Computer Networks 52 (2008) 3325–3341

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

On basic properties of fault-tolerant multi-topology routing

Tarik Čičić *,[1]

University of Oslo and Simula Research Laboratory, Oslo, Norway

ARTICLE INFO

Article history:
Received 22 December 2007
Received in revised form 22 April 2008
Accepted 27 August 2008
Available online 11 September 2008

Responsible Editor: G. Ventre

Keywords:
Multi-topology routing
Fault tolerance
Network protection
IP fast reroute
Separating sets
Heuristic algorithms
Routing state
Load distribution

ABSTRACT

Multi-topology routing has recently gained popularity as a simple yet efficient traffic engineering concept. Its basic purpose is to separate different classes of network traffic, which are then transported over disjoint logical topologies. Multi-topology routing is used as a basis for implementation of an IP fast reroute scheme called Multiple Routing Configurations (MRC).

MRC has a range of attractive properties, but they do come at a cost. In order to guarantee recovery from any single link or node failure in the network, MRC has to maintain several logical topologies and thus an increased amount of routing information. The number of the logical topologies in MRC need not be large; even simple heuristic algorithms often yield good results in practice. However, why this is the case is not fully understood yet.

In this paper, we introduce a theoretical framework for fault-tolerant multi-topology routing (FT-MTR). MRC is a practical implementation of FT-MTR in connectionless IP networks. We use FT-MTR to study how the internal topological structure of the communication network relates to two important problems. The first problem is minimizing the number of logical topologies and thus the routing state in FT-MTR. We show how to use the sets of nodes that separate the topology graph to devise an advanced heuristic for "intelligent" construction of the logical topologies. Finding the separating sets in a topology graph is computationally demanding; we present an algorithm that performs well in tested real network topologies. We evaluate the separation-set based heuristic for the logical topology construction and show that it outperforms the known MRC heuristics.

The second problem is the FT-MTR load distribution after a failure. We use the separating sets to devise a novel algorithm for failure load distribution. This algorithm does not require knowledge of the traffic demand matrix, still, our tests indicate that it performs as good as, or better than, known MRC load-distribution algorithms that do require the demand matrix as input.

1. Introduction

From the early days of the Internet, the ability to tolerate loss of network components has been one of the key goals in its design [1]. Internet routers include mechanisms that detect connectivity failures and topological changes, and convey this information to their routing protocols. The protocols distribute the change information network-wide, and the network gradually adopts the new routing paths and converges to the new stable routing state. Even within a single administrative network domain, this convergence process takes time to complete. All routers in the domain independently calculate a new valid routing table upon receiving the change notification. This process is not synchronized, and temporary instabilities in the form of packet loops and unreachable destinations can occur [2,3].

* Tel.: +47 93050249.
E-mail address: tarikc@ifi.uio.no
[1] The author is affiliated with the Department of Informatics, University of Oslo, P.O. Box 1080 Blindern, 0316 Oslo, Norway. Most of this work has been done during his engagement at the Simula Research Laboratory, P.O. Box 134, 1325 Lysaker, Norway.

1389-1286/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2008.08.021

Careful tuning of the routing mechanism can reduce the time scale of this reconvergence process to sub-second intervals [4]. Attempts to further improve the reconvergence process augment the routing instabilities to an unacceptable level. This reconvergence speed is not acceptable for time-critical and interactive Internet applications with stringent demands on network availability, like IP telephony. Such applications demand recovery times in the 50 ms range, which is traditionally achieved by Layer-2 protection mechanisms.

Fast recovery at IP level is desirable despite the existing options for Layer-2 protection. The Layer-2 protection mechanisms add complexity to the communication system, and often demand additional network resources. Additionally, logical IP failures cannot be detected below the IP layer.

A number of mechanisms for faster failure handling have been proposed for IP networks. Multi-Protocol Label Switching (MPLS) technology opened the possibility to use tunnels that avoid routing through the failed component [5]. MPLS is a versatile technology that has helped IP networks to adopt many protection schemes previously known from Layer-2 technologies [6]. Many networks, however, do not have MPLS mechanisms deployed. Recently, work on pure IP-layer fast reroute (IP Fast Reroute, IP FRR) has gained momentum [7].

1.1. IP fast reroute

Traditional IP reconvergence is slow because it is global and reactive. IP fast reroute mechanisms prepare backup IP routes *proactively*, i.e., before the failure occurs. This way the backup route can be used *immediately* after the failure is detected. IP fast reroute provides mechanisms that let the backup paths be selected *locally* by the node that detects the failure, further saving precious time.

IP FRR mechanisms have two very attractive properties. First, they respond quickly to a failure and prevent packet loss by allowing packet forwarding to continue on alternate routes while the routing protocol converges on the new topology. Second, they allow routers to delay the sending of a failure notification for a period of time while relying on the available repair path. This way, short-lived failures can be handled without triggering a global reconvergence. A large percentage of experienced network failures are short-lived [8], and handling such failures locally can improve network stability.

IP fast reroute should provide full protection against all single link and node failures in the network. The IETF IP FRR framework [7] distinguishes between different recovery schemes for use in IP networks. The simplest scheme is the Fast failure protection using Loop-Free Alternates (LFA, [9]). In case of failure, LFA redirects traffic to neighboring nodes which have a path to the destination that does not include the failed component. For example, if the node that detected the failure has two or more equal-cost paths to a destination, any of these paths can be used for packet forwarding to that destination without risking packets looping back to the detecting node. Other, less restrictive loop-free alternate routes are defined in the LFA document and can be calculated from the routing information base.

However, a loop-free alternate route does not exist in all failure cases: although the detecting node has multiple network neighbors, it is possible that all of them use the detecting node to reach the destination [10]. A more complex scheme is needed to provide 100% coverage from link and node failures, either as a complement to LFA or standalone. Several such schemes are proposed. "Not-via addresses" [11] uses IP tunneling to forward the packet to the "far side"of the failed neighbor, similarly to the MPLS fast reroute. This way, the packet gets around the failed component and continues its trip to the destination following its default route. Another IP FRR scheme that provides full coverage for all link and node failures is called Failure Insensitive Routing (FIR) [12]. Under stable conditions, every network node expects packets addressed to a given destination to reach the node from a specific subset of its interfaces. FIR provides fast recovery by inferring component failures from unusual link of arrival for the affected packets, and then forwarding the packets using proactively prepared "backwarding" tables.

"Multiple Routing Configurations" (MRC, [13]) is another well-known IP FRR scheme and represents the subject of this paper. MRC provides backup paths for all single link and node failures using the multi-topology routing.

1.2. Multi-topology routing

Multi-topology (MT) routing is a powerful traffic engineering concept based on introducing multiple logical topologies in the network. Each logical topology is intended to route a particular class of the network traffic. IP packets are classified and associated with their logical topology by analyzing their headers. For example, multicast or high-priority DiffServ traffic could be defined as a separate class of traffic and routed on a separate logical topology (Fig. 1). The separate logical topologies can be implemented by, e.g., providing multiple routing tables based on different link weights.

The IP community has recently shown a strong interest in this concept, and the standardization process is just completed [14,15]. For us, the MT routing is primarily interesting in the context of fault tolerance. The basic idea of fault-tolerant multi-topology routing (FT-MTR) is to construct the logical topologies so that certain components in the network are not used for packet forwarding, and to tag the recovered packet headers so that they can be identified with their logical topology. This is exactly what the

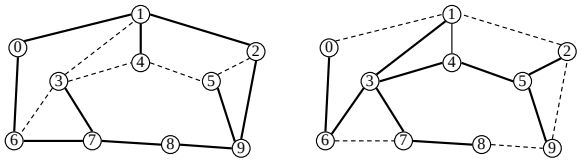

Fig. 1. Two logical topologies depicted by bold lines in a single network. All nodes are reachable in both topologies, while the shortest paths between different node pairs are highly disjunct.