Technical white paper

HP RTR Failover in multihome setup

Failover mechanism and timeout calculation

Table of contents

Introduction

This document elaborates on Network Interface Card (NIC) registration, failover mechanism, and failover time calculation in multihome setup of HP Reliable Transaction Router (RTR).

NIC registration

In a multihome setup, ensure that the `gethostbyname( )` API returns the list of all possible network addresses for the host. RTR determines the set of IP addresses to be used for a remote host when the host name is looked up using the `gethostbyname( )` API. The possible addresses include all the NICs of a host. In the event of a connection failure, RTR tries all possible addresses one-by-one to failover to the next node in the configuration.

Failover mechanism

Machines with multiple network adapters allow multiple IP connection targets for any host. With this capability, any pair of machines connected by multiple network paths can failover to an alternate path if the primary path becomes unusable. Additionally, multiple network paths can be established using multiple NICs. If the NIC of a node fails, RTR tries to connect to all the known NICs of the node.

The example below describes failover mechanism:

Node A (BackEnd) : 1.1.1.2

Node B (Router)   : 1.1.1.1   - Reachable

                        : 1.1.2.1   - Reachable (Different Subnet)

                        : 1.2.1.1   - Not reachable (Different Subnet)

Node C (Router)   : 1.1.2.2   - Reachable

If:

- RTR BackEnd on Node A is connected to RTR Router on Node B via IP address 1.1.1.1 of Node B.
- RTR on Node A knows all 3 addresses; 1.1.1.1, 1.1.2.1, and 1.2.1.1 of Node B (using `gethostbyname( )`).

Failover scenario:

    On Node B, NIC configured with address 1.1.1.1 fails.

Failing over to the next Router:

    If,

    Wait_Time = RTR_TIMEOUT_CONNECT + RTR_TIMEOUT_CONNECT_RELAX.

    RTR on Node A will try to connect to the IP address 1.1.1.1 and wait for Wait_Time number of seconds before it tries to connect to the next address, that is 1.1.2.1 of Node B. If this NIC is also down, after Wait_Time seconds, RTR will try to connect to the next IP address (1.2.1.1) and wait for Wait_Time seconds, even though it is not reachable. When the connection request to all the addresses of Node B fails, RTR BackEnd will failover to the next available Router node in the configuration, that is Node C with IP address 1.1.2.2.

Calculation of failover time

Link failover depends on the stipulated values of adjustable environmental timer parameters:

- RTR_TIMEOUT_CONNECT – The time (in seconds), within which the node should respond to the connect request. The default value is 60 seconds.
- RTR_TIMEOUT_CONNECT_RELAX – The time (in seconds), a node will wait/relax before sending connect request to the next known IP. The default value is 90 seconds.

In case there is a failure, RTR detects it within the stipulated timer parameters, and disconnects and retries the link according to router preferences. If a router fails to respond to the reconnect attempt, there will be a time lapse of RTR_TIMEOUT_CONNECT plus RTR_TIMEOUT_CONNECT_RELAX for the link failover to occur. In case of a multihome system, the total time required to failover to the next node will be proportionate to the number of unreachable NICs on the node.

In the scenario above, If RTR_TIMEOUT_CONNECT is set to 60 seconds and RTR_TIMEOUT_CONNECT_RELAX is set to 90 seconds, the total time needed for failing over to Node C will be

(60 + 90) seconds * 3(NICs) = 450 seconds or 7.5 minutes

This calculation is independent of the placement of the NICs (same or different subnet).

Restriction in polling NICs

During failover, RTR tries to connect every known IP address of a multihomed node, to identify if the node is reachable. Hence, RTR will try to connect to both the reachable as well as unreachable IP addresses during the failover. The failover is possible across subnets since each NIC can belong to different subnets. However, ensure that the IP addresses are known to RTR as mentioned above. RTR cannot dedicate IPs between nodes for communication. Also, RTR does not limit the failover when Routers have more IPs. It tries all the configured IPs, before attempting a failover.

RTR validates a link only by attempting a connection over the link. Until a connection request is attempted, RTR does not know which interfaces are dedicated to it. It gets all the addresses from the underlying network layer. During the failover, RTR tries all the known interfaces and failovers to next node only if it does not get response from these interfaces. Considering the RTR requirements toward having a consistent view of the network, restricting RTR to poll only specified NICs increases the complexity of RTR administration, makes it prone to errors, and may also lead to scalability issues.

For more information

To learn more about RTR, visit hp.com/go/rtr

To learn more about HP OpenVMS, visit hp.com/go/openvms

Sign up for updates
hp.com/go/getupdated

Rate this document