Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Networks 50 (2006) 2013–2027
Computer Networks
www.elsevier.com/locate/comnet

Efficient large-scale BGP simulations

Xenofontas A. Dimitropoulos *, George F. Riley

College of Engineering, Department of ECE, Georgia Institute of Technology, Atlanta, GA 30332-0250, United States

Available online 10 November 2005

Abstract

Simulation has been the method of choice for analyzing large, complex, and highly volatile systems. One of these systems is the inter-domain routing infrastructure of the Internet. Despite the need for high quality Border Gateway Protocol (BGP) simulation tools, traditional BGP simulators have limitations either in their modeling fidelity or in their scalability. In this work we introduce BGP++, a scalable BGP simulator that employs state-of-the-art techniques to address the abstraction-scalability trade-off.

BGP++ builds on high quality software in network simulation, routing and parallel-distributed simulation to deliver a detailed yet scalable implementation of BGP. Moreover, with respect to the needs of researchers and operators, BGP++ has a CISCO-like configuration language, a seamless partitioning engine for parallel-distributed simulations and a configuration toolset that expedites common simulation configuration tasks.

Keywords: BGP; Network simulation

1. Introduction

Modeling and simulation analysis has played a key role in the field of computer networks. Typically, vendors and researchers evaluate prospective architectures and perform comprehensive "what if" analysis using simulation. Simulation is also used for parameter-tuning, problem diagnosis and performance optimizations. It is immensely important in research on large, complex and heterogeneous systems, like the Internet, where analytical models and laboratory testbeds do not capture the detail or the sheer volume of the system. One of these systems is the BGP infrastructure of the Internet.

BGP is the de-facto inter-domain routing protocol in the Internet. It is the "glue" that interconnects more than 16,000 Autonomous Systems (AS) of diverse geopolitical nature. In contrast to Interior Gateway Protocols (IGP), BGP is a policy-based protocol. Business relationships and agreements between ASs determine how packets are routed.

During the last few years the inter-domain routing infrastructure has attracted substantial research interest. Some of the widely researched problems are its slow convergence [1–5], policy conflicts [6,7], instability [8,9], misconfigurations [10], lack of security [11], table growth [12,13] and path inflation [14–16]. These problems necessitate fixing or even replacing the current routing architecture. Several

* Corresponding author.
E-mail addresses: fontas@ece.gatech.edu (X.A. Dimitropoulos), riley@ece.gatech.edu (G.F. Riley).

1389-1286/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2005.09.033

examples of proposed new routing architectures can be found in the following references [17–23]. Yet, it is not clear how BGP should be modified.

The main reason the future of the routing architecture is dubious, is the lack of necessary tools to comprehensively understand the current infrastructure and to evaluate new alternatives. Measuring the performance of BGP is strongly prohibited by ownership and lack of measurement infrastructure. Analytical models of BGP are simplistic and do not capture the complexity, configurability and heterogeneity of the protocol. BGP simulators can only perform moderate-scale simulations of a few hundreds of routers, from which we cannot draw reliable conclusions.

In this work we are concerned with building a simulation tool that will help researchers shed light on the design flaws of the current BGP infrastructure and evaluate the performance of new architectures. BGP++ is designed along the *scalability-realism* diptych, enabling for the first time large-scale as well as detailed BGP simulations. BGP++ is not a new bottom-up implementation of BGP, but it capitalizes on and benefits from existing high quality software on network simulation, routing and parallel-distributed simulation. The following software are the basic components on which we build BGP++:

1. ns-2 [24] is a discrete-event network simulator that serves as a common platform on which researchers can test and compare their proposals. ns-2 has evolved into the most widely used simulator in networking literature. It includes numerous implementations of protocols with a special emphasis on TCP variants.
2. GNU Zebra [25] and its ancestor, Quagga [26],[1] is a detailed open-source implementation of BGP. It is used by a large community of providers, vendors and researchers for testing and experimenting with the protocol. It is also used for routing by small ASs that cannot afford buying expensive routers.
3. PDNS [27] is the parallel-distributed version of the ns-2 simulator. It enables large-scale simulations by distributing the simulation model on multiple workstations, thereby granting more physical resources. In contrast to its counterpart, SSFnet [28], PDNS supports both shared-memory multiprocessors and distributed-memory clusters of workstations. Support of distributed architectures offers more physical resources, overwhelming the limits inherent in parallel architectures. Remarkably, PDNS was recently used by Fujimoto et al. [29] to realize the largest network simulations ever, of more than 5 million network nodes.

Our contributions can be summarized as follows:

1. We develop and make publicly available [30] a packet-level BGP simulator on the widely used ns-2 simulation platform.
2. We integrate the BGP implementation of Zebra into ns-2, making the minimum possible changes to the original software. We realize an accurate BGP simulator that supports most of the details of Zebra's BGP implementation, including a CISCO-like configuration language.
3. We use and extend PDNS to support parallel-distributed BGP simulations. We also evaluate the performance of alternative model partitioning algorithms for parallel-distributed simulations.
4. We identify the representation of BGP routing tables as the main source of memory consumption in BGP simulations. We introduce a compact routing table data structure that exploits the redundancy of information in BGP routing tables and realizes significant memory savings.
5. We propose and develop a simple generic technique to speedup simulation trials using process checkpointing.
6. We survey recent advancements in measuring the Internet. We highlight measurement data and related models that should be explored to enhance the fidelity of BGP simulations.
7. We develop a seamless partitioning and configuration engine for parallel-distributed BGP simulations that hides the complexity of PDNS configuration and brings parallel-distributed simulation closer to the general ns-2 user. We also develop an automatic generator of CISCO-like configuration. We combine these tools in a toolset that expedites common BGP simulation and configuration practices.

The remainder of this paper is organized as follows. Section 2 elaborates on previous efforts on BGP simulation. Section 3 describes in detail the development of BGP++. Section 4 introduces three

[1] BGP++ project predates Quagga project, for this reason we use Zebra.