NAT Behavioral Requirements Updates

draft-penno-behave-rfc4787-5382-5508-bis-04

Reinaldo Penno <rpenno@cisco.com>
Simon Perreault <simon.perreault@viagenie.ca>
Sarat Kamiset
Mohamed Boucadair <mohamed.boucadair@orange.com>
Kengo Naito <kengo@lab.ntt.co.jp>

IETF 86, Orlando
BEHAVE meeting

2013-03-12

Purpose

- Clarify, update and fine-tune a very successful set of RFCs based on operational and implementation experience.
- It applies to all kinds of NAT44
  - Some issues identified also affect NAT64 but that is out of scope
- Collect requirements sprinkled over other documents and reference them here

Status

- Last presented at IETF 82 in Taipei
- From the minutes:
  - Alain: don't be religious; there is value in this document
  - Roberta: there is value, captures operational issues we've found
  - Lars: 1) address holes in current specs; 2) change requirements/recommendations we've already written; 3) describe security issues

Changes

- From -02 to -03:
  - Add reference to draft-naito-nat-resource-optimizing-extension
  - Nits
- From -03 to -04:
  - Merged with draft-naito-nat-resource-optimizing-extension

TIME_WAIT reduction

- It's a scalability problem: tons of TCP NAT state table entries in TIME_WAIT state
- Not to be confused with this from draft-ietf-behave-lsn-requirements:
  - REQ-8: Once an external port is deallocated, it SHOULD NOT be reallocated to a new mapping until at least 120 seconds have passed, with the exceptions being:
    - A. If the CGN tracks TCP sessions (e.g., with a state machine, as in [RFC6146] section 3.5.2.2), TCP ports MAY be reused immediately.
- A NAT either tracks TCP sessions (and therefore goes into TIME_WAIT state), or it does what REQ-8 says for TCP flows.
  - TIME_WAIT is per TCP session
  - REQ-8 is per port (transport protocol agnostic)
  - **DO NOT BE CONFUSED!**

Reducing TIME_WAIT with TCP timestamps

- Proposal: apply RFC 6191 to NAT
- NAT may do TCP sequence number or timestamp rewriting
- Seems straightforward. Any gotchas we should be aware of?

Port overloading behaviour

- RFC 4787 requires Endpoint-Independ Mapping
- Sacalability problem: one external port per 5-tuple uses many external ports
- Proposal: EIM by default, but MAY be non-EIM when the NAT **knows** it won't break the application protocol
  - e.g. HTTP, DNS don't need EIM and account for a lot of traffic
- draft-ietf-behave-lsn-requirements does something like this:
  - REQ-2: A CGN MUST have a default "IP address pooling" behavior of "Paired" (as defined in [RFC4787] section 4.1). **A CGN MAY provide a mechanism for administrators to change this behavior on an application protocol basis.**
- OK to proceed similarly in this case?

Next steps

- Adopt?
  - Does the WG want to work on updates to BEHAVE's core NAT behaviour documents?
- Once adopted, iterate with reviews and new revisions