Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: September 9, 2010

A. Ripke
J. Quittek
M. Brunner
NEC Laboratories Europe
March 8, 2010

Dynamic Port Range Re-Assignments for Address Sharing
draft-rqb-dynamic-port-ranges-02

Abstract

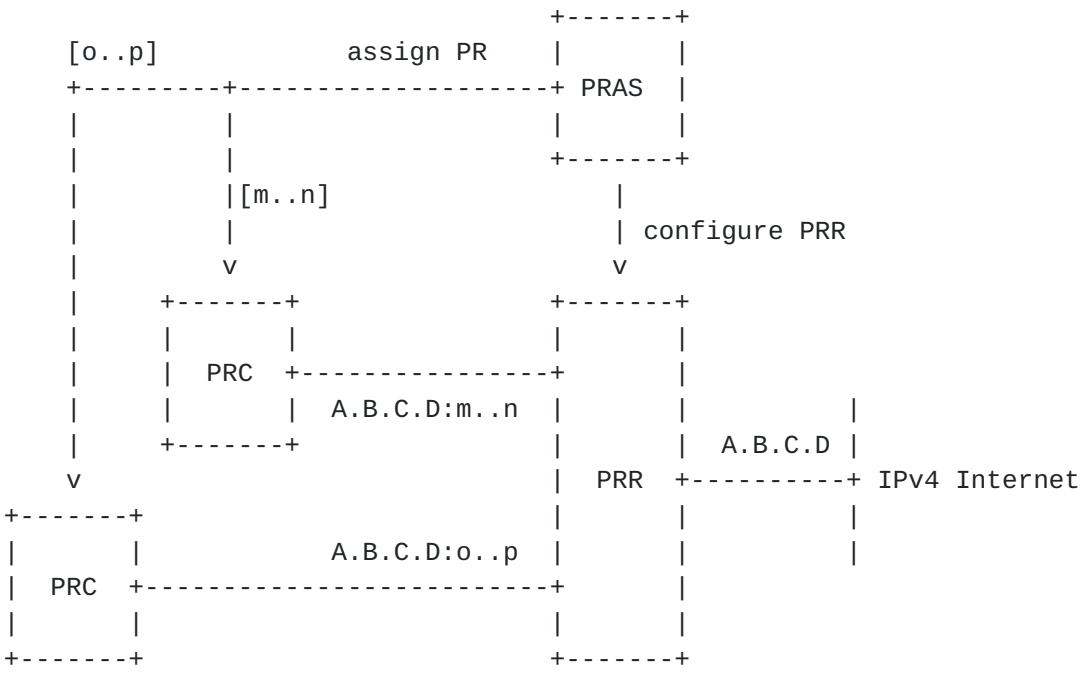
This document describes and extension of the port range assignment mechanisms used for the IPv4 address sharing framework (SHARA) that is based on port range routing. The extension provides means for dynamically changes port range assignments by allowing clients to smoothly migrate to a new port range before releasing the range that is currently in use. This way, the number of ports per client can be adjusted to actual usage patterns.

Table of Contents

1. Introduction

The IETF is discussing a scheme for enlarging the usable IP address space in [I-D.ymbk-aplusp] or [I-D.boucadair-port-range] using parts of the port numbers, similar to what Network Address Translators (NAT) do. This allows to assign the same IP address to several customers or hosts. The IP address together with the port bits extension differentiate the routing and forwarding of that communication.

A port range router (PRR) manages one or several IPv4 addresses that are to be shared among several port range clients (PRC). Each PRC gets a portion of one IPv4 address whereas this portion is defined by one or several port ranges that are assigned exclusively to a PRC. The PRR and the PRC can, for example, be a Broadband Remote Access Server (BRAS) and a Home Gateway (HGW), respectively. Alternatively, the client can be a home router, a single host, or the gateway of a large enterprise, A schematic sketch of sharing two non-overlapping port ranges is illustrated by the following figure:



A PRAS assigns two different non-overlapping port ranges [m..n] and [o..p] of an IPv4 address A.B.C.D to two PRCs.

Figure 1

DHCP extensions [I-D.boucadair-dhc-port-range] and PPP extensions [I-D.boucadair-pppext-portrange-option] that can be used by a Port

Range Assignment Server (PRAS) to assign IP addresses and port ranges to a PRC have already been proposed. However, since the deployments are very different for different users, customers with several users etc., further means for managing port assignments appear to be required. Measurements showed that different clients need different range sizes at different times [flow-counting].

This implies that dynamic port range assignment is needed for

- assigning clients larger port ranges when the current one becomes too small,
- assigning clients smaller port ranges, when the current ones are underused,
- changing clients port ranges for reducing fragmentation of the port space,
- balancing port consumption for a shared IPv4 address.

The existing means are sufficient to assign and re-assign port ranges (both contiguous and non-contiguous ones). However, a PRC cannot immediately switch from one port range to another one, because most applications cannot change port numbers while using them. Without interrupting existing connections, a PRC can only start allocating new ports in a new range and wait until ports in an old range are not used anymore. Consequently, a PRC needs to wait until applications have closed all ports in the old port range. Existing means allow to assign more than one port ranges to a PRC ([I-D.boucadair-port-range]), but not to identify one or more ranges that should not be used anymore by the PRC.

2. Dynamic port range re-assignments

This draft proposal provides a way for a Port Range Assignment Server (PRAS) to tag a port range with an attribute that signals the PRC not to allocate any more ports in this range. Such a signal can be sent when a server signals more than one port range to a PRC. A most simple implementation would be adding a flag to one or more port ranges during the (re-)assignment process that marks these ranges as not to be used anymore. A PRC receiving the signal would then stop allocating port numbers in the marked ranges. When the PRC does not use an address range anymore, it signals back that the port range is not in use anymore and can be re-assigned. This can be done individually for each range as soon as it is not used anymore or at once when all marked ranges are not used anymore.

The method can also be used for reducing (trimming) already assigned port ranges. For this purpose, the PRAS divides the single port range into two or more consecutive port ranges and re-assigns the single port range as a set of port ranges to the PRC with one or more of the port ranges marked as not to be used anymore. Again, the PRC would signal back that one or more ranges are not used anymore.

This new technique allows to postpone the de-allocation of port ranges until the respective ports are closed (lazy de-allocation). The PRC has the possibility to actively confirm the release of port ranges.

2.1. Detecting the change point

While the PRC is using the port range, several reasons may occur that make it desirable to change the port assignment.

- The PRC may observe that there are only few unused numbers left in the used range and that it may soon happen that no further ports would be available for requesting applications. In order to avoid this situation, the PRC requests an assignment of more port numbers at the PRAS.

- A user can actively close all ports in anticipation of an exceeding demand of ports from new applications to be started. All ports are released voluntarily in expectation of goodwill to get a larger port range assigned.

- The PRAS may monitor usage of port numbers by the PRCs and detect that there are only a few unused port numbers left in the range assigned to the PRC. It decides to assign a wider range to the PRC before port numbers run out.

- The PRAS may detect that the PRC is only using a small part of the port range assigned to him and decide to assign the PRC a smaller port range.

- The PRAS may identify a need to re-assign the port range of the PRC in order to reduce fragmentation of the port space.

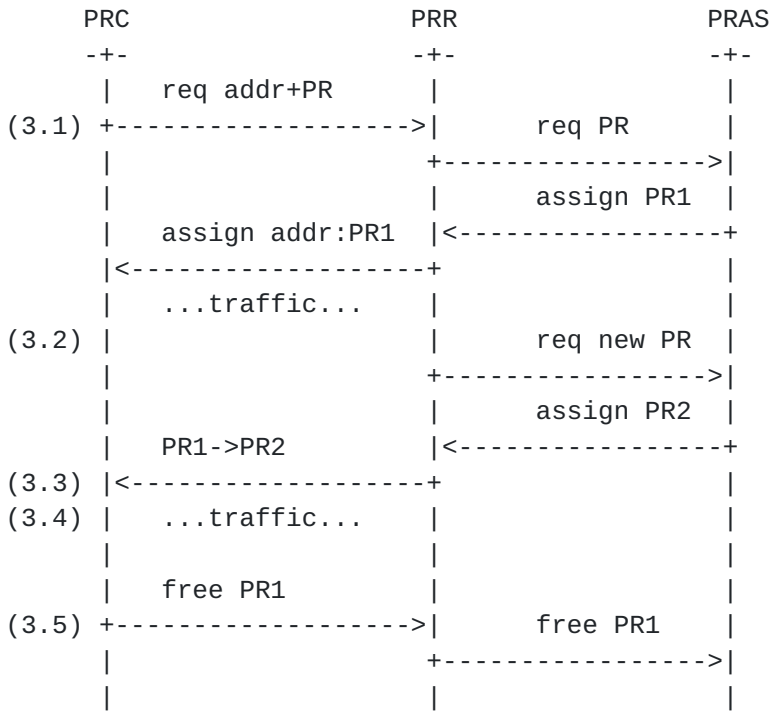
The port range change request could be both PRC and PRAS initiated.

3. Usage scenario

The following usage scenario describes the impact of the proposed method from the initialization phase till final de-allocation of port ranges.

In this scenario, a broadband provider's PRAS and PRR operates on a BRAS, which then manages IP addresses, according port ranges, and broadband access. The PRC is a home router, allocating port numbers when acting as NAT for the home devices.

Alternative scenarios have PRAS and PRR located at an MSAN (Multi Service Access Node), DSLAM (Digital Subscriber Line Access Multiplexer), an SGSN (Serving GPRS (General Packet Radio Service)) or GGSN (Gateway GPRS Support Node). The PRC can alternatively be located at various devices ranging from an enterprise gateway to a mobile terminal or a sensor. The message flow between PRAS, PRR, and PRC is illustrated by Figure 2:



A PRAS initially assigns port range PR1 to the PRC. Later, PR1 gets replaced by PR2.

Figure 2

3.1. Initialization

A home router requests an IPv4 address with optionally requesting a certain number of port numbers, a specific port range, or a specific set of not consecutive port numbers. The BRAS replies in his role as PRAS by assigning an IP address and a set of port numbers to the home router.

3.2. Detecting the change point

After some time the BRAS identifies a very high port consumption with one of its home routers. A certain threshold has been exceeded and still new connections are initiated from the home router side.

Alternatively: The home router identifies an upcoming shortage of available ports and sends a request for more ports to the server.

3.3. Assigning a new port range

The BRAS sends a message to the home router. The message contains two port ranges, the originally assigned one with a mark not to use it anymore and a new range to be used from now on. Optionally, the old port range may be tagged with a time stamp that indicates when this port range will definitely expire and cannot be used anymore by the home router.

3.4. Ongoing port consumption

The home router only allocates new port numbers of the new range and releases port numbers in the old port range.

3.5. Final de-allocation of a port range

The BRAS detects that no port number of the initial port range is not in use anymore (through monitoring) and signals to the home router that the assignment of the old range is expired.

Alternatively: The home router explicitly confirms the release by sending a signal to the BRAS that it is not using the initial range anymore and the BRAS can assign it to other PRCs. A more complicated option is a partial release of the old range agreed between BRAS and home router. This requires splitting the old port range into two sub-ranges, one to be released and one to be further used.

4. Signaling

The signaling between client and server can be done through different protocols including DHCP extensions, PPP extensions, Web Services, TR-069, or a novel protocol for address and port pool management.

5. Fragmentation

According to [I-D.boucadair-pppext-portrange-option] and [I-D.boucadair-dhc-port-range], it is possible to assign more than

one port range to a customer (using a port mask and a port locator). It is expected that contiguous port range allocation will be the preferred procedure. However, together with the introduced technique to enlarge or to reduce individual port ranges the port range manager might have to deal with heavily fragmented port mapping tables. Besides administration overhead this may lead to problems if new contiguous port ranges are requested. Dynamic port range re-assignment provides a technique that can both amplify and rectify this problem.

6. Port Range Swapping

Port randomization [I-D.ietf-tsvwg-port-randomization] is a mechanism to make it harder for "blind" attacks to spoof a system. However, having a smaller port range to choose from produces more port collisions. Local collisions can be easily detected by comparing a port against open connections. Remote collisions on the other hand are harder to detect unless recently closed connections are tracked like suggested in [I-D.ananth-tsvwg-timewait]. The problem is that an active closer of a TCP connection lingers in state TIME-WAIT for four minutes for the respective connection's five-tuple (local address, local port, remote address, remote port, protocol). Frequent connections to the same server might induce a situation where the client's ports are in said state on the server and no more connections are possible for a while. Clearly, the smaller the client's port range the more often this undesired effect may occur. One solution with dynamic port range management might be the possibility to exchange a used port range for a recently unused port range with the port range manager.

7. Service Management

As already mentioned in [I-D.levis-behave-ipv4-shortage-framework], a PRAS assigns the number of ports to the customer upon pre-configured policies which might depend on the individual contract with the customer or on the customer's usage profile.

7.1. Server policy

The process on the PRAS for deciding on how many ports to give away is based on policies configured into the PRAS from a management station. That might depend on the customer status. Premium customers paying a certain fee might request higher numbers. It can also depend on the current level of free addresses and ports. When there are only a few ports left the IP address and port range manager might be more restrictive with port allocations. In general, the

mechanisms described above in the usage scenario requires configuration on the PRAS to behave in one or the other way, also including the configuration of the client.

7.2. Client policy

The policies will also be configured into the client and can provide information about

- the amount of available space that can be requested.
- and what port consumption level triggers the dynamic mechanism of expanding or reducing the client's port range.

For example, the threshold for expanding the port range could be a port utilization of 80%. If the client exceeds this threshold a request for more ports is sent to the PRAS. Alternatively, the client only requests for more if its port range is entirely depleted.

8. Open issues

Dynamic port range re-assignment has several open issues to be solved or clarified:

- Modifications are required to both the DHCP and the PPP protocol in addition to the extensions described in [I-D.boucadair-dhc-port-range] and [I-D.boucadair-pppext-portrange-option] respectively.
- What strategy should be chosen to solve a potential port mapping table fragmentation?
- The constant port monitoring which the port range manager has to carry out might impose problems.
- How to handle expiration timers when requesting port ranges to be cleared?
- The processing of port overflow caused by exceeding port number requests might become a delicate problem. If available port numbers for a specific IPv4 address do not match a client's request it would be necessary to assign a new IPv4 address.

Eventually, the price to be paid for dynamic port range management is complexity.

9. Acknowledgements

The authors are supported by Trilogy (http://www.trilogy-project.org), a research project (ICT-216372) partially funded by the European Community under its Seventh Framework Program. The views expressed here are those of the author(s) only. The European Commission is not liable for any use that may be made of the information in this document.

Thanks to Mohamed Boucadair for his comments.

10. IANA Considerations

This document includes no request to IANA.

11. Security Considerations

TBD.

12. Informative References

[I-D.ananth-tsvwg-timewait]
Ramaiah, A. and P. Tate, "Effects of port randomization with TCP TIME-WAIT state.", draft-ananth-tsvwg-timewait-00 (work in progress), July 2008.

[I-D.boucadair-dhc-port-range]
Boucadair, M., Grimault, J., Levis, P., and A. Villefranque, "DHCP Options for Conveying Port Mask and Port Range Router IP Address", draft-boucadair-dhc-port-range-01 (work in progress), October 2008.

[I-D.boucadair-port-range]
Boucadair, M., Levis, P., Bajko, G., and T. Savolainen, "IPv4 Connectivity Access in the Context of IPv4 Address Exhaustion: Port Range based IP Architecture", draft-boucadair-port-range-02 (work in progress), July 2009.

[I-D.boucadair-pppext-portrange-option]
Boucadair, M., Levis, P., Grimault, J., and A. Villefranque, "Port Range Configuration Options for PPP IPCP", draft-boucadair-pppext-portrange-option-01 (work in progress), July 2009.

[I-D.ietf-tsvwg-port-randomization]
Larsen, M. and F. Gont, "Transport Protocol Port Randomization Recommendations", draft-ietf-tsvwg-port-randomization-06 (work in progress), February 2010.

[I-D.levis-behave-ipv4-shortage-framework]
Levis, P., Boucadair, M., Grimault, J., and A. Villefranque, "IPv4 Address Shortage: Needs and Open Issues", draft-levis-behave-ipv4-shortage-framework-02 (work in progress), June 2009.

[I-D.ymbk-aplusp]
Bush, R., "The A+P Approach to the IPv4 Address Shortage", draft-ymbk-aplusp-05 (work in progress), October 2009.

[flow-counting]
WAND, Network Research Group, "Flow Counting", <http://www.wand.net.nz/~salcock/someisp/flow_counting/result_page.html>.

Authors' Addresses

Andreas Ripke
NEC Laboratories Europe
Kurfuersten-Anlage 36
69115 Heidelberg,
Germany

Phone: +49 6221 4342 252
Email: andreas.ripke@nw.neclab.eu

Juergen Quittek
NEC Laboratories Europe
Kurfuersten-Anlage 36
69115 Heidelberg,
Germany

Phone: +49 6221 4342 115
Email: juergen.quittek@nw.neclab.eu

Marcus Brunner
NEC Laboratories Europe
Kurfuersten-Anlage 36
69115 Heidelberg,
Germany

Phone: +49 6221 4342 129
Email: marcus.brunner@nw.neclab.eu