Takeover Suggestion Flag for the ENRP Handle Update Message
draft-dreibholz-rserpool-enrp-takeover-21

Abstract

This document describes the Takeover Suggestion Flag for the
ENRP_HANDLE_UPDATE message of the ENRP protocol.

Status of This Memo

This Internet-Draft is submitted in full conformance with the
provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering
Task Force (IETF). Note that other groups may also distribute
working documents as Internet-Drafts. The list of current Internet-
Drafts is at https://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on September 7, 2019.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the
document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal
Provisions Relating to IETF Documents
(https://trustee.ietf.org/license-info) in effect on the date of
publication of this document. Please review these documents
carefully, as they describe your rights and restrictions with respect
to this document. Code Components extracted from this document must
include Simplified BSD License text as described in Section 4.e of
the Trust Legal Provisions and are provided without warranty as
described in the Simplified BSD License.
1. Introduction

Reliable Server Pooling as described in [RFC5351] defines protocols for providing highly available services. The management component used for pool administration is denoted as ENRP Server or Pool Registrar (PR). Since a single ENRP server constitutes a single point of failure, there must be multiple ENRP servers. Servers, denoted as Pool Elements (PE), use an arbitrary ENRP server for registration into the pool. The chosen ENRP server becomes the Home ENRP Server, also denoted as Home PR (PR-H), of the PE. It is responsible for making the PE identity known to the other ENRP servers (by using ENRP_HANDLE_UPDATE messages) and also to monitor the PE health (by using keep-alive messages).

As shown in [AINA2009], the following scenario leads to unbalanced ENRP server workload: consider a set of multiple ENRP servers with one subset being unreliable (for example, their network connection has problems) and some reliable ENRP servers. After a while, the reliable ENRP server will get the home ENRP server role for almost all of the PEs, which results in high workload for this ENRP server. Since the home ENRP server role is more computation-intensive (as shown by [IJHIT2008]), this leads to highly unbalanced workload for large RSerPool setups. This unbalanced workload remains, even when the unreliable ENRP servers become reliable again (for example, when the network problems have been solved).

1.1. Scope

The Takeover Suggestion Flag defined in this draft defines a flag for the ENRP_HANDLE_UPDATE message. If the flag is set, the receiving
ENRP server is suggested to take over the PE specified in the ENRP_HANDLE_UPDATE message.

1.2. Terminology

The terms are commonly identified in related work and can be found in the RSerPool Overview document RFC 5351 [RFC5351].

1.3. Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Takeover Suggestion Flag

2.1. Definition

In this subsection, only the differences to the ENRP_HANDLE_UPDATE message defined in [RFC5353] are explained. The following figure shows the ENRP_HANDLE_UPDATE message:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type = 0x04 |0|0|0|0|0|0|0|T|        Message Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Sending Server’s ID                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Receiving Server’s ID                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Update Action          |        (reserved)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Pool Handle Parameter                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Pool Element Parameter                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

T flag: 1 bit (boolean)

If set, the receiving ENRP server is suggested to take over the PE specified by the Pool Handle and Pool Element Parameters. It is RECOMMENDED for the receiving ENRP server to perform this takeover if it has the resources to do so.
3. Reference Implementation

The RSerPool reference implementation RSPLIB can be found at [RSerPool-Website]. It supports the functionalities defined by [RFC5351], [RFC5352], [RFC5353], [RFC5354] and [RFC5356] as well as the options [I-D.dreibholz-rserpool-asap-hropt], [I-D.dreibholz-rserpool-delay] and of course the option defined by this document. An introduction to this implementation is provided in [Dre2006].

4. Testbed Platform

A large-scale and realistic Internet testbed platform with support for the multi-homing feature of the underlying SCTP protocol is NorNet. A description of NorNet is provided in [PAMS2013-NorNet], some further information can be found on the project website [NorNet-Website].

5. Security Considerations

Security considerations for RSerPool systems are described by [RFC5355].

6. IANA Considerations

This document does not require additional IANA actions beyond those already identified in the ENRP and ASAP protocol specifications.

7. References

7.1. Normative References

[I-D.dreibholz-rserpool-asap-hropt]
Dreibholz, T., "Handle Resolution Option for ASAP", draft-dreibholz-rserpool-asap-hropt-23 (work in progress), September 2018.

[I-D.dreibholz-rserpool-delay]


7.2. Informative References


Authors’ Addresses

Thomas Dreibholz
Simula Metropolitan Centre for Digital Engineering
Pilestredet 52
0167 Oslo, Oslo
Norway

Phone: +47-6782-8200
Fax: +47-6782-8201
Email: dreibh@simula.no
URI: https://www.simula.no/people/dreibh