarajan, "Internet Group Management Protocol Version 3", Internet Draft <[draft-cain-igmp-00.txt](#)>, University of Delaware, Xerox PARC, August 1995

#### **24. Authors' Addresses**

Rick Boivie  
IBM Corp.  
17 Skyline Drive  
Hawthorne, NY 10532  
Phone: +1 914-784-3251  
Email: rboivie@vnet.ibm.com

Nancy Feldman  
IBM Corp.  
17 Skyline Drive  
Hawthorne, NY 10532  
Phone: +1 914-784-3254  
Email: nkf@vnet.ibm.com

Arun Viswanathan  
IBM Corp.  
17 Skyline Drive  
Hawthorne, NY 10532  
Phone: +1 914-784-3273  
Email: arunv@vnet.ibm.com

Richard Woundy  
Continental Cablevision  
The Pilot House - Lewis Wharf  
Boston, MA 02110  
Phone: +1 617-854-3351  
Email: rboundy@continental.com



