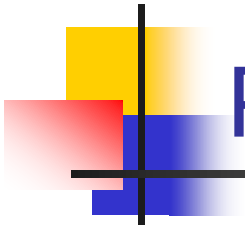
Collection and Characterization of BCNET BGP Traffic

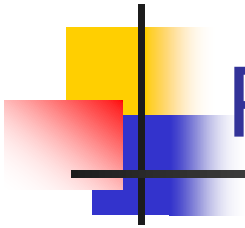
Sukhchandan Lally
lally@sfu.ca

Communication Networks Laboratory
http://www.ensc.sfu.ca/research/cnl
School of Engineering Science
Simon Fraser University



Roadmap

- Introduction
- BCNET packet capture
- Wireshark
- BGP attributes
- BGP update attributes
- Conclusions, future work, references

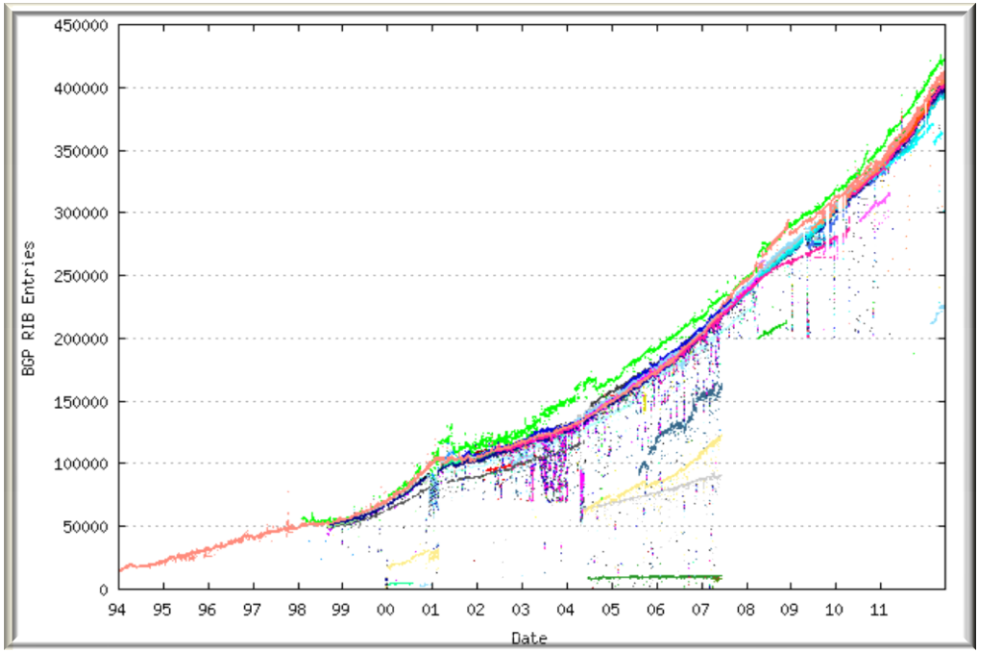


Roadmap

- Introduction
- BCNET packet capture
- Wireshark
- BGP attributes
- BGP update attributes
- Conclusions, future work, references

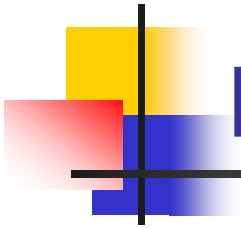
Motivation

- The size of BGP tables has exponentially increased since 1994, implying that timely analysis of BGP is important



Growth of the BGP table – 1994 to Present

http://www.potaroo.net/bgp/

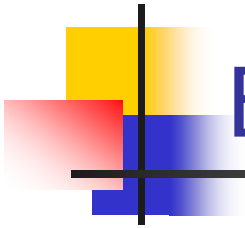


Border Gateway Protocol (BGP)

- BGP is de facto Inter-Autonomous System (AS) routing protocol
- Operates over the Transmission Control Protocol (TCP)
- Exchanges network reachability information among BGP systems
- Distributes route path information to peers
- Sends update messages as routing tables change
- Supports Classless Inter-Domain Routing (CIDR)

Project outline

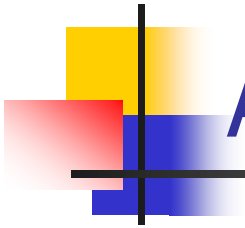
- We used a C# code to preprocess the readable MRT files
- We extracted update message attributes from BGP traffic and used MATLAB to generate the graphs
- The parser extracted the needed features from the BGP update messages received by a router from its peers
- We chose three dates in October, November, and December 2011 and compared different attributes to see if BGP data performs differently on these dates



BGP

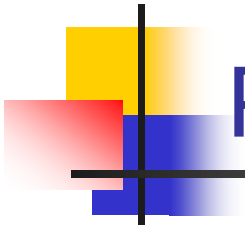
BGP routers exchange routing information using four types of messages:

- Open
- Update
- Notification
- Keepalive



Autonomous System (AS)

- AS is a group of networks sharing the same routing policy
- It is identified with Autonomous System Numbers (ASN)
- ASN assigned by Internet Assigned Numbers Authority (IANA)
- AS domain is represented as a node at inter-domain level or AS level

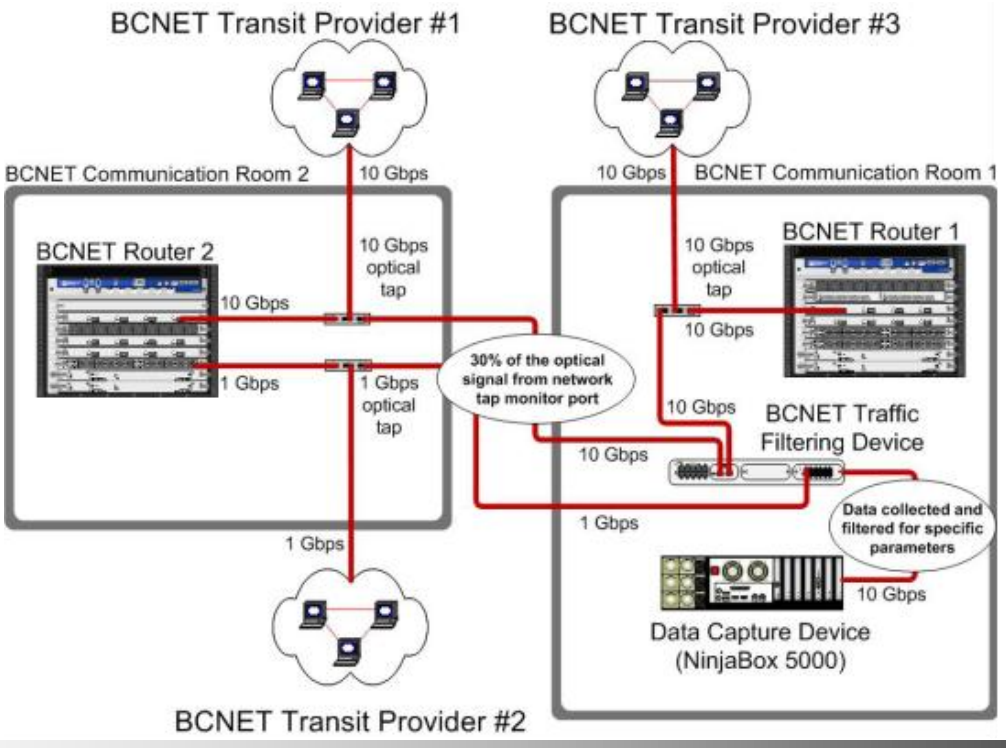


Roadmap

- Introduction
- BCNET packet capture
- Wireshark
- BGP attributes
- BGP update attributes
- Conclusions, future work, references

BCNET packet capture: physical overview

- BCNET is the hub of advanced telecommunication network in British Columbia, Canada that offers services to research and higher education institutions



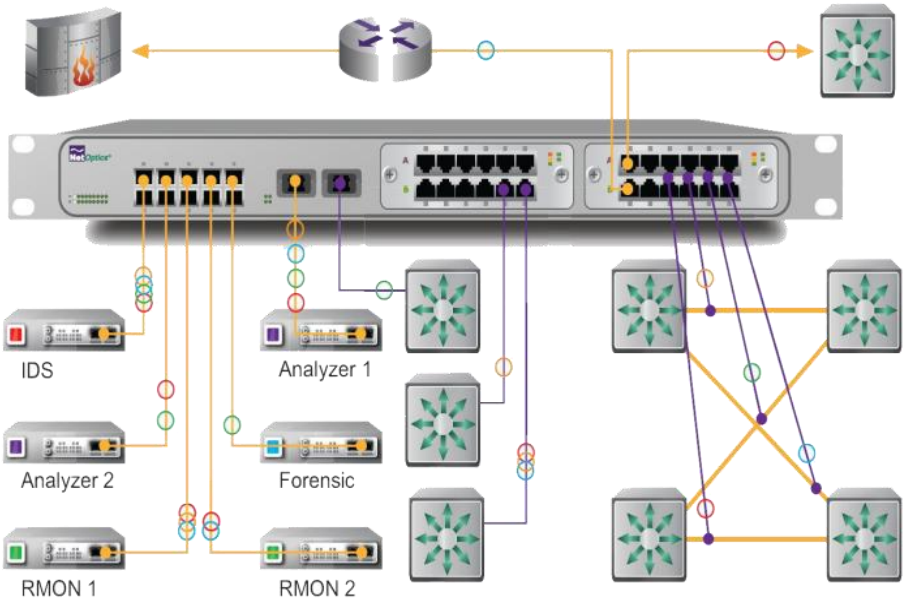
BCNET packet capture

- BCNET transits have two service providers with 10 Gbps network links and one service provider with 1 Gbps network link
- Optical Test Access Point (TAP) splits the signal into two distinct paths and the signal splitting ratio from TAP may be modified
- The Data Capture Device (NinjaBox 5000) collects the real-time data (packets) from the traffic filtering device

S. Lally, T. Farah, R. Gill, R. Paul, N. Al-Rousan, and Lj. Trajkovic, "Collection and characterization of BCNET BGP traffic," in Proc. 2011 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Victoria, BC, Canada, Aug. 2011, pp. 830–835.

Net Optics Director 7400

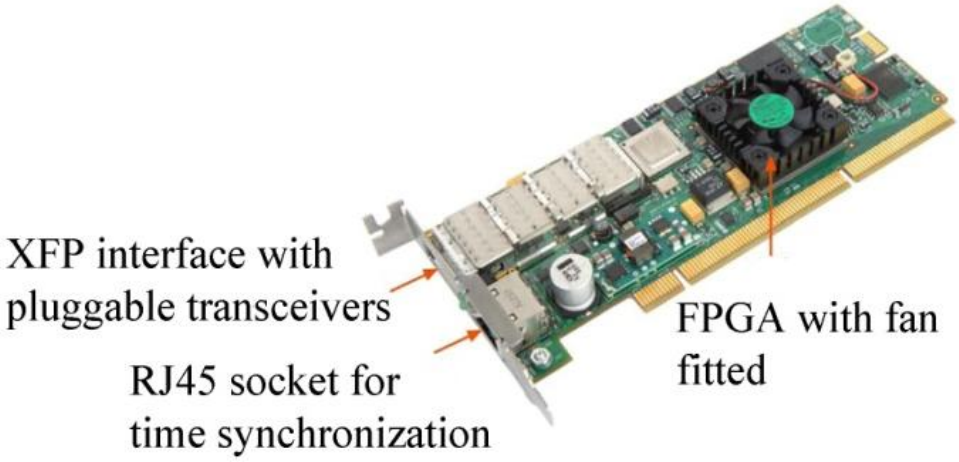
- Used for BCNET traffic filtering
- Directs traffic to monitoring tools such as NinjaBox 5000 and FlowMon



RMON: Remote Network Monitoring

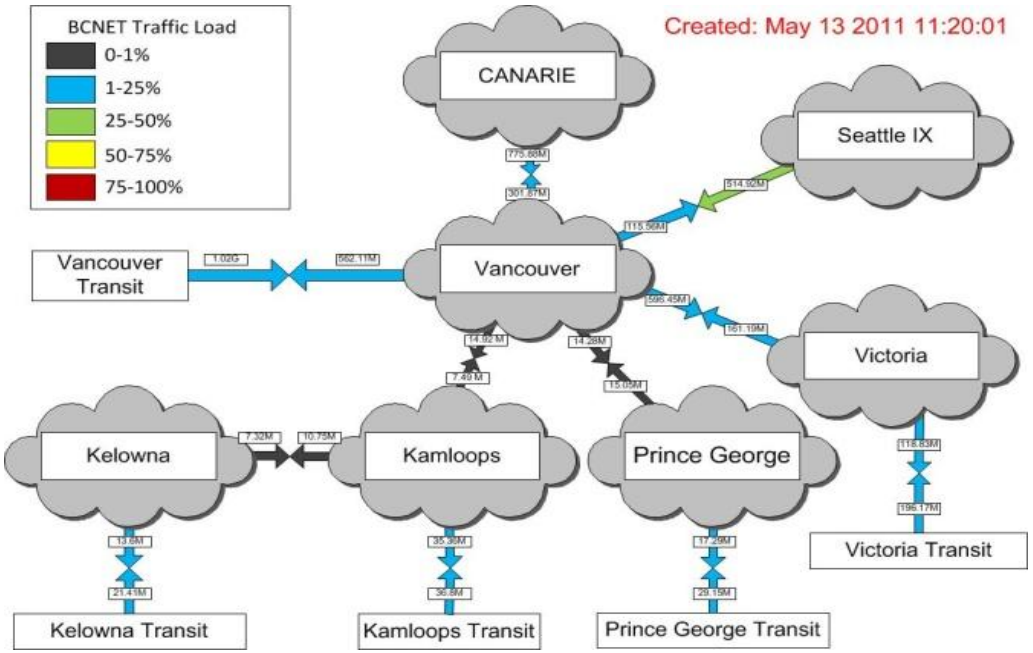
Endace DAG

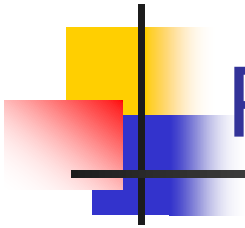
- Endace Data Acquisition and Generation (DAG) 5.2X card resides inside the NinjaBox 5000
- Captures and transmits traffic and has time-stamping capability
- DAG 5.2X is a single port Peripheral Component Interconnect Extended (PCIx) card
- Capable of capturing on average Ethernet traffic of 6.9 Gbps



Real time network usage by BCNET members

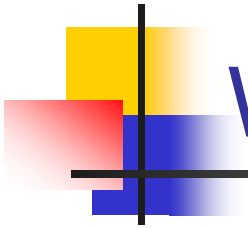
- British Columbia's network extends to 1,400 kilometres and connects Kamloops, Kelowna, Prince George, Vancouver, and Victoria





Roadmap

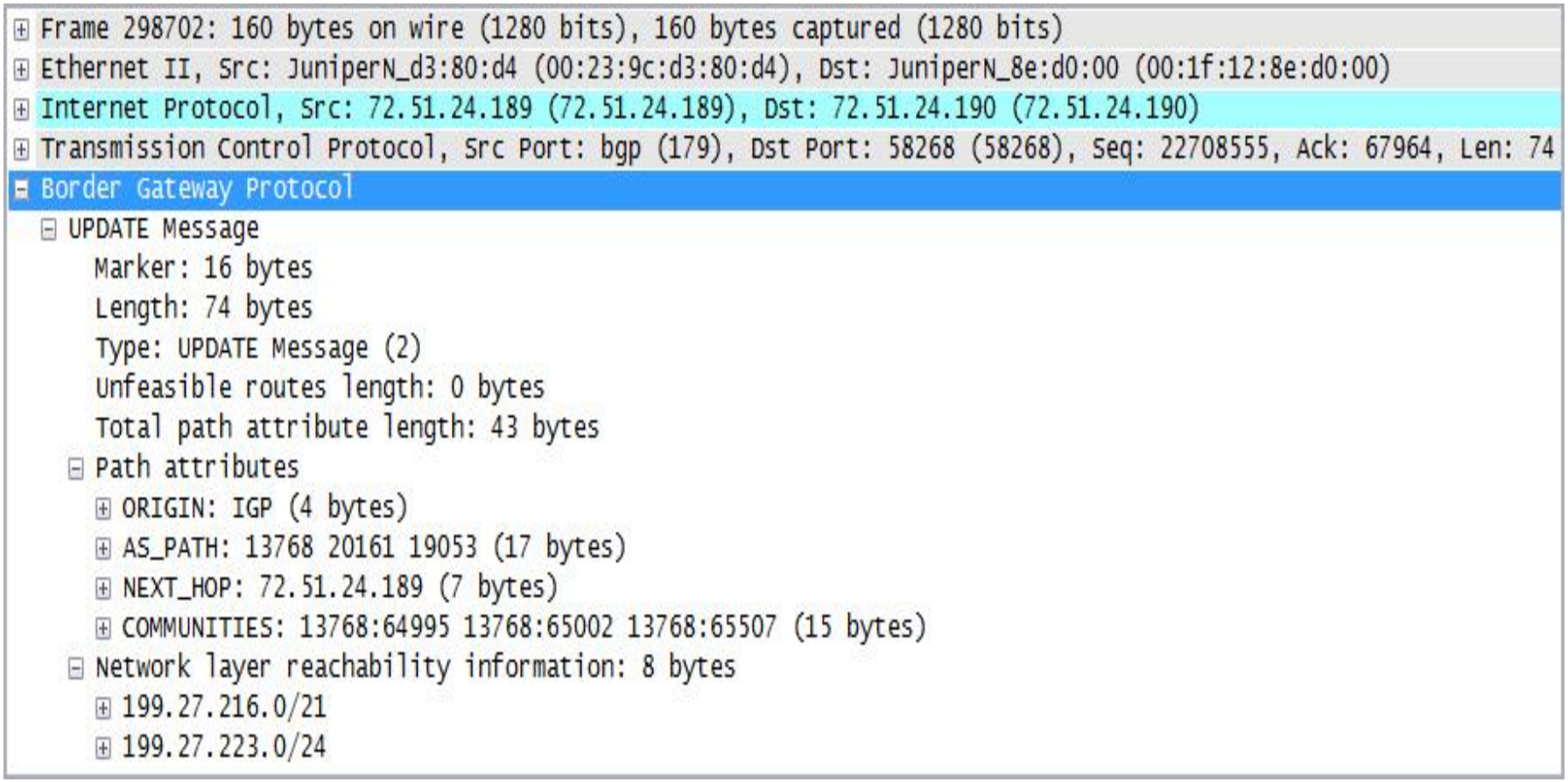
- Introduction
- BCNET packet capture
- Wireshark
- BGP attributes
- BGP update attributes
- Conclusions, future work, references



Wireshark

- Wireshark is an open source packet analyzer
- Provides comprehensive statistics such as the summary of traffic collected, input/output graphs, protocol hierarchy, and endpoints
- Opens and saves captured packet data, imports and exports packet data from and to other capture programs
- Captures network packet data from a network interface and displays those packets with detailed protocol information

Wireshark view of the collected traffic



BCNET traffic protocol hierarchy

- Each protocol has its statistical value (row) consisting of protocol's name, the percentage of protocol packets relative to total number of packets captured, number of packets, and number of bytes
- From 511,820 packets:
- 260,639 (50.9%) are BGP packets,
- 257,285 (50.3%) are TCP ACK packets
- 6,104 (1.2%) are piggyback ACKs

Protocol hierarchy of the captured packets

Protocols	Packets %	Packets	Bytes
Ethernet/IP/TCP	100	511,820	98,292,937
BGP	50.92	260,639	79,628,747

BCNET Network Endpoints

- Network endpoints are the source and destination addresses of a specified protocol layer
- Endpoints of the six BCNET transit exchanges (BGP peers) were captured
- There are various TCP connection statistics for each IP address of a BGP peer

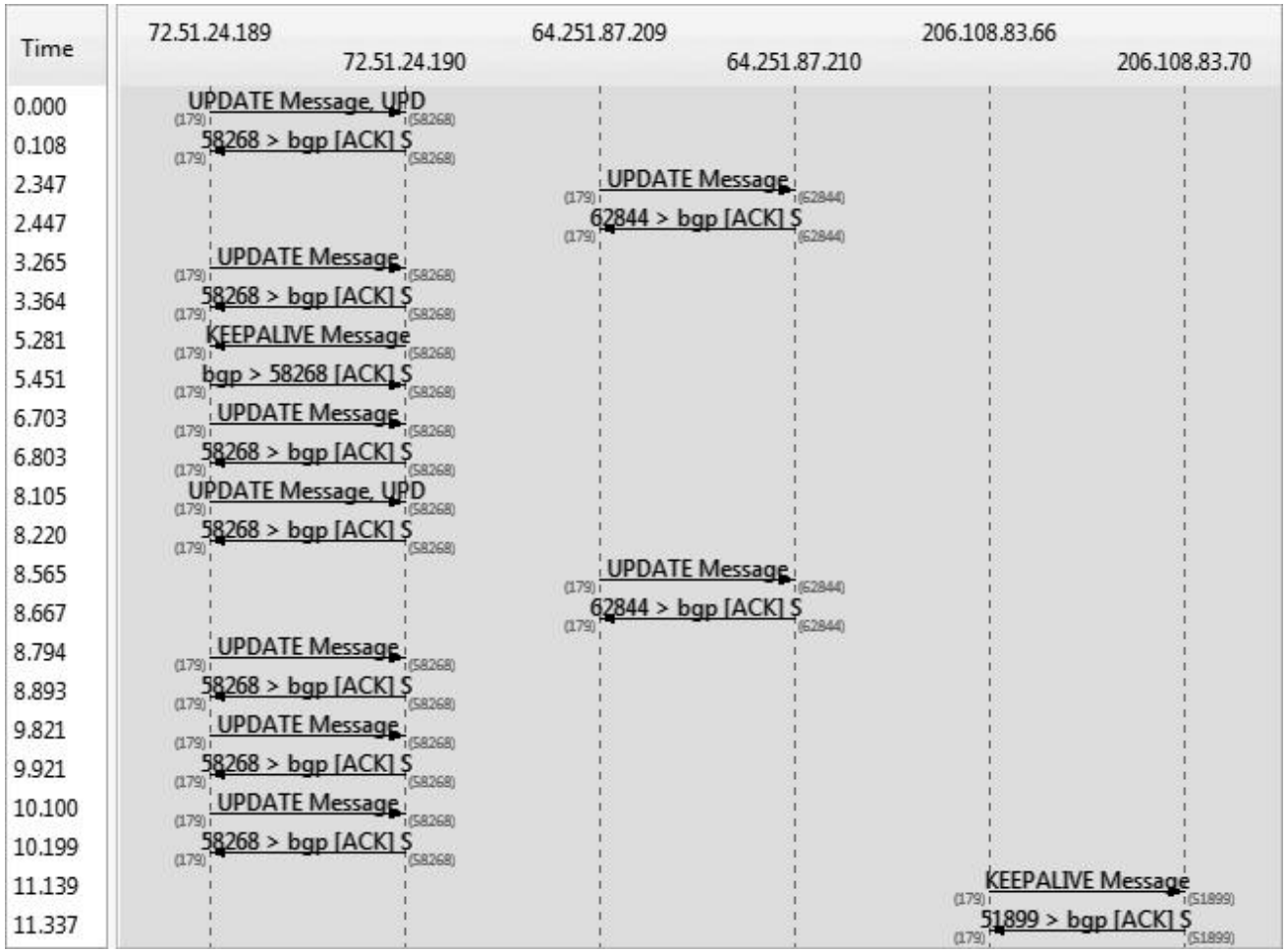
Network Endpoints

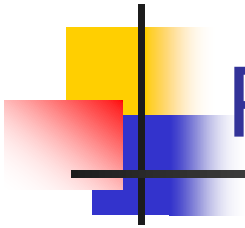
Address	Port	Packets	Bytes	Tx Bytes	Rx Bytes
72.51.24.189	bgp	401721	70836354	55894998	14941356
72.51.24.190	58268	401721	70836354	14941356	55894998
64.251.87.209	bgp	70069	14996289	12426684	2569605
64.251.87.210	62844	70069	14996289	2569605	12426684
206.108.83.66	bgp	40030	12460294	1500045	10960249
206.108.83.70	51899	40030	12460294	10960249	1500045

BCNET Traffic Service Response Time

- Time between a request and the corresponding response
- The flowgraph of the BGP peers includes the source address, destination address, TCP port number, TCP message (ACK), and type of the BGP message (open, update, notification, keepalive)

Flow graph of the BGP peers





Roadmap

- Introduction
- BCNET packet capture
- Wireshark
- BGP attributes
- BGP update attributes
- Conclusions, future work, references

BGP AS path

- AS path is a sequence of intermediate ASs between source and destination
- Neighboring ASes use BGP to exchange update messages and to reach different AS prefixes
- Three ASes were recognised in the BCNET traffic collection:
- AS 852 (Telus Advanced Communications)
- AS 6327 (Shaw Cable systems) and
- AS 13768 (Peer 1 Networks Inc.)

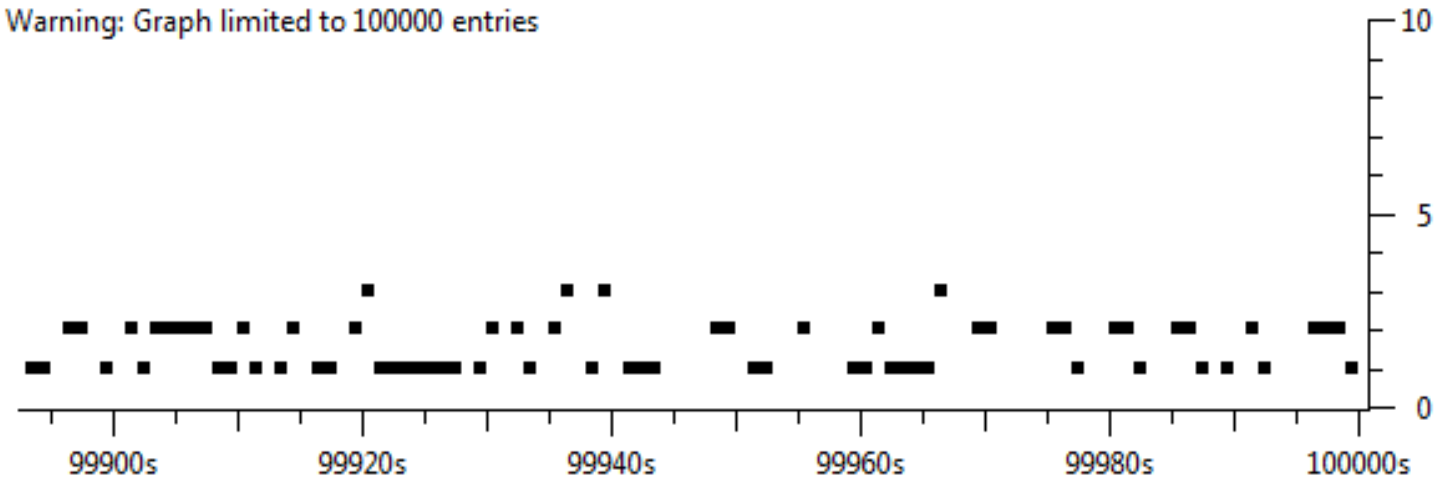
Number of connections for AS 6327

AS 6327 has 683 connections with other ASs

http://www.caida.org/home/

Origin attribute

- Origin attribute is set when the route is first introduced into the BGP
- It may have three values: IGP, EGP, or Incomplete
- The total number of packets in origin attribute are 245,168 and 210,414 of these are IGP packets

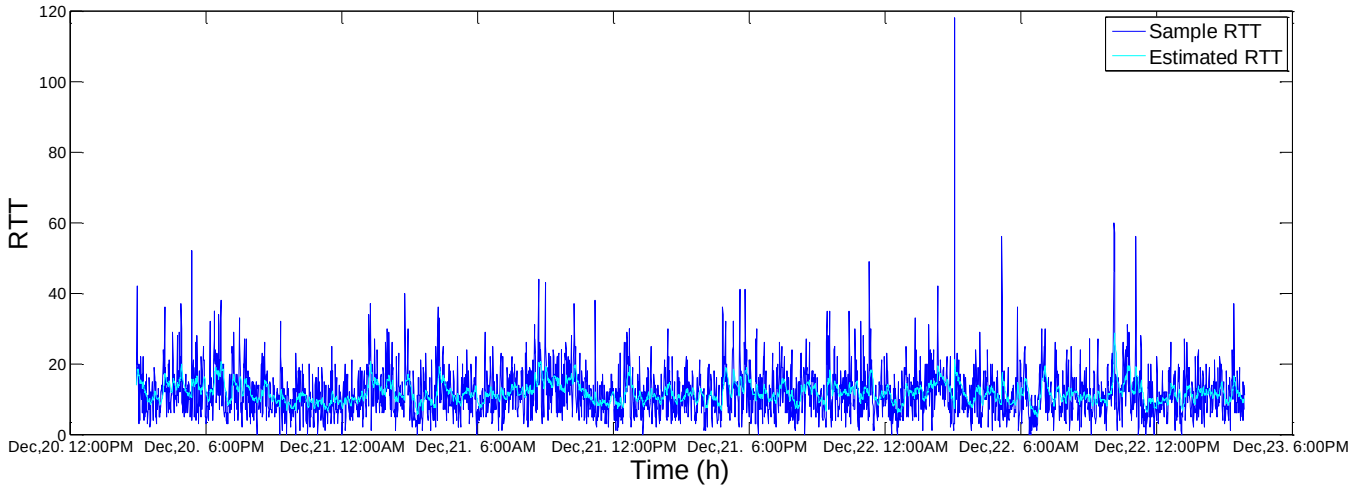


BGP message attribute

- Update messages transfer routing information between BGP peers
- Information in the update packet may be used to construct a graph describing the relationships of the various ASes
- A keepalive is a message sent by one device to another to confirm that the link between the two is operating and to prevent the link from being broken
- There were 230,424 update packets and 30,462 keepalive packets identified in the BCNET traffic

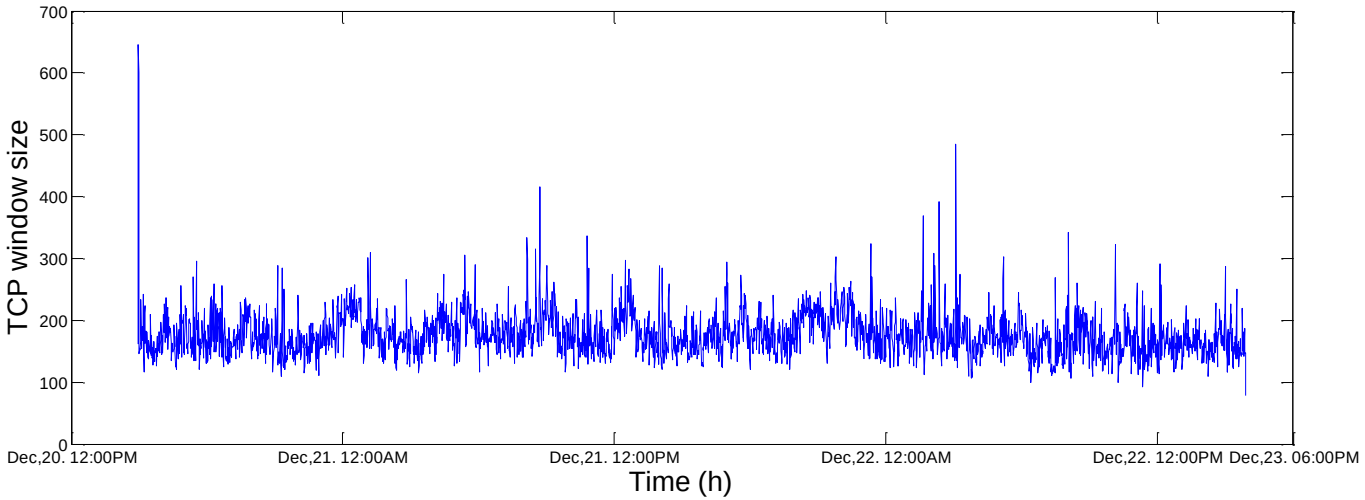
TCP round trip time

- RTT is the time measured from segment transmission until ACK is received
- The RTT average is approximately 11.7 ms
- The RTT standard deviation is 7.19 ms for sample RTT and 2.75 ms for the estimated RTT



TCP window size

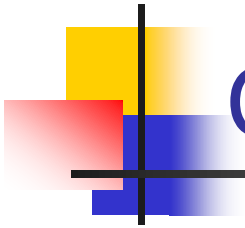
- The amount of data that a receiver may accept without acknowledging the sender
- TCP transmits data up to the window size before waiting for the acknowledgements
- The window size is for 200 samples of the data



Anomalies

- According to another study done on classification of BGP anomalies, 65% of the selected features were volume features
- The volume features are more relevant to the anomaly class than the AS path features
- Anomalies may be either of the two types: path anomalies or announcement anomalies.
- Statistical and machine learning techniques may be used to classify and detect BGP anomalies

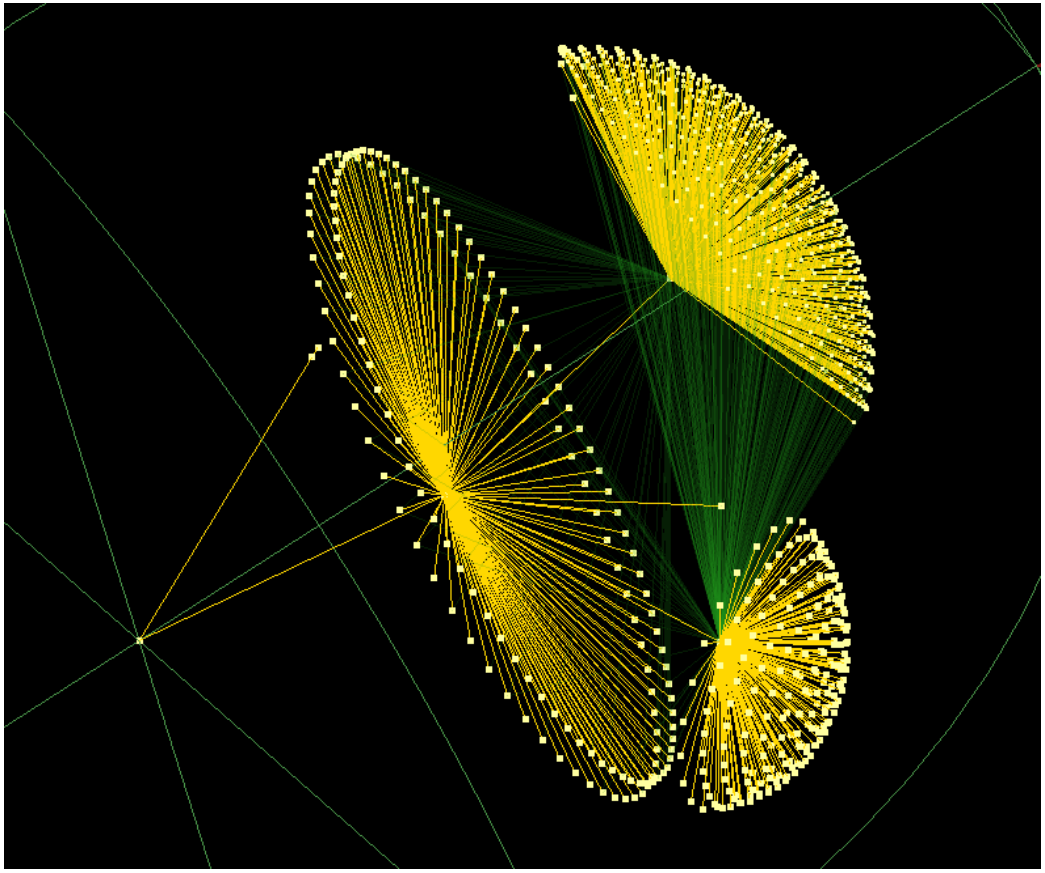
N. Al-Rousan and Lj. Trajković, "Comparison of machine learning models for classification of BGP anomalies," in Proc. HPSR 2012, Belgrade, Serbia, June 2012.



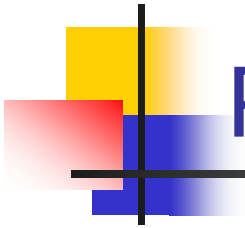
Clusters

- The centers of the three clusters correspond to BCNET transit providers
- The graph consists of 982 nodes, 981 tree-links, and 441 non tree-links
- Clusters consist of 683, 588, and 155 AS nodes
- They are created using the value of the BGP AS path attribute in BGP update messages

Walrus AS topology graph of the collected BCNET traffic



T. Farah, S. Lally, R. Gill, N. Al-Rousan, R. Paul, D. Xu, and Lj. Trajkovic, "Collection of BCNET BGP traffic," in Proc. 23rd International Teletraffic Congress, San Francisco, CA, USA, Sept. 2011, pp. 322–323.



Roadmap

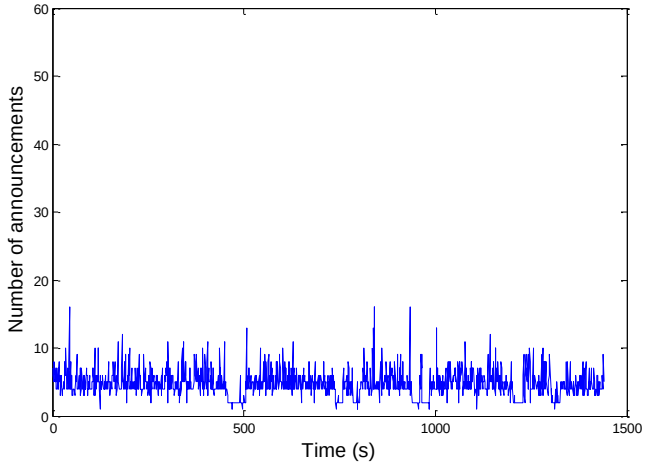
- Introduction
- BCNET packet capture
- Wireshark
- BGP attributes
- BGP update attributes
- Conclusions, future work, references

BGP update attributes

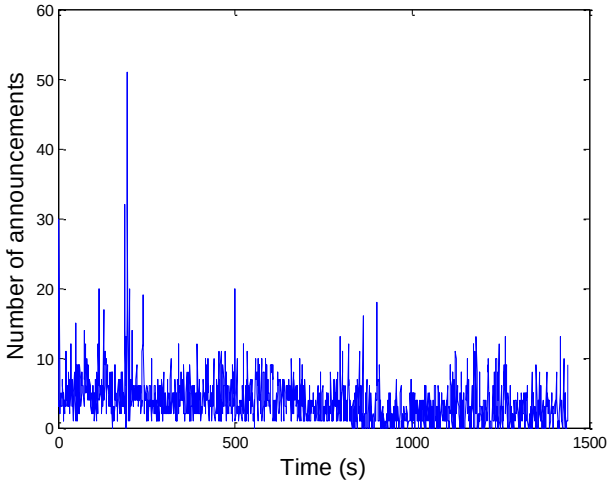
- Update messages can be either announcement or withdrawal
- The number of announcements and withdrawals exchanged by neighboring peers are an important feature during instability periods
- The features are categorized as volume (number of BGP announcements) and AS path (maximum edit distance) features

Number of announcements

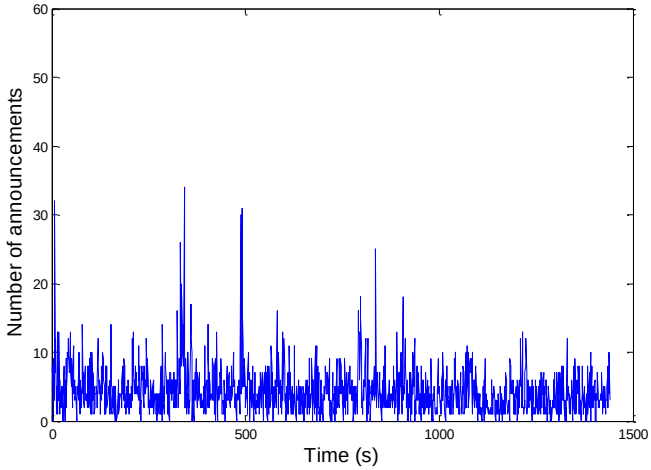
- The number of routes that are available for delivery of data from source to the destination
- They require a set of attributes to be described



October 2, 2011



November 2, 2011



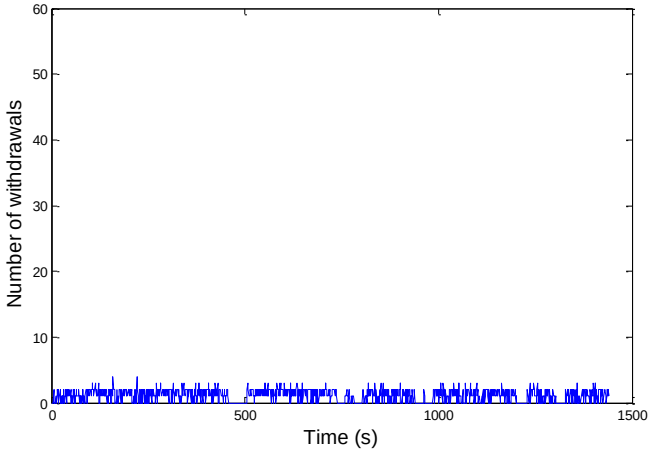
December 2, 2011

Number of withdrawals

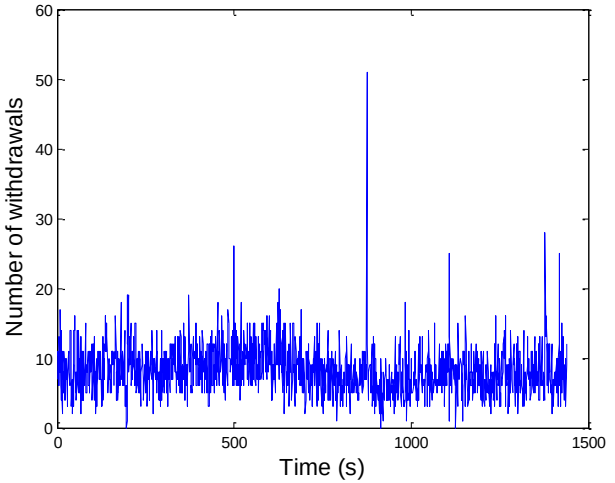
- A previously announced message that has become unreachable
- BGP withdrawal requires simply the address of the network for which the route is being removed
- An update message may advertise only one route but several may be withdrawn
- One BGP packet may contain more than one announced or withdrawn NLRI prefix

NLRI: Network Layer Reachability Information

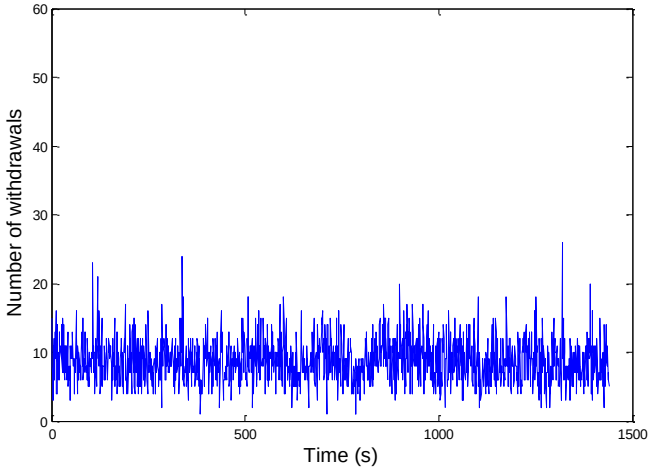
Number of withdrawals



October 2, 2011



November 2, 2011

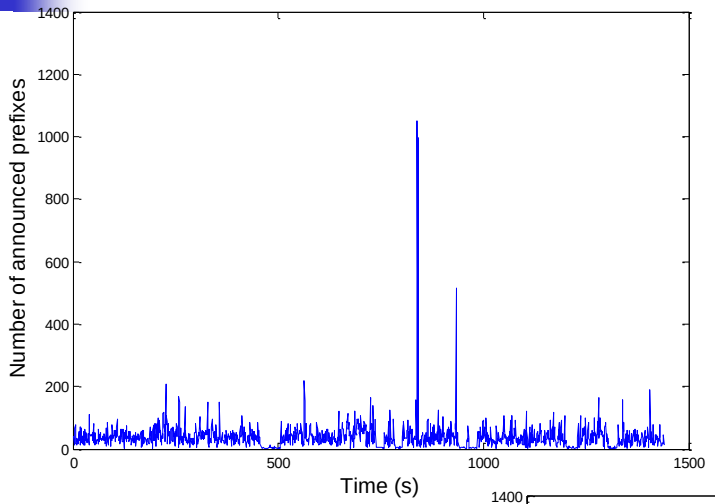


December 2, 2011

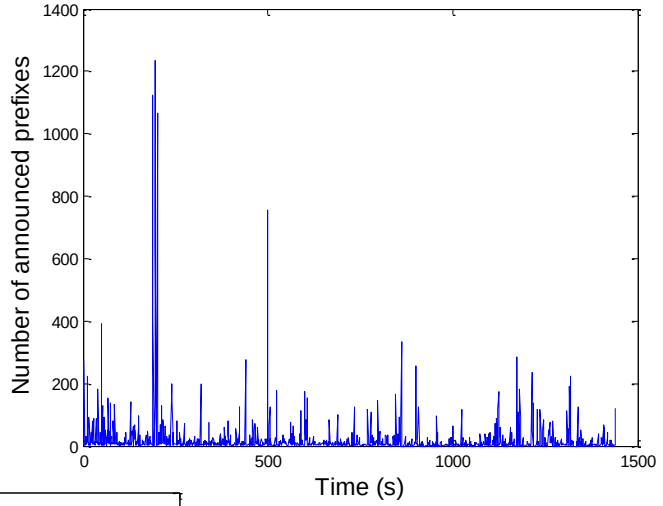
Number of announced prefixes

- BGP announces an IP prefix if a matching entry is found in the IP routing table
- Number of announced prefix represents the number of update NLRIs
- To advertise a classless prefix, the prefix and the mask in the BGP routing process needs to be configured

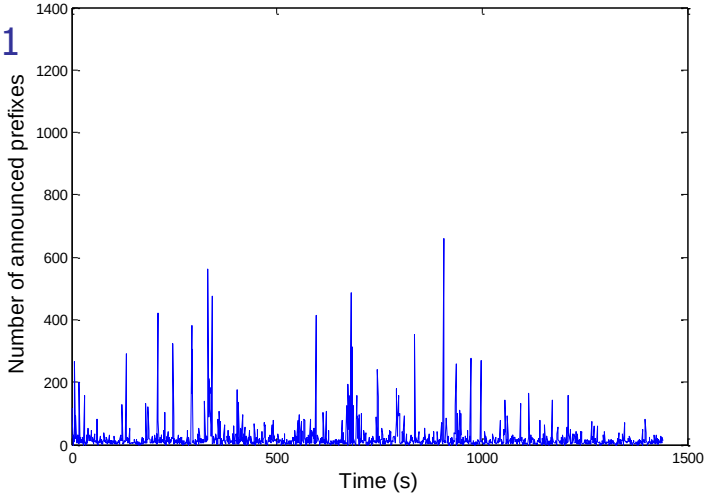
Number of announced prefixes



October 2, 2011



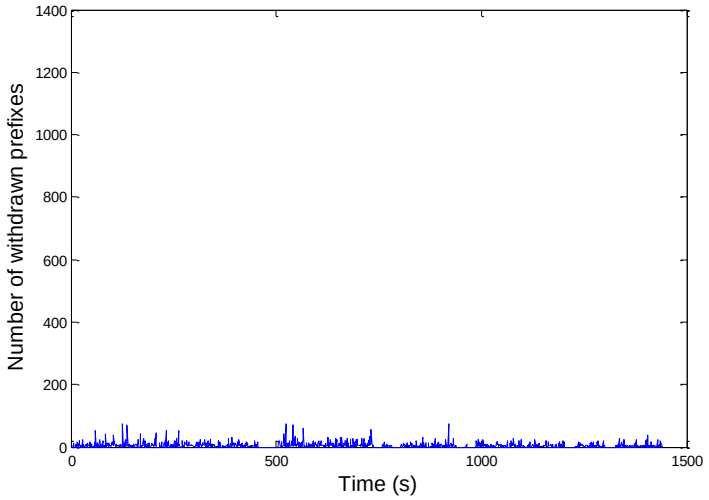
November 2, 2011



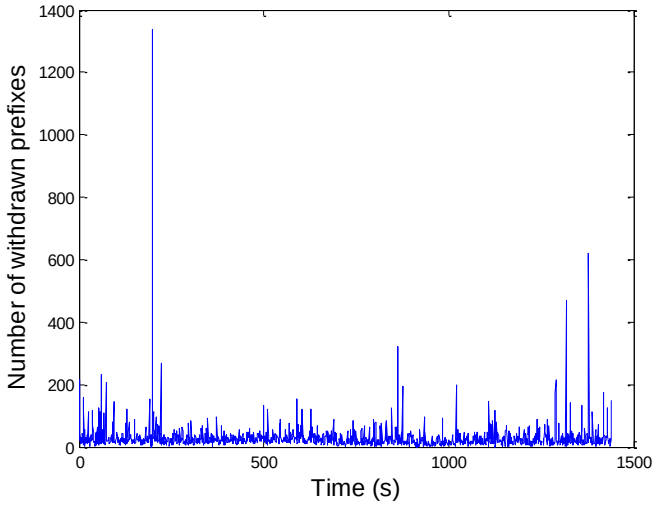
December 2, 2011

Number of withdrawn prefixes

- Number of prefixes withdrawn over a period of time
- Information, such as associated path attributes (AS Path) is not necessary for the routes being withdrawn



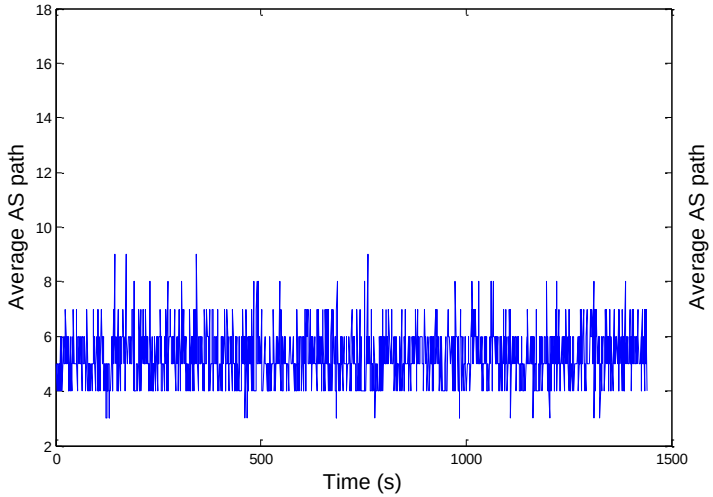
October 2, 2011



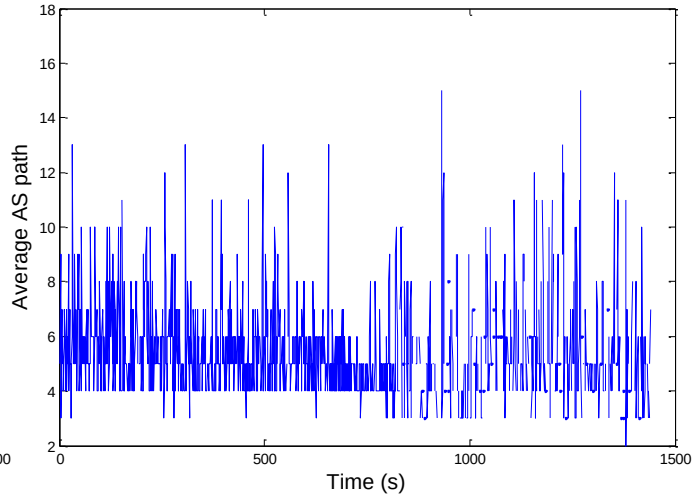
November 2, 2011

Average AS path

- The average number of AS peers on the AS path attribute of the BGP message
- It is calculated from packet flow and unique pair of AS to calculate the correlation between Flow Data and AS path length



October 2, 2011

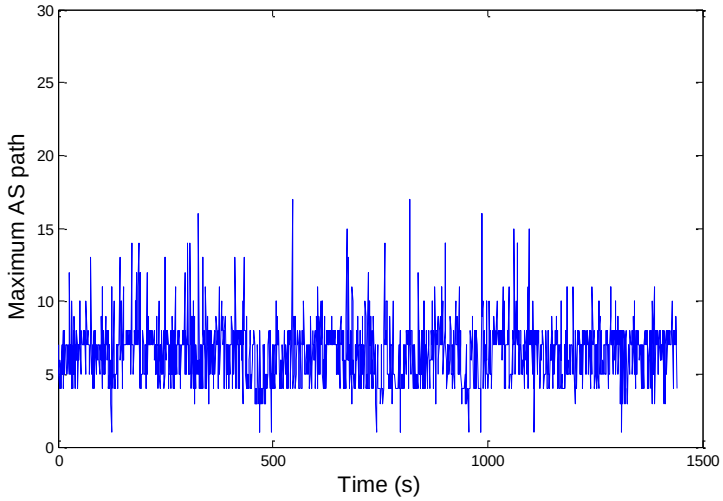


November 2, 2011

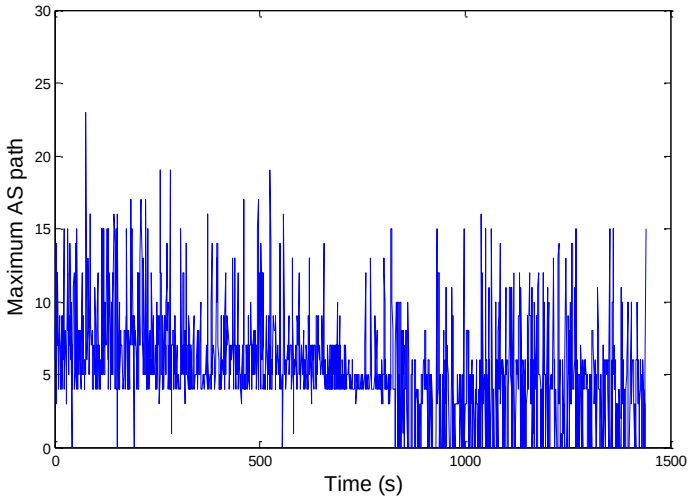
S. Deshpande, M. Thottan, T. K. Ho, and B. Sikdar, "An online mechanism for BGP instability detection and analysis," *IEEE Trans. Computers*, vol. 58, no. 11, Nov. 2009, pp. 1470–1484.

Maximum AS path

- It is the maximum number of AS peers during a pin duration (one minute) of BGP update messages



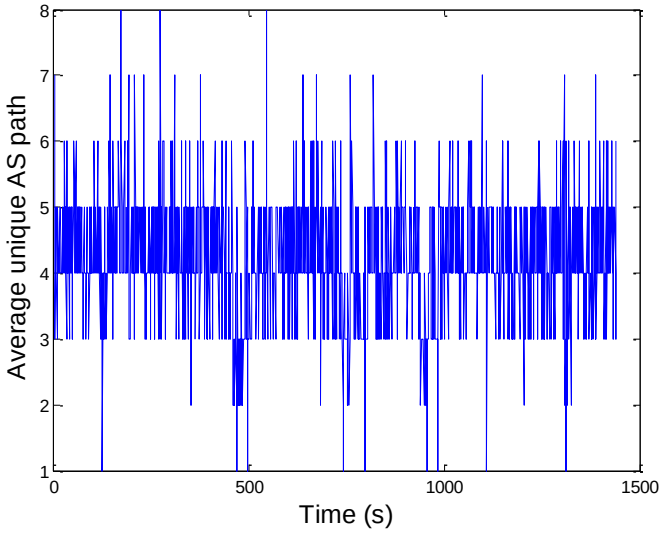
October 2, 2011



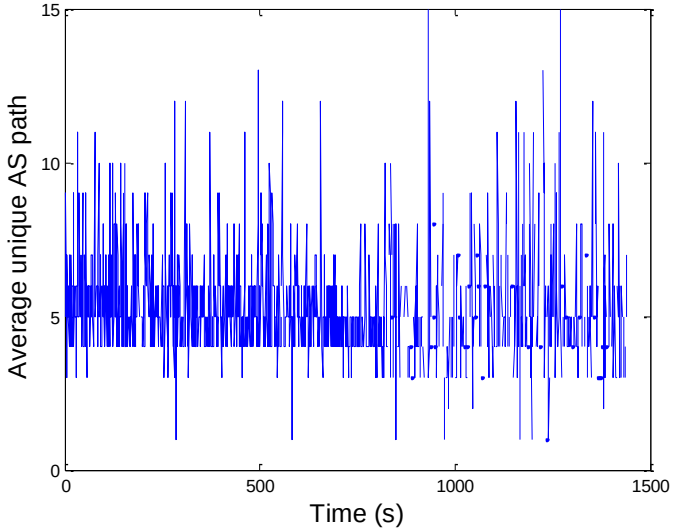
November 2, 2011

Average unique AS path

- It is the same as the average AS path length except that it considers the unique AS-path attributes during the one minute period



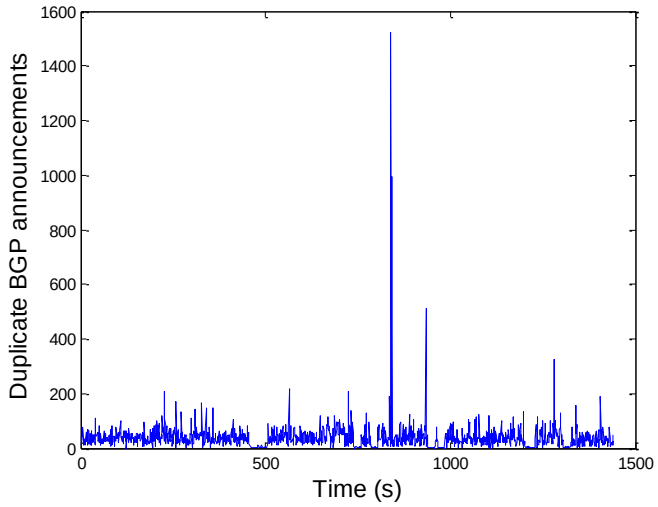
October 2, 2011



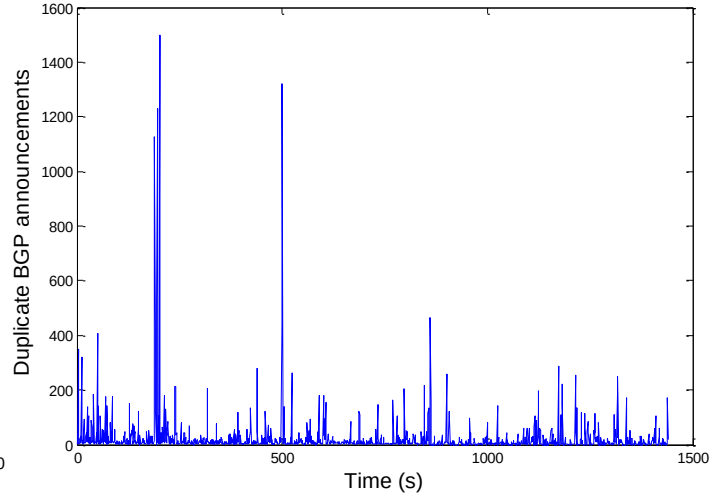
November 2, 2011

Duplicate BGP announcements

- It is an announcement that is identical to the last seen announcement for the same NLRI prefix
- There is no change in either the AS path or in any of the transitive route attributes



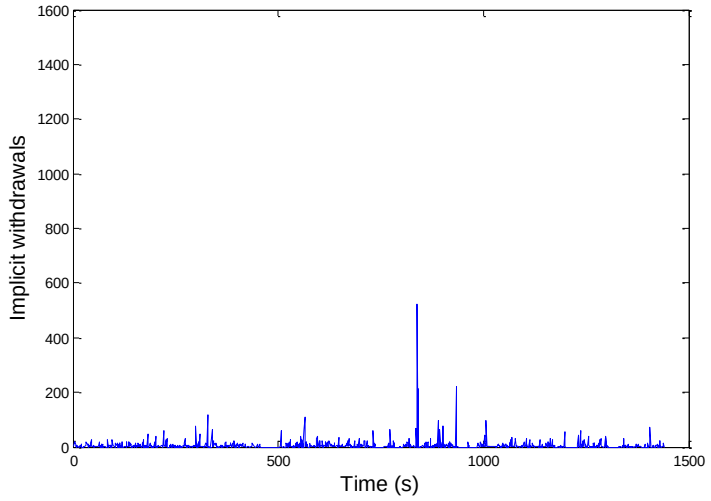
October 2, 2011



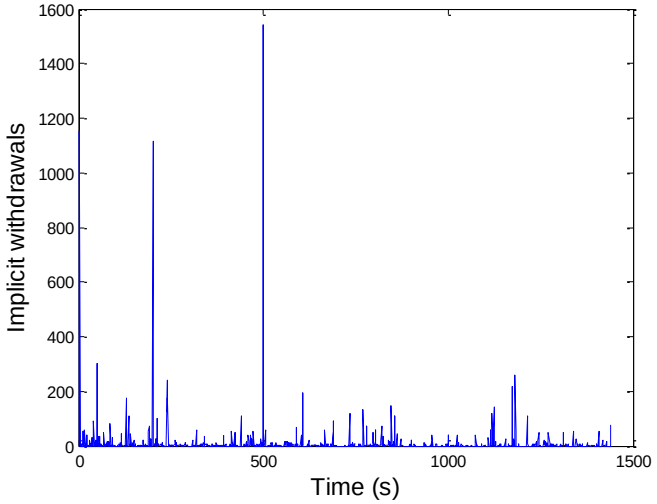
November 2, 2011

Implicit withdrawals

- Implicit withdrawal is a prefix that has been withdrawn and re-advertised
- This number is smaller than total prefix sent in that particular case



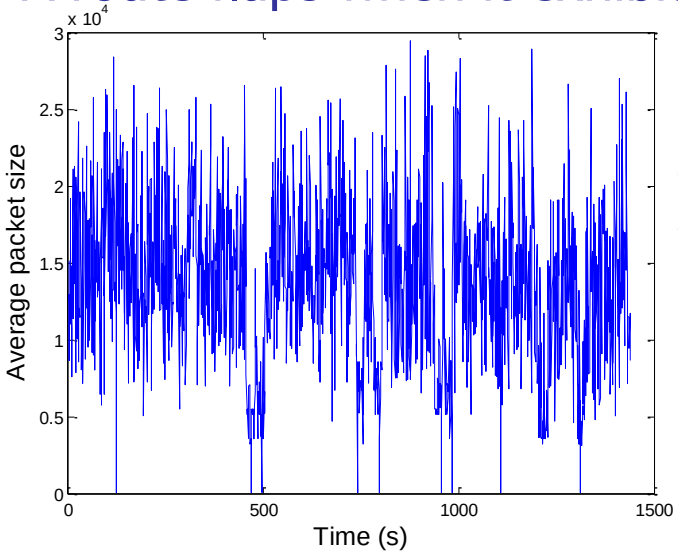
October 2, 2011



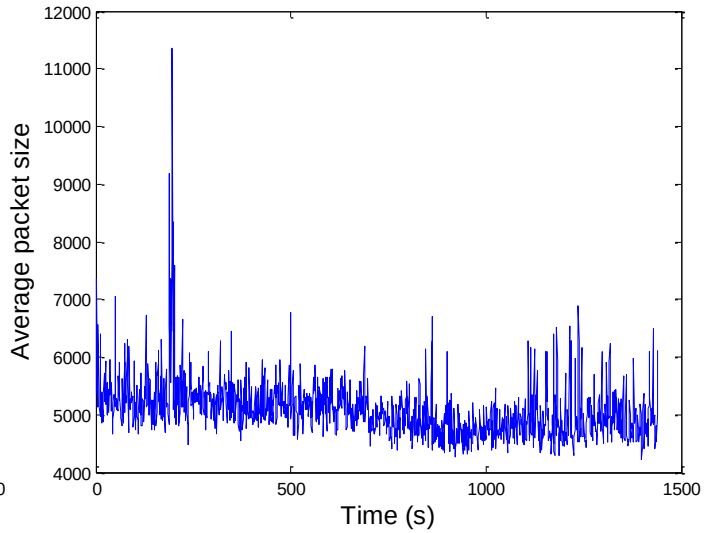
November 2, 2011

Average packet size

- Periodic stream of average packet size over a long period of time has a clothesline phenomenon which may be because of route flapping
- A route flaps when it exhibits routing oscillations

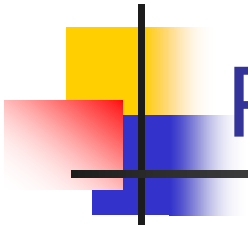


October 2, 2011



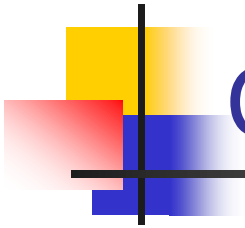
November 2, 2011

B. A. Prakash, N. Valler, D. Andersen, M. Faloutsos, and C. Faloutsos, "BGP-lens: Patterns and anomalies in internet routing updates," in ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 2009, pp. 1315–1324.



Roadmap

- Introduction
- BCNET packet capture
- Wireshark
- BGP attributes
- BGP update attributes
- Conclusions, future work, references



Conclusions

- The collected BCNET traffic was analyzed using Wireshark packet analyzer
- We provided details of the hardware used for traffic collection
- We described the testbed and BCNET measurements
- We can conclude that BGP traffic on November 2, 2011 showed better results as in almost all the features when compared to traffic on October 2, 2011 and December 2, 2011
- Number of announcements should always be greater than number of withdrawals and that was the finding in this case

Future work

- Collected BGP traffic data may be used to infer the Internet topologies and their historical development
- The collected data may be used use to analyze performance of BGP
- The effect of Route Flap Damping (RFD) and Minimal Route Advertisement Interval (MRAI) may be analyzed
- Classification and clustering of traffic may be done in future

References

- Y. Rekhter, T. Li, and S. Hares, "A border gateway protocol 4 (BGP-4)," IETF RFC 1771.
- B. A. Prakash, N. Valler, D. Andersen, M. Faloutsos, and C. Faloutsos, "BGP-lens: Patterns and anomalies in internet routing updates," in ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 2009, pp. 1315–1324.
- Autonomous System Numbers [Online]. Available: http://www.iana.org/assignments/as-numbers.
- BGP Routing table Analysis [Online]. Available: http://www.potaroo.net/tools/asns/.
- BCNET [Online]. Available: http://www.bc.net.
- Data Monitoring Switch [Online]. Available: http://www.netoptics.com/products/director.
- S. Lally, T. Farah, R. Gill, R. Paul, N. Al-Rousan, and Lj. Trajkovic, "Collection and characterization of BCNET BGP traffic," in Proc. 2011 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Victoria, BC, Canada, Aug. 2011, pp. 830–835.
- Wireshark [Online]. Available: http://www.wireshark.org.
- Wireshark User's Guide [Online]. Available: http://www.wireshark.org/docs/wsug_html_chunked/ChAdvTimestamps.html.

References

- Welcome to DAG [Online]. Available: http://www.endace.com.
- BCNET Traffic Map [Online]. Available: https://www.bc.net/atlconf/display/Network/BCNET+Traffic+Map.
- D. Meyer, "BGP communities for data collection," RFC 4384, IETF, 2006 [Online]. Available: http://www.ietf.org/rfc/rfc4384.txt.
- S. Deshpande, M. Thottan, T. K. Ho, and B. Sikdar, "An online mechanism for BGP instability detection and analysis," *IEEE Trans. Computers*, vol. 58, no. 11, Nov. 2009, pp. 1470–1484.
- W. Shen and Lj. Trajkovic, "BGP route flap damping algorithms," in Proc. SPECTS 2005, Philadelphia, PA, July 2005, pp. 488–495.
- L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "Observation and analysis of BGP behavior under stress," in Proc. 2nd ACM SIGCOMM Workshop on Internet Measurement, New York, NY, USA, 2002, pp. 183–195.
- J. Park, D. Jen, M. Lad, S. Amante, D. McPherson, and L. Zhang, "Investigating occurrence of duplicate updates in BGP announcements," in Proc. Passive and Active Measurement, Zurich, Switzerland, April 2010, pp. 11–20.

References

- J. Zhang, J. Rexford, and J. Feigenbaum, "Learning-based anomaly detection in BGP updates," in Proc. ACM SIGCOMM Workshop on Mining Network Data, Philadelphia, PA, USA, August 2005, pp. 219–220.
- V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions and reversals," in Soviet Physics Doklady, Technical Report 8, 1966, pp. 707–710.
- R. A. Wagner and M. J. Fisher, "The string-to-string correction problem," Journal of the ACM, vol. 21, no. 1, pp. 168–173, Jan 1974.
- C. Labovitz, R. Wattenhofer, S. Venkatachary, and A. Ahuja, "The impact of Internet policy and topology on delayed routing convergence," in Proc. IEEE INFOCOM, Anchorage, Alaska, April 2001, pp. 537–546.
- A. Elmokashfi, A. Kvalbein, and C. Dovrolis, "On the scalability of BGP: the role of topology growth," IEEE Journal on Selected Areas in Communications, Special issue: Internet Routing Scalability, October 2010, pp.1250–1261.
- N. Laskovic and Lj. Trajkovic, "BGP with an adaptive minimal route advertisement interval," in Proc. 25th IEEE Int. Performance, Computing, and Communications Conference, Phoenix, AZ, Apr. 2006, pp. 135–142.

References

- J. F. Kurose and K. W. Ross, "Transport layer," in Computer Networking: A Top-down Approach, 4th ed, New York: Pearson International, 2007, pp. 307–308.
- N. Al-Rousan and Lj. Trajković, "Comparison of machine learning models for classification of BGP anomalies," in Proc. HPSR 2012, Belgrade, Serbia, June 2012 (to be presented).
- T. Farah, S. Lally, R. Gill, N. Al-Rousan, R. Paul, D. Xu, and Lj. Trajkovic, "Collection of BCNET BGP traffic," in Proc. 23rd International Teletraffic Congress, San Francisco, CA, USA, Sept. 2011, pp. 322–323.
- Cooperative Association for Internet Data Analysis [Online]. Available: http://www.caida.org.
- S. Haeri, D. Kresic, and Lj. Trajkovic, "Probabilistic verification of BGP convergence," in Proc. IEEE International Conference on Network Protocols, ICNP 2011, Vancouver, BC, Canada, Oct. 2011, pp. 127-128 (students poster session paper).
- OpenFabrics Alliance Archive [Online]. Available: http://www.openfabrics.org/archives/spring2008sonoma/Wednesday/Endace-Wednesday.ppt.
- W. Feng and P. Tinnakornsrisuphap, "The adverse impact of the TCP congestion-control mechanism in heterogeneous computing systems," in Proc. The International Conference on Parallel Processing, Toronto, Canada, August 2000, pp. 299–306.

Acknowledgements

Committee members:

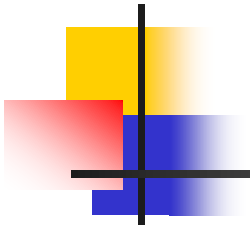
- Prof. Lesley Shannon
- Prof. Carlo Menon

Chair:

- Prof. Ash Parameswaran

Senior supervisor:

- Prof. Ljiljana Trajković



Thank You