IPv6 Source Routing for ultralow Latency
draft-foglar-ipv6-ull-routing-04

Abstract

  This Internet-Draft describes a hierarchical addressing scheme
  for IPv6, intentionally very much simplified to allow for very
  fast source routing experimentation using simple forwarding
  nodes. Research groups evaluate achievable latency reduction
  for special applications such as radio access networks,
  industrial networks or other networks requiring very low
  latency.

Revision Note for Version 02

  Reference to experimental verification of the concept is added in the

section "Acknowledgements".

Revision Note for Version 03

   Section 6 about Security Considerations has been inserted.

Revision Note for Version 04

   Section 7 about Redundancy has been inserted.

1. Introduction

   To achieve minimum latency the forwarding nodes must support
   cut-through technology as opposed to the commonly used store-
   and-forward technology. Cut-through means, that the packet
   header already leaves a node at the egress port while the tail
   of the packet is still received at the ingress port. This
   short time does not allow complex routing decisions.

   Therefore, a very simple routing address field structure is
   specified below. It should limit the complexity of the
   forwarding node used in the experiments. Therefore, in this
   text the term "forwarding node" is used instead of "router",
   although the device is operating in OSI Layer 3 and accordingly
   executes router functions such as decrementing the hop limit field.
   Redundancy issues are not considered.

2. IPv6 address prefix structure

   The following proposal uses the 64-bit IPv6 address prefix.

   Each forwarding node has up to 16 ports and hence needs 4 bits
   of the address field to decide to which port a packet should
   be forwarded. The 64-bit prefix is divided into 16 sub-fields
   of 4 bit, defining up to 16 hierarchy levels. A forwarding
   node is configured manually to which the sub-fields it should
   evaluate for the forwarding decision.

   A number n of leading 4-bit fields cannot be used for forwarding
   decisions, but must have a special value to indicate the
   'escape prefix' of the experimental forwarding mode.

   The 64-bit prefix of the IPv6 address has this structure:

   | n x 4-bit escape prefix |(16-n) x 4-bit address fields |

   The first 4-bit field following the escape prefix has the
   highest hierarchy level, the last 4-bit field has the lowest
   hierarchy level.

3. Forwarding node behavior

The forwarding node has up to 16 downlink ports and at least
one uplink port. Typically, the forwarding nodes are arranged
in a regular tree structure with one top node, up to 16 nodes
in the second hierarchy, up to 256 nodes in the third hierarchy
and so on for up to 16-n hierarchies.

A forwarding node must be configured to operate at a certain
position in the hierarchical network. For example, at third
hierarchy level, branch 4 of the first hierarchy and branch 12
of the second hierarchy.

The behavior of each forwarding node is depending on the
position of a node in a hierarchical network. For all
positions, the first step is to check the escape prefix. Only
packets with matching escape prefix are forwarded.

The top forwarding node with the highest hierarchy level
evaluates the first 4-bit field following the n x 4-bit escape
prefix. The value of the evaluation field determines the
output port of the packet. The remaining fields are don't
care:

```
| escape prefix | 4-bit | (16-n-1) x 4-bit |
<  mandatory   > <eval.> <   don't care   >
```

A forwarding node in a lower hierarchy first checks if the 4-
bit fields preceding the evaluation field match the configured
value. In case of match the value of the configured evaluation
field of the packet is used as downlink port number where the
packet is forwarded. The remaining 4-bit fields are ignored.
In case of mismatch the packet is forwarded to the uplink
port(s).

```
| escape prefix | m x 4-bit | 4-bit | (16-n-m-1) x 4-bit |
<  mandatory   > <  match  > <eval.> <    don't care    >
```

The parameter m indicates the hierarchy level with m=0
denoting the highest hierarchy.

Hence, when a packet enters a hierarchical network at the
lowest layer node it is forwarded in uplink direction until it
reaches a node where the m x 4-bit prefix matches the
configured value of the node. At latest, the highest-level
node will always match and forward the packet in the desired
downlink direction.

4. Numerical values

As mentioned, one pre-requisite of the simple forwarding
concept is to keep the complexity of the forwarding nodes low.
Also, the configuration of the nodes should be kept simple. In
particular industrial networks are operated by persons who are

not experts in communication. Configurations should be
intuitively understandable by all without long explication.
Therefore, for the first experimental forwarding node the
number of downlink ports is limited to 10 with numbers 0...9. 16
digits at the front panel of the forwarding device show the
configuration. Use of classical 7-segment digits make the
limits of the configuration obvious.

As escape code, the first two digits are fixed to the value
"AF" (binary '10101111'). These two characters contrast with
the following numerical digits, so that the escape code can be
clearly differentiated from the following configuration. The
display uses the 'H' character instead of the 'X' the usual
term for the variable.

The H specifies the digit of the packet prefix which is
evaluated for forwarding. When the H is selected all lower
digits are automatically set to '-' to indicate the don't care
nature.

To make the configuration still more obvious it is recommended
to configure the local telephone number. With that measure,
every local experimentation has unique numbers and can
potentially be interconnected via tunnels (IP, MPLS, VPN etc.)
with other experimental setups.

The length of 14 digits allows sufficient in-house
hierarchies, even for industrial applications where forwarding
nodes interconnect large numbers of sensor controllers.
Inhouse installations would be structured for example in
building, floor, fabrication unit, machine - with one sensor
controller per machine. For the sake of simplicity numbers are
deliberately wasted, for example if the building has only 3
stories the digits 4...9 are unused.

5. Example configuration

A small office in Munich with the telephone number +49-89-
45241990 configures its local top-level forwarding node to:

    AF49.8945.2419.90H-

Note that for the sake of simplicity this simplified notation
is introduced here as alternative to the usual notation
AF49:8945:2419:9000/56. With the new notation, the cabling
staff people can immediately check the hierarchy location of
the forwarding node and connect the cables to the floors at
ports 0...3.

The next hierarchy level is related to the floor. In case of a
3-story building only three next level forwarding nodes are

used with these configured values:

    AF49.8945.2419.900H at the ground level
    AF49.8945.2419.901H at the first floor
    AF49.8945.2419.902H at the second floor
    AF49.8945.2419.903H at the third floor.

In each floor, up to 10 sensor nodes can be connected.
Each of the sensor nodes can address several sensors/
actuators addressed via the interface identifier contained in
the second part of the 128-bit IPv6 address.

In the following a connection between sensors in this office to
otherIoT equipment located in Essex University is described. The
connection is realized with one additional forwarding node
installed at Essex University premises with the second level address
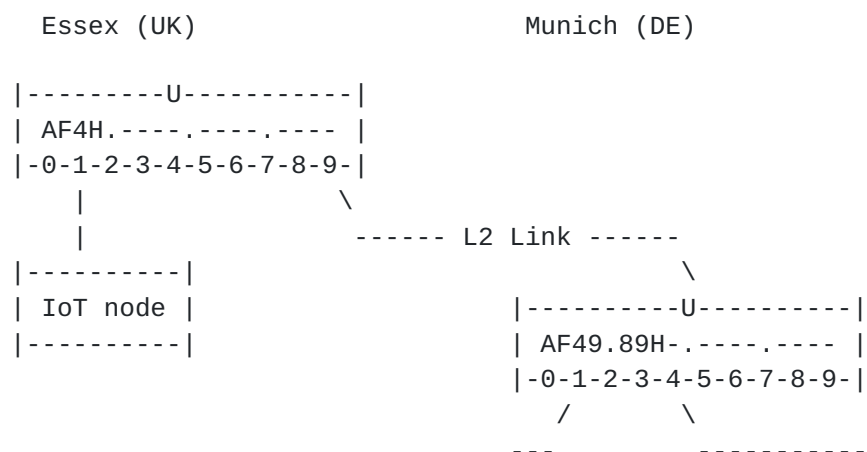
    AF4H.----.----.----.

This high level forwarding node can be used although the phone number
of the researcher is +44 1206 872413, as long as there is no further
node in UK.

At downlink port 9 the 13th level forwarding node in Munich is con-
nected via a Layer 2 link such as VLAN or SDH pipe or MPLS tunnel.
The levels in between must not be populated by forwarding nodes as

long as no other branch is needed at one of the two sides.
If for example another site in Munich center must be connected one
additional forwarding node must be installed with the 5th level
address

    AF49.89H-.----.----.

The small office mentioned above would be connected to downlink port
4 while the new site would be connected at downlink port 1, the
prefix for Munich center. The configuration is visualized in the
Figure below.

```
  Essex (UK)                      Munich (DE)

|---------U-----------|
| AF4H.----.----.---- |
|-0-1-2-3-4-5-6-7-8-9-|
    |                 \
    |                  ------ L2 Link ------
|----------|                               \
| IoT node |                  |---------U----------|
|----------|                  | AF49.89H-.----.---- |
                              |-0-1-2-3-4-5-6-7-8-9-|
                                  /         \
                                 ---         -----------
```

```
                         /                        \
           |----------U----------|    |----------U----------|
           | AF49.89H-.-----.---- |    | AF49.8945.2419.90H- |
           |-0-1-2-3-4-5-6-7-8-9-|    |-0-1-2-3-4-5-6-7-8-9-|
                                               |
                             |----------U----------|
                             | AF49.8945.5419.901H |
                             |-0-1-2-3-4-5-6-7-8-9-|
   U = Uplink                           |
                                  |----------|
                                  | IoT node |
                                  |----------|
```
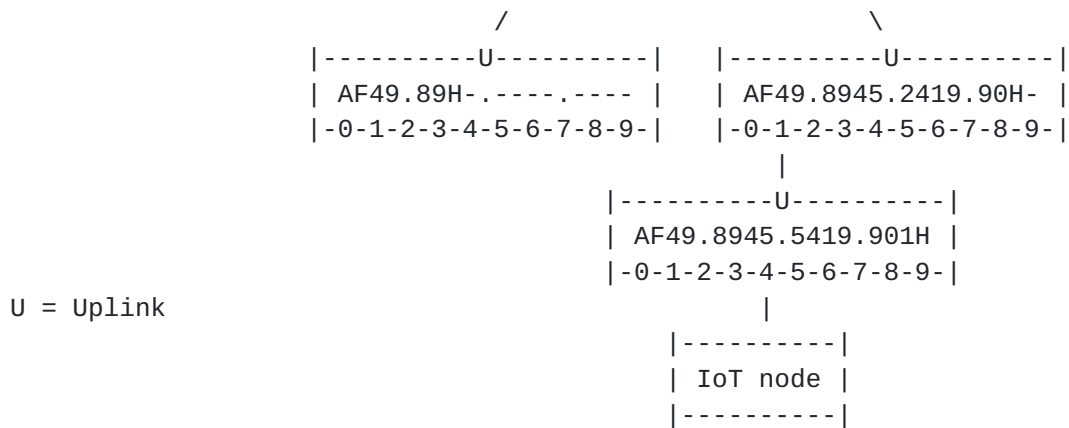Figure: Example Configuration

6. Security Considerations

In a hierarchical network as described above every forwarding node
can easily check a part of the source address of the packets. For
example, the IoT node in the above Figure will get the following
address assigned: AF49.8945.5419.9014. It will use this address as
source address when it sends packets in upstream direction. The node
above receives the packets at port 4 and expects that the digit at
the position of the 'H' will be a 4.

Moreover, a node in upstream direction will check if the prefix of
the packet's source address. For example the top Munich node in the
above figure will check the source address of packets received from
downlinks to be AF49.89.

Hence, the source addresses of all packets forwarded in upstream
direction will be checked repeatedly all the way up in the hierarchy.
In case of mismatch, either an error has occurred or an intentional

address falsification. Both cases should be avoided and therefore
such packets are discarded.

In consequence, a receiver in the hierarchical forwarding network can
rely on the correctness of the source addresses of the received
packets. This feature is ideal for white-listing, which in turn is
ideal for IoT applications. Other than in the 'normal' Internet
access to IoT devices is not needed for every host. In contrary, IoT
devices will have restricted access for authorized hosts only.

Note that the verified source address feature is orthogonal to other
security measures such as password authentication and encryption.

7. Reduncancy

The hierarchical structure implied by the addressing leads to the fact
that node failures have more implications the higher the hierarchy of
a node. Therefore, a node should be equipped with two redundant uplink

ports. Each of them is connected to a next higher hierarchy node, each of them having again two redundant uplinks.

Hence, with each hierarchy the number of uplinks doubles - and also the number of nodes. In the case of ten downlinks and two uplinks the number of nodes grows with the power of two and the number of terminals grows with the power of ten.

A three-dimensional network is constructed with up to n hierarchies and up to 2^n redundancy planes. With 14 hierarchies the number of redundancy planes becomes 16384. This number of top hierarchy nodes sounds very high, but distributed around the world would lead to well-balanced redundancy.

With the two uplinks (could also be more) a routing feature emerges in the network. In other words, each node has to take a routing decision in upstream direction, when forwarding packets to one the uplinks. This decision could be based on node-local information (autarkic) or based on routing protocols. This topic is for further study.

[8](). Acknowledgements

[9](). Authors' Addresses

Andreas Foglar
InnoRoute GmbH
Marsstr. 14a
80335 Munich
Germany
Email: foglar@innoroute.de

Mike Parker
Wivenhoe Park, Colchester
Essex, CO3 4HG
United Kingdom
Email: mcpark@essex.ac.uk

Theodoros Rokkas
Incites S.A.R.L.
130, Route d' Arlon
Strassen L-8008
Luxembourg
Email: trokkas@incites.eu