Increasing fault tolerance in service-level agreement constrained context-aware data networks

Ints Meijers

Faculty of Computer Science and Information Technology

Riga Technical University

Riga, Latvia

ints.meijers@edu.rtu.lv

Abstract— Attribute study and analysis of fault tolerant data networks. This work is aimed at introducing SLA constrain into fault tolerance and thus increasing overall network availability. Proposed model will evaluate given constraints and select best path that fits requirements. Fault tolerance is increased by adding multiple constraints and thus reducing available paths to best fitting ones.

Keywords—SLA, DMVPN, fault tolerance, computer networks

I. INTRODUCTION

Data networks experience a lot of interruptions during their lifecycle and mean time to repair (MTTR) serves as a basic expression of network unplanned maintenance time spent on particular asset divided by total number of failures that were experienced over a specific period, typically expressed in hours. Network downtime during unplanned maintenance has a direct impact on service-level agreement (SLA) [1]. For this paper assumption is that failure is inevitable and only means we have is to introduce redundancy via alternative paths through data network. We can have alternative paths at physical or data link level [2]. Multiple overlay networks such as dynamic multipoint virtual private network (DMVPN) may be built over same physical and data link connections thus introducing redundancy and path alternatives at routing level [3]. Using only one node with multiple overlay networks leaves us with a single point of failure at node and physical connection level. Redundancy may be at physical connection level where two network nodes are connected by multiple lines in parallel, employing a different data network access technology, for example ethernet line and cellular network. Next level of redundancy may be introduction of a separate network node within site for each separate connection and network access technology. Every node, connection, technology and overlay network we introduce into network, increases its fault tolerance.

In context of a data network SLA is an agreement that defines a level of service expected by a customer from a service provider, stating measurable metrics by which particular service is being measured and all penalties, if any, that apply if agreed-on service level is not achieved. Usually such agreement is honored by monitoring the data network, collection of performance metrics and creating live graphs of generating reports at preferred intervals. Monitoring is reactive and serves as a post factum proof that events occurred which is good for reporting and accountability. It can be performed by using simple network management protocol

(SNMP) and further detailed by introducing NetFlow analysis of data streams [4][5]. Problem with traditional SLA application lies in its isolation from data network and using it only as data network performance metric evaluation for historical data. Since SLA uses defined metrics that can be monitored by a separate monitoring server, it also is possible to use the same metrics at data network node as runtime information and feed this data to algorithm for evaluation and pro-active reaction to change of network state. Thus, by introducing SLA awareness into fault tolerance computation, aim is to bring the result closer to real world application and increase overall SLA completion.

Fault tolerance can be explained as data network ability to continue operating uninterrupted despite component failure, which at its lowest level leads to hardware redundancy with multiple chassis serving as one entity. At hardware level a single chassis system with multiple line cards may be used to achieve high level of redundancy. At next level is physical connection redundancy which leads to network level redundancy and building upon that are multiple overlay networks. Each level introduced increases fault tolerance and in combination helps to achieve a sustainable data network.

Context awareness in data networks is their ability to identify and separate multiple data flows, ensuring priority forwarding for real-time traffic and maintaining defined level of bandwidth over time for other traffic types. It is achieved by using quality of service (QoS) [6]. There are a set of QoS technologies that can be used in context-aware networks for bringing it closer to defined SLA. They include classification and marking tools to differentiate between applications and mark them according to defined policy. This marking can be applied at multiple network levels, from data link to network and used at each network node to separate traffic into different traffic classes. There are congestion management tools for packet scheduling which uses marking in place to determine which queue is appropriate for a received packet. Congestion avoidance monitor queue congestion and may use a random early detection (RED) or weighted random early detection (WRED) for low priority packet dropping when congestion occurs. RED only monitors queue Shaping and policing manipulate network traffic to maintain set level of bandwidth which is important when line speed is larger than contracted speed from service provider [6-7]. If customer sends traffic at line speed and provider drops excess traffic as per agreement, then SLA or fault-tolerance are unachievable. Shaping is used outbound from customer site and policing is used inbound from service provider side. All these tools may be used in any

combination to ensure SLA. Context-aware network will use this data to make traffic forwarding decisions.

The design details of the proposed model are presented in section 2. Analysis of the proposed model and concluded experiments are described in section 3 of this paper. The conclusion is presented in the last section.

II. DESIGN DETAILS OF THE PROPOSED MODEL

Proposed model will use DMVPN overlay networks to attain flexibility at physical, data link and routing level. It will measure network state by several metrics: round-trip time (RTT), path mtu discovery (PMTUD), available bandwidth and reserved bandwidth. Bandwidth reservation will be implemented by using resource reservation protocol (RSVP) [8-10]. It will ensure SLA by using QoS tools where applicable.

DMVPN has many advantages like the capability for creating a dynamic-mesh VPN network without need to statically pre-configure all possible tunnel endpoint peers. They can be introduced on demand and decommissioned without any extensive configuration. It is a hub-and-spoke network where hub is known to all spoke peers and is pre-configured on them. Another advantage is capability of forming on demand dynamic-mesh or spoke-to-spoke tunnels when traffic needs to be passed between spokes. By using at least two DMVPN tunnels, preferably one for each connection type we attain redundant and fault tolerant topology. For network path selection a dynamic routing protocol is used. In this case border gateway protocol (BGP) was selected due to its ability to support multiprotocol extensions and embedded support for BGP communities which may be used to transfer additional network parameters by mapping community values to specific QoS policy. In such way traffic can be directed over a different path, overriding destination-based routing. Other significant BGP benefits are stability, rich path metrics and scalability [11].

A new approach is to introduce multiple SLA metrics into DMVPN network and use them to determine best path for a given traffic, where it is defined by four SLA constraints: (1) bandwidth, (2) RTT, (3) jitter, (4) packet loss. This, in combination with a proper QoS marking at network edge, supports SLA end-to-end. Also, this gives a possibility to offload less important traffic to lower quality link. For example, non real-time traffic can be sent over cellular network and leased line with priority queueing would serve real-time traffic, either voice or video. So, destination-based routing can be overridden, by directing traffic to a different path which is selected by matching a traffic profile to available connections where these connections are evaluated by SLA metrics.

Relations between these modules are shown in Fig. 1

Fig. 1. Module relations

SLA module consists of traffic profile selector and performance metric evaluating algorithm. QoS policy is determined based on SLA profile. SLA performance metric evaluation includes metric measurement functionality for link quality monitoring. Metrics measured are RTT, jitter and packet loss. Path mtu is measured to ensure that it is the same end-to-end and packets sent will not encounter fragmentation in IPv4, which impacts performance and increases delaying of traffic, at any part of the network. Path mtu monitoring in IPv4 networks is performed by sending data packets with don't fragment (DF) bit set and in case of smaller mtu at any device along the traffic path Internet Control Message Protocol (ICMP) Fragmentation Needed (Type 3, Code 4) is sent back to originating host informing it of actual mtu available, thus enabling correction of packet size at sender side to alleviate this issue [12]. For transmission control protocol (TCP) it may be achieved by adjusting maximum segment size (MSS) [13]. In IPv6 networks fragmentation is not supported and initially hosts assume that mtu is the same along the path. In such network path mtu discovery can be performed by monitoring for ICMPv6 Packet Too Big (Type 2) messages received from an intermediate host on a traffic path which contains maximum mtu supported [14]. All measurements are done by each host individually and they rely on themselves for path quality verification.

SLA profiles are manually defined on each node separately and define bandwidth constraints as well as acceptable range of metrics. Typical SLA profiles that may be used in a network are shown in Table I. Gold traffic profile is best suited for real-time traffic such as VoIP or video where selected metrics adhere to ITU standard G.114 [15]. Silver and bronze are for business-critical traffic and important traffic. Basic SLA profile is best-effort treatment for traffic.

TABLE I. SLA PROFILES AND METRICS

SLA profiles	Metrics			
	Bandwidth	Latency	Packet Loss	Jitter
Gold	100 Mbit/s	<150 ms	<1%	<30ms
Silver	50 Mbit/s	<250 ms	<2%	30-50ms
Bronze	10 Mbit/s	<500 ms	<3%	50-70ms
Basic	n/a	n/a	n/a	n/a

With defined SLA profiles and measured path metrics appropriate QoS policy can be selected. Gold QoS policy must include a priority queue for real-time traffic forwarding and guarantee requested bandwidth. This queue must be last to experience any packet drops under normal operation. Only excess traffic may be dropped under tail drop condition when priority queue is full. Other queues will serve business critical and important traffic. Also, one queue is reserved for best effort traffic which holds no bandwidth guarantee. QoS policy is applied on all nodes to all interfaces in outbound direction and ensures traffic processing according to per hop behavior (PHB) under differentiated services (DiffServ) guidelines [16]. Silver and Bronze QoS policies may include priority queue and must include business critical and important queues with bandwidth guarantee. For each queue in policy appropriate RED or WRED parameters are set. They define a

point in queue, which fills with traffic if congestion starts to occur, at which the packets are being marked with explicit congestion notification (ECN) bit [17]. If queue still increases then packets are dropped randomly even before the queue is full, when tail drop would occur. Thus, tail drop is avoided, and TCP traffic slows down [18]. Such packet loss has no effect on user datagram protocol (UDP) traffic and for such traffic, packet drops will continue to occur, and it is up to the application to react to packet loss and reduce the traffic sent [19].

For network layer reachability BGP is used as dynamic routing protocol and provides one best route for every network in topology [20]. BGP has a separate table of all information received from its peers. It serves as source of alternative routes for injection into routing table if SLA calls for it. All routes are exchanged via DMVPN hub router which serves as BGP route reflector. That significantly simplifies spoke router configuration since only peering to hub for each DMVPN must be defined.

For each network next hop router is known. Next hop routers are reachable through DMVPN tunnels. For each destination SLA performance metric measurements are being calculated. If performance metric falls below SLA profile assigned to this network or individual host on this network, then alternative paths are evaluated. If alternative path has better metrics, traffic is switched over by injecting a local route. Since this is done on each peer individually, asymmetric routing may occur. Introduction of new network is shown in Fig. 2.

Fig. 2. New network introduction into topology

Monitoring of all available hub routers reachable via DMVPN tunnels is done by using BGP built in keepalive process. And separately SLA performance monitoring process is run to acquire data for evaluation. A small history bucket is stored since this data is not needed for reporting purposes however reporting functionality may be added as an improvement for detailed reports. If acquired metrics are within margins defined in SLA profile, then interface is marked as appropriate for that class. It is assumed that if interface satisfies Gold profile, it can satisfy any lower tier profile. This is shown in Fig. 3.

Fig. 3. Performance metric generation and evaluation

During network operation, event of network degradation may occur. This will be detected by SLA performance monitoring process and DMVPN tunnel will be remarked to reflect these changes as can be seen in Fig. 4.

Fig. 4. Performance metric update

If there is a directly connected network to the spoke with SLA requirements higher than current best route next hop DMVPN tunnel supports, BGP table is searched for alternative paths and for each alternative path its DMVPN tunnel is evaluated to determine if it may satisfy the demand. If such tunnel is found, traffic is redirected to this alternative path by introducing policy-based routing (PBR) that impacts only locally connected networks. All traffic coming inbound from that network is forwarded to alternative path. If no such path is found, network state is deemed unable to satisfy defined SLA and it must be signaled to monitoring server by generating an SNMP trap. This method ensures that network degradation is detected swiftly and before end user notices it. If network restores state, DMVPN tunnel is marked accordingly, traffic is shifted back to original best path and SNMP trap is generated to inform about state change.

For testing proposed method, following topology was built and can be seen in Fig. 5.

Fig. 5. Topology used for testing

Topology consists of 12 devices with following function – 4 end devices PC1 to PC4, two customer premises equipment devices CPE1 and CPE2, two DMVPN Hub devices HUB1 and HUB2 and 4 ISP backbone devices PE1 to PE4. CPE1 serves as default gateway for local area network devices (PC1 and PC2) and CPE2 does the same for PC3 and PC4. Each CPE has two separate connections to backbone, each connection to different PE device. There are two connection types – MPLS and 4G. MPLS is leased line service from ISP and 4G is mobile internet connection. MPLS connection is considered as primary line and 4G serves as backup. Each of these connections has different properties, such as bandwidth, speed and separately induced packet loss for 4G connection to simulate shared bandwidth experience. Every PE device has two connections to adjacent PE devices, providing redundant connection over the topology in case any individual PE device fails. If two PE devices fail at the same time, there can be three results – first case, when they fail diagonally (for example, PE1 and PE4), isolates two networks where local area devices (PC1 to PC4) are located from each other; second case, where they fail vertically (for example, PE1 and PE2), also isolates local area devices from each other; and third case, where failure is horizontal and has two separate results for each PE pair, (PE1,PE3) failure has impact on traffic since these PE devices are providing connection to primary lines and there is time needed to switchover from MPLS to 4G backup connection, (PE2,PE4) failure has no impact on traffic forwarding if primary lines are working.

This example network has all properties and elements that large-scale network solution may have. So, it can be scaled to show how proposed method solves issues in larger networks. Backbone uses OSPF, MPLS and BGP protocol combination, thus ensuring end device traffic isolation by employing MPLS Layer 3 VPN. Such backbone can easily support new devices for same VPN network and may be expanded with other VPN networks if needed. This model assumes that MPLS last mile connection is wired and 4G is wireless. Customer CPE has two separate last mile technologies and they are separated by using VRF (Virtual Router and Forwarding) functionality. Routing towards backbone is static in each separate VRF. Over these last mile connections two separate DMVPN tunnels are built. HUB1 is used separately as DMVPN Hub for tunnel over MPLS and HUB2 is used separately as DMVPN Hub for tunnel over 4G. This eliminates single point of failure at hub level. Within DMVPN tunnels BGP is used as dynamic routing protocol. At BGP routing level primary and secondary connections are identified by using Local-Preference value of 50 for secondary connection and

leaving it as a default of 100 (bigger is better). For all prefixes received from HUB1 the LP value will set to 100 and for prefixes received from HUB2 it will be set to 50. By default, failover between connections, in case of primary connection failure happens in 180 seconds. That is default hold-down timer for BGP protocol. So, it may cause three minute long downtime for the network.

Two separate DMVPN Hubs use their own IP ranges, used in configuring CPE devices for connection to Hub devices. Two protocol combination – GRE (Generic Router Encapsulation) and NHRP (Next Hop Resolution Protocol) is used for end device registration and passing this information onto other CPE devices. At VPN level dynamic routing is done by BGP protocol and CPE devices are route-reflector clients for hubs which are route-reflectors. That ensures received prefix propagation from one CPE to another without using full mesh of BGP peering.

IP SLA monitoring was configured between CPE1 and CPE2 devices over MPLS connection with following parameters: operation timeout (milliseconds): 10, type of operation to perform: icmp-jitter, operation frequency (seconds): 1, threshold (milliseconds): 10. IP SLA monitoring was combined with track object with parameter (delay up 60 seconds), which ensures a delay after connection quality is fully restored and protects from short term quality increases, thus lowering connection flapping. Track object was tied to static route for monitoring purposes. Static route is inserted in routing table only if tracking object is up, thus only active when SLA quality measurements are satisfying. BGP protocol monitors this static route and sends information about local prefixes only if this route is present in the routing table. If this route is not present in the routing table, local prefixes are withdrawn from advertising. Thus, default BGP timers (60 second hello and 180 second hold-down) can be used while providing prompt reaction to network connection quality decrease. By default, BGP protocol ignores network quality issues if at least one hello packet is received within hold-down (180 seconds) interval. In Fig. 6 BGP statistics are displayed and Fig. 7 shows route tracking statistics.

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
192.168.154.1	4	65001	1631	1622	225	0	0	1d00h	2
192.168.155.1	4	65001	1615	1617	225	0	0	1d00h	3

Fig. 6. BGP statistics

```
Track 101
IP SLA 101 state
State is Up
26 changes, last change 00:00:57
Delay up 60 secs
Latest operation return code: OK
Latest RTT (milliseconds) 1
Tracked by:
Static IP Routing 0
```

Fig. 7. Route tracking statistics

It shows that while BGP was working without peering interruption there were 26 cases when measured connection quality was determined to be not within operational parameters and routing was switched over to secondary path. Fig. 8 shows how network device reacts to quality decrease.

```
*Sep 04:10:09.955: TRACK-6-STATE: 101 ip sla 101 state up -> Down
*Sep 04:10:09.956: RT(vrt): del 172.16.255.255 via 0.0.0.0, static metric [1/0]
*Sep 04:10:09.956: RT(vrt): delete subnet route to 172.16.255.255/32
*Sep 04:11:00.205: BGP(4): Condition hgt-cond-exists changes to withdraw
*Sep 04:11:00.218: BGP(4): (base) 192.168.155.1 send UPDATE (format) 172.16.2.0/24, next 192.168.155.20, metric 0, path local
*Sep 04:11:00.218: BGP(4): (base) 192.168.154.1 send unreachable (format) 172.16.2.0/24
```

Fig. 8. Events when connection quality decreases

When quality is restored within operational parameters and is measured by SLA then routing switches over to primary connection. Route track update and BGP switchover events are shown in Fig. 9.

```
*Sep  8 21:12:53.217: ZIRACK 0 STATE: 101 In ip sla 101 state Down > Up
*Sep  8 21:12:53.217: RT(vbr): updating static 172.16.255.253/32 (ex1) :
    via 0.0.0.0 Nu0 0 1040578
*Sep  8 21:12:53.217: RT(vbr): add 172.16.255.255/32 via 0.0.0.0, static metric [1/0]
*Sep  8 21:12:58.798: RT(vbr): updating bgp 172.16.1.0/24 (ex1) :
    via 192.168.154.10 0 1048577
*Sep  8 21:12:58.798: RT(vbr): closer admin distance for 172.16.1.0, flushing 1 routes
*Sep  8 21:12:58.798: RT(vbr): add 172.16.1.0/24 via 192.168.154.10, bgp metric [200/0]
*Sep  8 21:12:42.962: 09%[4]: Condition bgp-exists checks out to advertise
*Sep  8 21:12:42.962: 09%[4]: (base) 192.168.155.1 send UPDATE (format) 172.16.2.0/24, next 192.168.155.20, metric 0, path local
*Sep  8 21:12:42.962: 09%[4]: (base) 192.168.154.1 send UPDATE (format) 172.16.2.0/24, next 192.168.155.20, metric 0, path local
```

Fig. 9. Events when connection quality restores

For monitoring how end devices would be impacted by such solution MTR (My traceroute v0.86) was used to gather statistics and see delays and switchover between primary and secondary path. Fig. 10. shows the situation when all links are performing well, and no quality issues are present.

Fig. 10. Network operates without issues

To simulate issues in network, a one second delay was applied to MPLS last mile connection. In Fig. 11. impact on end devices is shown when no SLA monitoring is configured and applied.

Fig. 11. Network quality decreased and is not monitored

With same one second delay on MPLS last mile connection we can see in Fig. 12. how end devices are impacted when SLA monitoring is configured and applied.

Fig. 12. Network quality decreased and is monitored

Thus we can see how proposed solution ensured switchover between primary and secondary path in accordance with set network connection quality parameters.

III. RESULTS AND DISCUSSION

By using separate process for continuous monitoring of performance metrics, possible traffic disruption causes are detected before they significantly impact actual traffic and thus SLA is fulfilled. This additional layer of monitoring increases fault tolerance, keeping the traffic flow through network as stable and uninterrupted as possible.

A further improvement would be dynamic SLA profile allocation from centralized repository to ease deployment and alterations of already implemented profiles. A further improvement of SLA process could be introducing new signaling method to notify either end user host application that network cannot satisfy required SLA level or notify central server of such state of network. That would improve overall transparency of network state.

A further improvement of routing would be to introduce separate signaling of locally overridden routes to peers to avoid asymmetric routing and shift traffic from both sides.

Further studies could improve proposed model by introducing signaling method for end user applications to ask for a specific performance metrics, such as bandwidth, delay and jitter, before sending any traffic into network. Since this should be done in a controlled manner, separate authentication, authorization and accounting processes must be implemented to avoid abuse of this service.

Fault tolerance is increased by proactively monitoring network for changes and shifting traffic to alternative paths before significant impact is noticed. Network is ready for changes and that ensures that users have available service.

Device built-in connection quality monitor can collect data but cannot interpret them if we need to connect this monitored data to action in another protocol. For example, when we want to suspend a link for a prolonged time to protect the network from flapping link consequences.

External module which collects statistical data and provides devices with commands when to switch from primary to backup and when to return to normal operation may increase customizing ability and may increase reaction times since relying on internal protocol timers, although works but may not be optimal time wise.

IV. CONCLUSION

A new model was proposed to increase fault tolerance when service-level agreement constraints are introduced in evaluation of network state. This approach improves over currently employed destination-based routing and is suited for multiple connections and last mile technologies, adapting to network changes dynamically. Further improvement opportunities were described, and it gives a room to expand on, enabling future research.

REFERENCES

- [1] M.Raza, "Service Level Agreement (SLA) Examples and Template," www.bmc.com/blogs/sla-template-examples/
- [2] International Organization for Standardization "ISO/IEC 7498-1:1994 Information technology — Open Systems Interconnection — Basic Reference Model: The Basic Model," www.iso.org/standard/20269.html
- [3] M. Al-shawi and A. Laurent, "Designing for Cisco Network Service Architectures (ARCH)," 4th ed., Cisco Press, 2020, pp.302-303.
- [4] [RFC1157] Case, J., Fedor, M., Schoffstall, M., and J. Davin, "Simple Network Management Protocol (SNMP)," RFC 1157, DOI 10.17487/RFC1157, May 1990, pp.2-3.
- [5] [RFC3954] Claise, B., Ed., "Cisco Systems NetFlow Services Export Version 9," RFC 3954, DOI 10.17487/RFC3954, October 2004, pp.1-2.
- [6] [RFC4594] Babiarz, J., Chan, K., and F. Baker, "Configuration Guidelines for DiffServ Service Classes," RFC 4594, DOI 10.17487/RFC4594, August 2006, pp.3-4.

- [7] O. Wendell, R. Healy, and D. Donohue. "CCIE routing and switching certification guide," Indianapolis, IN: Cisco Press, 2010, pp.546-549
- [8] [RFC1191] Mogul, J. and S. Deering, "Path MTU discovery," RFC 1191, DOI 10.17487/RFC1191, November 1990
- [9] [RFC8201] McCann, J., Deering, S., Mogul, J., and R. Hinden, Ed., "Path MTU Discovery for IP version 6," STD 87, RFC 8201, DOI 10.17487/RFC8201, July 2017
- [10] [RFC2750] Herzog, S., "RSVP Extensions for Policy Control," RFC 2750, DOI 10.17487/RFC2750, January 2000
- [11] [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)," RFC 4271, DOI 10.17487/RFC4271, January 2006.
- [12] [RFC0792] Postel, J., "Internet Control Message Protocol," STD 5, RFC 792, DOI 10.17487/RFC0792, September 1981, pp.4.
- [13] [RFC6691] Borman, D., "TCP Options and Maximum Segment Size (MSS)," RFC 6691, DOI 10.17487/RFC6691, July 2012, pp.3.
- [14] [RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification," STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, pp.10.
- [15] ITU-T recommendation G.114, telecommunication standardization sector of itu series, g: transmission systems and media, digital systems and networks, www.itu.int/rec/T-REC-G.114-200305-I/
- [16] [RFC4594] Babiarz, J., Chan, K., and F. Baker, "Configuration Guidelines for DiffServ Service Classes," RFC 4594, DOI 10.17487/RFC4594, August 2006
- [17] [RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP," RFC 3168, DOI 10.17487/RFC3168, September 2001, pp.60-61
- [18] [RFC5348] Floyd, S., Handley, M., Padhye, J., and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification," RFC 5348, DOI 10.17487/RFC5348, September 2008, pp.3
- [19] [RFC8085] Eggert, L., Fairhurst, G., and G. Shepherd, "UDP Usage Guidelines," BCP 145, RFC 8085, DOI 10.17487/RFC8085, March 2017, pp.21
- [20] [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)," RFC 4271, DOI 10.17487/RFC4271, January 2006