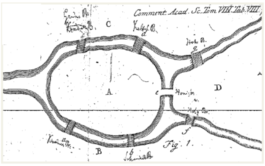
September 2013 No. 10



THE EULER NEWSLETTER

SEVENTH FRAMEWORK PROGRAMME

In this issue

EULER experimentation for a multicast-enable Internet

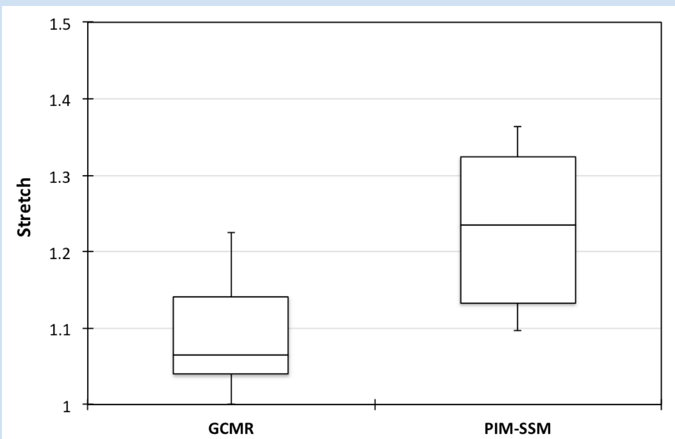
The most fundamental **issues** faced by the **Internet architecture** are the **scalability**, **convergence**, and **stability** properties of its **inter-domain routing system**. Solving them requires to address multiple dimensions together: i) the routing table size growth resulting from a larger number of routing entries, and ii) the routing system dynamics characterized by the routing information exchanges resulting from topological or policy changes. Worst-case projections predict that routing engines could have to process and maintain in the order of 1 million active routes within the next 5 years. Thus, while the Internet routing system prevents from any host specific routing information processing and maintenance (routing state), storing an increasingly large amount of network states in the routing system is expensive and places undue cost burdens on network administrative units.

These issues are even more evident if we consider multicast routing system. Originally defined in the 90's, its potential benefits have been verified by studies several times since then. By multicast routing, we refer to a distributed algorithm that, given a group identifier, allows any node to route multicast traffic to a group of destination nodes, usually called multicast group. To enable one-to-many traffic distribution, the multicast routing protocol configures the involved routers to build a (logical) delivery tree between the source and the multicast group, commonly referred to as the Multicast Distribution Tree (MDT). Although routing protocol independent routing schemes have been standardized during last decade, only the SSM variant of PIM (PIM-SSM) [1] has been deployed in the context of IPTV systems for routing multicast streams between VLANs, subnets or access networks (intra-domain multicast). Among other reasons, this failure could be attributed to the scaling limitations and relative complexity of the standard multicast protocol architecture, based on overlaying multicast routing on top of the unicast routing topology [2].

In EULER, we designed the Greedy Compact Multicast Routing scheme (presented in Newsletter No. 3) which is characterized by its independence from any underlying unicast routing topology; more specifically, the local knowledge of the cost to direct neighbor nodes is enough for the GCMR scheme to properly operate. During the past year of the project, we developed a prototype of the GCMR multicast scheme and experimentally evaluated its functionality and performance on the iLab.t virtual wall platform, which is a large-scale experimental Linux machine-based emulation testbed (presnted in Newsletter No. 7). The prototype of a GCMR routing engine has been developed using the libraries of the Quagga open source routing suite. The success in our endeavor, which was presented during the Hands-on-FIRE! Demonstration (collocated in the 2013 FIA Week in Dublin, Ireland), suggests a feasible multicast-enabled Internet.

In this letter, we present the set of experimental tests we executed in EULER to demonstrate the successful operation of GCMR in the context of inter-domain routing over a large-scale network topology compared to the standard PIM-SSM protocol. In particular, the following performance metrics are evaluated: 1) **Stretch**, defined as the sum of the weights of edges used in multicasting from the source to all receivers divided by the optimal such tree. Intuitively, the stretch of a routing scheme provides a quality measure of the path cost increase it produces compared to the optimal tree (which has clearly a stretch of 1); 2) **Routing Table** (RT) size, defined as the maximum number of memory-bits required to locally store the RT entries. The storage required by the algorithm is directly related to routing system scalability because the less memory a router needs to store its entries, the more scalable the routing system would be; 3) **Recovery Time**, defined as the maximum time needed to receive back a multicast transmission at the receivers once a failure occurs in a link of the MDT.

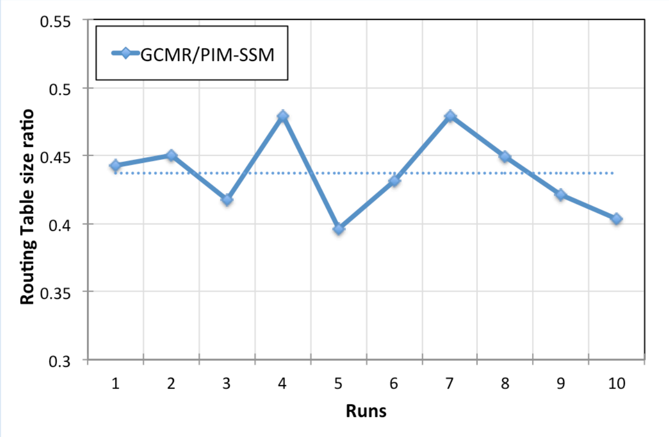
The experimentation tests of PIM-SSM and GCMR were performed in a network consisting of 207 Autonomous Systems (ASes). As we were interested in the inter-domain aspects, we represented the behavior of each AS with a single router with multiple interfaces. These 207 ASes emulate a portion of the Internet where one AS provides a multicast service to the rest of 206 ASes. In particular, we executed ten runs of the same experiment: one multicast server located in one AS is firstly selected and then ten receivers, located in ten different ASes, joined the MDT sequentially. Both the server and the receivers were randomly chosen.



Stretch

The figure on the left hand side shows the stretch of GCMR and PIM-SSM in a box-plot graph. **GCMR** presents in all runs **lower stretch** than **PIM-SSM** (0.17 better on average) and some deterioration (0.065 on average) against the optimal MDT (remind that stretch-1 is the reference). Using the information obtained from BGP, PIM-SSM establishes a shortest path tree (SPT) between the receivers and the source, and it is known that the SPT is not optimal from the stretch point of view. This result is consistent with the simulation results provided in Newsletter No. 3, where, in much more larger simulation scenarios (32k nodes and 500-4000 receivers), GCMR obtains approximately 0.1 better stretch than SPT

Subscribe to electronic EULER newsletter: https://sympa.inria.fr/sympa/info/euler-news



Routing table size ratio between GCMR and PIM-SSM

The comparison in terms of Routing Table (RT) size is presented in the figure above. To determine the size of the RT (in bits), we only consider the nodes involved in the MDT and use the RT formats defined in [3]: while GCMR only needs MRIB and TIB information, PIM-SSM, besides MRIB and TIB, also needs some unicast information from BGP to determine the shortest path towards the multicast source. The figure shows that **GCMR requires around 44% less bits than PIM-SSM**.

Finally, we emulated a failure in one link connecting two ASes and counted the maximum time elapsed to receive back the multicast transmission at the affected receivers. As **PIM-SSM** depends on the BGP, the obtained traffic interruption result quite high in the experiments, **2 minutes and half** approximately. In fact BGP tends to explore all alternatives (problem known as path exploration) before reaching a stable state and, as a consequence. In the case of **GCMR**, there is no need of reaching a unicast topology state convergence and thus the recovery time obtained in the experiments is of the order of **one second**.

[1] H. Holbrook, B. Cain, “Source-Specific Multicast for IP”, *IETF RFC 4607*, Aug. 2006.

[2] C. Diot, et al., “Deployment Issues for the IP Multicast Service and Architecture”, *IEEE Network*, vol. 4, no. 1, Jan. 2000.

[3] P. Pedroso, D. Papadimitriou, D. Careglio “Dynamic compact multicast routing on power-law graphs”, *Globecom 2011*, Houston, TX, USA, Dec. 2011.

Call for papers

Student Workshop at ACM CoNext 2013 **26/09/2013**
http://conferences.sigcomm.org/co-next/2013/workshops/student/
December 9, 2013, Santa Barbara, CA, USA

IEEE Int. Conf. ICC **30/09/2013**
http://www.ieee-icc.org/
June 10-14, 2014, Sydney, Australia

ACM Trans. Internet Technology **30/09/2013**
Pricing and incentives in networks and systems

IEEE J. Selected Areas in Communications **01/10/2013**
Energy-efficiency in optical networks

10th Int. Conf. DRCN **18/10/2013**
http://www.drcn2014.org/
April 1-4, 2014, Ghent, Belgium

13th IFIP Int. Conf. Networking **26/11/2013**
http://networking2014.item.ntnu.no/
June 2-4, 2014, Trondheim, Norway

IEEE Infocom Demo Session 2014 **15/12/2013**
http://www.ieee-infocom.org
April 27-May 2, 2014, Toronro, Canada

ACM Sigcomm 2014 **31/01/2014**
http://conferences.sigcomm.org/sigcomm/2014/index.php
August 17-22, 2014, Chicago, IL, USA

EULER Summer School 2013

Graph and routing dynamics: models and algorithms

A summer school sponsored by EULER project (EULERSS’13) took place on July 1-5, 2013 in Barcelona, Catalunya, Spain. This event focused on current research and related challenges on Internet routing paradigms for distributed and dynamic routing schemes applicable to the current Internet and its evolution. The goal of this summer school was i) to stimulate research in the interdisciplinary area that lies at the intersection of graph theory, distributed routing algorithmic and network dynamics modeling, and ii) to provide a forum for active discussions among teachers/researchers and students.

EULERSS’13 consisted of 9 courses of 3 hours each organised in 4 sessions:

- **Session 1: Internet routing system**. The purpose of this session was to expose the students to the foundations of Internet architecture, its routing system, and the latest research achivements and activities.
- **Session 2: Algorithmic graph theory**. The focus of this day was on understanding basic properties of graphs introducing most of the classical concepts of pure and applied graph theory. In addition, the use of effectively algorithmic techniques for various optimisation problems on graphs was presented and practical application discussed.
- **Session 3: Dynamic graph modelling**. One of the geometric properties of graphs underlying large-scale topologies (such as the Internet) is the hyperbolic nature of its topology, useful to enable navigation/routing. The main of goal this session centered in understanding this branch of graph theory in order to exploit its potential.
- **Session 4: Routing models and algorithms**. The topological structure of a network has particular mathematical properties that set them apart from random graphs. The courses of this session focussed on these properties of a network that marked influence on the processes (e.g. routing) that may take place on it. Mathamatical programming, combinatorial optimisation as well as graph and networks models for some routing problems were finally presented.

Finally, 50 students (master students, PhD candidates and researchers) participated to this succesful event. Among them, 8 students presented their work in progress research activities in an ad-hoc session and received comments/feedbacks from the lecturers and local assistants.

Forthcoming EC and FIRE events

XIFI Open Call Information Day **26/09/2013**
https://www.fi-xifi.eu/open-call.html
Brussels, Belgium

Future Networks Concertation meeting **22-23/10/2013**
https://ec.europa.eu/digital-agenda/en/future-networks-concertation-meeting
Brussels, Belgium

Shaping FIRE for Horizon 2020 **24/10/2013**
http://www.ict-fire.eu/events.html
Ghent, Belgium

ICT 2013 “Create Connect Grow” **06-08/11/2013**
http://ec.europa.eu/digital-agenda/en/ict-2013
Vilnius, Lithuania