Global Rendezvous Scale Issues for an Information-Centric Future Internet

Tuesday, September 11, 2012 2:09 PM (12 minutes)

The Past

In the 1990's we got involved with Internet traffic characterization [A, D] and in particular traffic-flow profiling [B], as well as Multicast and Continuous Media Dissemination [C, J, K, L] and Wireless Internet Multimedia Communications [F, G, H]. Those investigations involved vertical understanding (for example, continuous media dissemination feasibility and contributing issues throughout the layers), horizontal understanding (for example, considered long Internet paths with one or more wireless links at various points in the path and its impact on performance [F, G, H]), and large-scale computational modeling and analysis (for example, required the use of real-time sampling of the data to cope with its rate and size [E]) of the Internet or networks in general.

• Internet Traffic Characterization: We introduced the concept of IP flows and used it to characterize real Internet traffic at various levels of granularity. The notion of IP flows provides a flexible tool for bridging the gap between the connectionless/stateless networking model of the Internet's (inter)network layer and the connection-oriented/stateful model more appropriate for some applications (e.g., packet video).

• Multicast and Continuous Media Dissemination: We have contributed on various aspects of multicast protocols and multimedia multipoint communications. IP Multicast for point-to-point and wireless networks with mobility has been investigated. We have also developed efficient multimedia dissemination techniques that support heterogeneity (in both the terminals and network paths) and effective congestion control in packet switching networks using hierarchical coding of continuous media, such as real-time video.

• Wireless Internet Multimedia Communications: The Internet protocols were designed with wireline networks in mind and perform rather poorly in wireless environments. We contributed to the understanding of the problem and the awareness of the community, in addition to proposing a framework to address them in a realistic, effective, general, and efficient way.

The Future

It has now been realized for long that the Internet has evolved from an internetwork for the pairwise communication between end hosts to a substrate for the delivery of information. The users are increasingly concerned with the content they are accessing (or contributing), rather than the exact network end point providing it. This major shift has resulted in the emergence of a series of new technological enablers for the efficient delivery of content, ranging from application layer solutions (e.g. CDNs), to proposals for new, clean-slate designs for the Future Internet based on the Information-Centric Networking (ICN) or content-centric networking paradigm ([U-V], [1],[2],[3]).

In all these efforts, the act of locating the desired content in the network (e.g., through name resolution) has been regarded as an increasingly challenging task, facing serious scalability and complexity concerns. The huge volumes of available content in the Internet, especially with the advent of user generated content, have resulted in a correspondingly enormous name space challenging even the management of meta-data information and the act of locating the desired content. Considering that the current amount of unique Web pages as indexed by Google is greater than 1 trillion [4] and that some billions [5] of devices, ranging from mobile phones to sensors and home appliances are joining the network to offer additional content or information, we could be safely speaking that an ICN has to manage a number of Internet Objects (IOs) in the order of 1013. (Other studies raise this estimate to 1015 [6].) At the same time the large size of the Internet ecosystem adds to the scalability concerns, since the need for efficiently locating the desired content spans several thousand networks (more than 35K ASes reported in the latest CAIDA trace set), with hundreds of thousands of routers. Moreover, the vast number of (mobile) end host devices, is not only contributing to the huge volume of content, but also to a considerably high volume of requests for content.

Though major research efforts have been devoted to building highly scalable name resolution systems, locating information in the current (and future) Internet is further facing significant complexity challenges. The current Internet landscape is a mosaic of different actors. A multitude of different producers of content, ranging from simple end users to large content providers, is offering large volumes of content. Each content provider may require the establishment of different access rights and privacy policies. The provided content must be discovered and delivered across a multitude of distinct networks under different administrative authorities,

often following complex routing policies dictated by economic parameters. At the same time, the emergence of large CDNs introduces a further layer of complexity by allowing the replication and caching of content at several parts of the internetwork, driven by end user demand. In addition, different types of access networks (e.g., ADSL, Wi-Fi, 3G/LTE, 4G…) and end user devices (tablets, smartphones, laptops, etc.) introduce further complexities for the adaptation of content according to the current context.

It therefore becomes evident that locating the desired content in the current (and anticipated future) Internet is a task that has many dimensions that call for careful consideration. Until now most research efforts on this challenge were focused on particular aspects, often investigating a limited subset of the involved parameters in isolation e.g., using simplified inter-domain topologies, conducting small scale simulations, neglecting aspects such as content replication, etc. Hence, a horizontal understanding is required taking into account the entire set of the aforementioned aspects and the interactions among them. A series of important questions cannot be answered unless a holistic view on this landscape is taken e.g., how does the heterogeneity of the Internet impact on the mechanisms employed to locate the content? What is the impact on the performance of a name resolution system? How would the exchange of information (meta-data) between different actors affect the operation of such a system in terms of reachability of content?

In order to gain a better understanding of this issue, we need to simultaneously model a series of practical aspects and features stemming exactly from this diversity. Namely, we have to model aspects such as: (1) the generation of new content in the Internet, (2) the temporal evolution of the popularity of the different types of content available, (3) the locality characteristics of end user requests, (4) the (current) content replication/caching policies of CDNs, (5) both the inter-domain and intra-domain level topology characteristics, (6) (Inter-domain) Routing policies, (7) the implications introduced by wireless networks i.e., content tailored for mobile devices, smart phones, (8) the implications introduced by the Internet of Things (IoT) e.g., high volumes of information, geographical characteristics, access patterns, and socio-economic aspects.

This complicated set of interacting issues and models, is expected to impact the investigation of the various issues emerging from the interaction of the respective actors, contributing to the realistic investigation of currently available, as well as new mechanisms for locating content in the Internet. For instance, today it is difficult to assess the potential benefits and pitfalls stemming from the establishment of a synergy between CDNs, content providers, and ISPs expressed via the exchange of meta-data related to the discovery of the closest replica of some content.

Some first steps in these investigations have been undertaken in research projects we have or are participating, but we are only at the very beginning ([1, O-T]).

References in comments, below.

Primary author: Prof. POLYZOS, George (MMlab/AUEB)

Presenter: Prof. POLYZOS, George (MMlab/AUEB)