SCIENCE CHINA
Information Sciences

• RESEARCH PAPERS •

March 2011 Vol. 54 No. 3: 609–622
doi: 10.1007/s11432-010-4177-4

Self-healing routing: failure, modeling and analysis

XU MingWei[1,2*], LI Qi[1,2*], YANG Yuan[1,2], HOU MeiJia[1,2] & PAN LingTao[1,2]

[1]*Department of Computer Science, Tsinghua University, Beijing* 100084*, China;*
[2]*Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing* 100084*, China*

Received May 3, 2010; accepted September 16, 2010; published online February 4, 2011

Abstract Network failures occur frequently, and self-healing ability of existing routing protocols cannot guarantee fast route convergence under these failures without impacting packet forwarding. During routing convergence, network routes may be incorrect and even routing black holes and loops occur, which will result in extensive packet loss and thus influence network performance. To solve this problem, several improved routing solutions have been proposed. In this paper, we propose the concept and model of self-healing routing and analyze the key problems in current intra-domain and inter-domain self-healing routing protocols after briefly reviewing the characteristics of network failures. We classify the different self-healing solutions into two categories based on their design principles: routing restoration and routing protection, and systematically analyze these different typical solutions. Finally, we discuss several key issues in self-healing routing and propose a hybrid protection and restoration based routing scheme.

Keywords failure, intra-domain routing, inter-domain routing, self-healing routing

Citation Xu M W, Li Q, Yang Y, et al. Self-healing routing: failure, modeling and analysis. Sci China Inf Sci, 2011, 54: 609–622, doi: 10.1007/s11432-010-4177-4

1 Introduction

Routers exchange routing information through routing protocols, and achieve best paths by making routing decision based on the received information. Unstable factors are inevitable in networks, such as network failures and topology changes, causing route changes in different source-destination pairs. Route changes will further spread and be delivered to other routers in the Internet, which may interrupt network connectivity and impact network services. These incidents may cause extensive packet loss or even network disruptions. For instance, Taiwan 2006 earthquake caused long-term global network disruptions between China mainland and other areas, including Taiwan, North America and South-Eastern Asia.

Considering the impact of network failures or network robustness, different adaptive routing protocols have been designed and used. However, with the remarkable increase of Internet scale, network failures frequently occur nowadays. Especially, tens and thousands of Internet applications have emerged, e.g., VoIP and real time applications. Users have more and more strict demands for end-to-end network performance. At present, routing information protocol (RIP) [1], open shortest path first (OSPF) [2] and IS-IS [3] are typical intra-domain routing protocols, and border gateway protocol (BGP) [4] is the only inter-domain routing protocol Internet. It is not unusual for RIP to take hundreds of seconds to converge,

[*]Corresponding authors (email: xmw@csnet1.cs.tsinghua.edu.cn; liqi@csnetl.cs.tsinghua.edu.cn)

while OSPF may require tens of seconds and BGP may require several minutes or longer. Especially, RIP may have the count-to-infinity problem [5], and BGP may not converge because of policy conflicts [6]. During convergence process, routing black-holes and loops may be raised [7, 8], which may result in packet loss and large delay and thus influence performance of Internet applications.

Convergence time which is an important metric to evaluate routing performance denotes the transition time between two route stable states after network topology changes. Besides, we also need to consider routing stability and availability. Thus, we propose the concept of routing self-healing to analyze routing performance after network failures. Routing self-healing means that routing systems can automatically rebuild routes or switch to failover routes to guarantee packet forwarding under failures.

To address the slow convergence issue and packet loss and delay during convergence, researchers pay more attentions to improve routing convergence and routing availability. Existing solutions mainly focus on one of these problems, e.g., routing protection only considers to provide failover routes under short-term failure but does not consider the case under long-term failures and multiple failures. Systematical analysis of improved routing solutions should be essential from the view of self-healing routing. In this paper, we analyze characteristics of network failures and routing failures, and firstly propose the concept and the model of self-healing routing and systematically analyze the performance of different existing solutions.

The paper is organized as follows: Section 2 analyzes current typical network failures and failure handling. Section 3 proposes self-healing routing model and key issues. In section 4, we analyze several typical self-healing routing algorithms. Section 5 analyzes the key issues of self-healing routing and proposes an improved self-healing routing scheme. Section 6 concludes the paper.

2 Characteristics of network failures and failure handling

2.1 Characteristics of network failures

There are several measurements investigating the network failure events. Markopoulou et al. [9] have made a measurement study to understand the nature of intra-domain network failures recently, in which the study is carried out based on the early work proposed by Iannaccone et al. [10]. They collected IS-IS routing updates from Sprint network and classified these updates into different categories based on the update root causes, such as maintenance-caused failures, router-related failures, optical-layer-related failures and individual-link failures. The distributions of time between failures and time-to-repair of each class were also analyzed. Their studies have three important results: 1) Failures are hourly and common events. 2) Most failures are transient and only last for less than 100 s. 3) Unstable links have caused most individual link failures.

Although the above analysis provides detailed information on routing failures, it has a serious drawback: the data collected are filtered out by the IS-IS protocol. If IS-IS only uses hello packets to discover link failures, the delay between a failure occurred and the failure detected will be significant. For example, if the hello-interval is 10 s and three consecutive losses of hello packets tear down the adjacency, the delay will be randomly distributed between 20 and 30 s. The distribution of time-to-repair is also imprecise, because the link recovery will be reported only when full adjacency relation is established. Recently, Wang et al. [11] quantified the impact of BGP updates on control- and data-plane and analyzed failure log data in Tier-1 network. In addition, they also analyzed physical failure data, BGP session failure and traffic information when investigating the key issues caused BGP routing failures. Their work has some important findings: 1) The root causes of BGP routing failures are the single-link eBGP session failures. 2) Administrative session resets and link failures are two major causes of session failures, contributing to 46.1% and 30.4%, respectively.

2.2 Failure handling

Typical intra-domain routing protocols currently used are RIP, OSPF and IS-IS, and inter-domain routing protocol is BGP. When a failure occurs, traditional routing convergence may proceed in the following

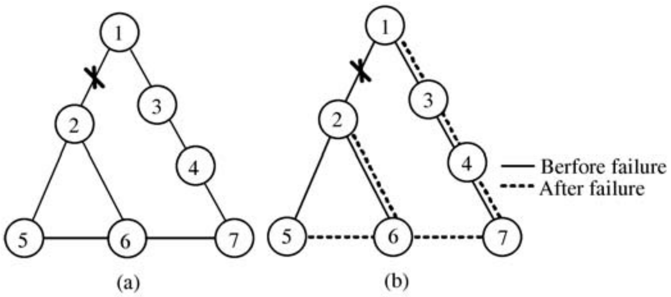

Figure 1 Network failure example. (a) Network topology; (b) routes.

steps:

- Failure detection. Routers exchange their routing information to check if neighbors work well. If there is no routing information received from a neighbor in a relatively long time, the router may deem that a node/link failure occurs.
- Failure advertisement. When a router detects a failure, the router will advertise it to all routers. Distance- and path-vector protocol (RIP and BGP) may advertise vector information hop by hop and link state protocols (OSPF) may flood the link information.
- Routing re-computation. A router calculates a new route after a failure, and updates its routing table.
- Forwarding information update. Routes are updated to forwarding information base (FIB) and packets will be forwarded based on the updated routes.

Routing convergence time denotes a delay during which routes are calculated to reach a consistent state. During routing convergence, routing black-holes and loops may be formed [7, 8], and thus packets may be forwarded to invalid or wrong routers. Figure 1(a) shows the network topology, in which node and edge denote router and link, respectively. Figure 1(b) denotes the chosen route after routing computation (the sink tree rooted in node 1)[1]. The solid lines indicate the routes before failures and the dash lines denote the routes after failures. When link 1–2 fails, nodes 5 and 6 may still forward packets to node 2 before node 2 adversities routing update to them, and then a routing black-hole is formed. Distance vector and link state protocols both have such black-holes. Moreover, different durations of routing loops may be formed in these two types of protocols. After node 2 detects the failure and finds that node 6 has a route to node 1, the packets will be sent to node 6. Unfortunately, the packet may be sent back to node 2, because node 6 does not finish route update, and then a routing loop is formed. When node 6 updates its forwarding table, the routing loop in link state protocol is eliminated, but another routing loop may exist in vector-based routing protocols. When node 6 receives a route withdrawal, node 6 may choose route as 6-5-2-1, however, node 5 may choose route as 5-6-2-1. Thus, a routing loop is formed between nodes 5 and 6. Finally, packets from nodes 2, 5 and 6 may be forwarded to node 1 through node 7.

The analysis above shows that routing black-holes and loops may be formed during routing convergence, which is caused by internal shortcomings of link state and distance/path vector protocols. Therefore, fast restoration cannot well solve the packet loss problem during routing convergence. From this point of view, routing protection should be a potential solution.

3 Model, evaluation and issues of self-healing routing

3.1 Self-healing routing model

In this section, we propose a concept of routing self-healing ability to evaluate and analyze routing systems.

Definition 1. Self-healing routing refers to the ability of routing systems to rebuild routes or switch failover routes to guarantee normal packet forwarding.

[1] For simplicity, we do not identify difference between routing protocols in this example.

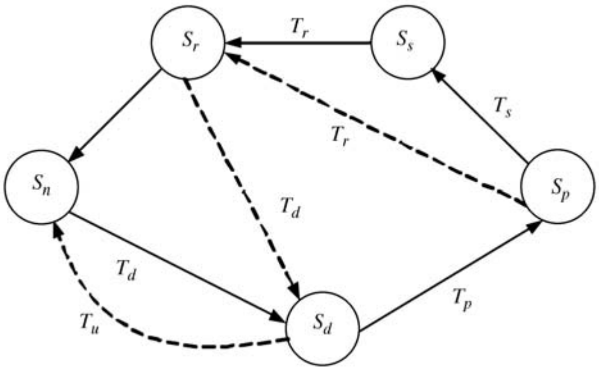

Figure 2 State machine of routing convergence in self-healing routing.

We analyze the self-healing process of routing protocol (including intra- and inter-domain routing). Since most of network change events are single network failures, for simplicity, we only analyze network state using the single failure model in this section. As a standard routing convergence process illustrated in Figure 2, there are five states: S_n denotes a normal state of a network, S_d denotes that the adjacent nodes detect the network failure, S_p denotes that the adjacent nodes finish route re-computation and propagate the update messages, S_s denotes that all the nodes finish route re-computation and the whole network stays in the sub-normal status, and S_r denotes that network failure recovers. After a network failure event, network status will get to state S_s via state S_d and S_p, so $T_d + T_p + T_s$ is the routing convergence time. During routing convergence process, packets sent through the failure node/link may be lost during T_d, and our-of-order packets may appear during T_p and T_s. Moreover, during T_p and T_s, some packets may still be lost because of transient route loops.

Packets will be sent to the failed links and then be dropped during $T_d + T_p$. Subsequently, the instable routes may result in packet loss or out of order packets. Actually, link recovery is also an event of topology change and the state transition is similar to what we discussed above.

3.2 The metrics of self-healing routing

The basic function of network routing systems is to find routes. It is infeasible to use metrics from the data flow view (i.e., reliability, delay, jitter and bandwidth) to evaluate the self-healing routing performance. First, different applications have different requirements. For instance, file transfer requires high reliability and is insensitive to delay, while video-meeting requires low delay and does not require high reliability. Second, these metrics are affected significantly by the traffic distribution and flow control, which is far beyond the current scope of self-healing routing research. We thus propose two metrics, availability and stability, to capture the characteristics of self-healing routing.

When a failure occurs, the chosen route impacted by the failure should be quickly switched to another route. Availability refers to the state of routing systems, which can be expressed as the ratio of mean time between failures (MTBF) to mean time to repair (MTTR). Nevertheless, availability can be described by a distribution curve. Given a network with N nodes, let $Avail(n_i, n_j, t)$ denote the availability between nodes n_i and n_j at time t.

Stability can refer to both route stability and routing protocol stability. Route stability refers to the frequency of path change. During the transition between two paths, there will be a transient timeout period. As the transmission time for a packet to traverse a routing system is in the order of ten milliseconds, the transmission can be viewed as an instantaneous event. If there is no transient timeout period, route stability is not a considerable issue—there is a route for packet transmission at any time. However, if the protocol is poorly implemented, the routing system may be unstable due to the underlying unstable physical components, or the frequent loss of protocol packets caused by the heavy load of traffic and CPU usage. Routing protocol stability is a basic requirement of routing systems. It refers to the resource consumption of a specific routing protocol. In general, there exist transient timeout periods after failures. Intra-domain routing requires tens of seconds to rebuild a route and BGP may require several minutes or longer. If the timeout period does not exist, routes are always available. Thus, route stability can be expressed by routing availability. In this paper, stability refers to control-plane stability of routing protocols. That is, it is used to evaluate the service consumption caused by network failures. For example, when a failure occurs, routing protocol #1 requires 20% CPU cycles to rebuild routes and protocols #2 requires 100% CPU cycles. Once all CPU cycles are occupied by one routing protocol, it is inevitable that packets in data-plane will be dropped and transmission delay will be enlarged. Therefore, we can draw the conclusion that routing protocol #1 has better stability than protocol #2.

Based on the metrics presented above, we can give another definition of self-healing routing.

Definition 2. Self-healing routing refers to the ability of routing systems to improve routing availability while maintaining routing stability.

The routing protocols that have such an ability are called self-healing routing protocols. To some extent, existing routing protocols have the self-healing ability. However, based on the discussion in subsection 2.2, we know that the self-healing ability of these protocols is not good enough to defend different failures. Thus, the goal of routing protocol design or improvement is to improve the self-healing ability. That is, we need to improve the routing availability while maintaining routing stability under failures. Routing availability is to reduce the number of route switch and the route switch delay, and routing stability is to reduce resource consumption of routes during route update. In the following, we will discuss the key problems in intra- and inter-domain routing protocols.

3.3 The self-healing problems of intra-domain routing

Traditional intra-domain routing protocols have some internal problems in the four convergence steps, which directly cause slow routing convergence even divergence.

- Long failure detection delay. Failure detection is restricted by timers which are always longer, so that routing system cannot always detect failures in time [12].
- Long time failure propagation. Failure advertisement requires flooding failure information to all routers within an AS. Distance vector protocol will advertise routes hop by hop, which suffers a long time. Especially, it may have the count-to-infinity problem. Link state protocol flood link state information, but it is constricted by different timers. However, we cannot simply shorten timer values because a short timer may cause routing oscillation.
- Long time route recomputation. Similarly, route recomputation is restricted by timers. Routing recomputation in different routers will not finish concurrently, which may lead to routing loops during convergence process.

3.4 The self-healing problems of inter-domain routing

The stability of intra-domain routing directly impacts that of inter-domain routing [13]. Besides similar self-healing problems in intra-domain routing, inter-domain routing protocol, BGP, has typical internal problems, which cause slow convergence or divergence.

- False failure detection. BGP routers exchange their routing information through Keepalive message over reliable TCP connections. However, the broken TCP sessions cause unnecessary routing re-computation [8]. This routing re-computation may cause global BGP routing change, and thus results in network instability.
- Long time routing computation. As a path vector routing protocol, BGP routers may undergo long-term path exploration to achieve best routing paths. This process is the root cause of BGP slow convergence. During the route selection process, a router may continuously choose many unavailable routes as its best route before identifying the real best route.
- Long time route propagation. Some improved mechanisms developed for BGP influence fast propagation of updated routes. For instance, MinRouteAdvertisementInterval (MRAI) determines the minimum amount of time that must elapse between advertisements of routes to a particular destination from a single BGP speaker. However, MRAI may significantly increase the convergence delay. Route flap damping (RFD) [14] is a widely deployed approach to limit route oscillation and improve routing stability. However, previous studies show that RFD will falsely suppress routes, and a valid route may be suppressed

for more than ten minutes [15]. Furthermore, when there is no interaction between different routers on route suppression, a route may be suppressed several times [15].

• Routing divergence. The flexibility and privacy issue of BGP policies may cause routing divergence. Since different ASes are controlled by different organizations, they need to configure routing policies based on their own interests and these policies are always not publicly available for their peers, which may cause policy conflicts and further result in routing divergence.

4 Analysis of self-healing routing schemes

To solve the routing problems, several improved solutions are carried out. These solutions can be divided into fast routing restoration and routing protection. Fast routing restoration refers to route recomputation and rebuilding after failures. This approach could be developed by changing parameters of routing protocols, improving routing stability, providing multipath and improving effectiveness of route computation. Routing protection approaches select failover routes in advance. When failures occur, route can be rebuilt by switching to failover routes. In addition, to ensure the effectiveness of fast routing restoration, a fast routing failure detection scheme is essential.

4.1 Routing restoration schemes

Existing fast routing restoration schemes include adjusting routing protocol parameters and weights, improving routing stability, providing multipath and other improvements. We will discuss these improved schemes applied in intra- and inter-domain routing protocols. The adjustment of routing parameters concentrates to accelerate the transmission of routing updates, that is, to shorten T_p in Figure 2. Improvement of routing stability and provision of multi-path focus on shortening the computation time T_s in Figure 2. The fast routing restoration scheme cannot eliminate the routing black holes and loops during convergence process.

4.1.1 *Parameter and weight change*

The change of intra-domain protocol parameters can be realized in the following aspects: 1) Failure detection: Hello exchange period can be shortened to sub-second to speed up failure detection. However, routing oscillation may be raised, e.g., links can sometimes recover after a short-term failure. The solution could send link failure news quickly and link recovery news slowly. 2) Failure propagation: Failure propagation can be speeded up by achieving higher priority than shortest path computation. 3) Shortest path computation: traditional Dijkstra algorithm can be improved by adopting the incremental computation algorithm in which only part of SPT is performed.

Moreover, the self-healing performance can be improved by changing link weights. Link weight is the critical factor of intra-domain routing. In general, the routing algorithm will choose a route who has the minimal link weights as the best one. The weight of intra-domain routing protocols (i.e., OSPF) can be set randomly by administrators in a range. Thus, weight change could be a potential solution to improve routing self-healing ability. One typical solution is proposed by Fortz et al. in [16]. They change a small portion of link weight to achieve load balance and avoid congestion. Generally, link weight is always determined by link capacity and load. They conducted some simulations according to American Nationwide Backbone Network and the results show that flow distribution can reach 90% of the optimal value only by changing the weight of three links after single link failure occurs.

Self-healing ability in inter-domain routing can be improved by adjusting the routing policies, e.g., change of routing update policy. Afek et al. [17] proposed a Ghost Flushing scheme to modify the timers of BGP protocol. In this scheme, different types of update messages are set with different timers. That is, routing withdrawals will be sent without delay and routing update will be sent with a delay. Besides, in this scheme, existing routes for a specific prefix will be withdrawn before advertising a new route of the prefix. Thus, fast convergence of inter-domain routing will be achieved by fast propagation of bad news and delay of good news.

The advantages of protocol parameters and weight change are that they can be easily deployed without changing protocol and infrastructures and deal with multiple failures. The disadvantages are as follows: there is no research result of convergence time under multiple failures and CPU consumption with short protocol timers. In addition, since link failures may frequently occur, network bandwidth will be wasted with these schemes. Thus, these schemes improve routing availability while decreasing routing stability.

4.1.2 *Stability improvement*

Failure-carrying packet (FCP) eliminates update flooding process to decrease the failure impact to networks [18]. The basic idea is as follows: When a node detects failures of links or nodes adjacent to it, it will cache the packets to these failed links/nodes, and perform SPT computation. Then, packets will be forwarded through the next hop of the new route, and carry all failure information in packets. In this way, the failure information is transmitted to other nodes with these packets, and thus usual failure flooding is eliminated. When a node receives packets with failure information, it performs the SPT computation. FCP makes real-time computation, and a routing path to destination can be found if it exists. Thus, FCP largely decreases packet loss.

Similarly, route flap damping (RFD) mechanism realized in inter-domain routing protocol can improve the routing stability and thus the self-healing ability of inter-domain routes. RFD tries to distinguish continuing unstable routes from routes with occasional failures, and suppresses unable routes to speed up routing convergence. In this mechanism, every node is set at a certain threshold. When penalty value of a route exceeds the threshold, the route will be punished and will not be used in a certain period of time except that the penalty value falls below the threshold. Unnecessary routing computation will be reduced by applying RFD mechanism to damp unstable routes.

Improvement of routing stability shortens the update process of route change[2)]. However, under certain circumstance, it will deteriorate the network performance. For instance, FCP will cause a long delay of packet forwarding, and RFD will exacerbate the BGP convergence by enforcing improper route damping. These schemes improve routing stability but decrease routing availability.

4.1.3 *Multipath*

Multipath schemes refer to provision of multiple paths or next hop addresses[3)]. Equal cost path in intra-domain routing and traffic engineering in inter-domain routing are both a kind of multipath. Generally, multipath schemes in intra-domain routing use different paths or next hops, e.g., Outdegree2 (O2) [19], to forward packets. The most common approach is to apply common link state protocols in general cases and use backup path or next hop when failure occurs, and routing convergence process starts at the same time. When the convergence finishes, routing systems switch to the usual link state protocol. The purpose of O2 [19] is to solve the traffic burst problem raised by failures. The basic idea is to let every node try to forward packets to two different next hops at the same time. Thus, traffic of each node will be shared by multiple paths if no failure occurs. When a single failure occurs, packets will not be lost but continually forwarded because there is another available next hop. Since traffic over failed link is switched into multiple paths, traffic burst will be eased. When failure occurs, each node re-computes the O2 routes, and then restores to the normal state under which packets are forwarded with two different next hops. In the network topology where nodes have high connection rate, most nodes usually can find two outdegrees. O2 algorithm forms multiple connection paths, and thus individual failure will influence the normal data transmission. Thus, it plays a key role in load balance for network traffic by traffic distribution to different links.

Gao et al. [20] proposed an inter-domain routing backup scheme, which improves the routing reliability without sacrificing routing stability. The premise of this scheme is that local AS routing policy should be consistent with business relationship of neighbors. Convergence is guaranteed under any node or link

[2)] Some routing protection schemes discussed in subsection 4.2 will improve routing stability, e.g., FIR (failure insensitive routing) [31] will hold the new route before advertising it.

[3)] The protection schemes presented in subsection 4.2 are special cases of multipath. Since there is no routing convergence involved in routing protection, we do not discuss routing protection schemes here.

failure by guidelines of BGP multipath. Moreover, Xu et al. [21] proposed a multipath inter-domain routing scheme (MIRO). In this scheme, dynamic backup path negotiation is introduced and then the default routing paths are still learned by default BGP protocols. Once backup path negotiation finishes, packets can be forwarded in backup path via tunnels.

Multipath is an effective approach to solve self-healing routing problems. However, some key issues concerning multipath are not well solved yet, such as routing loops and multiple failures. In addition, in BGP multipath schemes, multipath scheme can effectively improve the convergence performance [22] only under certain circumstances. However, when multiple failures occur, these schemes may delay the process of best route selection. Thus, these schemes neither improve the routing stability, nor effectively improve routing availability.

4.1.4 *Other schemes*

BGP will experience slow path exploration during routing computation, and the process is a critical factor that results in BGP slow convergence. This is also an important feature that distinguishes intra-domain from inter-domain routing protocols. This section will analyze the routing improvement schemes to address the BGP route selection issue. At present, there are three typical schemes: root cause in updates to identify invalid routes, e.g., RCN scheme [23]; examination of update effectiveness by comparing received updates, e.g., Consistency Assertion Scheme [24]; new inter-domain routing design, e.g., HLP [25].

(1) Pei et al. [23] proposed a root cause notification (RCN) algorithm to improve route selection efficiency. It tries to carry a new routing attribute in BGP update to indicate the root cause of routing update (failure cause of the BGP update message). These attributes include links which cause the route change, the current state and serial numbers of links. In traditional BGP protocols, when route changes, e.g., link failure, the BGP router which detects failures first puts the place and root cause of changed route to Update message. This can avoid or mitigate routing blackhole issue during routing convergence process. That is, some routers may use invalid path to send packets, which causes packet loss before all the routers reach the same route. Although RCN can effectively improve the convergence ability of routes, it has some limitations: firstly, the address and serial number introduces the communication overhead in the BGP convergence process; secondly, the BGP convergence performance in RCN scheme will depend on different connections between BGP speakers and network topologies, and BGP convergence in site networks may achieve significant improvements; lastly, RCN cannot support incremental deployment.

(2) Consistency assertion reduces path exploration to improve current BGP convergence performance. In consistency assertion, routers will compare received BGP update messages to distinguish invalid routing information. For example, when router S receives two routing messages to router D, paths are $S : I_1, x, D$ and $S : N_2, y, N_1, x, D$ respectively. If the first route to destination router D is withdrawn, then the second route will not be adopted, because the second path is unavailable. BGP convergence time and routing lookup times decreases greatly by recognizing and deleting unavailable path. But realizing invalid path lookup algorithm is difficult, and the algorithm will introduce heavy computation costs.

(3) Subramanian et al. [25] proposed a new inter-domain routing protocol called hybrid link-state and path-vector protocol (HLP). HLP uses infrastructure formed by provider-customer relations among ASes to restrain routing visibility among different layers so as to solve the slow routing convergence, global transmission of routing change and other issues. Specifically, existing BGP protocols take IP addresses of current network as routing granularity. But the granularity of HLP routing becomes an AS so as to reduce the routing amount in the inter-domain of the Internet. The HLP information hiding separates the local routing change, greatly improving the protocol scalability and routing convergence. Based on the AS topology, routing communication cost is reduced by 400 times compared with ordinary BGP. By limiting route lookup process, the routing convergence delay is linear. Although HLP outperforms existing BGP protocols, it still has some shortcomings. For example, routing churn in HLP may be still heavy, which causes poor routing availability.

In addition, intra- and inter-domain routing protocols both adopt scalability technology, e.g., area in OSPF protocol, route reflector [26] in BGP protocol. They improve routing convergence by realizing

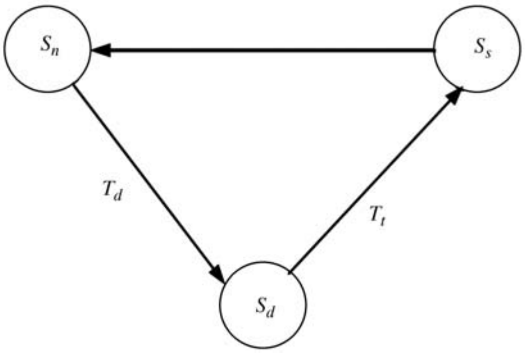

Figure 3 State machine of self-healing routing with routing protection.

scalability of routers. However, these mechanisms may bring side effect. For example, Route Reflector in inter-domain protocol may introduce the routing oscillation [27]. These schemes enhance routing availability to some extent, however, they cannot ensure or even reduce routing stability.

4.2 Routing protection schemes

Routing protection schemes refer to fast backup route switch in nodes adjacent to failures, thus these schemes localize failures. Self-healing model of routing protection is shown in Figure 3. Once neighbor detects a failure, it activates protection paths and network enters in S_s. This process finishes in T_t. After network failure recovers, routing node can directly return to S_n. In these schemes, routing black holes only exist within T_d, and routing self-healing time is T_d+T_t. Based on this, we know that route protection is an effective self-healing solution. The main existing routing protection schemes include general IP fast reroute scheme (IP-FRR) [28–33] and BGP fast reroute (BGP-FRR) [34–36].

4.2.1 *IF fast reroute(IF-FRR)*

Shand et al. [28, 29] proposed IP based reroute schemes, in which failure information will not be released when failure is detected. Failover paths are used to realize local failure recovery. With these schemes, routing interruption time is rather small, because only failure detection and backup route activation are involved. These schemes are different from MPLS based IP reroute. They also differ from traditional multipath schemes because no routing convergence process is involved and failure information is blocked. The rerouting performance depends on path switch. Therefore, it can be only used to recover from short-term failures.

IP based local reroute has many solutions [29–33]. Bryant et al. [30] assign a special address for every protected node, called not-via address. A not-via address not only indicates failure node P, but also indicates the destination node or a middle node B on the path. When failure occurs, adjacent nodes of failure will forward packets using a non-failure interface with a not-via address. Nodes receiving these packets will know that a node failure occurs in P, and then use backup route to forward the packet to node B. In addition, Nelakuditi et al. [31] use specific interface forwarding to realize data packet forwarding of next-hop. Bryant et al. [32] proposed a packet forwarding algorithm using tunnel under failures. Atlas et al. [33] proposed an algorithm to compute and choose non-loop backup tunnel.

4.2.2 *BGP fast reroute (BGP-FRR)*

Bonaventure et al. [34] proposed BGP fast reroute scheme. They pre-compute tunnels to provider networks for different types of stub network to achieve FRR. In this scheme, pre-configured protection is activated after failure detection with bidirectional forwarding detection (BFD) [14] to eliminate unnecessary path exploration. BGP-FRR dynamically adjusts local preference to avoid the route loop issues which may be raised by tunneling. Moreover, this scheme proposes a FIB fast update mechanism and different tunnel protection solutions according to different tunnel properties.

Table 1 Comparison of different self-healing routing schemes[a)]

	Self-healing routing schemes	Availability	Stability
Routing restoration	Parameter and weight change	increase	decrease
	Stability improvement	decrease	increase
	Multipath	maintenance	maintenance
	Other schemes	increase	decrease
Routing protection	IP-FRR	increase*	maintenance*
	BGP-FRR	increase*	maintenance*

a) * denotes that it can be achieved under specific conditions.

Although BGP-FRR scheme eliminates traditional path exploration in multipath schemes, it does not consider the multiple failure under which BGP-FRR scheme will lose effectiveness. BGP-FRR is only applicable for short-term routing failure. Long-term protection will introduce extra overhead to BGP routers, especially when a large number of data packets need to be segmented. So, only provision of BGP-FRR cannot effectively improve the self-healing performance. Thus, this scheme can only enhance routing availability and stability under certain conditions.

4.3 Comparison and analysis

Generally, existing fast routing restoration schemes cannot eliminate the constraints between routing availability and stability. These schemes improve routing availability at the cost of reducing routing stability, and vice versa. Similarly, routing protection scheme will have a better performance in most cases of short-term single link failure. At the same time, routing protection schemes effectively eliminate or mitigate the routing black holes and loops caused by network failures. However, when multiple link failures occur or protection fails, stability of routing protection scheme may decline. Routing availability and stability of different self-healing schemes are shown in Table 1. Since these schemes do not consider or solve the mutual constraint between routing availability and stability, simply solving the availability or stability cannot effectively improve the performance of self-healing routing. In addition, some routing protection schemes do not solve the routing stability issue under multiple link failures. Hence, these schemes cannot effectively improve the self-healing ability of routing protocols.

5 Design of self-healing routing scheme

Based on the analysis of self-healing routing algorithms in section 4, we know that existing schemes cannot fully satisfy the requirements of self-healing routing. That is, we should improve routing availability under the condition that routing stability is improved or maintained. We will analyze several key issues in self-healing routing, and then propose a scheme of self-healing routing.

5.1 Key issues in self-healing routing

- Mutual constraints between routing availability and stability. Availability and stability are two primary issues to address self-healing routing. Most schemes analyzed in subsections 4.1 and 4.2 only consider partial issues of routing availability and stability. On the one hand, failover paths are involved to guarantee packets forwarding under failures. In previous studies, failure coverage rate of failover path is an important metric to measure intra-domain self-healing routing schemes. However, it is hard for a single specific scheme to provide high availability and the cost to combine different algorithms is too high. Thus, an self-healing routing scheme with high availability is necessary. Existing studies mostly consider failure handling in routers adjacent to failures, and may not achieve effectiveness under large-scale failures. On the other hand, oscillation caused by frequent failure events should be avoided or alleviated. Convergence performance and stability are mutually constrained. Different failures have different features. If root cause of different failures can be identified, e.g., pre-configured recovery mechanisms can be proposed for routing failure caused by controllable router maintenance [11, 12], routing availability can be enhanced with guaranteed routing stability so as to improve routing self-healing ability.

• Adaptability to multiple failures. At present, there are several intra-domain self-healing routing schemes to recover from single link failures. However, failures in real networks are random and burst, and a decent model is required to describe routing under multiple failures and thus an improved self-healing routing solution is required to deal with multiple failures. Existing solutions could be further improved to adapt to multiple failures, e.g., routing protection schemes may fail under multiple failures. In addition, we need to further investigate parameter selection in various improved protocols under multiple failures. If combined with routing stability solutions, fast reroute could be a potential self-healing routing solution to recover from multiple failures.

• Other key issues. Simplicity and practicality are the principle of Internet and routing protocol design. Simple idea, easy implementation, and compatible deployment with existing protocols are requirements of an effective solution to routing self-healing problem. Most existing self-healing solutions cannot be deployed due to complexity [17]. From the point of view, self-healing scheme with routing protection could be a satisfactory scheme. Besides, most existing studies concentrate on routing convergence, rarely considering the routing availability issue which directly impact the packet forwarding performance. The development trend in network applications concentrates on user perspective of packet forwarding. Especially, current multi-media applications, such as VoIP, present strict demands of packet forwarding performance to routing systems, e.g., short delay and low packet loss, and thus provide user oriented self-healing routing. Thus, we should not only focus on routing issue, but on packet forwarding performance.

5.2 Improved self-healing routing scheme

According to the analysis of self-healing routing schemes in subsection 3.2, we know that self-healing routing should improve routing availability under the premise that stability is maintained or improved. To achieve this goal, the following key problems should be solved: 1) fast failure detection; 2) short black hole time; 3) elimination of routing loops during routing convergence process; and 4) improvement of routing stability. Fast routing failure detection can be achieved using BFD mechanism which is independent of any routing protocols. So, we will not discuss it in this section. In this section, we will analyze how to solve problems 2)–4) in self-healing routing. To solve problem 3), we propose a fast routing restoration scheme under routing protection, that is, a hybrid self-healing routing scheme by integrating the advantages of fast routing restoration schemes and routing protection schemes.

Figure 4 shows our proposed self-healing routing scheme. Different from the state machine shown in Figure 2, a protection state $S_{p'}$ is introduced in our scheme. The state indicates that adjacent node has detected routing failures and activated routing protection. When adjacent nodes detect routing failures, routing system can ensure the normal packet forwarding by protecting failed path after $T_d + T_{p'}$, and the protection activation time is much smaller than traditional routing convergence time $T_d + T_p + T_s$. After failure recovers, protection withdraws and routing system will return to the original stable state after T_u. Similarly, it also eliminates convergence process of traditional routing protocols, and the delay of convergence process is $T_d + T_p + T_s$. If long-term and unexpected failure occurs, network cannot always stay at S_p. This is because protection paths pre-computed are not the best ones. Remaining in this state may increase extra burden of routers and result in packet forwarding delay and network congestion. At this time, routers should apply fast routing restoration. In this process, because of the existence of failover routes, routing black holes and loops will not occur (we will discuss how routing restoration scheme with routing protection can avoid routing loop in a later section). After $T_p + T_s$, network will reach another stable state. In addition, our scheme considers fast routing restoration when routing protection fails. In this scenario, self-healing model is similar to Figure 2. In the analysis of our scheme, we will discuss how to improve self-healing routing in detail.

Figure 5 shows the key algorithms in the self-healing routing scheme. Routing protection can effectively reduce routing black hole time (Problem 2a) caused by simple failures. For instance, Figure 1 shows when the failure of link 1-2 occurs, protection path between nodes 2 and 7 is used after node 2 detects failure. Then, all packets sent to node 2 will be forwarded to node 1 through the failover route. At the same time, routing loop caused by simple network failure is eliminated with the failover route (Problem 3a).

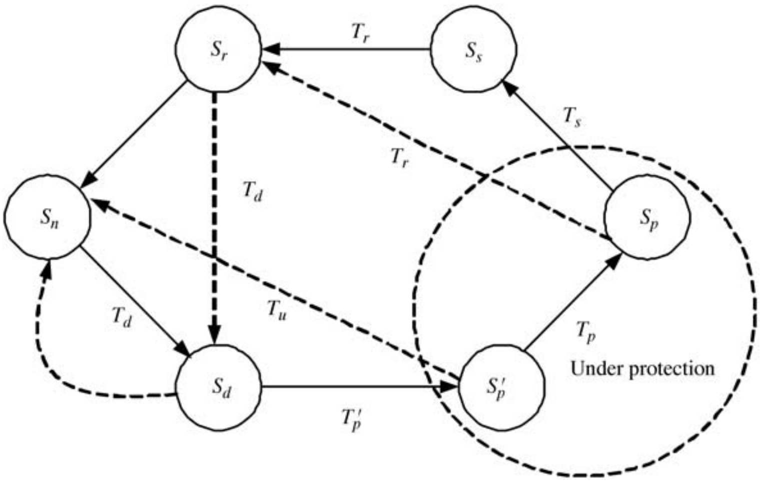

Figure 4 State machine of hybrid protection and restoration based scheme.

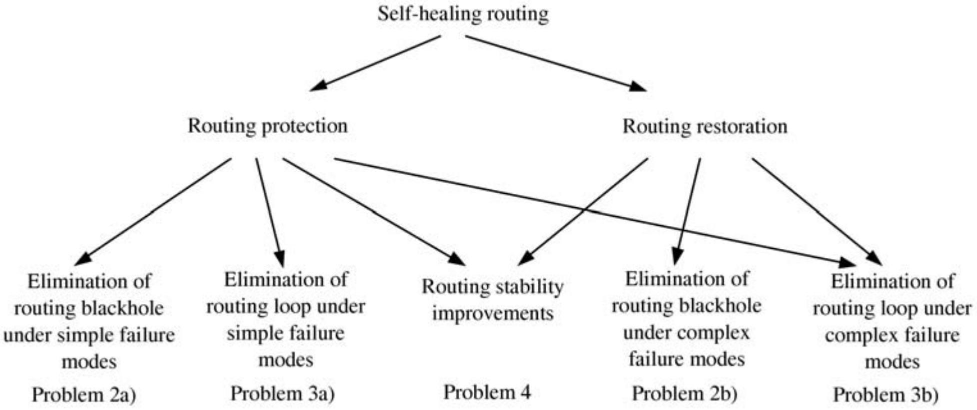
Self-healing routing
Routing protection
Routing restoration
Elimination of routing blackhole under simple failure modes
Problem 2a)
Elimination of routing loop under simple failure modes
Problem 3a)
Routing stability improvements
Problem 4
Elimination of routing blackhole under complex failure modes
Problem 2b)
Elimination of routing loop under complex failure modes
Problem 3b)

Figure 5 Improved self-healing routing scheme.

Similarly, by using routing protection, node eliminates many unnecessary routing computation, especially when short-term failure occurs. Therefore, routing protection can effectively improve routing stability (Problem 4).

As for short-term network failure, routing protection can effectively improve the self-healing ability of routing. But if network failure is a long-term and unexpected one, pure routing protection is not an effective self-healing solution because all routes are pre-computed and not the best ones. Routing protection for long-term failures may raise other packet forwarding problems, such as network forwarding delay and network congestion. Thus, fast routing restoration is a must in self-healing routing. We propose fast routing restoration under routing protection, i.e., routing systems still require route reselection. In such a context, we should consider possible routing loops introduced during routing convergence due to routing protection. As Figure 1 shows, different routing protocols adopt different routing computation algorithms, and routing loops formed also differ. Ordered FIB update can be applied in link state protocol to eliminate routing loop. As for routing vector protocol, attributes in routing update can be used to avoid routing loop [35].

However, when complex network failures occur, e.g., multiple failures, these failures may be not covered by a protection solution (Since protection paths are pre-computed, in general, the scheme cannot guarantee that protection paths can cover 100% failures and protect all links and nodes), and then routing protection will fail. In this context, simple fast routing restoration scheme can effectively improve self-healing ability of routing. Note that we should guarantee routing stability during convergence. Most existing schemes expedite route convergence but not consider routing stability during convergence. For example, we should filter or not choose some unstable links and routes in route selection so as to largely reduce routing protocol computation and probability of selecting unstable or invalid routing. To some extent, improvement of routing stability reduces the number of route switch and improve routing availability under failures as well.

6 Conclusion and future work

The widespread use of Internet applications imposes strict requirements on routing performance. However, existing routing protocols cannot well meet these requirements. Especially, routing protocols are unable to well adapt network failures and serve for different Internet applications under failures. We firstly review characteristics of network failures and then propose the concept and model of self-healing routing. Based on this, we analyze problems in different routing protocols, and we discuss self-healing problems of these protocols. Furthermore, we survey existing improvements and present qualitative analysis of advantages and disadvantages. In future work, we will quantify the performance of existing improved self-healing routing schemes.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 61073166), the National Basic Research Program of China (Grant No. 2009CB320502), the National High-Tech Research & Development Program of China (Grant Nos. 2007AA01Z2A2, 2009AA01Z205), and the Key Projects of National Science & Technology Pillar Program (Grant No. 2008BAH37B03).

References

1 Malkin G. RIP Version 2. RFC 1732, 1994
2 Moy J. OSPF Version 2. RFC 2328, 1998
3 Oran D. OSI IS-IS Intra-domain Routing Protocol. RFC 1142, 1990
4 Rekhter Y, Li T. A Border Gateway Protocol 4 (BGP-4). RFC 1771, 1995
5 Floyd S, Jacobson V. The synchronization of periodic routing messages. In: ACM SIGCOMM. New York: ACM, 1993. 33–44
6 Labovitz C, Ahuja A, Bose A, et al. Delayed Internet routing convergence. IEEE/ACM Trans Netw, 2001, 9: 293–306
7 Sridharan A, Moon S B, Diot C. On the correlation between route dynamics and routing loops. In: ACM IMW. New York: ACM, 2003. 285–294
8 Wang F, Gao L, Wang J, et al. On understanding of transient interdomain routing failures. In: ICNP. Washington: IEEE, 2005. 30–39
9 Markopoulou A, Iannaccone G, Bhattacharyya S, et al. Characterization of failures in an IP backbone. In: INFOCOM. New York: IEEE, 2004 2307–2317
10 Iannaccone G, Chuah C, Mortier R, et al. Analysis of link failures in an IP backbone. In: ACM IMW. New York: ACM, 2002. 237–242
11 Wang L, Saranu M, Gottlieb J, et al. Understanding BGP session failures in a large ISP. In: INFOCOM. New York: IEEE, 2007. 348–356
12 Katz D, Ward D. Bidirectional forwarding detection. In: RFC 5880 Internet Engineering Task Force, 2010
13 Teixeira R, Shaikh A, Griffin T, et al. Dynamics of hot-potato routing in IP networks. In: SIGMETRICS. New York: ACM, 2003. 307–319
14 Villamizar C., Chandra R., Govindan. R., BGP Route Flap Dampening. RFC 2439. 1998.
15 Mao Z M, Govindan R, Varghese G, et al. Route flap damping exacerbates Internet routing convergence. In: ACM SIGCOMM. New York: ACM, 2002. 221–233
16 Fortz B, Thorup M. Optimizing OSPF/IS-IS weights in a changing world. IEEE J Select Areas Commun, 2002, 20: 756–767
17 Afek Y, Bremler-Barr A, Schwarz S. Improved BGP convergence via ghost flushing. IEEE J Select Areas Commun, 2004, 22: 1933–1948
18 Lakshminarayanan K, Caesar M, Rangan M, et al. Achieving convergence-free routing using failure-carrying packets. In: ACM SIGCOMM. New York: ACM, 2007. 241–252
19 Schollmeiers G, Charzinski J, Kirstadter A, et al. Improving the resilience in IP networks. In: HPSR 2003. New York:

IEEE, 2003. 91–96

20 Gao L, Griffin T G, Rexford J. Inherently safe backup routing with BGP. In: INFOCOM 2001. New York: IEEE, 2001. 547–556

21 Xu W, Rexford J. MIRO: multi-path interdomain routing. In: SIGCOMM 2006. New York: ACM, 2006. 171–182

22 Feamster N, Andersen D, Balakrishnan H, et al. Measuring the effects of Internet path faults on reactive routing. In: SIGMETRICS. New York: ACM, 2003. 126–137

23 Pei D, Azuma M, Massey D, et al. BGP-RCN: improving BGP convergence through root cause notification. Comput Netw, 2005, 48: 175–194

24 Pei D, Zhao X L, Wang L, et al. Improving BGP convergence through consistency assertions. In: INFOCOM 2002, New York, 2002. 902–911

25 Subramanian L, Caesar M, Ee C T, et al. HLP: a next generation interdomain routing protocol. In: SIGCOMM. New York: ACM, 2005. 13–24

26 Bates T, Chen E, Chandra R. BGP route reflection: an alternative to full mesh internal BGP (iBGP). RFC 4456. 2006

27 Basu A, Ong C L, Rasala A, et al. Router oscillations in I-BGP with route reflection. In: SIGCOMM. New York: ACM, 2002. 235–247

28 Shand M, Bryant S. IP fast reroute framework. In: Internet Draft, draft-ietf-rtgwg-ipfrr-framework-08.txt, Internet Engineering Task Force, 2008

29 Shand M, Bryant S. A framework for loop-free convergence. In: Internet Draft, draft-ietf-rtgwg-lf-conv-frmwk-02. txt, Internet Engineering Task Force, 2008

30 Bryant S, Shand M, Previdi S. IP fast reroute using Notvia addresses. In: Internet Draft, draft-ietf-rtgwg-ipfrr-notvia-addresses-02. txt, Internet Engineering Task Force, 2008

31 Nelakuditi S, Lee S, Yu Y, et al. Fast local rerouting for handling transient link failures. IEEE/ACM Trans Netw, 2007, 15: 359–372

32 Bryant S, Filsfils C, Previdi S, et al. IP fast reroute using tunnels. In: Internet Draft, draft-bryant-ipfrr-tunnels-03. txt, Internet Engineering Task Force, 2007

33 Atlas A, Zinin A. Basic specification for IP fast-reroute: loop-free alternates. In: Internet Draft, draft-ietf-rtgwg-ipfrr-spec-base-12. txt, Internet Engineering Task Force, 2008

34 Bonaventure O, Filsfils C, Francois P. Achieving Sub50 milliseconds recovery upon BGP peering link failures. IEEE/ACM Trans Netw, 2007, 15: 1123–1135

35 Li Q, Xu M, Pan L, et al. A study of path protection in self-healing routing. In: IFIP Networking. Berlin: Springer-Verlag, 2008. 554–561

36 Li Q, Xu M, Wu J, et al. Achieving unified protection for IP routing. In: IEEE ICCCN 2010, New York, 2010. 1–6