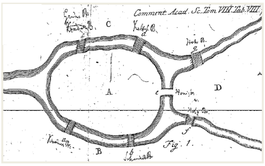


THE EULER NEWSLETTER

In this issue

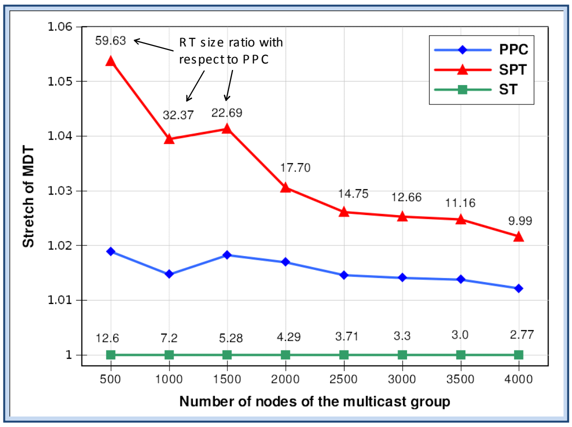
Compact multicast routing

With the advent of multimedia streaming/content, multicast distribution from a source to a set of destination nodes is (re-)gaining interest as a bandwidth saving technique competing with or complementing cached content distribution. Nevertheless, the scaling problems faced in the 90's when multicast received main attention from the research community remain unaddressed since so far. Indeed, overlaying multicast routing on top of unicast suffers from the same scaling limitations as current unicast routing with the addition of the level of indirection added by the multicast routing application.

To overcome the scalability of unicast routing, compact routing[1] schemes aim to find the best tradeoff between the memory-space required to store the routing table entries at each node and the stretch factor increase on the routing paths it produces. In this direction, a dynamic compact multicast routing algorithm, referred hereafter as AMR[2] (Abraham, Malkhi and Ratajczak), was proposed enabling the construction of point-to-multipoint (P2MP) routing paths (commonly referred to as the Multicast Distribution Tree (MDT)) from any source to any set of destinations (leaves). AMR is a name-independent, and root-initiated dynamic compact multicast routing scheme. In this context, "name-independent" means that nodes have arbitrarily chosen names, and "dynamic" refers to the on-line capability to timely process the join/leave requests as they arrive without re-computing and rebuilding the MDT from scratch. The AMR scheme employs a bundle of sparse covers (notion introduced by Awerbuch and Peleg[3]) constructed from a set of center node. To each node, a per sparse cover degree is associated that stores all labels for routing the vicinity ball of that node. In addition, each node stores the routing information for every tree cover containing that node. When a leaf node joins an MDT, it first determines: i) If one of the MDT node is already included in its local tree information table, it selects the center node with minimum degree cover that is associated to that MDT node; ii) Otherwise, the leaf node queries the source node to obtain the set of MDT nodes it currently includes and selects the MDT node that is part of the minimum intersection with its bundle. Once determined, in both cases the leaf node sends the join request to the center node which passes a label so that the selected MDT node can forward traffic to the newly joining leaf. The leaf node updates all nodes in its vicinity to allow them joining the MDT.

In the framework of EULER, ALB and UPC partners are collaborating in the design of PPC[4] (Pedroso, Papadimitriou and Careglio). PPC is a name-independent compact multicast routing scheme for leaf-initiated, distributed and dynamic construction of MDT. In this context, "leaf-initiated" means that the join/leave requests are initiated by the leaves and "distributed" implies that transit nodes process the join/leave requests and compute the routing table entries (no centralized processing by the root). The objective of PPC is to minimize the routing table sizes of each node. To this end, PPC reduces the local storage of routing information by keeping only (direct) neighbour-related entries rather than tree structures (as in Steiner Tree (ST) solution) or network graph entries (as in both Shortest Path Tree (SPT) solution and ST). The first version of PPC only allows dynamic join events, i.e. a leaf node joins the MDT by searching for the least cost branching path towards a node that belongs to MDT. The information needed to reach the MDT is acquired by means of a discovering search mechanism that returns the upstream node along the least cost branching path to the MDT. Since higher cost may hinder PPC applicability to large-scale topologies such as the Internet, the algorithm's search process is segmented in two different stages by executing first a local search covering the leaf's neighbourhood (the so-called vicinity B), and, if unsuccessful (no nodes belonging MDT are found), executing a global search over the remaining topology. The vicinity B is determined by the algorithm itself at each search process on-the-fly by means of a budget cost value; such value is decremented at each hop by the link cost. The rationale is to put tighter limits and search locally before search globally. Indeed, the likelihood of finding a node of the MDT within a few hops distance from the joining leaf is high in large topologies (whose diameter is logarithmically proportional to its number of nodes) and it increases with the size of the MDT.

The performances of this version of PPC are analyzed by its simulation on the CAIDA map of the Internet topology which comprises 16k nodes. We compare the performance of the proposed CMR algorithm to the SPT and the ST algorithms. The figure on the right-hand side illustrates the stretch ratio of the multicast routes (i.e. MDT) set up by the PPC and the SPT algorithms compared to the ST reference algorithm. The multiplicative stretch for the PPC is slightly higher than 1 with a trend curve decreasing as the multicast group size increases. The label in the figure indicates the RT size ratio (in term of number of RT entries) with respect to the PPC algorithm which presents the slowest reference value. PPC obtains outstanding performance. Finally, the table shows the communication cost in term of number of messages. The results for PPC are relatively high compared to the SPT even if much lower than the communication cost implied by the ST. However, this communication cost does account for the evolution of the routing topology. Work is ongoing to further reduce this cost (realistically up to a factor 10.).



Communication cost (in terms of number of messages) vs. multicast group size

	Multicast group size							
	500	1000	1500	2000	2500	3000	3500	4000
PPC/SPT	13.86	15.52	16.62	17.24	17.52	17.67	17.77	17.82
ST/SPT	939.2	1760.6	2494.1	3188.5	3851.3	4485.4	5059.7	5599.6

1 D. Peleg et al., "A trade-off between space and efficiency for routing tables," J. ACM, Jul. 1989.
2 I. Abraham et al., "Compact multicast routing," 23rd DISC'09, Sep. 2009.
3 B. Awerbuch et al., "Sparse partitions," 31th FOCS, Oct. 1990.
4 P. Pedroso et al., "Dynamic compact multicast routing on power-law graphs," accepted paper, IEEE Globecom 2011, Dec. 2011.

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

How the EULER project is progressing?

The EULER project, started on October 1st, 2010, completed the second quarter of the 3-year project duration.

A technical project meeting took place on 9-12 May 2011 at Villa Cynthia, Frejus, France. It consisted of three-days meeting in the form of "retreat". The first full day has been dedicated to topology models, generation, properties verification techniques and tools. Routing system architecture, procedures/algorithms, and formal verification and associated tools have been the topics of the discussion of the second day. Finally, the last day focussed on the experimental evaluation and comparison of routing protocols.

Regarding the WPs activities, WP2 started on M04 (January 2011) with the task T2.1 dedicated to *Routing system architecture*. The work plan for this task comprises two parts: Part 1 "Routing system architecture specification" (from M04 to M07) and Part 2 "Routing models" (from M07 to M12). During Q2, work concentrated on Part 1 and in particular on the Functional modelling and the Information modelling. A functional model based on Enhanced Functional Flow Block Diagram has been considered as a common framework from which any specific routing scheme will be derived. An information model based on Enhanced Entity-Relationship model has been considered as the model representation of concepts, relationships, properties, constraints, rules, and operations to specify data semantics for routing systems.

In T3.1, the main efforts have been dedicated to the preparation of deliverable D3.1. Such deliverable will include a overview of models and properties. Regarding models, a large number of existing models, mostly static but also dynamic, have been documented such as power-law/degree based models, structural models, positive-feedback-preference model, etc. Regarding properties, a set of them useful for compact routing schemes (such as chordality, treewidth, etc.) as well as a set of properties useful for greedy routing on hyperbolic space (such as clustering, centrality, hyperbolicity, etc.) have been identified.

In T3.2, the work on 4 different subtasks progresses. In subtask 1, UPMC continued to work on the design and implementation of a method for accurate and reliable measurement of the Internet topology using PlanetLab monitors. In subtask 2, UPMC focused on analyzing the dynamics observed by the Radar tool. In subtask 3, UPMC worked on event detection while UPC and ALB on determining stability of the routing system and defining pertinent stability metrics. In subtask 4, UPMC and UCL started the analysis of statistics of interests of P2P traffic.

Concerning WP4, the efforts on deliverable D4.1 were finalized under the lead of editor ALB. The deliverable is responsible for describing the performance objectives, evaluation criteria, and metrics for the routing schemes to be experimented in the context of the defined routing system architecture in T2.1. The resulting content will serve as input for the definition of the experimental methodology, scenarios and tools as planned for T4.2.

In the context of WP5, a newsletter topical program until December 2011 has been established. Dimitri Papadimitriou (ALB) and Albert Cabellos (UPC) participated to the IETF80 held in Prague, Czech Republic on March 27-April 1, 2011. Dimitri Papadimitriou (ALB) and Florin Coras (UPC) participated to the FIRE week held in Budapest, Hungary on May 16, 2011.

For more information: **http://www.euler-fire-project.eu.**

Upcoming events

FIRE Open Calls Information Day **14/09/2011**
http://www.ict-fire.eu/events/meetings/2ndfireopencallsinformationday.html
Brussels, Belgium

25th International Symposium DISC **20/09 - 22/09/2011**
http://disc2011.dis.uniroma1.it
Rome, Italy

Future Internet Conference Week **24/10 - 28/10/2011**
*** Future Internet Assembly** **25/10 - 26/10/2011**
*** FIRE Conference** **26/10 - 27/10/2011**
http://www.fi-poznan.eu/
Poznan, Poland

Presentation of the TERA-NET workshop

EULER will organize the 2nd EULER International Workshop (named TERA-Net). This event will be held during the 25th International Symposium on DIStributed Computing (DISC), on September 20-22, 2011 in Rome, Italy.

The Internet routing system is facing performance challenges in terms of scalability (growth rate of the Border Gateway Protocol (BGP) routing tables) and in terms of dynamics of the routing information exchanges (convergence, and stability/robustness) that result into major cost concerns for network designers but also protocol designers. There is a growing consensus among the scientific and technical community that the current practice of "patching" the forwarding and BGP routing protocol of the Internet will not be able to sustain its continuous growth at an acceptable cost and speed. On the other hand, the Internet size and scope make the deployment of new routing scheme(s) extremely challenging. Recent advances in routing algorithmics take benefit of the statistical and structural properties of the Internet topology and better characterization of its dynamics. The research domain dedicated to new routing paradigms aims to design distributed routing schemes that are specialized for the Internet while taking into account its dynamics and its continuous evolution. From this perspective, the goals of the TERANET full-day workshop are 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 speakers and participants.

Agenda

The second edition of the TERANET full-day workshop structures around two sessions, each dedicated to a cross-disciplinary topic.

- Introduction by Workshop Chairs
- Session 1 (morning): Network dynamics modeling
  Three talks (about one hour each including questions)
- Lunch Break
- Session 2 (afternoon): Distributed routing algorithmics
  Three talks (about one hour each including questions)
- Panel / Concluding talk

Organizing Committee

- David Ilcinkas (University Bordeaux 1, Bordeaux, France)
- Dimitri Papadimitriou (Alcatel-Lucent, Antwerpen, Belgium)
- Josep Lluís Marzo (University of Girona, Girona, Spain)

Full details about the workshop are available on **http://disc2011.dis.uniroma1.it/workshop.php?lang=eng**.

Call for Papers

IEEE/IFIP NOMS 2012 **22/08/2011**
http://www.ieee-noms.org/
April 16-20, 2012, Maui, Hawaii, USA

IEEE ICC 2012 **06/09/2011**
http://www.ieee-icc.org/2012/
June 10-15, 2012, Ottawa, Canada

13th PAM 2012 **14/09/2011**
http://pam2012.ftw.at/
Mar. 25-30, 2012, Orlando, Florida, USA

Computer Networks journal **15/09/2011**
Green Communication Networks

Computer Networks journal **01/10/2011**
Challenges in High Perf. Switching and Routing in the Future Internet

IEEE Communications Magazine **01/11/2011**
Information-Centric Networking

11th IFIP Networking 2012 **10/12/2011**
http://networking2012.cvut.cz/
May 21-25, 2012, Prague, Czech Republic