POPI: Inferring Router User-Level Tool for Packet forwarding priority

V.NARASIMHAM[#1], Y.M.S.D SASTRY[#2]

(#1 M.Tech Student, B.V.C. Engineering College, Odalarevu)

(#2 Associate Professor, Department of CSE, BVCEC, Odalarevu)

Abstract

Packet forwarding prioritization (PFP) in routers is one of the mechanisms commonly available to network operators. PFP can have a significant impact on the accuracy of network measurements, the performance of applications and the effectiveness of network troubleshooting procedures. Despite its potential impacts, no information on PFP settings is readily available to end users. In this paper, we present an end-to-end approach for PFP inference and its associated tool, POPI. This is the first attempt to infer router packet forwarding priority through end-to-end measurement. POPI enables users to discover such network policies through measurements of packet losses of different packet types. We evaluated our approach via statistical analysis, simulation and wide-area experimentation in PlanetLab. We employed POPI to analyze 156 paths among 162 PlanetLab sites. POPI flagged 15 paths with multiple priorities, 13 of which were further validated through hop-by-hop loss rates measurements. In addition, we surveyed all related network operators and received responses for about half of them all confirming our inferences. Besides, we compared POPI with the inference mechanisms through other metrics such as packet reordering [called out-of-order (OOO)]. OOO is unable to find many priority paths such as those implemented via traffic policing. On the other hand, interestingly, we found it can detect existence of the mechanisms which induce delay differences among packet types such as slow processing path in the router and port-based load sharing.

Index Terms—Network inference, network neutrality, packet forwarding priority

---

Introduction

Packet forwarding prioritization has been available in off-the- shelf routers for quite a while, and various models from popular brands, such as Cisco and Juniper Networks offer support for it. Network operators have come to rely on these mechanisms for managing their networks, for example as a way of rate limiting certain classes of applications.PFP can have a significant impact on the performance of ap- plications, on the accuracy of measurement tools' output, and on the effectiveness of network troubleshooting procedures. There are a couple of challenges for designing and implementing POPI. First, background traffic fluctuations canseverely affect the end-to-end inference accuracy of router properties. Second, probe traffic of a relatively large packet bursts are neither independent nor strong correlated. Most existing inference methods have to assume certain independence (i.e., i.i.d. processes) or strong correlation models for inference (e.g., back-to-back probe packets). However, as for the relatively large packet bursts sent by POPI, a good mathematical model is

needed to determine whether the loss rates difference between two packet types is the consequence of a random effect or being treated really differently. Third, we want to measure more than two packet types at the same time, so simply determining whether they are treated differently is not enough. To overcome these challenges, POPI takes the following three steps to infer packet forwarding priority inference. First, it sends a relatively large amount of traffic to temporarily saturate the bottleneck traffic class capacity, which gives POPI better resistance against background traffic fluctuations. Second, we apply a robust nonparametric method based on the ranks instead of pure loss rates. Thirdly, we assign a rank-based metric to each packet type and use a hierarchical clustering method to group them when there are more than two packet types.

Related Work

To the best of our knowledge, this is the first attempt to infer router packet-forwarding priority through end-to-end measurement. Perhaps the efforts most closely related to this work are those identifying shared congestion . Such efforts try to deter- mine whether two congested flows are correlated and share a common congested queue along their paths. If we consider the flows of different packet types along a same path, our problem becomes to identify whether these flows do not share a common congested queue. While both problems are related clearly, we usually need to simultaneously consider a much larger number of packet types (e.g., 26 packet types in the Planet Lab experiment).Note that the correlation based method used for shared congestion identification methods requires back-to-back probing which, in our case, translates into $O(n^2)$ pairs probing for packet types. In addition, those efforts focused on flows which experience congestion (ignoring uncongested ones), so their probe traffic rate is low and not busty. To identify packet forwarding prioritization in routers, one must send relatively large amounts of traffic to temporarily force packet drops (by saturating the link). Thus, for better scalability and accuracy, our problem requires different measurement and statistical interference methods. Kuzmanovic and Knightly proposed a framework for enabling network clients to measure a system's multiclass mechanisms and parameters. The basic idea is similar to ours, i.e., to inject multiclass traffic into the system and use a statistical method to infer its scheduling types and parameters based on the output. However, the technique did not consider cross-traffic effects and only simulation results were presented. PFP inference also has some goals in common with efforts on network tomography. However, unlike in network tomography where loss information and topology information are combined to infer link losses, we look to identify if different packet types (based on protocol or port numbers) experience different loss rates. In addition, while probes used for network tomography are always nonintrusive in order to get accurate link loss/delay, our problem requires that we saturate links in order to uncover the configuration of the routers.

Inferring Packet-forwarding Priority

Background on Priority Mechanisms:

Network administrators can enforce priority/link-sharing mechanisms in a router by defining a traffic class (usually IP protocol and TCP/UDP port number) and associating with it a particular queuing/scheduling mechanism. Some of the commonly available mechanisms are as follows.

- Priority Queuing (PQ). This allows users to assign arbitrarily defined packet classes to queues with different priorities. Since queues are served based on their priority, this allows specified packet types to be always sent before other packet types.
- Proportional Share Scheduling (PSS). With PSS each traffic class is given a weight. Bandwidth is allocated to classes in proportion to their respective weights. There is no strict priority difference between classes. There are dif- ferment ways to implement this scheduling mechanism, e.g., Weighted Fair Queuing (WFQ), Weighted Round-Robin(WRR).In Cisco routers, the CBWFQ is Class-Based WFQ and the Custom Queuing is WRR based.
- Policing. This restricts the maximum rate of a traffic class. Traffic that exceeds the rate parameters is usually dropped. The traffic class cannot borrow unused bandwidth from others.

Only the first mechanism sets absolute priorities between traffic classes. There is no absolute priority difference between the other two classes, and the loss experienced by one class depends on whether its traffic rate exceeds its allocated band- width.

Probing the Path:

Illustrates our link probe method. We want to test k packet types. POPI sends several bursts() $n_r \times$ from a source to a destination. The interval between bursts is Δ. Each burst consists of rounds, in which k packets, one for each packet type studied, are interleaved in random order. So, there are k back-to-back packets in each burst. There are three parameters for the probe method, Δ,and In order to achieve independence between bursts, i.e., to ensure the router's queuing busy period caused by one burst does not interfere with the following one, Δ should not be too small. On the other hand, in order not to experience large background traffic fluctuation duration the probe, we need to keep the whole probe duration within a real- timely short period. In practice, Δ is set to one tote seconds to keep overall probe duration within several minutes.

Deriving Ranks:

For every burst, loss rate ranks are computed by first sorting packet types in ascending order according to their packet loss rates in that burst and then assigning ranks in order, i.e., the packet type with the largest loss rate has rank 1, the one with the second largest loss rate has rank 2 and etc.1Similar to packet loss rates, due to randomness of packet losses, the ranks of different packet types are like random ar- arrangements over the all bursts when the packet types are treated equally. On the other hand, the ranks of certain packet types are always small when they are treated with low priority.2 However, the advantage of using ranks is that we have a theory to bound the variance of loss ranks caused by the random effects whereas we do not have that bound for loss rates when the loss model is unknown.

PlanetLab Experiments

POPI and Its Two Probe Modes

POPI works in two probe modes, End-to-End Probe (EEP) and Hop-by-Hop Probe (HHP). In both modes, the sender sends multiple packet types toward the receiver. The receiver feed- backs certain information of every received packet to the sender, which is used by the sender

to measure the end-to-end losses and reordering events along the path.

HHP mode is used to locate the configured router or device by measuring the losses and reordering events to every router on the

PROT	Type/Port Number
ICMP	ICMP_ECHO
TCP	20, 21, 23, 110, 179, 443 (well-known app) 1214, 4661, 4662, 4663, 6346, 6347, 6881 (P2P applications) 161, 1000, 12432, 38523, 57845 (random)
UDP	110, 179, 161 (SNMP) 1000, 12432, 38523, 57845 (random)

PACKET TYPES CONSIDERED FOR PLANETLAB EXPERIMENTS

In each cycle, it sends bursts whose TTLs increase from one to, which is the total hop count from sender to receiver. For, POPI measures the end-to-end packet losses and reorderings based on the feedback from the receiver. For, POPI calculates the losses and reorderings up to a hop by counting time-exceeded ICMP responses from that router. When packets do not traverse the configured box, we will not observe packet loss or reordering difference. After packets traverse the box, the loss or reordering difference will be similar to that observed at the receiver, and will exhibit over the remaining hops. Once we observe such phenomenon, the configured box should be around the spot of difference, the hop at which the difference begins to show.

To note, the losses and reorderings of ICMP responses actu- ally include round-trip effects. However, as the response packets were all "ICMP time exceeded" packets of a same packet size,

it is very unlikely that any router on the reverse path is going to treat them differently. Hence, even when there are losses or reorderings on the reverse link, the effects are unlikely to introduce bias against a specific packet type.

While it may seem necessary to test all packet types of different protocol/port number combinations to validate our approach, in practice there is only a small number of packet types that network administrators may want to treat differently. We selected 26 packet types as listed in Table IV. For UDP and TCP packets, 30002 is used as the destination port, because it is very unlikely that ISPs will set an explicit priority policy based on it. The port numbers listed in Table IV are used as source ports to measure the source port based priority policy. (Destination port based policy can be measured in a similar manner.) These packet types are selected to check:

- Whether ICMP, TCP and UDP packets are handled with equal priority.
- Whether some well-known applications are granted higher priority. This set includes ftp (port 20, 21), telnet (port 23), POP3 (port 110), BGP (port 179), and HTTPS (port 443). Port 80 is not included because it is used by PlanetLab maintenance.
- Whether P2P traffic is treated with lower priority. The seven ports tested are used by four major P2P applications, Fasttrack, eDonkey, Gnutella, and BitTorrent.

Comparing Priority Inference with Different metrics

In this section, we compare the inference results of the loss-based and OOO-based method using PlanetLab experiments.

We do not use packet delay because the reordering metric is more robust than the delay metric although they both reflect the packet delay differences. When the delay variation generated by the non configured devices is large, a packet with a shorter delay at the configured box can have a larger end-to-end delay than
a packet with a larger delay at the configured box. Hence, the delay differences between different packet types introduced by the configured box are overwhelmed by the large delay variation introduced by the non configured devices along the path. Large delay variation can often be observed for congested routers. However, routers usually do no reorder packets. Hence, the reordering events introduced by the configured box are usually observed by the receiver without any distortion.
OOO-based method is generally more accurate than the delay-based method.

	Total	P2P	ICMP	T179	Load Share	Unknown
# OMGP	56	3	15	11	17	12
# HHP	36	2	12	10	12	0
Validation	11/0	2/0	4/0	4/0	1/0+8	0
# LMGP	19	5	6	3	3	4
# HHP	10	4	4	2	0	0
Validation	7/0	4/0	1/0	2/0	0	0
# Overlap	7	0	1	2	3	1

GROUP PATTERNS GIVEN BY THE OOO-BASED AND LOSS-BASED METHODS FOR . THE VALIDATION RESULT IS (THE NUMBER OF POSITIVE CONFIRMED PATHS)/(THE NUMBER OF NEGATIVE CONFIRMED PATHS). FOR LOAD SHARING PATTERN, PLEASE REFER TO THE RELATED TEXT FOR DETAILS

Results

Table above shows the number of Multi-Group Paths (MGPs), the number of paths that had their spots of difference identified by HHP and the number of validated paths. We sent more than 20 e-mails and got 10 replies. All of them are positive confirmations. To note, one reply can confirm several paths, e.g., two unidirectional paths between a pair of nodes. Therefore, the number of validated paths is slightly larger than the number of replies received.

There are 56 OMGPs and only 19 LMGPs in $\mathcal{N}_2$. The overlap between the OMGPs and LMGPs is small, i.e., only seven paths are both flagged as OMGP and LMGP. In $\mathcal{N}_1$, P2P and ICMP are two main priority patterns configured by the ISPs. However, only one of 11 LMGPs of these two patterns in $\mathcal{N}_2$ is identified by the OOO-based method. On the other hand, OMGPs mainly concentrate on the ICMP, T179 and load sharing pat- terns whereas very few LMGPs show the latter two patterns. Such findings substantiate our analysis in Section III-B that the OOO-based method may fail to discover many multipriority paths but to flag many paths caused by the mechanisms other than QoS.

CONCLUSION

In this paper, we have demonstrated that POPI, an end-to-end priority inference tool, is able to accurately infer the router's packet forwarding priority. The contributions of this work are the findings over Internet as well as the methodology.

REFERENCES

[1] "Cisco Ios Quality of Service Solutions Configuration Guide Release12.2," Cisco Systems [Online]. Available: http://www.cisco.com/ univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/

[2] "Filter-based forwarding," Juniper Networks, 2001 [Online]. Avail- able: http://www.juniper.net P. Grant and J. Drucker, "Phone, cable firms rein in consumers'

[3] Internet use," 2005 [Online]. Available: http://online.wsj.com/ar- ticle/SB112985651806475197.html

[4] J. Cheng, "Evidence mounts that Comcast is targeting Bittorrent traffic," 2007 [Online]. Available: http://arstechnica.com/news.ars/post/20071019-evidence-mounts-that-comcast-is-targeting-bittor- rent-traffic.html

[5] V. Kumar, "Comcast, Bittorrent to work together on network traffic," 2008 [Online]. Available: http://online.wsj.com/article/SB120658178504567453.html

[6] G. Lu, Y. Chen, S. Birrer, F. E. Bustamante, C. Y. Cheung, and X. Li,"End-to-end inference of router packet forwarding priority," in Proc. IEEE INFOCOM, 2007, pp. 1784–1792.

[7] K. Harfoush, A. Bestavros, and J. Byers, "Robust identification of shared losses using end-to-end unicast probes," in Proc. IEEE ICNP,2000, pp. 22–36.

[8] D. Rubenstein, J. Kurose, and D. Towsley, "Detecting shared conges- tion of flows via end-to-end measurement," IEEE/ACM Trans. Netw., vol. 10, no. 3, pp. 381–395, Jun. 2002.