Over Provisioning-Centric QoS-Routing Mechanism for the Communication Paradigm of Future Internet 4WARD Proposal

Augusto Neto[1], Sérgio Figueiredo[2], Rui Aguiar[2], Leandro Marçal[1], Leandro Freitas[1]

[1] Federal University of Goiás, Informatics Institute, Bloco IMF I, Campus II, Samambaia
74001-970 Goiânia-GO, Brazil

[1] Institute of Telecommunication, University of Aveiro, Campus Universitário de Santiago
3810-193 Aveiro, Portugal

{augusto, leandro, leandrocc}@inf.ufg.br
ruilaa@ua.pt, sfigueiredo@av.it.pt

Abstract. The FP7 4WARD clean-slate Project envisions overcoming the limitations of current Internet by redefining it to efficiently support complex value-added sessions and services, such as location-based, health-care, critical-mission, and geo processing. The list of networking innovations from 4WARD's *Future Internet* (FI) proposal includes a new connectivity paradigm called *Generic Path* (GP), a common representation for all communications. From the networking point of view, a GP is mapped to a communication path for data propagation. For that, GP architecture relies on routing mechanism for selecting best communication paths. In order to assure reliable communications, the routing mechanism must efficiently provision QoS-aware multi-party capable paths, with robustness functions, while keeping network performance. Therefore, this paper proposes the *QoS-Routing and Resource Control* (QoS-RRC) mechanism to deal with the hereinabove requirements by means of an over provisioning-centric (bandwidth and paths) approach. QoS-RRC achieves scalability by avoiding per-flow operations (e.g., signaling, state storage, etc.). Initial QoS-RRC performance evaluation was carried out in Network Simulator v.2 (NS-2), enabling drastic reduction of overall signaling exchanges compared to per-flow solutions.

Keywords: Future Internet, GP Architecture, QoS routing, Over-provisioning.

1 Introduction

The FP7 4WARD Project [1] is a clean slate proposal based on the lessons learned from the past which, along with a visionary perspective, takes into account the need for support of different types of sessions and ubiquitous access with guaranteed resources over the time via multi-party (unicast, multicast or both) connection. Among its features (e.g. network of information, autonomous management), the *Future Internet* (FI) architecture proposed by 4WARD includes a new connectivity paradigm called *Generic Path* (GP). The GP architecture aims to optimize connectivity complexity (introduced by heterogeneity) via a generic communication representation. The GP architecture is specified with native functions for reliability, robustness, Quality of Service (QoS), multi-party, mobility, etc., independent of technology or architectural layer. Thus, GP expects to drastically facilitate the specification and inclusion of new applications and services. In order to allow data transport over GPs, the architecture support native functions to map sessions into physical resources. IN the network scope, such resources can be links, network interfaces, routers, etc. The GP architecture relies on routing mechanisms to provide best network resources for reliable and pervasive communications over the time.

The time-sensitiveness characteristic of some applications, such as (real-time) multimedia sessions, demands requirements that cannot be efficiently fulfilled by current Internet routing approach. Among other aspects, resource-constrained sessions require enough bandwidth guaranteed over the time and robustness for intermittent data propagation beyond the traditional best-effort transport paradigm. Thus, QoS support is of paramount importance for routing, and network resources must be provisioned so that sessions can experience low rates of propagation delay, jitter and packet loss end-to-end (E2E). The literature shows that current QoS-routing solutions (e.g., [2][3]) are inefficient, mainly in terms scalability, where per-flow approach place excessive signaling load, energy consumption and CPU/memory with control functions, most of them in a centralized way [4].

Therefore, this paper proposes the *QoS-Routing and Resource Control* (QoS-RRC) mechanism to improve GP architecture in the networking scope. QoS-RRC aims to allow dynamic GP setup fulfilling session requirements in terms of QoS and multi-party connectivity over heterogeneous environment. It's over provisioning-centric approach faces scalability issues, and takes distributed decisions without per-flow resource control operations. Thus, QoS-RRC is expected to allow the support of many types of time-sensitive sessions while keeping overall network performance in comparison with per-flow solutions.

This paper initially introduces the GP architecture, describing its mechanisms and elements. Afterwards, a database for storing and accessing GP information, the GP Resource Management Database, is described. Further on, we described in details QoS-RRC mechanism. The output of this work is presented by a performance evaluation of QoS-RRC on Network Simulator v2 (NS-2). Finally, resulting conclusions and future work are presented at the end.

2 Generic Path Architecture Overview

In order to generalize and abstract communication without being held by current Internet model, GP goals are [5]: *i*) to develop foundations for both describing and prescribing any network communication; *ii*) provide a generic communication service model not limited by any communication paradigm; *iii*) design architectural constructs and primitives from this framework. As GP architecture is object-oriented, its features include: *i*) excursiveness, allowing description of a communication in a self-similar way; *ii*) controlled opacity and virtualization, with state and functionality at any level accessible from any level; *iii*) agnostic, relatively to technology, platform and communication; *iv*) and modularity, providing the possibility to establish communication contexts and federations of network services, among others, such as polymorphism or overloading. The combination of these and other characteristics make GP architecture a powerful one for network design and description, while selectively hiding the management complexity if desired.

2.1 GP architecture and its elements

In GP architecture, the clients of a GP are Entities (ET). In order to communicate, Entities must use the same communication protocol, established within a Compartment (CT), which is also associated with a determined namespace.

Nodes (terminal or network ones) are represented by a Node (vertical) CT, where ETs exist. An ET may model a running service at any level and generalize a communication data processing function. ETs communicate with each other horizontally or vertically, using, respectively, a GP or a Hook (the combination of 2 Ports). Mapping to a typical process-based system like UNIX, an ET abstracts a process, Hooks abstract inter process communication and Ports abstract process IDs/file descriptors.

ETs access GPs services through *End Points* (EP). While ETs are related to the control and management of GPs (service discovery, routing, name resolution, .etc), EPs are related to data transfer and control (error control, flow control, encryption, coding, etc). The basic GP architecture can be depicted from Fig. 1. A GP comprehends the set of necessary resources that provide end-to-end communication between two or more ETs. A diversity of services may be represented by a GP, so it needs to be contextualized: the scope of the communication is delimited by the CT in which the GP is running, and may for example correspond to a network, an application or procedure (local or remote).

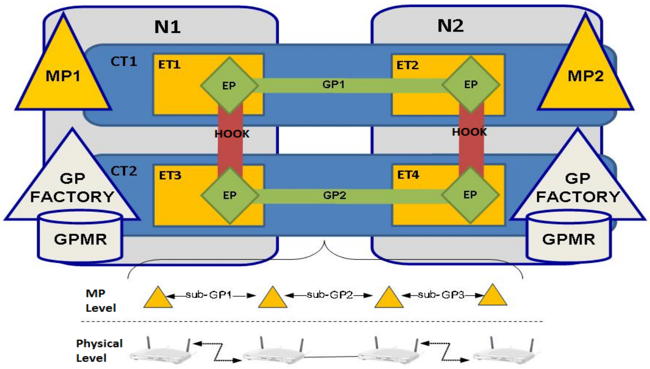

Fig. 1. GP Network Example Scheme

The GP representing the communication at its highest level of abstraction, named the *end-to-end* (E2E) GP, allows the ubiquitous view of the path, and is composed of other GPs, each implementing a different service technology. For instance, a VoIP call between 3 users would be represented by a single E2E GP composed from (for example) an UDP GP, 2 Wi-Fi GPs, a VoIP application GP, and any other necessary GPs for that session to exist. All GPs requests are sent to a GP Factory (one per Node CT), which begins by checking the existence or not of a previous GP, and is responsible for node CT-internal information exchange, such as ETs instantiation for a particular CT. The CT works as an application framework, enhancing the GP with the GP class specific structure. A *Mediation Point* (MP) is necessary for operations, as data pipelining between GPs (GP interconnection).

2.2. GP Resource Management Database

In GP-aware networks, data is transported by managing (creation, modification and destruction) of multiple GPs. Each GP is characterized by the following Items: 1) associated EPs; 2) sub-GPs: and 3) resources / attributes (QoS metrics, privacy settings, or any interesting attribute). The access to the information is possible through Dials, for information inspection, typically state information, and Knobs, related to configurable information.

In order to provide a proper resource management, according to the communication context (deployed service, centralization, connection-orientation, etc), a framework for collecting and controlling the path diverse information is necessary. The use of specific records fills the gap by creating a database for storing and inspecting GP information. A record maintaining the high-level GP information is introduced in [6] – the *Master Record* (MR) -, depicting the relationship between records holding information of composing (sub) GPs – the GPMRs. MR holds a view of all lower-level GPs running associated to the same communication.

Adding the concept of *Compartment Record* (CTR), a record listing the GPMRs running in a CT allowing the establishment of a distributed hierarchical scheme, may refine the previously proposed solution. This way, MR points to a list of CTRs, one for each CT in which the Node CT participates. The GP needs proper control, and as the element that stores all the relevant information about it, the GPMR can achieve such a necessary feature, associated with the EP. Additionally, such framework brings an innovative approach to today's networks through a unified framework to control the path.

At the highest abstraction level, a GP is the result of the composition of lower level GPs, one for each of the technologies that the communication depends on (TCP, Ethernet, Optical Fiber, etc). In order to allow GP selection/comparison, each E2E GP is described with very generic attributes, namely throughput and E2E delay. In sub-GPs, the managed resources range from statistics such as throughput, SNR and end-to-end delay, to privacy and policies data, such as packet priority level and identification keys. The nature of the stored information and the record structure depends on the class of the GP, and therefore, one GPMR class exists for each GP class. All GPMR classes derive from a Base GPMR class, which means, an instantiated GPMR structure will always depend on the type of associated GP attributes Class, which will be sub-classed according to such characteristics as connection-orientation (stream vs. datagram), physical related properties (wireless vs. wired) or number of destinations (unicast vs. multicast vs. broadcast). Hence, a GP will have a number of very specific GPItems that are related to the technology or service it refers to, as in Fig. 2.

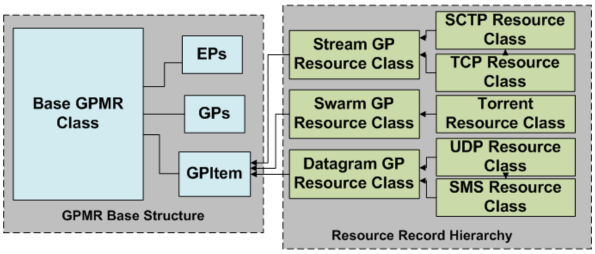

Fig. 2. GPMR and Resource Classes

The bootstrap of these records is as follows: *i*) the MR is initialized at the Node CT's creation; *ii*) a GPMR is initialized, being filled with the GP's characteristics; *iii*) if the GP is of a non-existing type to the node CT, a CTR is also initialized, otherwise the existing CTR is updated, i.e., the GPMR is added to its list. The level of management centralization is set according to the communication needs by the way these MRs are used: the basic model idea is to implement a distributed database, with each node having a view of GP resources of the GPs it is part of. On top of this model it will be possible to build any other. For some CTs, it will make sense to hold a more centralized database (e.g. Windows NT domain CT); another option is to use a more distributed database with partial views on each node (e.g. Delay Tolerant Networks or Wireless Mesh Networks). Another case may be maintaining a replicated database (e.g. unstructured P2P network). A strong point in such method is avoiding the limitations of a “one-model fits all” approach.

Matters such as scalability can therefore be assured by associating to each service a specific record organization, thus providing stability to the network and enabling presently dysfunctional mechanisms such as multipath routing or multihoming.

3 QoS-RRC Overview

The QoS-RRC proposal addresses the support for GP factoring to fulfill the requirements of interested ETs, while preventing system performance degradations. This way, QoS-RRC complements the GP architecture with a routing-enabled QoS control approach with operations following a dynamic network resource over-provisioning strategy for scalability. Such strategy associates admission control to setup multiple sessions without per-flow signaling. Moreover, QoS-RRC maintains a path state table bootstrapped with surplus shortest QoS-aware multi-party paths inside the network for routing. Therefore, a broadly network view associated with metrics describing QoS (bit-rate, tolerance to delay and loss, and destination) and transport (unicast and multicast) capabilities allow routing more efficiently than current proposals. In the latter, current QoS-routing proposals mainly keep next hop information (limited topology knowledge) with resource capabilities classified by inefficient metrics and maintained by constant flooding operations. QoS-RRC follows our previous works [7][8], which distinguishes itself by scaling and resource control capabilities.

3.1 Over provisioning-centric QoS-routing

The over provisioning-centric QoS-routing strategy follows our patented ideas described in [9]. The main idea consists in allowing QoS-RRC maintaining a routing table with surplus shortest QoS-aware multi-party paths. The routing table carries information about the paths (IP address of each router's interface along the path) and QoS capabilities (per-class bandwidth available and rates of delay, jitter and loss of bottleneck's link). The routing table is composed at the system bootstrap, by means of a flooding-controlled scheme. To that, each ingress node floods the network, and each visited deploy following operations. Firstly, each class of service of the outgoing interface is assigned with an over-reservation based on pre-configured factors (e.g., ½, ¼ of link capacity). Over-reservations are dynamically controlled according to session demands and network conditions (described in section 3.2). Next, the outgoing interface's IP (only if not present, to avoid infinity looping) and associated QoS capabilities (if local information is worst than those derived from the message) are added in the message. Finally, the message is flooded downstream. This flooding collects best QoS-aware unicast paths (one ingress, a set of core and one egress router per path).

After finishing the flooding cycle, each ingress router takes the unicast path information to over-provision shortest multicast path information. Thus, QoS-RRC implements an algorithm that matches paths from each ingress router towards reachable egress routers, where a limited number of nodes along the tree can be assigned. Pre-established matching's are discarded by filters to keep best multicast trees. For instance, trees with ingress routers in the middle are discarded to keep only downstream possibilities for efficiency. Afterwards, QoS-RRC signals each path to install multicast state. As depicted in Fig. 3, QoS-RRC is hosted at nodes implementing MPs.

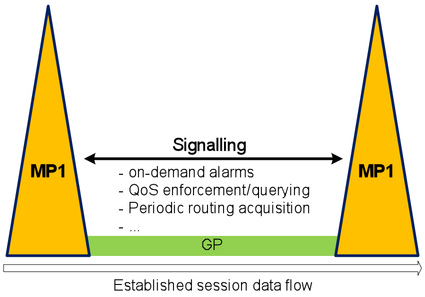
GP Factory
QoS-RRC
Connection request
EP
Routing Table
GPMR
Signalling and QoS enforcement
- Available paths
- QoS bottleneck capacbilities
- GP information
MP1
Signalling
- on-demand alarms
- QoS enforcement/querying
- Periodic routing acquisition
- ...
MP1
GP
Established session data flow

Fig. 3. QoS-RRC architecture

Furthermore, a connectivity control mechanism of QoS-RRC allows reducing state and processing burden of legacy IP multicast per-flow operations, by aggregating multiple GPs into the same path (tree). Moreover, such mechanism allows fast resilience for intermittent communications, where GPs are re-routed seamlessly to cope with dynamic network events (e.g., link failure). Next section describes how QoS-RRC manages over-reservations on-demand.

3.2 QoS Over-provisioning Mechanism

The QoS-RRC suite provides QoS support by associating admission control and dynamic bandwidth provisioning. In this sense, aggregated bandwidth control (per-class) overmatches per-flow, since the latter is well known as inefficient due to the exceeding control overhead. QoS-RRC controls bandwidth over-reservations following our resource control apparatus patent [10]. As described in Section 3.1, initial per-class over-reservations are bootstrapped at the flooding cycle. Afterwards, QoS-RRC controls per-class over-reservation re-sizing dynamically, taking into account session demands, statistics about network conditions and QoS policies. The integration of admission control and per-class over-reservation foresees network performance improvements, where multiple sessions (GPs) can be accommodated without per-flow signaling, as much as possible.

After having initiated all over-reservations and supplied the routing table, the ingress routers keep waiting for GP setup requests. GP setup operations succeed without network signaling as log as selected paths have available resources to accommodate the requesting GP. Whenever QoS-RRC realizes that the potential path mapping a GP has no bandwidth in the demanding class for its accommodation, it attempts to calculate a new over-reservation for the demanding class. If none of the potential paths can accommodate the GP (no available bandwidth for re-sizing), the request is denied.

4 QoS-RRC Performance Evaluation

The performance evaluation of QoS-RRC was carried out in the *Network Simulator* v2 (NS-2) [11] to examine expected benefits. NS-2 was extended with both GP and QoS-RRC functionalities accordingly. The simulation model uses a network topology, generated by BRITE (NS-2 topology generator based on real scenarios), of 14 routers with links of 100Mb/s and varying propagation delay. For class-based traffic differentiation, DiffServ was extended with WFQ scheduling discipline. Classes are configured the following way: one EF-alike (Premium); two AF-alike (Gold and Silver); and one best-effort. In order to avoid class starvation, each class was assigned with a maximum reservation threshold of 20% of the link capacity, thus leaving best-effort with 40% of the resources ensured. For multi-party support, a PIM-SSM implementation was used. In what concerns traffic patterns, the simulation model was scaled with 1,000 UDP multicast sessions with 256 Kbps of constant bit-rate for different classes (333 for Premium-alike, 333 for Gold-alike and 334 for Silver-alike). The corresponding session setup and session releasing events has been placed following a poison distribution. Three set of analysis have been deployed to study the impact of using QoS-RRC to establish sessions in the simulation model, where averaging results are plotted following a confident interval of 95%.

The signaling load is an important measure to study the impact of QoS-RRC in the overall system performance. Processing burden overloads with the increasing signaling load, because visited entities must at least inspect messages to further take any decision. The methodology applied in this set of experiments was to compare QoS-RRC behavior against a relevant per-flow solution. Therefore, we chose the Multi-service Resource Allocation (MIRA) [12] proposal, because it demonstrated several benefits in comparison with the current Internet *de facto* QoS approach, the Resource Reservation Protocol (RSVP) [13]. Fig. 4 shows the signaling load results of QoS-RRC and MIRA.

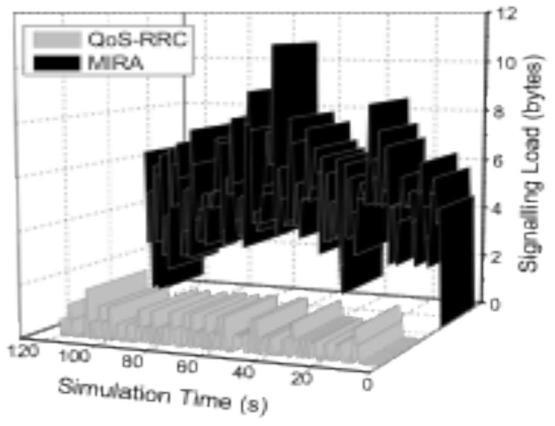

Fig. 4. Numerical results of signaling loading analysis

Results of Fig. 4 show that whereas MIRA averages 4.72 signaling Bytes during the simulation time, QoS-RRC places only 0.82 Bytes. Thus, the integration of admission control and per-class over-reservation demonstrates performance optimizations, with QoS-RRC reducing in more than 5 times the overall bandwidth consumption of signaling in the simulation model. Such reduction is achieved since multiple sessions have been admitted without any signaling exchange. MIRA introduced much more signaling load due to its per-flow basis for setting up each session. Thus we can conclude that QoS-RRC drastically improves overall network performance over MIRA.

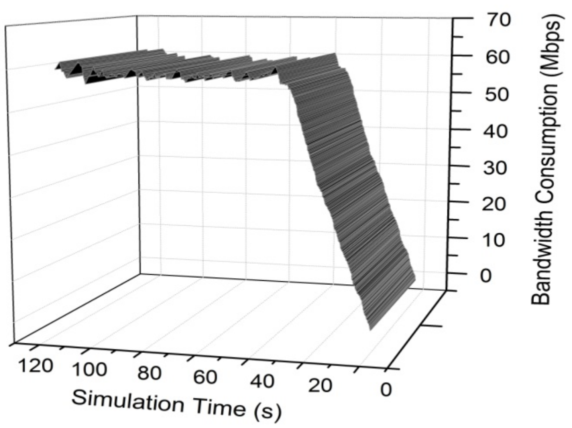

Fig. 5. Overall Bandwidth Consumption in the Simulation Model

Fig. 5 shows that the admission control function of QoS-RRC prevented congestions by not allowing excessive traffic of sessions, never exceeding 60Mbps. The quantitative results traced in simulation proved that even under congestion periods of some classes (not shown in Fig. 5), the resource management strategy of QoS-RRC allows efficiently controls over-reservations. Therefore, we can conclude that QoS-RRC resource over-provisioning strategy allows better bandwidth utilization than per flow analysis by avoiding waste of resources even during congestion periods or presence of residual bandwidth.

5 Conclusion and Future Work

The success of GP architecture depends upon the specification of routing and QoS control mechanisms, in order to correctly map physical resources (links, interfaces, etc.) for supporting the expected types of sessions. This paper presented a proposal to fulfill the above requirements, by extending the GP architecture with a suite of mechanisms addressing routing–enabled QoS approach, QoS-RRC, a mechanism capable of deploying integrated routing decisions and QoS control for allowing efficient GP mapping into physical resources. QoS-RRC is based on network resource over-provisioning to allow its operation without per-flow signaling, thus optimizing overall system performance while establishing resource-constrained sessions. Evaluations in NS-2 have shown the accomplishment of its expected benefits in terms of signaling load, blocking probabilities and bandwidth control.

In near future, the QoS-RRC is expected to be deployed through prototyping, by integration in the GP architecture, along with the GP Resource Management Database, allowing evaluation in terms of session setup times, latency of resilience and QoS control operations, and traffic analysis for user experience.

References

1. The FP7 4WARD Project, http://www.4ward-project.eu/
2. D.H. Lorenz, A. Orda, D. Raz, and Y. Shavitt. Efficient QoS Partition and Routing of Unicast and Multicast. In Proceedings IEEE/IFIP IWQoS, Pittsburgh, PA, June 2000.
3. Q. Ma and P. Steenkiste, Quality of Service Routing for Traffic with Performance Guarantees. IWQoS97, May 1997.
4. Manner, J., and Fu, X. (2005), "Analysis of Existing Quality-of-Service Signaling Protocols", IETF RFC 4094, May 2005
5. Randriamasy, S. et al., Mechanisms for Generic Paths, 4WARD deliverable D-5.2, December 2009
6. Figueiredo, S., Lourenço, J., Aguiar, R. L., and Neto, A., "Taxonomy for GP-aware mobility". Proc. of the First International ICST Conference on Mobile Networks and Management, 2009
7. Neto, A., Cerqueira, E., Monteiro, E., Mendes, P., (2008), "Scalable Resource Provisioning for Multi-user Communications in Next Generation Networks", In: IEEE Globecom 2008 Next Generation Networks, Protocols, and Services Symposium, 2008, New Orleans, LA, USA
8. Neto, A., Cerqueira, E., A., Monteiro, E., Mendes, P., (2008), "Scalable Multimedia Group Communications through the Over-provisioning of Network Resources", In: 11th IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services Management of Converged Multimedia Networks and Services (MMNS 2008), 2008, Samos Island
9. Neto, A., Curado, M., Monteiro, E., Mendes, P., (2009), "Method and apparatus for multicast tree allocation", European Pattent Officer EP2031796 (A1), March 2009
10. Neto, A., Curado, M., Monteiro, E., Mendes, P., (2009), "Method and apparatus for QoS resource reservation and configuration of multicast network resources", European Pattent Officer EP2037636 (A1), March 2009
11. The NS-2 Home Page, http://www.isi.edu/nsnam/ns/
12. Neto, A., Cerqueira, E., Rissato, A., Monteiro, E., Mendes, P., (2007), "A Resource Reservation Protocol Supporting QoS-aware Multicast Trees for Next Generation Networks" In: 12th IEEE Symposium on Computers and Communications (ISCC), 2007, Aveiro
13. Brades, R., Zhang, L., Berson, S., Herzog, S., Jamin, S. (1997) "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", IETF RFC 4094, September 1997