

Quality of Service and Denial of Service

Stanislav Shalunov
shalunov@internet2.edu

Benjamin Teitelbaum
ben@internet2.edu

ABSTRACT

We argue that most forms of internet QoS must provide protection against DoS attacks. To date, the most cost-effective way of improving the treatment of some traffic has been to improve the treatment of all traffic. In a well-provisioned network, protection from DoS is essentially what defines QoS. QoS is not about the typical case; it is about the worst case.

Researchers have all too often assumed “normal network conditions” and developed forms of QoS that do not take into account adversarial considerations.

We argue that adversarial analysis needs to be performed on any QoS technique. The questions that need to be answered include: What are the consequences of a compromise of a host? A router? If no routers are compromised in a given domain, will it be always able to fulfill its QoS promises? If this is not the case, how many neighbors (and with what peering link capacities) does an adversary need to compromise to deny service to a specific pair of communicating hosts or, for a specific host, to deny service to some or all destinations? How can the operator control the possibility of DoS? How can it react to an ongoing attack?

We call for more research on the security aspects of QoS, especially the prevention of DoS.

Categories and Subject Descriptors

C.2.1 [Computer-Communication Networks]: Network Architecture and Design—*Packet-switching networks*; C.2.5 [Computer-Communication Networks]: Local and Wide-Area Networks—*Internet*; C.2.0 [Computer-Communication Networks]: General—*Security and Protection*

1. INTRODUCTION

Network researchers have dedicated significant resources to studying Internet quality of service (QoS). Most of the work has focused on network engineering techniques to regulate congestion. Many factors might affect the outcome of a

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM SIGCOMM 2003 Workshops August 25 & 27, 2003, Karlsruhe, Germany.

Copyright 2003 ACM 1-58113-748-6/03/0008S ...\$5.00.

network transaction: network congestion, routing stability, physical connectivity, router reliability, end-host capacity for performing the transaction in question, etc. QoS research has concentrated on the first of these factors, and so do we: in this paper, we consider network congestion—whether it arises naturally or through malicious action. When we discuss the worst-case scenario, we only consider the worst case of offered traffic load—and not, for example, router malfunction or a fiber cut.

2. QOS PROBLEM STATEMENT

The best-effort packet delivery service provided today by the Internet is susceptible to network denial of service (DoS) attacks. Although well-provisioned networks deliver very good typical performance [1] [5], they will, in general, deliver unpredictable service and, in the worst case, no service. At the same time, best-effort service has excellent efficiency properties when there is no attack.

What we need to understand is the problem statement for QoS.

If the problem is to provide a good service to a certain class of traffic (for some definition of *good*, such as near-zero loss and near-constant delay when a traffic profile is matched), then, in a modern world of inexpensive fiber, the solution is clear: by supplying sufficient capacity, provide the same good service to all traffic. As demonstrated in part by the lack of QoS deployment, service providers have concluded that, in today’s fiber-based networks, it is cheaper to provision generously than to deploy any form of QoS that gives better treatment to a privileged traffic class. Thus, very good best-effort service with abundant capacity wins over rationed capacity with only select traffic getting better than average treatment.

If the problem is to allow use of the full capacity of the network with some traffic still getting good service, then *elevated priority services*¹ are not the answer. There exist simpler solutions that are much easier to deploy and don’t require dramatic changes to operational practices or pricing models.

Best-effort service has very strong advantages. QoS needs to offer more than an optimization of capacity use if it is to be considered useful. The optimization, at least in today’s fiber networks, is illusory. The essential property of QoS that is desired by customers is that their preferred traffic is insulated from worst-case offered load scenarios.

¹See section 3 for an explanation of this term.

3. ELEVATED AND NON-ELEVATED PRIORITY SERVICES

We distinguish between *elevated priority services* and *non-elevated priority services*.

An elevated priority service delivers treatment that is better than the treatment afforded to the default best-effort service class. Examples of elevated priority services are Premium service, based on Expedited Forwarding (EF) per hop behavior [4], and services based on assured forwarding [2].

A non-elevated priority service delivers treatment that is either worse than the treatment afforded to the default best-effort service class or, in some way, equal but different. Examples of non-elevated priority services include the Alternative Best Effort service [3], which lets users trade off delay against loss, and scavenger-like services [6], which allow users to take advantage of otherwise unused network capacity in a way that does not affect the performance of the default best-effort service.

In this paper, we only consider elevated priority services. This is not because we believe that the other kind is less useful (in fact we believe that non-elevated priority services stand a better chance of deployment), but because non-elevated priority services are of limited utility in protecting against DoS attacks, and because most QoS research has concerned elevated priority services. On the other hand, we argue that protection against DoS is a defining characteristic for elevated priority services.

4. TRAFFIC FORECASTING AND STATISTICAL PROVISIONING

Traffic forecasting and statistical provisioning work well in circuit-switched voice networks with the notable exception of periods of panic dialing, as on and immediately after 2001-09-11. Data networks are, however, harder to predict. This is especially true when there is no per-bit charge for Internet traffic, as is the case with many research and education networks and some ISPs. Without pricing disincentives, individual users, both legitimate and malicious, can affect network utilization very significantly and very suddenly. Pricing feedback creates an incentive for the legitimate user to keep his offered load predictable. Moreover, it indirectly affect the success of malicious users who might rely on compromised hosts, as host owners would be concerned about the possibility of an expensive compromise. Statistical provisioning only works if the potential impact of (perhaps adversarial) individuals is small.

Note that individuals might use more than a single host. Using compromised hosts, a single person can quickly offer an amount of traffic that is as large or even significantly larger than ‘normal’ traffic load [7]. Such actions of individuals cannot be effectively predicted using statistical observations because they might happen only infrequently (but perhaps at the most inconvenient times), use a different toolset, and, consequently, have a different traffic signature each time they happen. Such attacks can target specific links in order to deny service to a specific site or, perhaps, a pair of sites.

Statistical provisioning works well under normal network conditions, but will fail catastrophically during attacks that introduce maliciously out-of-the-ordinary traffic.

5. WHAT IS A NETWORK DOS ATTACK, EXACTLY?

It is important to note that the line between network DoS attacks and legitimate traffic is blurred. This is why fighting DoS attacks is so hard. If a user automatically tests connectivity to some site every hour, is it DoS? What if someone runs some aggressive throughput tests? What if the tests generate so many small packets that a router becomes overloaded? Is intent what defines a DoS attack? If so, then it is impossible to recognize a DoS attack by inspecting traffic. Intent is in a user’s mind, not in the network. As far as we know, there exists no broad enough definition of DoS in the literature, for the reasons just outlined. With no definition of DoS, it would be difficult to prevent DoS by a general and automatic method that does not involve human judgment, which is, of course, expensive and does not operate on the time scale of electronic packet networks.

If it is hard to prevent or quickly cancel the effects of a DoS attack, a network operator might wish to have mechanisms that protect certain traffic that is known to be legitimate.

6. QOS MUST PROTECT FROM DOS

Protection of important traffic from unexpected changes in network utilization is highly desirable. Again, the changes might be due to deliberate malice or to natural changes in usage. Such protection is also the acid test of whether a guaranteed service is functioning correctly.

If an elevated priority service offers excellent statistical average-case properties, it still will not be deployed in today’s fiber-based networks. This is because these networks already offer excellent average-case treatment.

What networks cannot offer today is a guarantee, though such guarantees are sought by customers who perceive that they are necessary and are willing to pay extra for them. Customers are not willing to pay extra for a service that is indistinguishable from already excellent best-effort in the average case and offers no additional assurance of delivering the quality they need in the worst case.

Having an adversarial model in mind helps to understand worst-case properties better.

7. SERVICE EVALUATION CHECKLIST

We offer the following list of questions to consider when designing or studying an elevated priority service:

7.1 What are the consequences of a compromised host?

Are hosts expected to follow a code of honor in marking their traffic? What happens to both the default service and the elevated priority service(s) when a host misbehaves? What is the worst misbehavior scenario for a single host?

Is there a concept of a reservation? Are reservations placed and provisioned manually or automatically? Can a single compromised host allocate all or most of the capacity to an elevated service and then flood using a better-than-default class of service?

7.2 What are the consequences of a compromised router?

What classes of service could be denied across the complete domain (and how is a “domain” defined for the purposes of the particular service definition)? Is the situation

any worse than what can be done in a non-QoS-enabled world by supplying maliciously incorrect routing information?

How is the damage contained? Can a router place reservations with neighboring domains? How much capacity can it allocate?

7.3 If no routers are compromised in a given domain, will it be always able to fulfill its QoS promises?

It is fair to assume that routing is stable because route pinning is a separable problem. However, one still needs to ask whether the service implemented across a domain truly provides a hard guarantee? Do the topology of the domain and the scheduling disciplines employed make the domain vulnerable to carefully timed and synchronized attacks?

7.4 How many neighbors does the adversary need to compromise to deny service to a specific pair of communicating hosts?

Not all services provide hard guarantees. When they do not, it is important to understand what the promise means and under what circumstances it would be broken. Is there a total access capacity that compromised hosts need to have to accomplish DoS? How many compromised neighbors and with what combinations of access link capacities does an adversary need to deny service to a specific pair of communicating hosts (or, for a specific host, to some or all destinations)?

7.5 How can an operator control the possibility of DoS?

Ideally, the choice to provide or not to provide a hard guarantee should be an operator's business decision. On the technical side, the operator would then need the ability to control the degree of feasibility of a successful DoS attack. Can the operator change the percentage of compromised hosts required for an attack by changing its provisioning? Is a hard guarantee ever provided within the framework of changing provisioning parameters? How realistic is it to use hard guarantees with the implementation? For example, what percentage of total network capacity can be used for the elevated priority service?

7.6 Can administrative domains be compartmentalized?

To contain damage from compromised hosts or routers, can a single administrative domain be split into several sub-domains? How will this affect the percentage of total network capacity that can be used for the elevated priority service? How does this affect the complexity of the reservation system?

7.7 How can an operator react to an ongoing attack?

If an ongoing attack is identified, how can its sources be identified? Is the situation better for sources that use an elevated priority service than for sources that use the default service?

Other than by blackholing traffic based on the source and other than by using an *ad hoc* traffic signature identification

mechanism to filter traffic², how can an operator respond? Are there response mechanisms specific to the service in question?

7.8 How is a DoS attack detected?

What, if anything, constitutes a DoS attack? When should the defenses that the provider has be used? How can collateral damage be minimized? Should collateral damage be minimized or used to create peer pressure on the network that is the attacker's haven?

8. SERVICE VERIFICATION

8.1 Service Verification for the Customer

If QoS is insurance against DoS and an elevated priority service might deliver the same or similar average-case performance as the best-effort service while improving the worst-case performance, how can a user verify that the requested elevated priority service is, in fact, delivered?

Consider a typical modern service level agreement (SLA). It usually specifies service parameters such as average loss. These parameters can be computed when a traffic flow is observed and compared to the specified value. If they are within tolerances, the SLA is satisfied.

With DoS protection, however, an assurance that it is "on" in a technical sense cannot be obtained from observation of normal traffic flow, because normal traffic flow is not supposed to be very much different from that of best-effort in the average case.

It would appear that the only way a user could ensure that DoS protection is functioning correctly would be to initiate a DoS attack and verify that it has no effect on the elevated priority service in question. As this is clearly unacceptable, we should think about ways to demonstrate to a user's satisfaction that DoS protection is actually provided. Any sustainable solution to this problem is likely to include an economic element (such as, perhaps, something similar to a "triple your money back if the QoS promise is broken" warranty).

8.2 Service Verification for the Provider

More interestingly, the provider is not in a much better position to make sure that the QoS promise is actually guaranteed to be kept in case of an arbitrary DoS attack. One might argue that the provider might inspect the router configuration and deduce from it the guarantees that are technically offered. For reasons such as mismatches between configuration and actual router behavior, providers would never use this technique to make sure that connectivity is established after a change of configuration. They would exchange actual traffic (perhaps by running *ping* or a more sophisticated network measurement tool).

Configuration inspection alone is never sufficient proof of connectivity. If QoS guarantees are to be taken seriously, they need to be verified with at least the same degree of rigor as basic connectivity. That is, they need to be subjected to active measurement. A guarantee is not tried by running under normal operating conditions (because, under

²These two mechanisms are well-understood and much-used for combating DoS in the best-effort service class today. They are, however, human-labor-intensive, and, if QoS is to be viewed as protection from DoS, should be automated and augmented by additional mechanisms.

normal operating conditions, there might be no difference between a QoS-guaranteed service and best-effort). Like the customer, the provider also needs to observe the service under abnormal conditions, i.e., under attack. Thus, while the provider has better knowledge of the network, it does not necessarily know whether a guarantee is indeed in effect without mounting a network capacity starvation DoS attack on its own network and its peering points (an idea that few providers are likely to embrace).

9. CONCLUSION

Best-effort service on today's fiber-based networks typically provides near-zero loss and near-constant delay determined by speed of light. Well-provisioned networks that deliver good service in the default class have become and (with little dependency on the future price of capacity) are likely to remain the norm. The claim about the future is, naturally, hypothetical and hinges on user demand: once network users become accustomed to good service in the default class, expectations become entrenched. Competition with such a service in the average or typical case is hopeless. In fact, QoS might be at a technical disadvantage with respect to best-effort service because of its extra complexity; the additional mechanisms (such as policers) might actually increase the typical-case probability of failure for QoS *vs.* best-effort.

QoS can, however, improve the worst case. Worst case is best thought of in adversarial terms: *What could happen if a determined attacker wanted to deny service to a particular host or pair of hosts?*

Network research should devote more attention to the properties of QoS that allow for protection from DoS. A good goal would be an elevated service that could use a sizable fraction of network capacity and offer hard guarantees for admitted reservations in the case when no routers inside a domain are compromised.

We conclude with a question that we think needs to be answered by further research: *If QoS is insurance against DoS, how, short of mounting test DoS attacks, does one verify a guarantee?*

10. REFERENCES

- [1] *Abilene measurements using OWAMP*. <http://owamp.internet2.edu/abilene/>.
- [2] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski. Assured Forwarding PHB Group. RFC 2597, June 1999.
- [3] P. Hurley, J.-Y. Le Boudec, P. Thiran, and M. Kara. ABE: Providing a Low-Delay Service within Best Effort. *IEEE Network Magazine*, v 15, no 3, May/June 2001.
- [4] V. Jacobson, K. Nichols, and K. Poduri. An Expedited Forwarding PHB. RFC 2598, June 1999.
- [5] D. Newman. Core competency: ISP backbones stand up in grueling 30-day performance test. *Network World*, Dec 16, 2002, <http://www.nwfusion.com/research/2002/1216isptest.html>.
- [6] S. Shalunov and B. Teitelbaum. QBone Scavenger Service (QBSS) Definition. *Internet2 Technical Report*, Proposed Service Definition, Internet2 QoS Working Group Document, March 2001, <http://qbone.internet2.edu/qbss/>.
- [7] S. Staniford, V. Paxson, and N. Weaver. How to Own the Internet in Your Spare Time. *Proceedings of the 11th USENIX Security Symposium*.