Dynamic Routing Protocol Performance in a Fault-Tolerant Ethernet-based IP Network

Jan Baranski, Marshall Crocker, Georgios Lazarou
Department of Electrical and Computer Engineering
Mississippi State University
Mississippi State, MS 39762
{jdb2, mac1, glaz}@ece.msstate.edu

ABSTRACT

Using a hybrid star-ring shaped Ethernet network with a fault-tolerant reconfigureable Ethernet proxy (FREP), we experimentally evaluated the performance of the RIP, OSPF, and EIGRP dynamic routing protocols. The protocols were analyzed in their ability to provide fast convergence rates in the FREP environment with a minimum trade-off in CPU utilization and network overhead. Our results show that all three protocols can be finely tuned to provide the desired convergence rates and are well suited for operation in a simple ring-shaped Ethernet-based network. RIP is well-suited for this application due its inherent support of routing loops. It's main drawback is the broadcasting of $\sim$1kB packets at the specified update interval. OSPF and EIGRP also performed well with their main disadvantage being configuration complexity and increased CPU overhead. These two protocols, however, possess additional advanced features that allow them to scale well into larger and more complex networks. Convergence rates as low as 3.3 sec. were achieved with all three routing protocols as opposed to an average of 40 sec. using each protocol's default settings.

Keywords: Communication system performance, Communication system routing, Ethernet, Fault-tolerance, IP Network, Proxy, and Reconfigurability

1. INTRODUCTION

Typical Ethernet-based IP networks often do not provide fault tolerant capabilities, and more expensive technologies are required to maintain network connectivity in the event of a failure. In [1], we proposed a network proxy device mated with a hybrid star-ring topology as a means of providing an automatically reconfigurable fault-tolerant solution for low-cost Ethernet-based IP networks. The implementation uses standard off-the-shelf equipment and does not require specialized network hardware. We tested the functionality of our solution based on its operation in an OSPF routing environment. The routing protocol timers/parameters were, however, not optimized, and the tests simply provided a proof-of-concept.

*This work was supported by the Office of Naval Research (ONR).

In this paper, we evaluate three dynamic routing protocols, Routing Information Protocol (RIP), Open Shortest Path First (OSPF), and Enhanced Interior Gateway Routing Protocol (EIGRP). Although previous work exists in which routing protocols have been compared ([2], [3]), no comparisons of the protocols' performances in response to a link failure in a small Ethernet-based network have been documented. Additionally, prior studies have generally employed simulation-based techniques and do not provide an experimental basis for the comparisons.

Each of the three above-mentioned dynamic routing protocols behaves differently upon detection of a failed link, and each has distinct advantages and disadvantages. Some comparisons of their performance and associated overhead traffic, however, can be made. Using test scripts and our previously designed fault-tolerant Ethernet-based IP network proxy (FREP), we measured the duration of network outage after a link failure while running one of the aforementioned dynamic routing protocols.

The design and operation of RIP, OSPF, and EIGRP has been described in detail in various sources ([4], [5], [6]). Their distinct characteristics and advantages/disadvantages are also often cited. In this paper, however, we have presented an approach to the evaluation of the routing protocols based on their performance and ability to reconfigure quickly after a link failure. We tested the protocols in a simple Ethernet-based network and evaluated their performance based on convergence rates and additional network and CPU overhead.

The paper is organized as follows. In section 2, we describe our testbed network in which the analysis was performed. Section 3 describes the test protocol used to test each of the routing protocols. We present our results in Section 4, and we conclude in Section 5.

2. TEST ENVIRONMENT

The testbed network used in our analysis was previously constructed for the research performed in [1]. It is a

switched Ethernet network based on a hybrid star-ring topology as shown in Fig. 1. The network consists of three router/subnet pairs. Thus, each subnet is connected to the backbone via two routers. One router is designated as the default gateway for each node on the subnet and is referred to as the primary router. The other router is referred to as the backup router and does not route any outbound traffic during normal operation. Additionally, one subnet contains the FREP described in [1] to provide fault-tolerance. The FREP monitors connectivity to the primary router and reroutes traffic to the secondary router in the event of a failure. The subnet containing the FREP was used as the origin for the performance tests of the routing protocols. Identically configured Cisco 1720 modular access routers and Linksys EtherFast 10/100 Ethernet switches were used for the physical network implementation. Servers running RedHat Linux 7.3 reside on each subnet and were used to perform the measurements.

3. TEST PROTOCOL

The performance of the RIP, OSPF, and EIGRP protocols was analyzed in our testbed network with the aid of the FREP. By using the FREP to forward packets to the secondary router after a link failure, we were able to measure the convergence rate of the dynamic routing protocols. The measurements were made from an end-user point of view by calculating the duration of network outage using TCP traffic with the RIP protocol and ICMP traffic with OSPF and EIGRP. The link failures were simulated by disconnecting the primary router from the network. Operation of the FREP and the TCP- and ICMP-based tests are described below.

Frep Operation

The FREP ensures connectivity by monitoring the echo port of the primary router. Reconfiguration into fail-over mode takes place if it is unable to establish a connection.

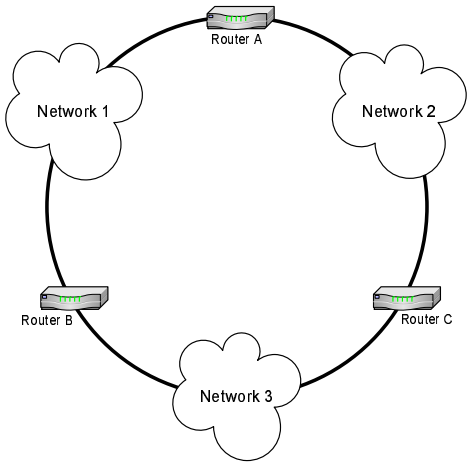

Fig. 1. The analysis of the dynamic routing protocols was performed in a tested network that was constructed based on the star-ring topology.

Once in fail-over mode, the FREP takes over the IP address of the primary router (via ARP spoofing) and forwards all traffic to the secondary router. Thus, all outbound traffic from the network nodes that is intended for the primary router is transparently forwarded to the secondary router. Previously, the FREP was measured to have a reconfiguration time of 1.55 sec [1]. Since the FREP reconfigures faster than any of the routing protocols studied, no additional delays in the routing of packets are introduced by the FREP after a link failure and prior to reconfiguration. That is, the FREP is able to forward packets to the secondary router before the secondary router is able to properly route packets to their next hop along the backbone ring.

RIP

The IOS software used in Cisco routers always performs packet switching when two equal paths to a particular destination exist. The `ip route-cache` command controls the use of the high-speed switching caches in the routers. By default, switching occurs on a per-destination basis with the router keeping a cache of the preferred route to a specific destination. When two equal paths exist to a particular destination and a non connection-oriented protocol such as ICMP or UDP is used, the IOS software alternates the path that the packets travel. When ICMP or UDP packets are forwarded to the backup router after a link failure, it instantly sends packets to their destination via an alternate route. Thus, it is not possible to accurately measure the convergence rate of RIP using one of these protocols. By establishing a TCP connection before the link failure occurs, however, we are able to prevent the router from using the alternate route until convergence of the RIP protocol takes place.

In order to test the performance of the RIP protocol, a Perl application that communicates with a server via the TCP protocol was used. The program consists of a client and server and is responsible for both the generation of traffic via a TCP stream and the measurement of network outages. Upon connection to the server, the client program performs a fork and spawns a child process. The child captures network traffic via the tcpdump utility and calculates inter-packet arrival times of traffic from the server. The parent process, meanwhile, maintains a TCP connection with the server by sending 2-byte packets to the server at a 0.01 sec. interval. The server also responds with an identical stream of traffic to produce a bi-directional stream. Thus, by establishing a TCP connection with a remote machine and measuring the inter-packet delay for incoming traffic, we are able to measure the duration of network outage after a link failure and prior to convergence of the RIP protocol.

OSPF and EIGRP

A similar utility to the TCP-based version described above was written for ICMP-based testing of the OSPF

and EIGRP protocols. Due to the nature of ICMP echoes, however, a separate client and server were not required. The program behaves identically to its TCP-based counterpart, with the exception of the type of traffic that is generated. The ICMP-based test program initiates a periodic ICMP ping at a rate of 1 packet/0.01 sec. It also captures the replies and calculates the duration of time between subsequent responses.

Attempts to use TCP-based tests with the OSPF and EIGRP protocols yielded results that matched the operation of the TCP retransmission timer (RTO). If a TCP connection was established prior to the link failure, the TCP protocol would make attempts to reestablish connectivity based on the RTO. Thus, regardless of when convergence of the routing protocol took place, the measured duration of network outage was always a result of the RTO. The RTO is calculated as described in [7]. This behavior forced us to use ICMP-based tests during analysis of the OSPF and EIGRP protocols.

4. RESULTS

The duration of network outages while running the RIP, OSPF, and EIGRP protocols was measured using custom-designed utilities as described above in 3. Cisco IOS C1700 Software version 12.2(11)T2 was used and, unless otherwise mentioned, the parameters of the routing protocols were left at their default values. The results for analysis of each protocol are outlined individually below.

Routing Information Protocol (RIP)

The RIP protocol allows four timers to be controlled by the user: update, invalid, holddown, and flush. These timers control the behavior of the protocol and have a direct effect on performance. The update timer controls the rate at which routing updates are sent. The invalid timer specifies the interval of time from the last update after which a route is declared invalid. In default configurations, the invalid timer is typically set to six times the update interval. The holddown timer specifies how long the hold-down phase lasts after a route has been declared invalid. During the hold-down phase, the router will not process any additional updates it receives regarding the particular route. The flush timer controls how long a router waits after receiving the last update before removing a route from the routing table. The flush timer overrides the hold-down timer and can be set such that the hold-down phase is never entered. By adjusting these three parameters and using a TCP-based test program, we measured the performance of the RIP protocol. Twenty trials were performed for each test case and the results are shown in Table I.

In all test cases, the hold-down and flush timers were set to a value that would avoid the hold-down phase of the RIP protocol altogether. Additional data also showed that the update timer will affect the deviation of the network outage duration, and overhead associated with the TCP protocol will add additional delays. In all five cases, connectivity was re-established within 1.60 sec. of the invalid timer (before the expiration of the flush timer).

TABLE I

AVERAGE DURATION OF NETWORK OUTAGE AFTER LINK FAILURE MEASURED VIA A TCP-BASED ANALYSIS OF THE RIP PROTOCOL

RIP Timers (sec.)				**Outage Duration (sec.)**
Update	Invalid	Holddown	Flush	
1	3	3	6	3.29
2	6	6	8	7.35
3	9	9	12	10.60
4	12	12	16	12.84
5	15	15	20	15.52

TABLE II

AVERAGE DURATION OF NETWORK OUTAGE AFTER LINK FAILURE MEASURED VIA AN ICMP-BASED ANALYSIS OF THE OSPF PROTOCOL

OSPF Timers (sec.)[1]		**Outage Duration (sec.)**
Hello	Dead	
1	3	2.94
1	5	5.01
1	10	10.04
1	15	14.86
5	20	18.46

[1] The SPF delay and SPF hold timers were set to 0 sec. for all test cases.

OSPF

Tests of the OSPF protocol were performed using ICMP traffic in the same manner as TCP traffic was used to analyze the RIP protocol. Four timers exist that affect the performance of the OSPF protocol, the hello timer, dead timer, SPF delay timer, and SPF hold timer. The hello timer controls how often a router sends “hello” packets to any listening routers in the same routing area. The dead timer specifies the amount of time after receiving the last hello packet after which it declares its neighbor “dead.” The SPF delay timer specifies the amount of time that the router waits before performing the SPF calculation after receiving a routing update. Finally, the SPF hold timer defines the amount of time a router waits between performing subsequent SPF calculations. The SPF delay and SPF hold timers were both set to 0 during the analysis. In a small network, these two timers can typically be set to a low value because large SPF calculations will not be performed often. Thus, we adjusted the hello timer and the dead timer to measure the convergence rates of the OSPF protocol. Twenty trials were performed for each test case, and results of the analysis are shown in Table II.

In all test cases, connectivity was reestablished within 1.54 sec. of the dead timer. OSPF, however, does not

TABLE III
AVERAGE DURATION OF NETWORK OUTAGE AFTER LINK FAILURE MEASURED VIA AN ICMP-BASED ANALYSIS OF THE EIGRP PROTOCOL

EIGRP Parameters		Outage Duration (sec.)
Hello interval (sec.)	Hold time (sec.)	
1	3	2.73
1	5	4.58
1	10	9.51
5	15	12.65
5	20	17.54

cache multiple routes in the routing table. Thus, unlike RIP, the problematic routes were completely flushed from the routing table before packets were properly routed around the failed link. A higher hello interval also increased the deviation of the duration of network outage.

EIGRP

Our analysis of EIGRP was identical to the ICMP-based analysis performed on the OSPF protocol. Two parameters can be tuned by the user to control the performance of EIGRP, the hello interval and the hold time. The hello interval specifies the frequency at which a router sends "hello" packets. The hold time defines how long a router will wait before flushing a route from its table after receiving the last hello packet. By adjusting these two parameters, we were able to perform an analysis in which we measured the convergence rates of the protocol. Once again, twenty trials were performed, and the results are shown in Table III.

Network connectivity was reestablished faster than the hold time in all test cases. Increasing the hello interval also increased the resolution of the duration of network outage as seen with the OSPF and RIP protocols.

5. CONCLUSION

RIP

Results of the RIP analysis show that, in all test cases, the flush timer of the protocol does not need to expire in order for routing to resume after a link failure. Prior to a link failure, the router possesses two valid routes to a particular network; both routes are used in an alternating fashion when "fast switching" is enabled (default setting) in the IOS software. The router alternates use of the two paths based on a per-connection policy. Thus, if a particular TCP stream is using a route in which a link failure occurs, the router will not reroute traffic to the alternate route until the invalid timer for the failed route expires. The analysis results verify that only the invalid timer must expire, and the route does not need to be completely flushed before the alternate route will be effectively used for all traffic. Additionally, the update timer can be controlled by the user to provide a finer resolution of the convergence rate. For example, with an invalid timer setting of 15 sec. and an update timer setting of 5 sec., convergence will occur in the range of 10-20 sec. after a failure. With an invalid timer setting of 15 sec. and an update timer setting of 1 sec., however, convergence will occur within 14-16 sec. of a failure.

The routing information protocol (RIP) is one of the more widely used dynamic routing protocols today. It is a relatively simple Distance Vector protocol and is often used to provide inter-network routing for small WANs/LANs. A RIP version 2 has also been developed and adds advanced features to the protocol. RIP inherently supports an environment in which routing loops are present, which makes it well-suited for networks with a ring-shaped backbone topology. In this type of network, only the invalid timer of the protocol must expire before proper routing is resumed after a link failure. The additional overhead associated with an increased RIP update frequency must be considered carefully for low bandwidth networks, however. The maximum size of a RIP update is 1,080 bytes. Thus, given the worst case, 1,080 byte packets will be broadcast to each connected network every time the update timer expires. Although RIP is able to provide a fast convergence rate in ring-shaped networks, a significant amount of available bandwidth may be sacrificed to the routing protocol in such an environment. It should also be considered that in order to avoid the "count to infinity" problem, RIP distances are limited to 15 hops. Thus, network rings routed using the RIP protocol are limited to 16 nodes.

OSPF

Results from the ICMP-based analysis of the OSPF protocol indicate that convergence in a ring-shaped network takes place after the dead timer for a particular route expires. The duration of the outage can be further controlled via the hello timer. The effect of adjusting the OSPF hello timer is identical to that of adjusting the RIP update timer – the resolution of the duration of network outage is affected accordingly. Unlike RIP, OSPF routers do not broadcast their entire routing table each time the hello timer expires. OSPF sends only partial updates (Link State Advertisements – LSAs) when the effective topology of the network changes. Furthermore, the hello packets are minimal in size (68 bytes for a router with 20 neighbors) and do not contain any routing information. The OSPF protocol also supports multicasting, which allows it to keep from flooding the entire network with LSA traffic. A major trade-off to fault-tolerance, however, is the amount of processor time required by the protocol. Because OSPF is a Link State protocol that uses Djikstra's algorithm, shortest path first (SPF) calculations must be made each time an LSA is received. In a complex network, these calculations can become processor intensive and affect the ability of the router to properly route traffic. In a simple ring-shaped network, however, this issue does not pose a

problem. Additionally, OSPF supports advanced features such as separate routing areas and variable length subnet masks, which allow it to scale well to large and complex networks.

EIGRP

As similarly observed with the OSPF and RIP protocols, results of our analysis show that the convergence rate of the EIGRP protocol can be tightly controlled by two parameters, the hold time and the hello interval. The hold time directly affects the duration of network outage, while the hello interval can be used to alter the possible range of convergence rates. EIGRP is a proprietary Cisco protocol that supports many advanced features, which allow it to scale well into complex modern networks. Additionally, EIGRP supports multicasting and uses small hello packets to maintain neighbor relationships. It is an advanced Distance Vector protocol and offers vast improvements over more traditional protocols such as RIP. EIGRP decreases convergence rates by employing a diffusing update algorithm (DUAL). Although the EIGRP protocol provided the fastest convergence rates and lower processor usage than OSPF, its main drawback during deployment is the requirement of Cisco routers.

Summary

Through our analysis, we have experimentally demonstrated that RIP, OSPF, and EIGRP can be customized to provide the desired level of performance by adjusting the appropriate parameters. Specifically, the EIGRP protocol provided the fastest convergence rates in relation to its hello interval. Generalizations such as "a slow convergence rate" are often encountered regarding some of these protocols. We have shown, however, that the convergence rates can be adjusted by the user. Trade-offs, such as increased processor time and additional network traffic, however, must also be carefully considered during deployment. Due to the variety and possible complexity of today's networks, a routing protocol must be chosen based on a number of factors, only one of which is the convergence rate. Additionally, we have shown that RIP, OSPF, and EIGRP can be used to effectively provide fault-tolerance in a simple Ethernet-based high-speed network. Although our tests were performed within a hybrid star-ring topology, the principles presented here can be applied to any high-speed network based on an arbitrary 2-connected topology to achieve similar levels of performance from all three protocols.

6. REFERENCES

[1] J. Baranski, M. Crocker, and G. Lazarou, "Fault-Tolerant Reconfigurable Ethernet-Based IP Network Proxy," **Proceedings of The 2nd IASTED International Conference on Communications, Internet, and Information Technology (CIIT)**, 2003, pp. 214-220.

[2] W. Zaumen and J. Garcia-Luna-Aceves, "Steady-state Response of Shortest-path Routing Algorithms," in **Proc. IPCCC**, 1992, pp. 323-332.

[3] Y. Zhao, X. Yin, B. Han, and J. Wu, "Online Test System Applied in Routing Protocol Test," in **Proc. Ninth International Symposium on Analysis and Simulation of Computer and Telecommunication Systems**, 2001, pp. 331-338.

[4] R. Malhotra, **IP Routing**. Sebastopol, CA: O'Reilly & Associates, pp. 10-156.

[5] L. Petersen and B. Davie, **Computer Networks: A Systems Approach, Second Edition**. San Francisco, CA: Morgan Kaufman Publishers, pp. 280-309.

[6] A. Leon-Garcia and I. Widjaja, **Communication Networks: Fundamental Concepts and Key Architectures**. Boston, MA: McGraw Hill, pp. 590-618.

[7] Internet Engineering Taskforce and Internet Engineering Steering Group: Network Working Group. RFC 2988: Computing TCP's Retransmission Timer. [Online] Available: http://www.rfc-editor.org/rfc/rfc2988.txt