November 2000

Label switching using the IPv6 address hierarchy

P. Boustead
University of Wollongong, boustead@uow.edu.au

Joe F. Chicharo
University of Wollongong, chicharo@uow.edu.au

Recommended Citation

Boustead, P. and Chicharo, Joe F.: Label switching using the IPv6 address hierarchy 2000.
https://ro.uow.edu.au/infopapers/205

Label switching using the IPv6 address hierarchy

Abstract

Current label switching protocols can use routing, address, and address hierarchy information to group flows for cut-throughs that bypass IP forwarding. This paper examines a label switching solution that uses the IP version 6 (IPv6) address structure to classify and cut-through flows based on address hierarchy. The performance of this approach is examined using actual backbone traffic traces with associated hierarchical address information obtained from Internet address registries, routing arbiter databases and route servers. This hierarchical address information is used to map a hierarchical address structure over the packet level trace. We investigate the relationship between aggregation bit-mask size versus label switching performance. We show that aggregation greater than IPv6 destination site address does not significantly improve performance. Our trace driven simulation studies show that it is possible to reduce the network layer packet forwarding requirements to below 0.15% of total packets at boundary routers within the core network by using IPv6 address hierarchy linked label switching.

Keywords

Internet, asynchronous transfer mode, packet switching, telecommunication network routing, telecommunication traffic, transport protocols

Disciplines

Physical Sciences and Mathematics

Publication Details

This paper originally appeared as: Boustead, P & Chicharo, J, Label switching using the IPv6 address hierarchy, IEEE Global Telecommunications Conference, GLOBECOM '00, 27 November - 1 December 2000, vol 1, 500-504. Copyright IEEE 2000.

Label Switching Using the IPv6 Address Hierarchy

Paul Boustead, Joe Chicharo
Switched Networks Research Centre
University of Wollongong
Wollongong, Australia, 2522

Abstract— Current label switching protocols can use routing, address, and address hierarchy information to group flows for cut-throughs that bypass IP forwarding. This paper examines a label switching solution that uses the IP version 6 (IPv6) address structure to classify and cut-through flows based on address hierarchy. The performance of this approach is examined using actual backbone traffic traces with associated hierarchical address information obtained from Internet address registries, routing arbiter databases and route servers. This hierarchical address information is used to map a hierarchical address structure over the packet level trace. We investigate the relationship between aggregation bit-mask size versus label switching performance. We show that aggregation greater than IPv6 destination site address does not significantly improve performance. Our trace driven simulation studies show that it is possible to reduce the network layer packet forwarding requirements to below 0.15% of total packets at boundary routers within the core network by using IPv6 address hierarchy linked label switching

I. INTRODUCTION

Label switching combines the flexibility and robustness of IP routing with ATM's label swapping forwarding paradigm. Routers are under strain due to an increase in the size of internal routing tables resulting from the dramatic increase in size of the Internet over the last few years [1], and the need for higher packet forwarding rates. Label switching protocols such as Multi-Protocol Label Switching (MPLS) [2] and IP Switching [3] attack this problem by reducing the number of datagrams that need to be routed by the network layer. Forwarding decisions are 'cached' in fixed length labels that are carried with the data. In the case of cell based label switching, this fixed length label is the VCI field in the ATM cell. Some approaches (e.g. MPLS) do not restrict themselves to ATM, they are designed to run over all types of data link layer. This paper concentrates the use of label switching in IP over cell networks, however many of the results also apply to IP networks without an underlying cell network.

Existing and proposed label switching protocols can use routing, address, and hierarchy information to define cut-through flows which bypass network layer forwarding. This paper concentrates on examining the performance of address hierarchy linked label switching. An example of address hierarchy label switching is Destination-Site Label Switching (DSLS)[4]. DSLS aggregates flows using the IPv6 hierarchical address. The performance of this approach is examined using actual backbone traffic traces with associated hierarchical address information obtained from Internet address registries, routing arbiter databases and route servers. This information is used to map a hierarchical address structure over the packet level trace. We investigate the relationship between aggregation bit-mask size versus label switching performance. We show that aggregation greater than IPv6 destination site address does not significantly improve performance. Our results show that it is possible to reduce packet forwarding requirements to below 0.15% of packets at boundary routers within the core network by using address hierarchy linked label switching.

Section II introduces routing table, address, and address hierarchy cut-throughs. An example of address hierarchy label switching called DSLS is examined in Section II-A. Section III introduces a novel framework that enables an examination of hierarchy linked label switching. Results are presented in Section IV. Section V concludes the paper.

II. CUT-THROUGH FLOW DEFINITION

Cut-through flows can be defined using routing, address and hierarchy information. This section examines each of these cases and provides examples.

IP Switching is an approach that uses IP addresses to define cut-through flows. The first packets in a source-destination pair flow are forwarded at the network layer by each switch/router. If the number of packets exceeds a threshold (called Packet Threshold or PT) in a certain time then a cut-though path is negotiated with an adjacent switch/router. If this happens at each switch/router then an source-destination cut-through is created. Published results [3] show that 80-85% of packets will bypass network layer forwarding.

Linking routing table entries to cut-through routes, which we will call table linked cut-through, is a common approach used by MPLS. Each routing table entry is assigned a label. This label represents a cut-through flow that has been negotiated with an adjacent switch. When a packet enters an MPLS network the correct routing table entry is found by traditional network layer forwarding techniques. The packet is then placed upon the cut-through route associated with that routing table entry. The packet will follow this cut-through route towards the destination. If a route exists from source to destination in all routers then the cut-through route will forward the packet to the destination. However to improve the scalability, robustness and convergence time of IP routing protocols it is desirable to aggregate routes within backbone routers. When routing table entries are aggregated they will need to be de-aggregated at some point. At this de-aggregation point, the cut-through route must end and packets will again require network layer forwarding to determine which route to follow. Thus the cut-through route will not extend from source to destination. It is likely that route aggregation will rise with the introduction of IPv6 to enable

0-7803-6451-1/00/$10.00 © 2000 IEEE

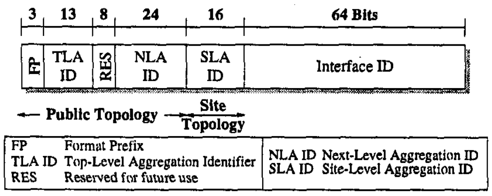

Fig. 1. Aggregatable Global Unicast Address

routing protocols to scale to larger and faster networks [1]. This will lead to shorter cut-through paths and an increase in the amount of network layer forwarding required. The requirement for network layer forwarding at de-aggregation points may be removed by the use of label stacks [2], however this is not possible in IP over cell based networks since the size of the stack is limited (ATM only allows two labels which are the Virtual Path and Virtual Circuit identifiers).

Defining flows by address hierarchy allows a higher level of aggregation than address linked flows. Cut-through routes are independent of routing information and are dependent instead on the address hierarchy. In order for address-hierarchy aggregation to be possible each switch must be able to determine address hierarchy information for each packet. It is also necessary that provider based address allocation is adhered to and customer sites are re-numbered if they change service providers. For example, if a service provider X owns the block of addresses 110.120.0.0/16, then customers of X must use a subset of this block (e.g. 130.120.130.0/24). If this customer changes to service provider Y then they must change addresses to a subset of Y's address block. Address hierarchy label switching is not possible in an IPv4 network, because a fixed address structure does not exist, and provider based addressing with mandatory renumbering is not enforced or practical.

The strict provider based IPv6 hierarchical addresses will allow address hierarchy flow aggregation. It will be possible to determine address hierarchy information from the address structure. Several types of addresses are supported. This paper will concentrate on the use the best-effort unicast address called "aggregatable global unicast address" [5]. This address format was designed to "facilitate scalable Internet Routing" [6]. The address format has a fixed structure as shown in Figure 1 and is organized into a three level hierarchy: Public Topology; Site Topology; and Interface Identifier. The strict nature of the address hierarchy results in a strict provider based address structure. Site networks that change from one service provider to another will obtain a different public topology address, and keep the same interface IDs.

The public topology consists of a two level hierarchy of service providers with a Top-Level Aggregation Identifier (TLA ID) and a Next-Level Aggregation Identifier (NLA ID). The TLA ID is initially to be restricted to 13 bits which translates to 8192 routers in the core IPv6 network. This was done to constrain core routing table sizes. However, there is provision for expansion by 8 bits if necessary [6]. The NLA ID is 24 bits long and allows for a flat or hierarchical allocation of the NLA address space. The Site-Level Aggregation Identifier (SLA ID) is 16 bits long. The SLA ID is used by an individual organization to define its local address hierarchy, and subnets.

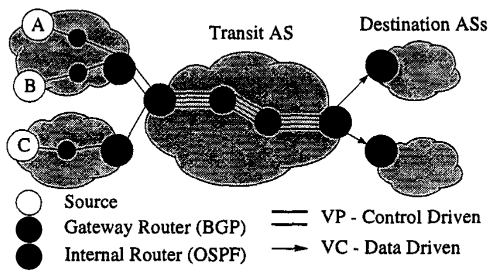

Fig. 2. Destination-Site Label Switching

An example of an approach that uses address hierarchy aggregation is Destination Site Label Switching (DSLS) [4]. This approach aggregates all traffic destined for the same IPv6 destination site addresses, using the SLA ID.

A. *Destination Site Label Switching*

DSLS uses both table linked cut-throughs, and address hierarchy cut-throughs. Two levels of labels are used, and stored in the ATM cells VPI and VCI fields. Table linked cut-throughs are created within ASs. VP paths are set-up from ingress to egress nodes. The hierarchy linked cut-throughs represent VC paths to the destination and are created using information from the IPv6 header. The purpose of this approach is to maintain the ability of many nodes to switch all packets (as in Tag switching), while eliminating the effect of IP route aggregation on performance by not assigning labels based on routing table entries for inter-AS traffic.

The address hierarchy cut-throughs are negotiated between egress routers of ASs. A cut-through is negotiated between ingress and egress nodes across the VP pipes. A destination based cut-through is created, based on the destination site network address. Using the aggregatable global unicast address format with fixed hierarchical address fields, it will be possible to create cut-throughs based on destination-site address. Cut-through routes will be created using a packet threshold technique similar to that used by IP Switching. Due to statistical multiplexing of traffic destined for the same destination site address we predicted that a high percentage of packets will be switched through egress routers of the core ASs to the gateway router of the destination site. Preliminary results in [4] show 96% of packets are switched at AS boundary label switches with 100% of packets switched by label switches within the AS. However these results use the class based address structure, which is outdated by CIDR, to estimate subnet boundaries.

Within ASs routing table linked VP pipes are created from ingress to egress nodes. An example of this can be seen in Figure 2. Routers internal to transit ASs forward all transit packets at the data-link layer. Only routers at the edges of transit ASs will be required to forward a small percentage of packets at the network layer. These VP pipes are destination based VPs from multiple ingress nodes to the one egress node. Using destination based VPs removes the need for hardware VC merge internal to ASs.

Unlike other data driven label switching techniques such as IP Switching, destination-site label switching will not allow a single cut-through from source to destination. It is necessary to route all packets in the ingress router at the destination site. This is not a significant disadvantage since most large organizations maintain IP firewalls around their sites and packets need to be forwarded by IP at this point irrespective of which label switching technique is used.

III. FRAMEWORK FOR PERFORMANCE EVALUATION

Previous studies have estimated destination-site address using the obsolete class based structure [4]. This paper presents a novel approach to estimating the address hierarchy information by obtaining information from Internet address registries and route arbiter databases. This information is used in conjunction with a packet level trace obtained from a major Australian telecommunication company backbone.

A. Hierarchical Trace Collection

Four traces were obtained from a major Australian Internet backbone. These traces contained 40 million packets each. The traces were obtained during work hours on weekdays at 11am 19/10/99, 2pm 12/11/99, 3pm 16/11/99 and 4pm 18/11/99. The trace durations were 3387, 1380, 1510 and 1608 seconds respectively. Trace bit-rates vary from 65Mbps for the short trace to 11.8Mbps for the longest trace. The traces contained packet level and hierarchical address information.

For each address in the trace hierarchy information was obtained from a hierarchical address database. This database contains information from route servers, routing arbiter databases, and Internet address registries.

Internet address registries maintain databases of registered sub-nets and is used to obtain more specific destination subnets than provided in route information. We collect Internet number entries with subnet sizes 4 bits or greater. An example of an Internet registry entry from APNIC is:

Internet number	202.6.91.0 - 202.6.91.255
Description	National Library of Australia
Description	Parkes Place
Description	Canberra ACT 2600
Source	APNIC

TABLE I

HIERARCHY INFORMATION

Internet Number Registries	
Asia Pacific NIC	European Internet
Taiwan NIC	Australian NIC
American Registry	Japan NIC
Australia NIC	
Routeing Arbiter Databases	
ANS Communications	Merit-Routing Arbiter
Bell Nexxia	Advanced Cutting Edge
AAPT - Australia	Carynet Communications
Cable and Wireless	GlobalCrossing
Verio Communications	Planet Online Europe
Planet Online Europe	MCI WorldCom
Route Servers	
Planet Online Europe	Planet Online Europe
ATT Cerfnet	Oregon Internet Exchange
ATT WorldNet Service	

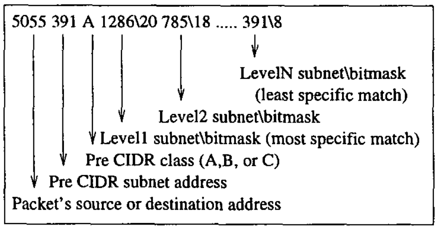

Fig. 3. Example Address Hierarchy

Route information is also collected. Route information gives larger aggregates, and is used to examine the relationship between aggregation levels and label switching performance. Route-servers maintain a superset of routing information required by peering service providers. We collect "route" objects from these databases. An example of a collected entry follows:

Route	202.61.224.0/19
Description	Digital non portable CIDR space
AS Number	AS7586
source	APNIC

Routing-Arbiter Databases contain collections of route information used for operation-purposes by network service providers. An example of the entries collected is:

B	170.143.0.0/16	via	198.32.176.25
B	207.25.252.0/24	via	134.24.88.55
B	204.179.85.0/24	via	134.24.88.55
B	204.145.119.0/24	via	134.24.88.55

In the above database entries B represents a BGP routing entry. The second column is the only information we are interested in and represents aggregated routes in core routers.

Address and route information is collected from the databases in Table I into our hierarchical database. After removal of many duplicate entries the database contains 321805 entries from the Internet address registries, and 135827 entries from the routing databases. The IP addresses in the packet level trace are anonymised by the re-numbering method. Before the addresses are anonymised the hierarchical database is searched for all matching entries, these entries are then anonymised and stored in conjunction with the packet level trace. An example of a hierarchy determined for a particular address can be seen in Figure 3.

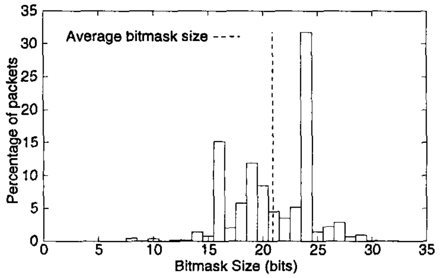

Fig. 4. Subnet sizes

B. Determining Site Address

The site address is determined as the most specific match in the hierarchy database. For example the most specific match for the example in Figure 3 is 1286\20 where 1286 is an anonymised address and 20 is the CIDR bit-mask. This site address is in many cases more specific than the most specific route in a full core routing table because of the addition of address registry information to give more specific customer sites. The levels 2 to N shown in Figure 3 represent aggregations higher than the site-address.

IV. RESULTS

A. Performance of Destination-Site Label Switching

The hierarchical trace allows us to examine the performance of destination-site label switching as well as other levels of aggregation. Site addresses are calculated by finding the most specific match in the hierarchical database, for a particular address. A histogram of the site bit-mask sizes for all the addresses in the trace is shown in Figure 4.

The performance of DSLS with varying levels of aggregation as well as IP Switching performance is shown in Figure 5. IP Switching forwards 13.5% of packets at the network layer. Aggregating traffic between source and destination sites reduces this to 8%. The highest level of aggregation shown is destination-site where all packets destined for a single site address are aggregated, this results in less than 2% of packets requiring network layer forwarding. The number of VCs used also reduces as aggregation increases from 16500 VCs for no aggregation to 5000 VCs for destination-site aggregation. The VC set-up reduces from 170 VCs/sec with no aggregation to 20VCs/sec for destination site aggregation.

The results for percentage of packets routed are an average of all four traces with 95% confidence intervals generated using the batching method. The VC results are also a combination of results from each of the traces. In order to average the VC results they were first normalised with respect to the packet-rate of each trace, the results are then scaled to the average packet rate of the traces for readability.

B. Higher levels of aggregation

The previous section examined use of site address to determine stream aggregation. In this section we examine the relationship between aggregation and label switching performance.

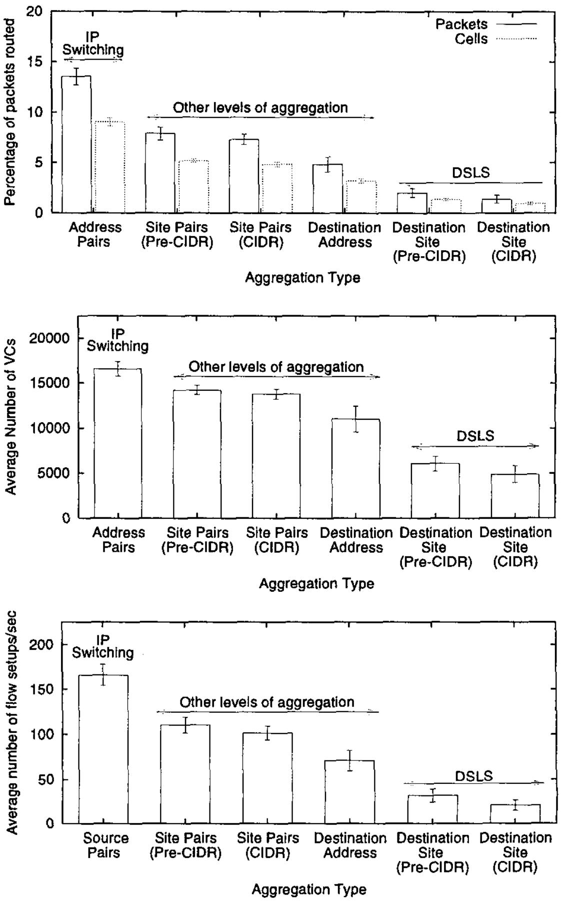

Fig. 5. Site based aggregation results

Figure 6 examines the relationship between aggregation level and percentage of packets forwarded at the network layer. In order to obtain this relationship aim aggregation levels (bit-masks) were selected between a bit-mask of 16 to 32. For each packet in the trace the hierarchy database is searched for a subnet that contains the address with a bit-mask closest to the aim bit-mask. The x-axis in Figure 6 shows average bit-mask size. A bit-mask size of 32 represents source-destination flows, a bit-mask size of 21 (see Figure 4) represents the site address as examined in the previous section. A bit-mask size of less than 21 represents flow aggregation greater than site-address. It can be seen that increasing aggregation to site-aggregation significantly increases performance. Further increasing aggregation particularly for destination-site flows does not significantly improve performance. These results are for the first trace, results for the remaining traces show a similar trend.

C. Parameter Adjustment

A packet threshold of 10 was used in the previous tests. This has been shown to be a sensible choice [3], [7] for source-

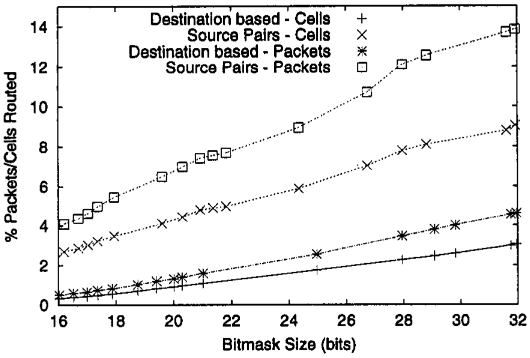

Fig. 6. Percentage of packets routed vs. Aggregation

destination flows as a compromise between VC usage and percentage of packets switched. This section examines the effect of aggregation on sensible packet threshold choice.

The results presented in Figure 5 use a packet threshold of 10 packets in 60 seconds to trigger creation of cut-through flows. This is frequently used for non-aggregated source-destination flows [7] as a compromise between a high percentage of packets switched and low VC usage. Higher aggregations may have a different optimum VC usage. To examine this we use the hierarchical trace to plot a surface showing the relationship between aggregation, packet threshold and average VC usage. This is shown in Figure 7. The aggregation level is expressed in terms of bit-mask size, a small bit-mask represents a high level of aggregation. It is clear that as aggregation increases (decreasing bit-mask size) the sensitivity of VC usage to changes in packet threshold decreases. This indicates that for higher levels of aggregation a lower bit-mask size will increase the percentage of packets switched with less effect on VC usage.

The performance of data-driven aggregated label switching protocols with reduced packet threshold is shown in Figure 8. Reducing the packet threshold to 5 lowers the percentage of packets switched for the destination subnet case from 1.38% to 0.8% a reduction of 42% which results in 6% increase in average VC usage. A further decrease in packet threshold to 1 reduces the network layer forwarding requirement to 0.13% with minimal increase in VC usage.

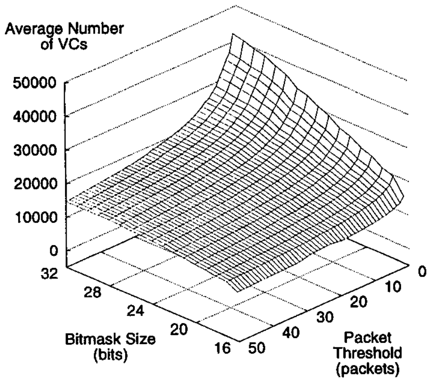

Fig. 7. Average VC Requirement: Source Pair

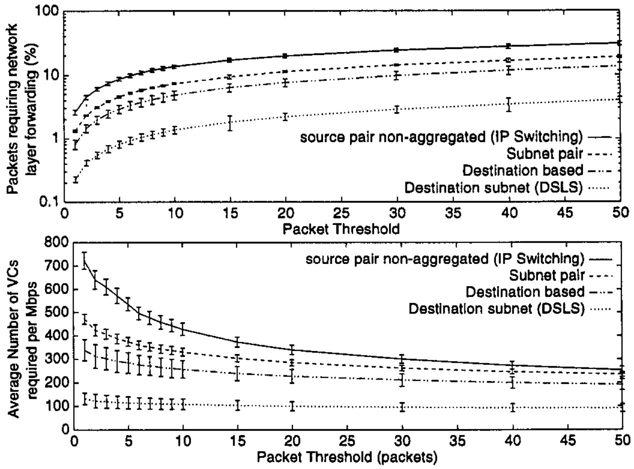

Fig. 8. Subnet aggregation with reduced packet threshold

V. CONCLUSIONS

This paper examines the performance a label switching approach that uses IPv6 address hierarchy information instead of routing information. We provide a framework that facilitates an estimation of the performance of such a protocol using current backbone traffic traces and information obtained from Internet address registries, route servers and route arbiter databases. This framework is used to examine the relationship between aggregation bit-mask size versus label switching performance. We show that aggregation greater than IPv6 destination site address does not significantly improve performance. The performance of destination-site label switching is then examined. We show that it is advantageous to use a packet threshold less than 10. Using a low packet threshold we show that the percentage of packets that are require network layer forwarding at gateway routers can be reduced to below 0.15% of total packets.

REFERENCES

[1] Tim Bass. Internet exterior routing protocol development: problems, issues, and misconceptions. *IEEE Network*, 11(4):50–55, July 1997.

[2] Arun Viswanathan, Nancy Feldman, Zheng Wang, and Ross Callon. Evolution of multiprotocol label switching. *IEEE Communications Magazine*, pages 165–173, 1998.

[3] Peter Newman, Greg Minshall, and Thomas Lyon. Ip switching - atm under ip. *IEEE/ACM Transactions on Networking*, 6(2):117–129, 1998.

[4] Paul Boustead, Scott Barnett, Joe Chicharo, and Gary Anido. Label switching and ip version 6. In *7th International Conference on Computer Communications Networks*, volume 1 of *IPv6 over ATM Networks*, pages 561–565, Lafayette, Lousiana USA, 1998. IEEE Computer Society.

[5] R. Hinden and S. Deering. Ip version 6 addressing architecture. Internet draft, IETF, January 22 1998. Expires July 22 1998.

[6] R. Hinden, M. O'Dell, and S. Deering. An ipv6 aggregatable global unicast address format. Internet draft, IETF, 13th March 1998.

[7] Steven Lin and Nick McKeown. A simulation study of ip switching. In *ACM Sigcomm 1997*, 1997.