ASIA-PACIFIC INSTITUTE OF INFORMATION TECHNOLOGY

Computer Networks and Distributed Systems

(*Research Paper*)

Resilient Overlay Networks

By:

Tan Wee Khoon

KL003844_HF0272SE

1. Introduction

The objectives of this paper are to discuss the problems and issues that occur in Resilient Overlay Networks (RON). This paper is divided into ten sections:

1) Introduction

2) Background
   - This topic is important to provide the reader a subtle hint of what is RON and how does it functions.
   - It also mention some problems with current wide area routing protocol
   - It also discusses the goals of RON.
   - The problems of approaching this topic are there are a lot of research needed to be done in RON and wide area routing protocol.

3) Main Design Goals
   - This topic is important to provide the reader a broader knowledge of RON design goals and its slightly technical workings.
   - It cited the types of failures, how does it occur and how it affects current transport level protocols.
   - This section further discusses how BGP handle these failures.
   - The problems of approaching this topic is further research needs to be done in policy routing.

4) Problems with RON
   - This topic is important to let the reader knows the drawbacks of RON.
   - It tells the reader that RON is lacking in security and is incompatible with some networking devices.
   - The problem of approaching this topic is research needs to be done whether the specified drawbacks can be solved by the author.

5) Existing Efforts to Find Solutions
   - This section is essential for the reader to know the effort of the RON researchers in finding a solution to the specified problems.
   - The problem of approaching this topic is words must be carefully chosen to describe the current situation.

6) Issues of RON
   - This segment is vital for the reader to gain insight into the issues circulating RON.
   - The problem in approaching such topic is the solutions devised must be logical and feasible.

7) Evaluation and Comments
   - This topic is valuable for the reader to know other reviewers' opinions including the author himself.
   - The problem in approaching this topic is the comments must be logically deduced in accordance to author's knowledge.

8) Conclusion

- This topic is vital for the reader to recall the important points and arguments presented.
- The problem in approaching this topic is author must try to trigger reader's interest to carry further research on RON.

9) Future Work
   - The problem in approaching this topic is the author must scour through numerous research papers for the original researchers view.

10) References

2. Background

RON is an architecture that allows a small group of distributed Internet applications to detect and recover from path outages and periods of degraded performance within several seconds, improving over today's wide-area routing protocols that take at least several minutes to recover. A RON is an application-layer overlay on top of the existing Internet routing substrate [1]. The RON nodes monitor the functioning and quality of the Internet paths among themselves, and use this information to decide whether to route packets directly over the Internet or through path that exists between RON nodes by utilizing chosen routing metrics.

RON evolution is based on previous studies that measure end to end network reliability and performance on Internetworking Protocol (IP) based routing for fault tolerance and on overlay network techniques to enhance performance.

According to the measurement results on March 2001 of a working RON deployed at sites that is scattered across the Internet which is sampled for 64 hour period across the twelve nodes RON, 32 significant outages that lasts over thirty minutes were discovered over the 132 measured paths [2]. However, RON was able to detect, recover and reroute around all of the outages in less than twelve seconds on average. Moreover, RON was perceived to improve the loss rate, latency or throughput in data transfers.

RON is conceived due to the problems with the current wide area routing protocol namely Border Gateway Protocol (BGP). The current version of BGP is known as BGP-4 [9]. It is developed for use with Transmission Control Protocol (TCP)/IP suite and has become the preferred Exterior Routing Protocol (ERP) which is also known as Exterior Gateway Protocol (EGP) for the Internet. TCP (which is connection oriented) and User Datagram Protocol (UDP - which is connectionless oriented) are the two common transport level protocols in use as part of the TCP/IP suite.

Connection oriented protocol has features such as flow control, error control, sequenced delivery et al. While connectionless oriented protocol are fundamentally unreliable. ERP is a protocol used to pass routing information between routers in different Autonomous System (AS). An AS is a set of routers and networks managed by a single organization which has path between any two pair of nodes. It consists of a collection of routers exchanging information via a common protocol [9].

In this BGP architecture, detailed routing information which is usually shared between network service providers and AS is extensively filtered and summarized at the border routers. Therefore, it is able to allow the Internet to scale into a million of networks. Its scalability is by means of aggressive aggregation and information hiding which is achieved via peering and transit relationships.

Nevertheless, this scalability is at the expense of reduced fault tolerance in end to end communication between Internet hosts. This happens because BGP hides many topological details in the scalability and policy enforcement terms, has little information about network traffic conditions and lessens routing updates when problem arises to prevent costly large scale route oscillations. Thus, the Internet is vulnerable to router and link faults, configuration errors and malice which affects the connectivity provided by Internet Service Provider (ISP). Furthermore, BGP fault recovery method takes ten minutes or more to connect affected routes reliably when significant outages happen.

Consequently, the Defence Advanced Research Projects Agency (DARPA) and Space and Naval Warfare Systems Centre, San Diego sponsored the Massachusetts Institute of Technology Laboratory of Computer Science for a research in this area [2].

Interestingly, RON is able find paths between its nodes if the underlying topology has physical path redundancy even when wide area routing protocol like BGP cannot. This is because some of the path is legal but is forbidden by BGP transit policy. Essentially, RON routing control is moves towards the end systems hence providing end to end monitoring of the network traffic. This means path measurements information is exchanged among RON nodes via a routing protocol and forwarding tables is build according to the path metrics that consists of latency, packet loss rate and available throughput. Each RON node obtains the path metric through aggressive probing and passive observations of on going data transfers. RON is designed to support between 2 and 50 nodes due to excessive bandwidth overhead incurred for the aggressive probing experiments which leads to faster problems recovery in several seconds rather than several minutes.

RON main goal is to facilitate a group of nodes in communication with each other though there are problems with the underlying Internet paths connecting them. RON detects problems by aggressively probing and monitoring the paths connecting its nodes. If the underlying Internet path is not the best, then RON will forward data to its other nodes. Else, other RON nodes do not get involved. RON requires mostly only an intermediate hop to route around failures instead of the current hop by hop model is a worthy of note characteristic. This is because hop by hop model is error prone.

The second goal of RON is to integrate routing and path selection that is tightly coupled with distributed applications which is more tightly than have been done conventionally. This integration provides the ability to utilize application specific metrics in path selection and also the flexibility that allows application specific measures to determine what network conditions constitute a “fault”. Thus, RON can be used in a variety of ways.

The third goal of RON is to provide a framework for the implementation of expressive routing policies which govern the choice of paths in the network besides

the application specific metrics. As an example, RON facilitates grouping packets into categories that assists in implementing application specific measures of acceptable use which can also be used to enforce forwarding rate controls.

RON design is based on the overlay network and indirect routing techniques. Overlay network is a reused old idea that stems from the Internet which was developed as an overlay on the telephone network. Indirect routing is a technique inspired by the Detour framework which is a project of University of Washington [2].

3. Main Design Goals

These are the three main design goals of RON:

1) Failure detection and recovery in less than 20 seconds
2) Tighter integration of routing and path selection with the application
3) Expressive policy routing

3.1 Fast Failure Detection and Recovery

Current Internet routing protocol i.e. BGP-4 does not handle failure satisfactorily. For this reason, there are two kinds of failures associated with network namely: link failures and path failures. Link failures happen when a router or a link connecting two routers fails because of software error, hardware malfunction, configuration error, operational error or link disconnection. Path failures happen due to bursts of traffic that trigger a high degree of packet loss or variable latencies i.e. Denial of Service (DOS) attacks, severe congestion et al.

Unlike network, applications perceive failures in two ways: outages or performance failures. Link failures and severe path failures cause outages which affect most protocols including TCP to degrade by several orders of magnitude. It happens when the average packet loss rate over a sustained period of several minutes is high which is about 30% or higher [2]. At this rate, most packets times out which causes TCP to work at lower performance or becomes largely unusable. Outages at a packet loss rate of 10% or more causes UDP without good packet level error correction not to work at all. In contrast, performance failures are less extreme which affects throughput, communication latency and degradation of packet loss rate by a factor of 2 or three et al [2].

When a link failure causes an outage, BGP takes several minutes to converge to a new valid route. Furthermore, IP layer protocols like BGP cannot detect problems such as packet flooding and persistent congestion on links or paths which greatly degrade end to end network performance. This is because BGP will persist to route packets to the faulty path as long as the BGP session is still active which do not provide an acceptable performance for application that is using it. As a result, RON goal is to detect and recover from these outages and performance failures within several seconds.

3.2 Tighter Coupling with Application

Albeit failures and faults occur, adaptive applications such as TCP may still work but at inferior performance. Therefore, RON is designed to allow applications independently define and react to failures. In this way, applications may specify their metrics and priority in path selection. As an example, the application may give higher priority to packet loss rate over throughput depending on the interactivity. On the contrary, a router may not be able to utilize all of these metrics and priority simultaneously. Currently, RON is designed to allow application influences the path selection using a single metric [2]. Thus, user can select their preferred routing metrics which cannot be achieved at the moment.

3.3 Expressive Policy Routing

These days, there are demand for policy routing and enforcement of Acceptable Use Policy (AUP) or other policies. AUP is a policy that a user must agree to follow in order to be provided with access to a network or to the Internet. It is a common practice among educational institution, ISP etc.

Nevertheless, BGP is incapable of expressing fine grained policies targeted at users or hosts. As a result, this reduces the set of paths available especially in the case of failure. Moreover, it also inhibits innovation in the use of carefully targeted policies such as end to end per user rate policies rate controls or enforcement of AUP based on packet classification. Besides, the network will be used efficiently if only it could enforce routing policies such as providing different routing priorities and metrics in conformance to the traffic types i.e. video, voice and data. However, RON could violate AUP as a result of its flexible policy routing.

RON will usually run on fairly powerful end points. So, the creators believe RON is more suitable to provide fine grained policy routing. According to the present RON research, the creators have designed and implemented a system that achieves fine grained connection level failover across both local and wide area server replicas which is without a front end transport or application layer switch [5]. It uses some form of redundancy that is already in used by some reliable services that is available on the web and also the recently proposed end to end connection migration for transport protocols such as TCP.

4. Problems with RON

These are the two problems of RON that will be discussed:

1) Misuse of an established RON by an authorized user.
2) RON hosts behind Network Address Translator (NAT).

4.1 Misuse of an Established RON

RON is designed to be deployed between a small numbers of cooperating nodes which is between 2 and 50 nodes. For this reason, RON is able to secure its network from cyber attacks by means of cooperation among nodes. Hence, they cannot be used to find “backdoors” into networks without the permission of an authorized user of that network [2]. Nevertheless, these so called authorized users can actually abuse the use of the network.

4.2 RON Hosts behind NAT

Usually, hosts that are behind NAT do not have a globally reachable Internetworking Protocol (IP) address or a Domain Name System (DNS) name. If both RON hosts are behind NAT, they cannot communicate directly. Then, RON will perceive this as an outage which leads to reroute attempts around it. Sooner or later, this will establish connectivity between the involved hosts. However, this may result in inferior routing performance.

5. Existing Efforts to Find Solutions

The abuse of an established RON problem is deemed can be dealt at the administrative (human) level by the RON creators because of the small scalability. Furthermore, participation in a RON implies trust that the other hosts will not misuse the network.

The creators think that users who violate network policy should be resolve at the human level as in today's enforcement. Therefore, the creators have not implemented cryptographic authentication and access control which can prevent misuse of an established RON. However, they do agree that these methods in complement of other security mechanisms are required in order to detect misbehaving users if RON do achieves widespread deployment [2]. Currently, the creators are more concerned with the handling of DOS attacks by RON which uses the distributed filtering techniques.

Next, the problem where RON hosts are behind NAT do not appear to have a progressive development as the creators are more concerned whether the performance of RON are affected in a widespread development due to the implementation of aggressive routing algorithm. However, they had proposed to use the caching technique to solve the problem where "one" RON host not both is behind NAT. It works by caching the reply address or port for the host behind NAT which uses the RON aggressive probing characteristic to allow this port mapping technique [2].

6. Issues of RON

These are the 2 issues of RON that will be discussed:

1) Network security in RON
2) Compatibility of RON with existing network

6.1 How Secure is RON?

RON is touted as a platform for cooperative security among its nodes. It is shown that RON can provide several protections based on its end to end monitoring characteristic. Firstly, it can protect data over RON virtual links. Secondly, it provides protection of routing and network traffic. This is because RON can utilize policy routing where the users can choose which sites to trust. Thirdly, protection against DOS attacks as a result of end to end authentication. According to the latest research, RON can further absorb the impact of DOS attacks by using the distributed filtering

techniques. Nevertheless, RON is not totally invulnerable to DOS attacks. However, it can handle these attacks better than the current system.

After all these security claims, it is unfortunate that RON creators overlook the importance of security within. As there goes a saying, “Day prevention, night prevention, thief within is difficult to prevent (Chinese proverb translated)”. Unfortunately, this saying is true even till today. Furthermore, there are numerous malicious “authorized” users reported by mass media i.e. TV, newspaper etc. Though, some of them are still largely undetected. Additionally, the infamous hacker Kevin Mitnick mentioned how he got the information he needed for his numerous computing giant firm systems break-ins from duping unsuspecting authorized users. This means the so called authorized users can also actually be unauthorized users that are exploiting the inside information he/she got.

I am not a security expert therefore I will reuse the latest and time tested security measures instead of reinvent the wheel. Other than the proposed yet not implemented cryptographic authentication and access control, the network administrators will surely implement security features such as traffic padding, firewall in complement with security guidelines such as frequent changes of authentication codes etc. Moreover, they can also implement message authentication and public key encryption as needed. In the current modern computer world, digital signatures and conventional encryption had been implemented in most system. Soon, most system will utilize IPSec; security features issued in IPv6 which is compatible with IPv4 (the currently used IP) and IPv6.

My suggestions are there should be a system at each RON node that alerts any administrator when there are authorized users including administrators themselves abuse the use of the network, record its offence and report it to other RON nodes. Hence, there should be several administrators that have equal authority monitoring the network. From then on, the RON administrators can decide what to do i.e. forbid the particular user, place restriction upon offender etc. Furthermore, there should be a strict code of conduct in disseminating information especially those that are derogatory in every authorized user.

6.2 RON Compatibility with Existing Architecture

Currently, RON has problems operating across NAT and firewall. This is because both RON hosts can be behind NAT meaning they do not have a global IP address. Therefore, the RON node receiver would have to redirect the packet to the destination port address. However, the firewall may regard the access to the port as unauthorized therefore preventing its entry.

My suggestions are whenever a RON host is established; the IP address including the port address (if any) of the host should be announced to each RON node and the particular network administrator should configure the node’s firewall to allow access to the specified port. Then, the sender of the packet should include its port address (if any) in addition to the IP address. In this manner, RON would not have to redirect the packet when they reach the IP address as in the proposed solution for one host behind NAT. Furthermore, the addresses of the both RON hosts behind NAT are globally reachable within the RON.

7. Evaluation and Comments

Other than the discussed RON internal security and incompatibilities with current architecture, most critics think that RON cannot sustain its operation in widespread deployment because of the bandwidth incurred as a result of aggressive probing to explore alternate paths. The RON creators are well informed about this fact and they do not expect it to change. However, they think this is an area for future work [2].

RON is mentioned to be violating AUP and BGP transit policies. Nevertheless, the creators are ignorant to the fact that RON seriously violates security policies. In fact, RON could be regarded as an active attack i.e. DOS attack by some system due to its aggressive probing nature if the security thresholds are low. This could be true if there are many RON that aggressively probe the same site. Moreover, its passive observation of on going data transfers could be deemed as a passive attack i.e. traffic analysis which will prevent RON from executing properly by some security mechanisms. Passive attacks are difficult to detect therefore the security mechanisms emphasized on preventing the success rates of these attacks. Hence, RON may produce inaccurate forwarding table for path selection. In addition, these RON applications could be used as a security breaching tool with/without modifications if it falls into the wrong hands. As an example, a hacker (isn't it an unauthorized user?) would be grateful for the traffic analysis data that RON provides. Since RON had been deployed at several sites but none reported these problems. Hopefully, these security violations do not materialize.

8. Conclusion

This paper states that RON is a reliable application layer overlay network architecture that allows small group of distributed Internet application detect and recover from failures in less than 20 seconds, tightly integrated with application to provide user selected routing metrics and priorities in path selection and expressive policy routing that allows enforcement of network policies such as AUP. Furthermore, RON can route packets through virtual links that exists between RON nodes but restricted by BGP transit policy instead of the Internet. In addition, RON is shown to be able to reroute around failures using an intermediate hop.

BGP provides substantial scalability at the cost of inferior fault tolerance which is due to significant summarization and filtering of detailed routing information. Therefore, RON is devised to handle the limitations of BGP. RON uses aggressive probing experiments and passive observations of data transfers to build its forwarding table that consists of several path selection metrics and detect problems in existing paths.

RON is designed to handle outages and performance failures better than it is handled by current wide area routing protocol BGP. In addition, it allows application specific metrics to influence RON path selection. These metrics are usually chosen by user that responds to different traffic type or applications. However, BGP cannot accommodate these dynamic routing metrics. RON is designed to be able to utilize expressive policy routing which will make available more alternate paths especially in

the case of failures which is deemed illegal by BGP transit policy. Furthermore, RON is designed to provide fine grained policy routing.

There are two problems of RON discussed: misuse of an established RON by authorized users and both RON hosts behind NAT. As we know, an authorized user can also be an unauthorized user. Therefore, it is important that security mechanisms are fully utilized in addition to adhering to strict security guidelines. If Kevin Mitnick could easily dupe the people for the information he need, probably the authorized network administrators should have some mental toughness in dealing with these psychology deceptions. Maybe these administrators should be trained to handle these trickeries. It is true that the current security mechanisms are sufficient to prevent intruders? I do not think so because hardware prices keep falling coupled with rising processor speed which provides the opportunity for crackers to decipher the encrypted codes. However, it may take some time in order for them to do so.

Whenever a RON host is established; the global address including the local address (if any) of the host should be announced to each RON node and the particular network administrator should configure the node's firewall to allow access to the specified port. What if RON does achieve widespread deployment? It is probably not feasible anymore to announce addresses to each RON node. Maybe the Internet Engineering Task Force (IETF) should revise its IPv6 definition to allow inclusion of local address to its 128 bit addresses. In future, maybe the RON creators should developed a system that automates the tasks of configuring each node's firewall to allow access to specific port when a new RON node is established.

As expected, RON creators do not think the tradeoffs between scalability and reliability to change if it does achieves extensive deployment. If there are numerous RON on the Internet, will the bandwidth overhead increase up to a performance degradation extent? I think it could happen if there is node of one RON that is also a node of another RON and so on because the connectivity will increase the scales of the RON. Maybe the creators should set a policy that a RON node cannot be a node of another RON. However, extra effort may be needed to ensure this does not happen.

RON characteristic perceived as an attack! How true could it be? I believed it can happen if many RON actively probing the same site which will eventually caused traces of DOS attacks. Next, the traffic analysis data could be used by authorized (or is it unauthorized?) users for its malicious purpose and so on.

(4137 words including titles but excluding references link)

9. Future Work

Currently, the creators are trying to understand the interactions between co-existing RON that will compete on Internet paths. They are also investigating the routing stability in the Internet where many RON co-existing.

They also plan to explore multiple criteria path selection in the future. Currently, RON allows only a single criterion metric in path selection. Also, they viewed RON as the catalyst for future Content Delivery Network (CDN) development by providing the routing components required by these services that are already functions of RON.

10. References

1) **Article in HTML on the Internet (from a Home Page):**
Resilient Overlay Networks, Available on the WWW at: http://nms.lcs.mit.edu/projects/ron/

2) **Research Paper (downloaded from the Internet):**
David Andersen, Hari Balakrishnan, Frans Kaashoek and Robert Morris, Resilient Overlay Networks, SOSP 2001, Available for download at: http://nms.lcs.mit.edu/papers/ron-sosp2001.html

3) **Research Paper (downloaded from the Internet):**
David G. Andersen, Resilient Overlay Networks, Master of Science Thesis, May 2001, Available for download at: http://nms.lcs.mit.edu/projects/ron/

4) **Research Paper (downloaded from the Internet):**
David G. Andersen, Hari Balakrishnan, M. Frans Kaashoek and Robert Morris, The Case for Resilient Overlay Networks, Available for download at: http://nms.lcs.mit.edu/papers/ron-hotos2001.html

5) **Research Paper (downloaded from the Internet):**
Alex C. Snoeren, David G. Andersen and Hari Balakrishnan, Fine-Grained Failover Using Connection Migration, Available for download at: http://nms.lcs.mit.edu/papers/migrate-failover.html

6) **Research Paper (downloaded from the Internet):**
Nick Feamster, David G. Andersen, Hari Balakrishnan and M. Frans Kaashoek, Measuring the Effects of Internet Path Faults on Reactive Routing, Available for download at: http://nms.lcs.mit.edu/papers/failures-sigm2003.html

7) **Research Paper (downloaded from the Internet):**
Jaeyeon Jung, Emil Sit, Hari Balakrishnan, *Member, IEEE* and Robert Morris, DNS Performance and the Effectiveness of Caching, Available for download at: http://nms.lcs.mit.edu/papers/dns-imw2001.html

8) **Research Paper (downloaded from the Internet):**
David Andersen, Nick Feamster, Steve Bauer and Hari Balakrishnan, Topology Inference from BGP Routing Dynamics, October 2002, Available for download at: http://nms.lcs.mit.edu/projects/ron/clustering-imw2002.html

9) **Book:**
William Stallings, Data & Computer Communications Sixth Edition, Prentice Hall, 2000.