RSERPOOL Redundancy-model Policy
draft-xie-rserpool-redundancy-model-00.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on September 25, 2003.

Copyright Notice

Copyright (C) The Internet Society (2003). All Rights Reserved.

Abstract

This document defines RSERPOOL Redundancy-model Member Selection Policy parameter and the related procedures. This policy is designed to be flexible and capable of supporting a wide range of advanced redundancy models found in some high availability systems. The design uses the extensibility in RSERPOOL pool load sharing policy.
Table of Contents

1. Introduction ............................................. 3
1.1 Definitions ............................................... 3
2. Conventions ............................................... 3
3. Redundancy-model Member Selection Policy Parameter ....... 3
4. Member Selection Procedures under Redundancy-model Policy .. 5
5. Examples for Implementing Various Redundancy Models ....... 6
5.1 2N Redundancy with Active/Standby Protection ........... 6
5.2 N+M Redundancy with Active/Standby Protection .......... 6
5.3 N-way Redundancy with Active/Standby Protection ........ 7
5.4 N-way Redundancy with Active/Active Protection .......... 7
6. Security Considerations ................................... 7
   Normative References ...................................... 7
   Authors’ Addresses ........................................ 8
   Intellectual Property and Copyright Statements .......... 9
1. Introduction

The pool member selection policy in RSERPOOL design [2][3] is both flexible and extensible. By observing the pool’s member selection policy a pool user (PU) can accomplish a wide range of different load distribution schemes and communications strategies across the pool elements (PE) in the pool.

Some most basic load distribution schemes, such as the round-robin and least used policies, are already defined in ASAP [1]. In this document, we define another member selection policy – Redundancy-model policy – in order to support a variety of more advanced redundancy models used in some high availability systems.

1.1 Definitions

This document uses the following terms:

Pool (or server pool): See [3];

Pool element (PE): See [3];

Pool user (PU): See [3];

ENRP namespace (or namespace): See [3];

ENRP namespace server (or ENRP server): See [3];

2. Conventions

The keywords MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, NOT RECOMMENDED, MAY, and OPTIONAL, when they appear in this document, are to be interpreted as described in [5].

3. Redundancy-model Member Selection Policy Parameter

This defines a new Member Selection Policy TLV structure that SHOULD be used to indicate the role of a PE in a high availability redundancy model. This definition follows the member selection policy extension rules described in Section 3.6? of [6]
The Type, Length, and Policy fields are defined in Section 3.6? in [6], with the addition of the following new Policy value:

<table>
<thead>
<tr>
<th>Value</th>
<th>Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x5</td>
<td>HA Redundancy-model support</td>
</tr>
</tbody>
</table>

Other fields are defined as follows:

**High Availability (HA) role:** 8 bits (unsigned integer)

Indicating the role of the corresponding PE in the HA redundancy model. The following HA roles are currently defined:

<table>
<thead>
<tr>
<th>Value</th>
<th>HA state</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x1</td>
<td>No role</td>
</tr>
<tr>
<td>0x2</td>
<td>Standby</td>
</tr>
<tr>
<td>0x3</td>
<td>Active</td>
</tr>
<tr>
<td>0x4</td>
<td>Dual role</td>
</tr>
</tbody>
</table>

Other values are reserved by IETF and MUST NOT be used.

**Service State:** 8 bit (unsigned integer)

Indicating the current service state elected by the corresponding PE. The following service states are currently defined:

<table>
<thead>
<tr>
<th>Value</th>
<th>Service state</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x1</td>
<td>In service</td>
</tr>
<tr>
<td>0x2</td>
<td>Stopping</td>
</tr>
</tbody>
</table>

Other values are reserved by IETF and MUST NOT be used.
Backup PE Identifier Parameter:

This TLV is used in some redundancy models to indicate the designate backup of this PE. The format of PE Identifier Parameter is defined in [6]

More detailed description of the HA states and Service states are given in the following sections.

4. Member Selection Procedures under Redundancy-model Policy

When sending to a pool of Redundancy-model policy, a PU MUST use the following rules to select the recipient of the message:

1. When sending a new message for the first time to the pool, the PU SHOULD select a PE whose HA role is either "Active" or "Dual role" and service state is "In service." If more than one PEs in the pool qualify these criteria, the PU SHOULD select one of them as the recipient of this message in a round-robin fashion.

2. When the PU is re-sending a message which was previously sent to a PE (PE_A) in the pool but failed because PE_A is found unreachable (i.e., a fail-over case), the PU SHOULD select a qualified backup PE as the message recipient, by following the steps below:

   A. First check if PE_A has a designated backup PE, whose presence and PE Id are indicated by the presence of the "Backup PE Identifier Parameter" in the Member Selection Policy TLV of PE_A.

   B. If no backup is designated by PE_A, the PU SHOULD select a PE whose HA role is either "Standby" or "Dual role", and service state is "In service." If no such PE can be found in the pool, the PU SHOULD abort re-sending the message and notify the application about the failure. If multiple PEs exist in the pool that meet these criteria, the PU SHOULD select one of them in a round-robin fashion.

   C. If a backup (PE_B) is designated by PE_A and exists in the pool and its service state is "In service", the PU SHOULD select PE_B as the recipient of the message being re-sent, regardless of the HA state of PE_B.

   D. If a backup (PE_B) is designated by PE_A but does not exist in the pool (i.e., the PE Id in the "Backup PE Identifier Parameter" in PE_A does not match any PE in the pool), the PU SHOULD abort re-sending the message and notify the application...
about the failure.

E. If a backup (PE_B) is designated by PE_A and exists in the pool but its service state is NOT "In service", the PU SHOULD attempt to find a qualified backup for PE_B, by recursively invoking the above steps.

In the case where the failed PE does not have a designated backup and the sending PU has succeeded in finding a qualified backup PE for the failed PE using the above procedures, the PU MAY keep a record of its finding and use the same backup for all subsequent fail-over events that may occur on the same failed PE.

5. Examples for Implementing Various Redundancy Models

5.1 2N Redundancy with Active/Standby Protection

In this redundancy model there are two PEs registered in the pool; one has its HA role set to Active and the other’s set to Standby. The active PE is providing the service while the standby PE is prepared to take over the active role should the active PE fail.

In this redundancy model, the active PE does not have to have a designate backup, i.e., it does not need to include a "Backup PE Identifier Parameter" in its selection policy TLV when registering.

5.2 N+M Redundancy with Active/Standby Protection

This redundancy model allows flexible active-standby relationships to be defined between the PEs in a pool. A pool with the N+M redundancy model consists of N PEs registered with their HA role set to Active (active PEs) and M PEs registered with their HA role set to Standby (standby PEs).

Normally, the value of M is less than or equal to N, which means that a single PE can act as a standby for several active PEs at a time. A common use of the N+M redundancy model is N+1 redundancy. This is a special case of the N+M redundancy model. In the pool using N+1 redundancy, a single PE acts as a standby for all active PEs.

In operation, the active PEs are provided the service while the standby PEs are prepared to become backup to any of the active PEs, should one or more of them fails.

In this redundancy model, the active PEs do not have to have a designate backup, i.e., they do not need to include a "Backup PE Identifier Parameter" in their selection policy TLV when registering.
5.3 N-way Redundancy with Active/Standby Protection

In this redundancy model, there are N (an even number) PEs in the pool. The PEs are paired to form N/2 active/standby pairs. In each pair, the active PE is providing the service while the standby PE is prepared to take over the active role should the active PE fails. Load is distributed across the N/2 active PEs.

To implement such a pool, at registration each active PE needs to set its HA role to Active and indicate its designated backup by including a "Backup PE Identifier Parameter" in its selection policy TLV that contains the PE id of the backup. The standby PEs simply register with their HA role set to Standby and omit the "Backup PE Identifier Parameter."

5.4 N-way Redundancy with Active/Active Protection

In this redundancy model there are one or more PEs in the pool, and are registered with their HA role set to Dual role. All PEs in the pool are actively providing service and are prepared to backup others at the same time.

Note, this redundancy model can also be implemented (in some cases more efficiently) by using other active/active member selection policies such as the Round Robin and Weighted Round Robin policies (see Section 4.5.2? [1]).

6. Security Considerations

TBD

Normative References


Authors’ Addresses

Qiaobing Xie
Motorola, Inc.
1501 W. Shure Drive, 2-F9
Arlington Heights, IL 60004
US

Phone: +1-847-632-3028
EMail: qxie1@email.mot.com

La Monte H.P. Yarroll
US Army Corps of Engineers
ERDC-CERL
P.O. Box 9005
Champaign, IL 61826
USA

Phone:
EMail: piggy@acm.org
Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF’s procedures with respect to rights in standards-track and standards-related documentation can be found in BCP-11. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

Full Copyright Statement

Copyright (C) The Internet Society (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION
HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.