

# Collection and Characterization of BCNET BGP Traffic

Sukhchandan Lally, Tanjila Farah, Rajvir Gill, Ravinder Paul, Nabil Al-Rousan, and Ljiljana Trajković

*Simon Fraser University*

*Vancouver, British Columbia, Canada*

*E-mail: {lally, tfarah, rajvirg, rpa28, nalrousa, ljilja}@sfu.ca*

## Abstract

*Border Gateway Protocol (BGP) is an Inter-Autonomous System routing protocol that operates over the reliable Transmission Control Protocol (TCP). The most important function of a BGP speaker is to exchange network reachability information with other speakers. In this paper, we describe collection and preliminary analysis of BCNET traffic. The traffic was collected using special purpose hardware: the Net Optics Director 7400 and the Endace DAG 5.2X card. Collected data was analyzed using the Wireshark open-source packet analyzer.*

## 1. Introduction

Traffic measurements in operational networks help understanding traffic characteristics in deployed networks, developing traffic models, and evaluating performance of protocols and applications. Traffic analysis provides information about patterns of user behavior and enables network operators to understand the behavior of network users.

BCNET is British Columbia's advanced research and innovation network. It is a dedicated high-speed fiber optic network that spans across the province of British Columbia, Canada. Border Gateway Protocol (BGP) provides mechanisms for supporting classless inter-domain routing (CIDR) [1], which makes it a dominant Internet routing protocol and allows the aggregation of routes. BGP supports policies conforming to the “hop-by-hop” paradigm.

In this paper, we provide a description of the special purpose hardware and software used to collect the BCNET BGP traffic. The equipment was deployed in an operational testbed and was used to collect a preliminary traffic trace from December 20, 2010 to December 22, 2010. The collected traffic will be used to analyze performance of routing protocols.

This paper is organized as follows. In Section 1, we also provide a brief description of BGP and routing policies. The BCNET architecture and the equipment used for traffic measurements are described in Section 2. BCNET traffic details are given in Section 3. Wireshark views of collected BCNET traffic and preliminary traffic analysis are described in Section 4. We conclude with Section 5.

## 1.1 Border Gateway Protocol

BGP is de-facto Inter-Autonomous System (AS) [2] routing protocol. An AS comprises of groups of routers that are administrated by a single administrator and use Interior Gateway Protocol (IGP). Each AS is responsible for carrying traffic to and from a set of customer IP addresses. The AS numbers are used by various routing protocols and are assigned to the regional registries by the Internet Assigned Numbers Authority (IANA).

BGP speakers that participate in a BGP session are called neighbors or peers. Peer routers exchange four types of messages: *open*, *update*, *notification*, and *keepalive*. The main function of BGP is to exchange reachability information among BGP systems. This information is based on a set of metrics: policy decision, the shortest *AS path*, and the closest *Next hop* router.

BGP operates over a Transmission Control Protocol (TCP) (port number 179), which has an advantage over User Datagram Protocol (UDP) connections. BGP does not need to implement fragmentation, retransmission, acknowledgment, and sequencing.

## 1.2 Routing Policies and BGP Routing Tables

To select the routing path, BGP utilizes a path vector algorithm called the best path selection algorithm. This algorithm enables BGP to select the best path to be included in routing tables for destinations with multiple paths. BGP applies policies to the information contained in routing updates and accepts or rejects update information based on attributes [3]. These attributes have

the capability of imposing policies based on various routing preferences and constraints. BGP allows a wide range of routing policies to control the exchange of traffic.

After receiving an update, a router decides whether or not to use the path according to import policies and updates the neighbouring ASs according to export policies. An AS uses import policies to transform incoming route updates. The import policies include denying or permitting an update and assigning a local preference to indicate how favourable the path is [4].

BGP routing tables are publicly available and may be retrieved from the Route Views server in Oregon [5]. The University of Oregon Route Views project was originally envisioned as a tool for Internet operators to obtain real-time information about the global routing system from the perspectives of several distinct backbones and locations around the Internet. The Réseaux IP Européens (RIPE) database also contains information about allocations and assignments of IP address space, routing registry information, reverse Domain Name System (DNS) designation, and related objects [6], [7]. The size of BGP tables have exponentially increased since 1994, as shown in Figure 1, implying that timely analysis of BGP is important.

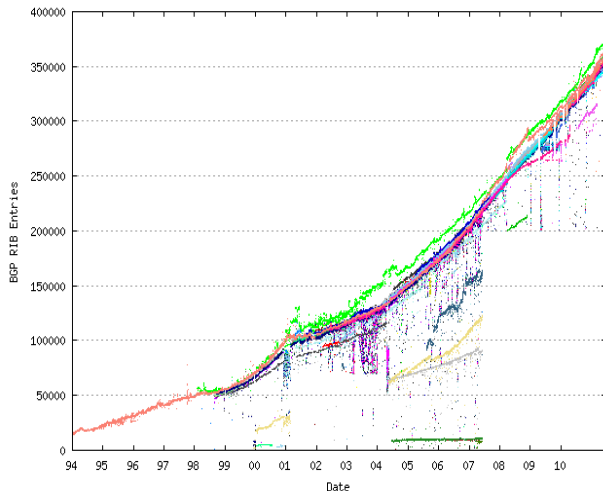


Figure 1. Growth of the BGP tables: 1994 to present [7].

## 2. BCNET Packet Capture

BCNET is the hub of advanced telecommunication network in British Columbia, Canada that offers services to research and higher education institutions. This advanced network offers unconstrained bandwidth to research and innovation centers making it suitable to address unique research requirements [8]. It is used for

collaboration among researchers across institutions in British Columbia. A map of the BCNET is shown in Figure 2. BCNET transits have two service providers with 10 Gbps network links and one service provider with 1 Gbps network link. These links are connected via two routers. BCNET Router 1 and Router 2 are placed in two separate physical rooms. Router 1 is connected via 10 Gbps link services while Router 2 connects both 10 Gbps and 1 Gbps providers. BCNET transit exchange interconnects the participants and provides local peering and multi-hopping services to open and private data exchanges.

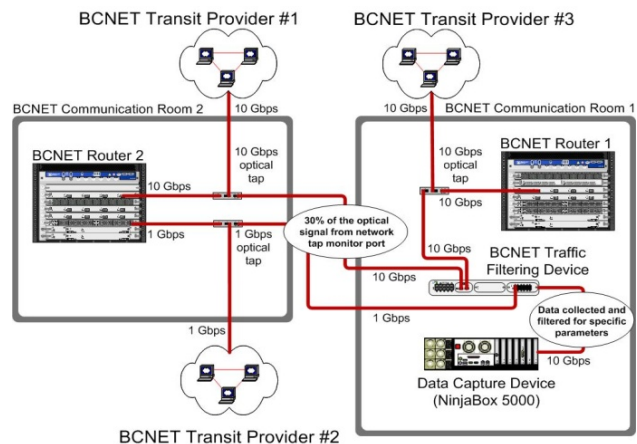


Figure 2. Physical overview of the BCNET packet capture [8].

Optical Test Access Point (TAP) shown in Figure 2 is used to split the signal into two distinct paths. Signal splitting ratio from TAP may be modified. In the BCNET transit exchange, 30% of the optical signal is directed to a traffic filtering device while the remaining 70% of the signal is sent to routers for processing.

The Data Capture Device (NinjaBox 5000) collects the real-time data (packets) from the traffic filtering device. NinjaBox 5000 relies on Linux operating system to capture data at line rates using conventionally made network monitoring. NinjaBox 5000 comes preconfigured with Endace DAG monitoring interface technology. Conventional capture systems cannot capture data at high rates due to the overhead in processing the captured packets, which results in lost packets.

Net Optics Director 7400 is used for BCNET traffic filtering. All three BCNET service provider links are connected to the device, as shown in Figure 2. Captured data are filtered for a specific set of parameters by the filtering device and then directed to the NinjaBox 5000. The filtering device operates as a data monitoring switch [9]. It directs traffic to monitoring tools such as NinjaBox 5000 and FlowMon. The Net Optics Director application diagram is shown in Figure 3.

A data monitoring switch provides access to traffic from network links. It provides functionalities that include monitoring traffic from multiple links, regenerating traffic to multiple tools, pre-filtering traffic to offload tools, and directing traffic according to one-to-one and many-to-many port mappings. It assists organizations to use efficiently their monitoring tools, to centralize traffic monitoring functions, and to share tools and traffic access between groups.

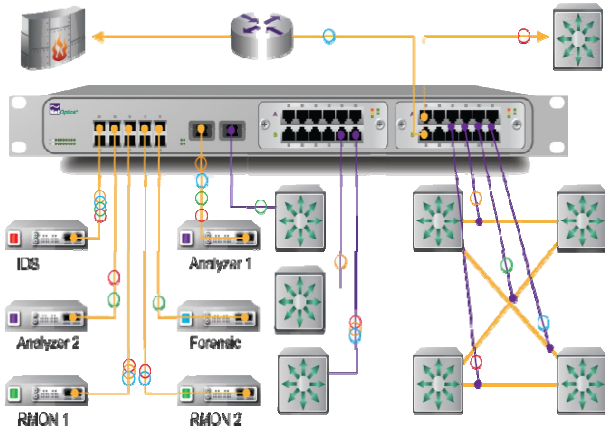


Figure 3. Net Optics Director application diagram [9].

The Endace Data Acquisition and Generation (DAG) 5.2X card shown in Figure 4 resides inside the NinjaBox 5000. It captures and transmits traffic and has time-stamping capability. Resolution of Endace DAG time stamp is 10 ns. In contrast, software based captures such as Wireshark support 1  $\mu$ s resolution [10]. The fine granularity of hardware based captures ensures credibility, accuracy, and reliability of measured BGP traffic and its subsequent analysis, characterization, and modeling [11], [12].

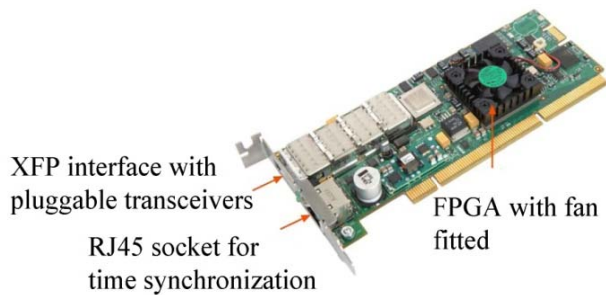


Figure 4. Network monitoring and analyzing Endace card [13].

Endace specializes in high performance network monitoring and analysis. “DAG” is a technology that combines hardware design using field programmable gate

array (FPGA) technology and software based on a programmable chip [13]. It also supports Data Stream Manager (DSM) feature that allows discarding or routing packets to a particular stream based on the packet contents, physical port, and the output of two load balancing algorithms.

DAG 5.2X is a single port Peripheral Component Interconnect Extended (PCIx) card. The card provides capture and transmission of full duplex optical 10 Gbps Ethernet network data. It is capable of capturing on average Ethernet traffic of 6.9 Gbps.

### 3. BCNET Traffic

The BCNET network facilitates high-definition videoconferencing, remote research, virtual laboratories, distributed computing, distant learning, and large-scale data transfers.

The BCNET network is high-speed fiber optic research network that is used to transmit telephone signals, cable television signals, and Internet communication. This network has been primarily installed for long-distance applications, where it can be used to its full transmission capacity. Network interconnections are accommodated by BCNET transit exchanges that implement peering between associates and access data exchanges with local peering and multi-hopping services. The peering requires an exchange of routing information and physical interconnection of networks through the BGP routing protocol. The BCNET offsets the increased Internet transit cost and improves network performance due to high-speed fiber network and peering services.

British Columbia's network extends to 1,400 kilometers and connects Kamloops, Kelowna, Prince George, Vancouver, and Victoria. The BCNET network span expands to Prince George and Victoria through Vancouver and contributes up to 72 wavelengths of capacity at 10 Gbps. The BCNET network connects over 140 provincial universities and institute campus sites, provincial health centers, research facilities, federal and provincial research labs, and colleges and schools that use the Provincial Learning Network.

BCNET is associated with the network alliance Canada’s Advanced Research and Innovation Network (CANARIE), which links Canada to the United States through Internet and to Europe through Delivery of Advanced Network Technology to Europe (DANTE). The BCNET traffic map shown in Figure 5 displays the real time network usage by BCNET associates. The arrows in the map show the traffic bound for CANARIE, the commercial Internet (Transits), and peering traffic at the Seattle Internet Exchange (Seattle IX).

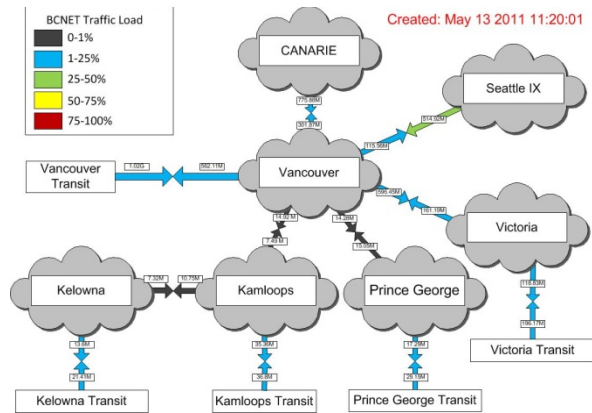


Figure 5. Real time network usage by BCNET members, collected on May 13, 2011 [14].

#### 4. Views of BCNET Traffic using Wireshark

Wireshark is an open source packet analyzer that captures network packet data from a network interface and displays those packets with detailed protocol information [10]. It is used as a measuring tool to examine the inside of a network cable. It opens and saves captured packet data, imports and exports packet data from and to other capture programs, filters and searches packets based on various criteria, colorizes packet display based on filters, and creates various statistics.

Wireshark provides comprehensive statistics such as the summary of traffic collected, input/output graphs, protocol hierarchy, and endpoints. A view of the collected traffic is shown in Figure 6. It illustrates the protocol structure for a randomly selected BGP *update* message, which contains path attributes for the advertised Network Layer Reachability Information (NLRI). Wireshark automatically detects relationship to another packet in the capture file and generates a link to that packet, as shown by the third line in Figure 6.

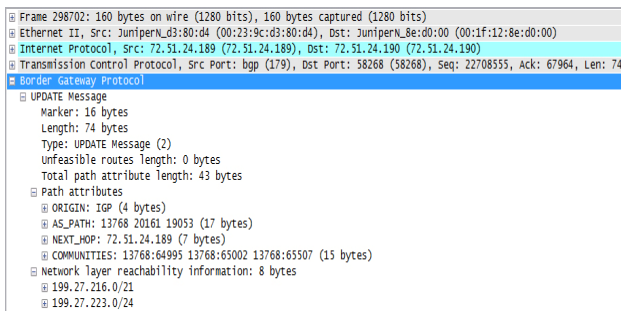


Figure 6. General Wireshark view of the collected traffic.

BCNET *Traffic Summary* of the collected data is shown in Figure 7. The timestamps show the first and the

last packets. There were 511,820 packets collected over the period of 48 hours. An example of summary statistics for a specific filter is shown in the Displayed column.

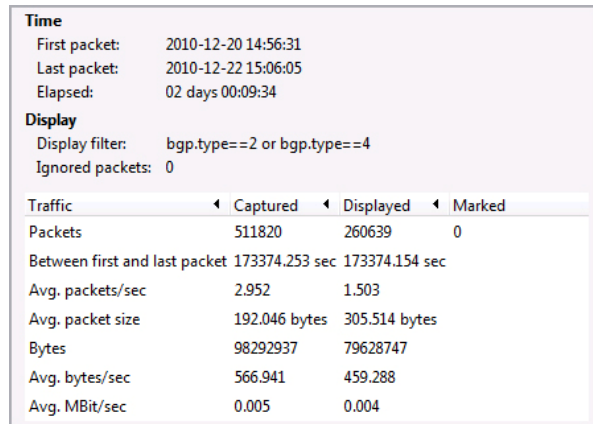


Figure 7. Summary of BCNET traffic collected over a period of 48 hours.

BCNET *Traffic Input-Output Graphs* define up to five filters. The number of samples is limited to 100,000. A sample graph is shown in Figure 8.

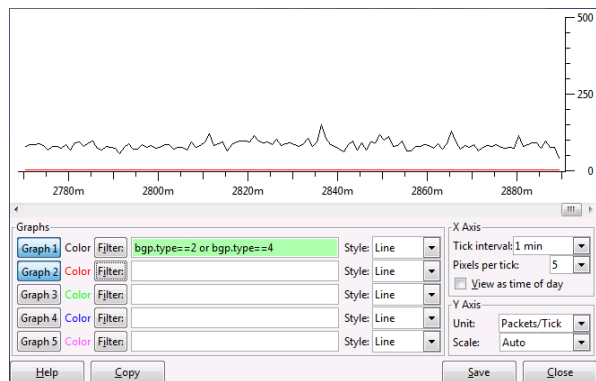


Figure 8. Input-Output graph of the packets captured. The x-axis: tick interval = 1 s, 5 pixels/tick. The y-axis: unit = packets/tick, scale = 10.

BCNET *Traffic Protocol Hierarchy* statistics of collected traffic is shown in Table 1. 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, and 6,104 (1.2%) are piggyback ACKs. Packets originate from multiple protocols and, hence, a packet may be attributed to more than one protocol. Protocol layers may consist of packets that do not contain any higher layer protocol and therefore the sum of all higher layer packets may not add to the protocols packet count. A single

packet may be counted more than once if it is encapsulated by multiple protocols.

Table 1. Protocol hierarchy of the captured packets.

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

BCNET *Network Endpoints* are the source and destination addresses of a specified protocol layer. The case of TCP endpoints is shown in Table 2. Endpoints of the six BCNET transit exchanges (BGP peers) are captured. For each IP address of a BGP peer, various TCP connection statistics are shown.

Table 2. Statistics of the captured TCP 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* is defined as the time between a request and the corresponding response. The flow graph of the captured BGP peers traffic is shown in Figure 9. It includes the source address, destination address, TCP port number, TCP message (ACK), and type of the BGP message (*open*, *update*, *notification*, *keepalive*).



Figure 9. Flow graph of collected traffic. Shown are time stamps of correspondence between BGP peer routers.

#### 4.1. Preliminary Analysis of BCNET Traffic

TCP characteristics of BCNET traffic are illustrated in Figure 10. The throughput of the collected data is shown in Figure 10 (top). Various factors such as link speed,

propagation delay, window size, link reliability, and congestion of network and intermediate device affect the throughput of TCP. The throughput graph illustrates that the average throughput of the collected data is 177.1 packets/min and that the maximum throughput is 645 packets/min. The window size is shown in Figure 10 (middle) for 200 samples of the data. TCP round trip time (RTT) shown in Figure 10 (bottom) is the time measured from segment transmission until ACK is received. Sample RTT and the estimated RTT of the collected TCP segments are shown. The RTT average is approximately 11.7 ms. The RTT standard deviation is 7.19 ms and 2.75 ms for the sample and estimated RTT, respectively.

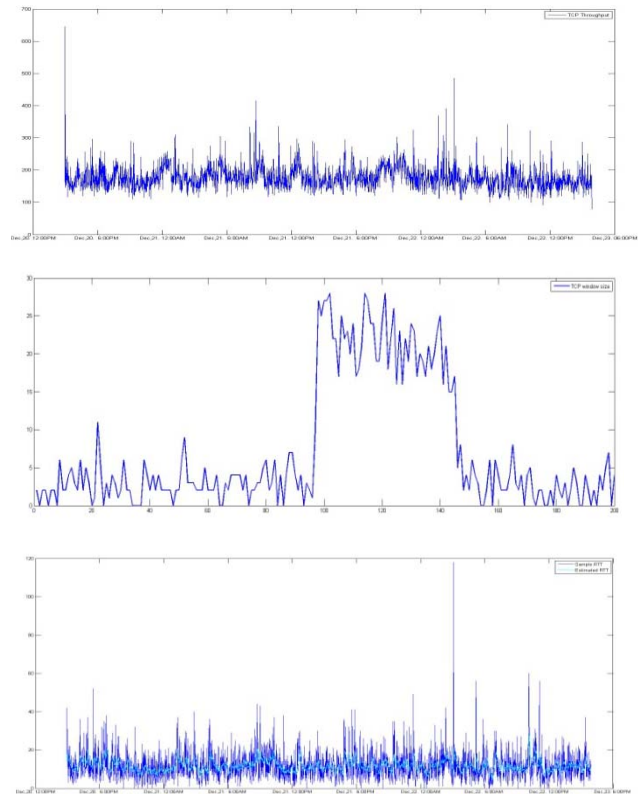


Figure 10. TCP throughput of the BCNET traffic with an average of 177.1 packets/min (top); TCP window size of the BCNET traffic for 200 samples (middle); TCP RTT of the BCNET traffic with an average of 11.7 ms (bottom).

BGP characteristics of BCNET traffic are captured by applying Wireshark BGP display filters in the Input-Output Graphs for the collected packets. Sample of display filters are: *bgp.type*, *bgp.next\_hop*, *bgp.origin*, *bgp.local\_pref*, *bgp.community\_as*, *bgp.as\_path*, and *bgp.multi\_exit\_disc*. All filters except *bgp.type* are BGP *update* attributes filters. Traffic of the overall, *update*, and *keepalive* BGP messages are shown in Figure 11. The

three graphs are obtained by applying filters: “*bgp.type==2 or bgp.type==4*” (left), “*bgp.type==2*” (middle), and “*bgp.type==4*” (right). In the collected BGP traffic, 88% are BGP *update* messages. The remaining are *keepalive* messages.

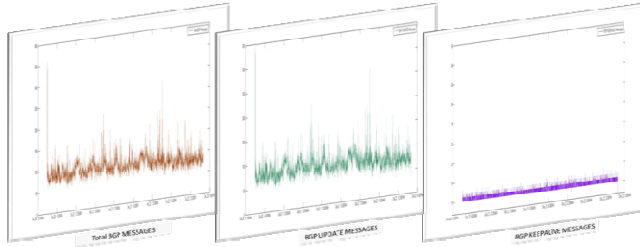


Figure 11. BGP traffic: total (left), *update* (middle), and *keepalive* (right).

An example of BGP *update* message attribute is the BGP *origin*, which defines the origin of the path. It may have three values: Interior Gateway Protocol (IGP), Exterior Gateway Protocol (EGP), or Incomplete. It assumes IGP when NLRI is interior to the AS of origin, EGP when NLRI is exterior to the origin, and Incomplete when NLRI is unknown or learned via other means [15]. The IGP announced prefixes (NLRIs) are preferred over the EGP or Incomplete prefixes. The EGP, Incomplete, and IGP messages shown in Figure 12 account for 0.003%, 13.84%, and 85.82% of the total number of BGP update messages, respectively.

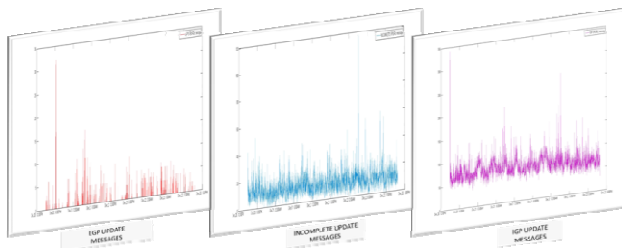


Figure 12. BGP *update* traffic: EGP (left), Incomplete (middle), and IGP (right).

## 5. Conclusions

In this paper, we described collection of BCNET BGP traffic and provided details of the hardware used for the traffic collection. The preliminary collected traffic from BCNET was analyzed using the Wireshark packet analyzer. We reported preliminary statistics of traffic data collected over a 48-hour period.

Future analysis will employ longer collections of generated traffic and performance evaluation of routing protocols. The collected data will be used to analyze performance of the BGP protocol and the effect of route

flaps and parameters such as the minimal route advertisement interval (MRAI). BGP route flaps refer to persistent routing oscillations caused by network instabilities such as configuration errors, transient data, link failures, and software defects. The BGP convergence time is affected by duration of the MRAI and the implementation of MRAI timers [16]. The default MRAI value (30 s), used in majority of routers, may need to be revised and optimized [17]. BGP traffic data collected from BCNET will be compared to Internet topologies generated from the publicly available Route Views and RIPE datasets [18].

## References

- [1] Y. Rekhter, T. Li, and S. Hares, “A border gateway protocol 4 (BGP-4),” *IETF RFC 1771*.
- [2] Autonomous System Numbers [Online]. Available: <http://www.iana.org/assignments/as-numbers>.
- [3] BGP Best Path Selection Algorithm [Online]. Available: <http://www.cisco.com/en/US/tech/tk365>.
- [4] L. Gao, “On inferring autonomous system relationships in the Internet,” *IEEE/ACM Trans. Networking*, vol. 9, no. 6, pp. 733–745, Dec. 2001.
- [5] BGP capture datasets [Online]. Available: <http://archive.routeviews.org>.
- [6] Réseaux IP Européens [Online]. Available: <http://www.ripe.net/ris>.
- [7] BGP Routing Table Analysis Reports [Online]. Available: <http://www.potaroo.net/bgp/>.
- [8] BCNET [Online]. Available: <http://www.bc.net>.
- [9] Data Monitoring Switch [Online]. Available: <http://www.netoptics.com/products/director>.
- [10] Wireshark [Online]. Available: <http://www.wireshark.org>.
- [11] Wireshark User's Guide [Online]. Available: [http://www.wireshark.org/docs/wsug\\_html\\_chunked/ChAdvTimestamps.html](http://www.wireshark.org/docs/wsug_html_chunked/ChAdvTimestamps.html).
- [12] OpenFabrics Alliance Archive [Online]. Available: <http://www.openfabrics.org/archives/spring2008sonoma/Wednesday/Endace-Wednesday.ppt>.
- [13] Welcome to DAG [Online]. Available: <http://www.endace.com>.
- [14] BCNET Traffic Map [Online]. Available: <https://www.bc.net/atlconf/display/Network/BCNET+Traffic+Map>.
- [15] BGP Case Studies [Online]. Available: <http://www.cisco.com/application/pdf/paws/26634/bgp-toc.pdf>.
- [16] W. Shen and Lj. Trajković, “BGP route flap damping algorithms,” in *Proc. SPECTS 2005*, Philadelphia, PA, July 2005, pp. 488–495.
- [17] N. Lasković and Lj. Trajković, “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.
- [18] Lj. Trajković, “Analysis of Internet topologies,” *IEEE Circuits and Systems Magazine*, vol. 10, no. 3, pp. 48–54, Third Quarter 2010.