Experimenting with Real Application-specific QoS Guarantees in a Large-scale RINA Demonstrator

Jordi Perelló
Department of Computer Architecture
UPC BarcelonaTech
Barcelona (Spain)
perello@ac.upc.edu

Albert López
Department of Computer Architecture
UPC BarcelonaTech
Barcelona (Spain)
alopez@ac.upc.edu

Davide Careglio
Department of Computer Architecture
UPC BarcelonaTech
Barcelona (Spain)
careglio@ac.upc.edu

Abstract—**This paper reports the definition, setup and obtained results of the Fed4FIRE+ medium experiment ERASER, aimed to evaluate the actual Quality of Service (QoS) guarantees that the clean-slate Recursive InterNetwork Architecture (RINA) can deliver to heterogeneous applications at large-scale. To this goal, a 37-Node 5G metro/regional RINA network scenario, spanning from the end-user to the server where applications run in a datacenter has been configured in the Virtual Wall experimentation facility. This scenario has initially been loaded with synthetic application traffic flows, with diverse QoS requirements, thus reproducing different network load conditions. Next, their experienced QoS metrics end-to-end have been measured with two different QTA-Mux (i.e., the most accepted candidate scheduling policy for providing RINA with its QoS support) deployment scenarios. Moreover, on this RINA network scenario loaded with synthetic application traffic flows, a real HD (1080p) video streaming demonstration has also been conducted, setting up video streaming sessions to end-users at different network locations, illustrating the perceived Quality of Experience (QoE). Obtained results in ERASER disclose that, by appropriately deploying and configuring QTA-Mux, RINA can yield effective QoS support, which has provided perfect QoE in almost all locations in our demo when assigning video traffic flows the highest (i.e., Gold) QoS Cube.**

Keywords—RINA, large-scale experimentation, QoS

I. Introduction

RINA [1] is a clean-slate recursive multi-layer architecture that models computer networking as distributed Inter-Process Communication (IPC). In contrast to the fixed five-layer model of the current Internet, RINA is based on a single type of layer implementing only two protocols, called Error and Flow Control Protocol (EFCP) and Common Distributed Application Protocol (CDAP), which repeats as many times as required by the network designer. This layer is called a Distributed IPC Facility (DIF), which essentially is a distributed application that provides IPC services over a given scope to the distributed applications above (which can be other DIFs or regular applications). These IPC services are defined by the DIF Application Programming Interface (API), providing operations to allocate flows to other applications by specifying an application name and a set of characteristics for the flow (such as delay, loss, capacity, etc.), read/write data from/to the flows, and de-allocate flows and free the resources associated to them.

A key characteristic of RINA is its design based on the separation of all functions in mechanisms and policies, which dramatically simplifies networking. While all DIFs implement the same two protocols (EFCP and CDAP), the specific operation of each DIF can be customized to its particular scope via programmable policies. In this way, routing or packet forwarding policies configured in a backbone DIF can differ from those in a datacenter DIF, as their topological characteristics and dynamicity of the supported traffic can differ significantly. All in all, the recursive programmable architecture, together with the detailed knowledge of application requirements, make RINA a smart Internet model better accommodating new applications with heterogeneous requirements, as each layer knows the expectations of the applications (or DIFs) using that layer, and the Quality of Service (QoS) that it can expect from the layers below.

In order to provide application-specific QoS support, RINA incorporates the configurable Quantitative Timeliness Agreement (QTA)-Mux policy in the IPC Process (IPCP), namely, the application process that is instantiated in a node to enroll to a certain DIF (i.e., that IPCP represents that node in that DIF). This policy was proposed and initially prototyped in the FP7 PRISTINE Project [2] and further refined in the H2020 ARCFIRE Project [3]. QTA-Mux builds upon the ΔQ framework [4], defining a Cherish/Urgency matrix for QoS class (QoS Cube in RINA) differentiation, thus enabling inter-flow resource contention based on both delay and loss requirements. Moreover, Policer/Shaper (P/S) modules can also be contemplated per QoS Cube, addressing intra-flow contention, e.g., by spacing incoming packets or even changing the traffic distribution.

QTA-Mux is seen by the RINA research community as a key enabler of its QoS support. Nevertheless, despite the DIF recursivity present in all RINA scenarios, with several DIFs always stacked one on top of another, most studies so far have evaluated its capabilities on single isolated DIFs (e.g., see [5][6][7][8]). Therefore, further effort is needed for giving insight into its most appropriate placement on the stack of DIFs, quantifying the effects of these decisions on the QoS experienced end-to-end by higher level flows.

In this context, the Fed4FIRE+ medium experiment ERASER has pursued 4 key objectives:

1. To define different QTA-Mux deployment scenarios and QoS cubes that this policy has must enforce in a large-scale metro/regional RINA network scenario, with multiple DIFs stacked one on top of another.

2. To emulate a realistic large-scale metro/regional RINA network scenario by injecting synthetic application traffic flows reproducing diverse network applications and load conditions.

3. To evaluate the RINA QoS support in the emulated large-scale metro/regional RINA network scenario by measuring the perceived QoS metrics by the injected synthetic application flows, analyzing whether their

initial QoS expectations are met in the QTA-Mux deployment scenarios previously defined.

4. To perform a real HD video streaming demonstration on the emulated large-scale metro/regional RINA network scenario, so as to better illustrate the outcomes of the RINA QoS support under high congestion (i.e., the perceived Quality of Experience, QoE, upon reception).

All experimentation activities in ERASER have been conducted on the Fed4FIRE+ Virtual Wall test-bed [9] hosted and operated by imec in Ghent (Belgium), which has kindly provided us with the physical resources that we required in a wired networking environment.

Specifically, ERASER has not been focused on proposing and developing new ideas and protocols from scratch to achieve the goals pointed out above. Conversely, the aim has been to use open source implementations and tools previously made available by the RINA research community, such as the IRATI RINA Stack [10], the evaluated QTA-Mux scheduling policy, or the Rumba framework [11] allowing to easily configure the targeted RINA network scenario in the Fed4FIRE+ Virtual Wall. From the very beginning, we identified this situation as win-win. ERASER has undoubtedly taken profit from such implementations and tools to achieve the planned goals. This being said, it has also experimented with them and provided feedback to their original authors on encountered bugs and additional functionalities that were still necessary to realize the experiment.

The remainder of this paper continues as follows. Section II introduces the setup of the experiment, detailing the definition of the QTA-Mux deployment scenarios under evaluation and enforced QoS Cubes, as well as the characteristics of the synthetic traffic flows injected into the network. Section III presents the evaluation of the RINA QoS support measuring the perceived QoS metrics by the synthetic traffic flows end-to-end. Later on, section IV reports the conducted HD video streaming demo over RINA. Finally, section V concludes the paper.

II. EXPERIMENT SETUP

The large-scale metro/regional RINA network scenario that ERASER has emulated for the experimentation activities consists of 37 nodes in total, spanning from the end-user terminal to the server where applications run in a datacenter. Fig. 1 gives a high level perspective of this scenario, where IR stands for Interior Router, PE for Provider Edge router, HR for Home Router, BR for datacenter Border Router, S for Server, and C for Client.

As seen, the central part of the metro/regional Service Provider network configures 9 IRs in a two-dimensional meshed 3x3 grid. Moreover, 8 PE routers are attached to 8 of these IRs, offering connectivity to customers in the region, like residential users, business companies or datacenters, or to other Service Provider networks. Specifically, we consider 5 PE routers connected to 3 HRs each, 3 PE routers connected to other Service Provider networks and one PE router connected to a datacenter in the region. For the QoE evaluation, we consider 1 S within the datacenter premises, where a video streaming test application server (VLC VideoLAN [12]) will run. Moreover, in 3 HRs we connect one client that will request a

video streaming session from that server (delivered in a unicast manner for simplicity). We have placed one client close to the datacenter, another at mid distance, and another one far from it, testing in this way the experienced QoE across different route lengths. To implement each one of these 37 network nodes, a physical machine has been requested in the Virtual Wall, where a Linux image with the IRATI RINA Stack has been installed.
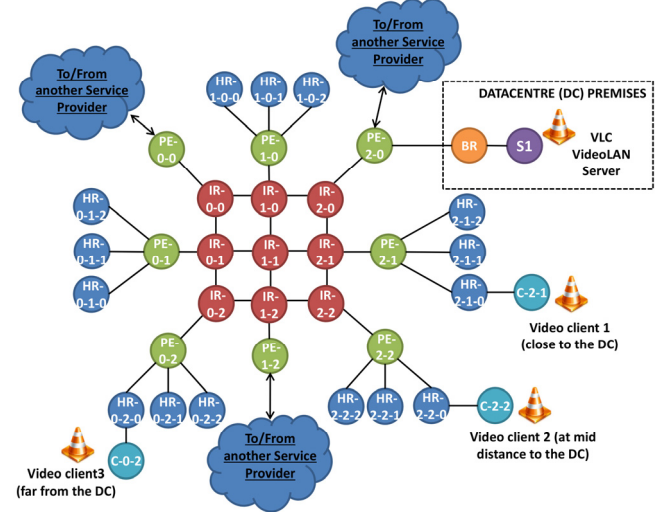

Fig. 1. ERASER large-scale metro/regional RINA network scenario.

Fig. 2 depicts the configuration of DIFs in the ERASER scenario. An Ethernet Shim DIF has been configured over the Ethernet links interconnecting the physical nodes, allowing the use of RINA over this legacy communication technology. This type of Shim DIF is already supported by the IRATI RINA Stack. On top of these Shim DIFs, a Metro-Regional (MR) DIF interconnecting PE routers has been configured, as well as a Service Provider Network (SPN) DIF extending the communication between HRs and the datacentre border router, a Datacentre Network (DCN) DIF inside the datacentre, a Home DIF inside the end-user home, and an upper level HD Video Streaming Application DIF supporting the delivery of the video streaming sessions to end-users across S, datacentre BR and HRs. As VLC VideoLAN is an application that runs over IP, iporinad daemons (also included in the IRATI RINA Stack) are required on the video streaming session endpoints, in order to create IP point-to-point tunnels over the RINA network.
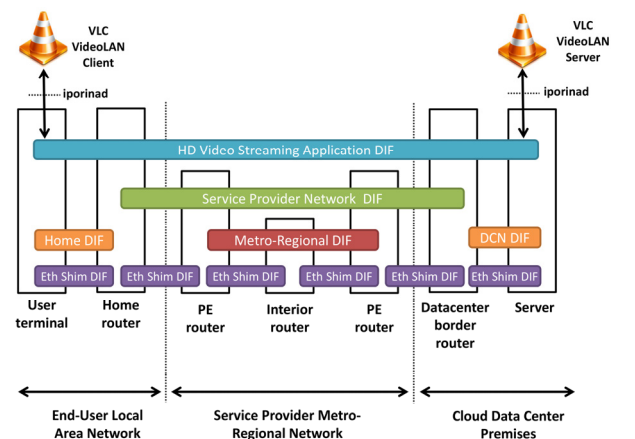

Fig. 2. Configuration of DIFs in the ERASER RINA network scenario.

A. QTA-Mux deployment scenarios & enforced QoS Cubes

In this scenario, ERASER has defined the QTA-Mux deployment scenarios to be experimentally evaluated, as well as the QoS Cubes that this policy will have to enforce. In this regard, the configuration of differentiated QoS treatment at both SPN and MR DIFs has been contemplated, whereas no QoS differentiation is applied in the rest of DIFs. The rationale behind this decision was to restrict the number of scenarios to be evaluated, while still getting the most interesting and illustrative results, as both DIFs support the vast majority of injected traffic in the network.

This being said, starting with the definition of QoS Cubes, four QoS Cubes (A1, A2, B1, B2) have been considered at the MR DIF, according to the 2x2 Cherish/Urgency (C/U) matrix depicted in Fig. 3 (right). Please note here that the C/U matrix describes the operation of the C/U multiplexer within the QTA-Mux policy, able to enforce a bi-dimensional relative QoS Cube differentiation based on delay and loss requirements. For instance, flows over the MR DIF assigned to QoS Cube A1 will be prioritized with respect to losses (i.e., they will be more cherished) and delay (i.e., they will be served with higher urgency) requirements. In contrast, flows assigned to QoS Cube B1, for example, will still be prioritized with respect to losses, but un-prioritized with respect to delay (i.e., they will experience, and thus should tolerate, higher delays).

Moving up to the SPN DIF, one additional QoS Cube has been considered in order to better differentiate among heterogeneous application flows (i.e., with finer granularity). Specifically, the five QoS Cubes offered here are: Gold, Silver, Bronze, Sensitive Best Effort (BE) and BE, as described in the 3x2 C/U matrix depicted in Fig. 3 (left). It is important to remark here that flows assigned to QoS Cubes at the SPN DIF have to be transmitted over A1, A2, B1 and B2 flows across the MR DIF. So, an adequate mapping of SPN DIF to MR DIF QoS Cubes becomes crucial in order to provide the expected QoS to end-user applications. In this scenario, both Gold and Silver flows in the SPN DIF are mapped to A1 flows in the MR DIF, Sensitive BE flows to A2 ones, Bronze flows to B1 ones and, finally, BE flows to B2 flows in the MR DIF.
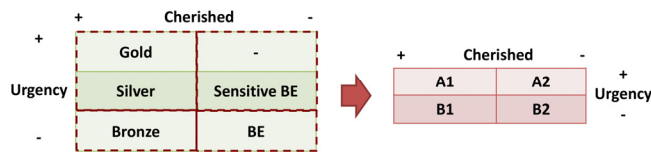

Fig. 3. QoS Cubes defined at the SPN (left) and MR (right) DIFs.

The definition of QoS Cubes at both SPN and MR DIFs has been followed by the definition of the QTA-Mux deployment scenarios to be evaluated and compared. Two different scenarios have been contemplated:

- QTA-Mux only at the SPN DIF (SPN DIF Only): IPCPs at the SPN DIF differentiate packets of flows assigned to the 5 QoS Cubes. However, when injected into the MR DIF, all packets are treated equally. This scenario is interesting to appreciate if QoS differentiation only at the SPN DIF is enough to meet the QoS requirements of flows end-to-end.

- QTA-Mux at both SPN and MR DIFs (SPN+MR DIFs): IPCPs at both SPN and MR DIFs differentiate packets as defined in the C/U matrices depicted in Fig. 3. This scenario is the ideal one. However, it is also interesting to observe QoS improvements versus the previous scenario to conclude if it is really necessary to configure QTA-Mux at both DIFs.

B. Synthetic application traffic flow characteristics

The next activity in ERASER has been to define the characteristics of the synthetic traffic flows injected into the emulated metro/regional RINA network. In this regard, the IRATI RINA Stack comes with two native RINA applications for generating synthetic traffic: RINA-tgen and RINA-perf. The former one, RINA-tgen, can generate flows describing Poisson and CBR traffic distributions, whereas RINA-perf is an equivalent of the well-known iperf application, able to measure the available bandwidth between to application processes by injecting a CBR flow at the maximum bit-rate that the network can tolerate. However, neither RINA-tgen nor RINA-perf gathers the entire set of end-to-end QoS metrics that ERASER aims to take for every synthetic traffic flow, namely, minimum delay, maximum delay, average delay and average packet losses. Apart from that, bursty traffic patterns cannot be generated. Hence, it was decided to extend RINA-tgen source code for ERASER, so as to come up with three different traffic generator types: CBR, Poisson and ON-OFF. The first two ones include minor extensions to RINA-tgen in order to gather the entire set of desired QoS metrics. The third one, in addition to gathering all QoS metrics is an application able to generate ON-OFF bursty traffic patterns, with the duration of ON-OFF traffic periods being exponentially distributed and CBR traffic distribution during ON traffic periods (no traffic is sent during OFF periods). Using them, 5 different types of synthetic application flows have been injected in the ERASER scenario, whose characteristics and QoS Cube assigned are detailed in Tab. I.

TABLE I. SYNTHETIC TRAFFIC FLOW CHARACTERISTICS

Application flow	Traffic distribution	Traffic direction	Details	QoS Cube assigned
HD Video Call	CBR	Bidirectional	CBR bitrate: 1.5 Mbps	Gold
Online Gaming	ON-OFF	Bidirectional	ON-OFF period avg. duration: 4s – 2s / CBR bitrate during ON period: 4 Mbps	Silver
VoIP	ON-OFF	Bidirectional	ON-OFF period avg. duration: 3s – 3s / CBR bitrate during ON period: 64 kbps	Sensitive BE
File Sharing	ON-OFF	Bidirectional	ON-OFF period avg. duration: 2s – 1s / CBR bitrate during ON period: 5 Mbps	Bronze
Interactive traffic	Poisson	Bidirectional	Avg. bitrate: 2 Mbps	BE

Specifically, we have decided to setup synthetic application flows between PE-0-0, PE-2-0 and PE-1-2 (i.e., PE routers providing connectivity to other Service Provider networks) and some HRs in the network (see Fig. 1). Specifically, one synthetic application flow of each type (HD Video Call, Online Gaming, VoIP, File Sharing and Interactive) is setup between PE-0-0 and HRs HR-0-1-0, HR-0-2-0, HR-2-2-0, HR-2-1-0 and HR-1-0-0. Another synthetic application flow of each type is setup between PE-2-0 and HRs HR-0-1-2, HR-0-2-2, HR-2-2-2, HR-2-1-2 and HR-1-0-2. Finally, one synthetic application flow of each

type is setup between PE-1-2 and HRs HR-0-1-1, HR-0-2-1, HR-2-2-1, HR-2-1-1 and HR-1-0-1. As a result, the offered synthetic traffic matrix will initially be composed of 75 bidirectional traffic flows in the 37-Node network scenario.

Injecting this traffic matrix, the procedure to generate different load conditions in the network has been as follows: 1) Launch an experiment allowing each network link interconnecting every pair of MR routers in the MR DIF to take benefit from the entire underlying Gigabit Ethernet Network Interface Card (NIC) capacity; 2) Measure the average bit-rate per link between IRs the entire experiment; 3) Launch experiments limiting the capacity of each of these links (i.e., NICs) at the beginning using Linux Traffic Control (TC), so that the average traffic previously measured becomes the 80% or 90% of the capacity.

III. RINA QoS SUPPORT EVALUATION

As mentioned before, 37 physical machines have been requested in the Virtual Wall, in which an Ubuntu 16.04 LTS Linux image with the IRATI RINA Stack has been installed. Fed4FIRE+ provides jFed [13] as a tool for provisioning and managing experiments over the federation of test-beds, which works well already for the allocation of a moderate number of nodes and links interconnecting them. However, RINA-specific configurations (e.g., creation of DIFs, enrolment of nodes in DIFs, allocation of flows, etc.) are hard to be performed manually, even in a small-scale scenario. Hence, a scripting framework like Rumba for automating the configuration of RINA scenarios over Fed4FIRE+ becomes really helpful and, thus, we used it in ERASER from the very beginning.

To start, we have focused on the SPN+MR QTA-Mux deployment scenario, configuring QTA-Mux at both SPN and MR DIFs. In this scenario, we have injected 75 bidirectional application traffic flows over the SPN DIF, as described in previous section. Moreover, 4 different bidirectional flows (one A1, another A2, another B1 and another B2) are set up over the MR DIF between PE routers exchanging traffic over the SPN DIF (i.e., 4 different flows between PE-0-0 and PE-0-1, 4 more between PE-0-0 and PE-0-2, etc.), so as to allow QTA-Mux enforcing QoS differentiation at the MR DIF as well, thus having 60 bidirectional flows over this DIF in total.

Injecting this synthetic traffic, we have executed an initial 10-minute experiment to measure the average traffic carried by all NICs between IRs, in order to reproduce later on different load conditions by limiting their capacity, so that the measured average traffic becomes the 80% or 90% of them. Next executions have served to find appropriate Cherish level values, providing desirable outcomes in the most loaded scenario. Specifically, we have decided to set Cherish level values at the MR DIF of 1250 packets for QoS Cubes A1 and B1, 50 for QoS Cube A2, and 1100 for QoS Cube B2. Similarly, at the SPN DIF we have set Cherish level values of 1250 packets for Gold, Silver and Bronze QoS Cubes, 50 for Sensitive BE, and 1100 for BE one.

Fig. 4 depicts the Cumulative Distribution Function (CDF) of the end-to-end delay experienced by synthetic traffic flows in the 80% and 90% load scenarios. In all cases, we have set a 5% probability to skip the most urgent queue at the MR DIF, which is configurable in QTA-Mux, in order to avoid starvation of the least urgent one.
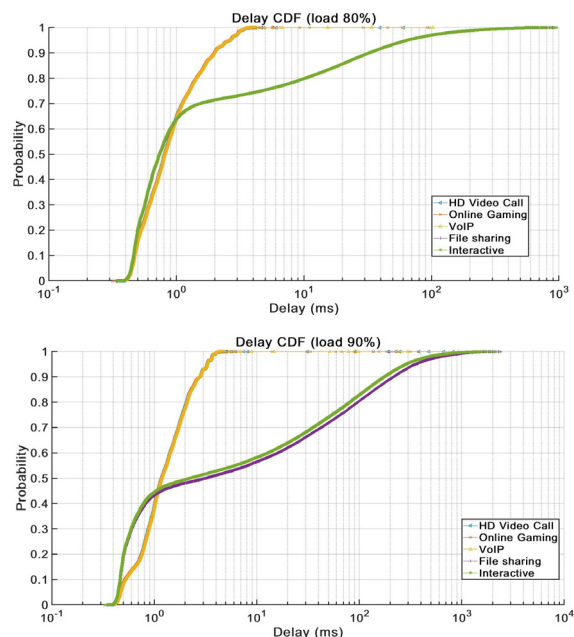

Fig. 4. CDFs of the end-to-end delay experienced by synthetic traffic flows in 80% (top) and 90% (bottom) load scenarios, when QTA-Mux is configured at MR and SPN DIFs.

Looking at these results, noticeable differences start appearing under 80% offered load. Indeed, CDF curves of HD Video Call, Online Gaming and VoIP application traffic reach probability almost equal to 1 in less than 5 milliseconds (almost 100% of the generated packets experience such latencies at maximum), although a very small subset of packets experience end-to-end delays up to 100 milliseconds due to transient congestion in high urgency queues. Conversely, CDF curves of File Sharing and Interactive traffic struggle to reach probability equal to 1, reaching maximum end-to-end delays around 1 second.

The fact that HD Video Call, Online Gaming and VoIP CDFs overlap, as well as File Sharing and Interactive ones, means that most QoS differentiation in terms of end-to-end delay is introduced at the MR DIF. This is comprehensible, as we disabled the congestion control functionality in all DIFs to stress the QTA-Mux policy performance. So, the MR DIF tends to always accept packets offered by the SPN DIF, concentrating congestion there. As a result, urgency queues at the SPN DIF become much emptier, leading to uniform packet treatment at the SPN DIF level.

Looking at the results under 90% offered load, QoS differentiation is much more pronounced there. Besides, we can see that the CDF curve of Interactive application traffic stays slightly above that of File Sharing until reaching probability equal to 1. This is reasonable as File Sharing traffic is ultimately assigned to QoS Cube B1 at the MR DIF, which has assigned a higher Cherish level than QoS Cube B2 (QoS Cube ultimately assigned to Interactive traffic). So, File Sharing traffic packets are accepted (and stored) in the associated urgency queue even under high occupancy, trading higher queuing delay for lower losses.

Moreover, we can observe that CDF curves of File Sharing and Interactive traffic start growing faster than those of HD Video Call, Online Gaming and VoIP traffic. While quite surprising at first sight, we found that this happens due to the following reason. Even very slight clock

synchronization inaccuracies among physical machines could alter our latency measurements, in the order of the millisecond. Therefore, we have finally employed rina-echo-time for measuring end-to-end latencies, instead of computing them from the synthetic traffic packet timestamps. To this end, we have set up 5 rina-echo-time bidirectional flows between all pairs of nodes exchanging traffic, one per QoS Cube offered at the SPN DIF, obtaining end-to-end latency measurements as the measured Round Trip Time (RTT) divided by two, which is reasonable as injected traffic is completely bidirectional. Specifically, a rina-echo-time client node starts sending echo-request messages to a rina-echo-time server node every certain period of time, 100 milliseconds in our experiments. However, if the associated echo-reply message does not arrive within this period of time, the client node waits for it before sending the next echo-request. We have observed that rina-echo-time clients generating flows associated to Gold, Silver and Sensitive BE QoS Cubes (as HD Video Call, Online Gaming and VoIP flows) send/receive close to 6000 echo-request/reply messages during the 10-minute experiment. However, flows associated to Bronze and BE QoS Cubes (as File Sharing and Interactive ones) end sending/receiving only around 3500 echo-request/reply messages during the 10-minute experiment, due to the high latencies. Therefore, although having measured a similar number of messages experiencing delays below 1 ms, they end representing a higher % in the case of File Sharing and Interactive traffic, growing their CDFs faster. However, this does not mean that they end receiving better outcomes in terms of delay.
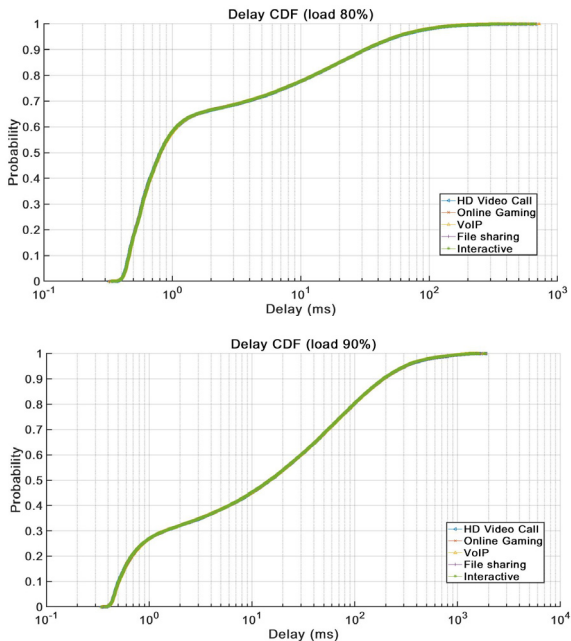

Fig. 5. CDFs of the end-to-end delay experienced by synthetic traffic flows in 80% (top) and 90% (bottom) load scenarios, when QTA-Mux is configured at the SPN DIF only.

Fig. 5 details such end-to-end delay CDF functions, but in the SPN-Only QTA-Mux deployment scenario, also for offered loads of 80% and 90%. Therefore, in the MR DIF we only have 1 single flow between pairs of PE routers exchanging traffic over the SPN DIF in this case, as the default FIFO scheduling policy is configured. Regarding the Cherish levels assigned to QoS Cubes at the SPN DIF, these

have been exactly the same as in the previous QTA-Mux deployment scenario, namely, 1250 packets for Gold, Silver and Bronze QoS Cubes, 50 for Sensitive BE, and 1100 BE one. Besides, for fair comparison purposes, the size of the FIFO queue per output port at the MR DIF has been set to 2500 packets, that is, the aggregated buffering capacity of the two urgency queues present in the previous SPN+MR DIF QTA-Mux deployment scenario.

Looking the obtained results, no QoS differentiation in terms of end-to-end delay is provided among all synthetic application traffic flow types, no matter the offered load scenario (all CDF curves overlap). We detect that the MR DIF tends to always accept packets offered by the SPN DIF, thus concentrating all congestion there. However, in this scenario, the default FIFO scheduling in the MR DIF equally treats all flows, losing the QoS differentiation previously enforced at the SPN DIF. Hence, we suggest discarding this QTA-Mux deployment in any RINA network with a similar configuration of DIFs as here.

TABLE II.   LATENCY AND PACKET LOSS (90% OFFERED LOAD)

QTA-Mux deployment scenario	Traffic type	Min. latency (ms)	Max. latency (ms)	Avg. latency (ms)	Avg. packet loss (%)
SPN + MR DIFs	HD Video Call	0.345145	830.5	1.483987	0.001659
	Online Gaming	0.3626	375.76	1.468362	0.002339
	VoIP	0.345375	924.95	1.469585	0.001342
	File Sharing	0.354075	2408.85	72.310218	0.231222
	Interactive	0.32893	2276	58.650972	3.469220
SPN DIF Only	HD Video Call	0.342600	1889.6	69.764187	0.139793
	Online Gaming	0.343865	1893.9	70.126824	0.173666
	VoIP	0.345180	1846.05	69.738218	0.150003
	File Sharing	0.351630	1876.3	70.223907	0.155559
	Interactive	0.346925	1865.15	69.441732	0.185567

Tab. II presents all latency (minimum, maximum, average) and packet loss measurements collected in both QTA-Mux deployment scenarios. Due to the lack of space, only measurements under 90% offered load are shown. Looking at the latency measures, they reflect the same behaviour as CDF curves depicted in Fig. 4 and 5. As for the experienced packet losses, QTA-Mux at both SPN and MR DIFs allows for an effective differentiation, reaching a maximum value of 3.46% for BE (i.e., Interactive) traffic. Packet losses are significantly lower, around 0.23%, for Bronze (i.e., File Sharing) traffic, having the same Urgency but a higher Cherish level. Lastly, configuring QTA-Mux only at the SPN DIF neither succeeds in providing QoS differentiation in terms of end-to-end latencies nor in terms of packet losses, experiencing all types of applications similar loss levels under every offered load scenario.

IV. HD VIDEO STREAMING DEMONSTRATION

To realize the final HD video streaming demonstration, the setup depicted in Fig. 6 has been employed. We have established HD video streaming sessions over UDP from server node S1 to the 3 client nodes C-0-2, C-2-2 and C-2-1, transmitting the same 1080p HD video file using VLC VideoLAN v3.0.1. Using iporinad, we have been able to transmit the IP traffic of the video streaming session over the HD Video Streaming Application DIF. For these purposes, iporinad sets up two TUN IP interfaces, between which a RINA flow operates as an IP tunnel. So, VLC VideoLAN traffic is initially sent to the remote IP address of

the iporinad tunnel. But, instead of reproducing the video there and exporting the display over Secure Shell (i.e., ssh – X), we directly forward it over the Internet to our PC in Barcelona, where VLC VideoLAN has been installed. Thus, we avoid undergoing the Secure Shell data encryption, which would consume a lot of resources at the client node.
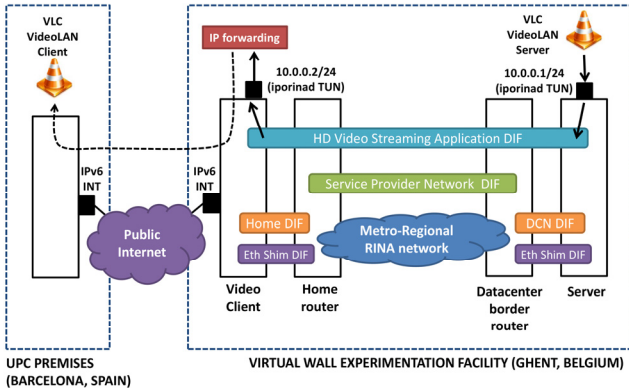

Fig. 6. Setup of the HD video streaming demonstration over RINA.

For this demo, QTA-Mux has been configured at both SPN and MR DIFs. Moreover, it has also been configured at the HD Video Streaming Application DIF (offering Gold, Silver, Bronze, Sensitive BE and BE QoS Cubes), so as to allow the subsequent mapping of traffic flows tunnelling the HD video sessions onto Gold, Silver, Bronze, Sensitive BE and BE flows across the underlying SPN DIF. Tab. III shows the perceived QoE (ranked from 1 to 5) by the 3 clients when loading the network with synthetic traffic flows reproducing the previous 80% and 90% load scenarios.

TABLE III. QoE OF THE RECEIVED VIDEO STREAMING

Offered load	QoS Cube assigned	S1 -> C-2-1 (1 hop at MR DIF)	S1 -> C-2-2 (2 hops at MR DIF)	S1 -> C-0-2 (4 hops at MR DIF)
80%	Gold	5 (perfect quality)	5 (perfect quality)	5 (perfect quality)
	BE	2 (severe stuttering & frames lost)	1 (completely frozen)	2 (severe stuttering & frames lost)
90%	Gold	5 (perfect quality)	3 (moderate stuttering & frames lost)	5 (perfect quality)
	BE	2 (severe stuttering & frames lost)	1 (completely frozen)	1 (completely frozen)

As seen, perfect QoE has been observed in the 80% load scenario for all video clients when assigning the RINA flow supporting the HD video session the Gold QoS Cube. In contrast, severe video stuttering or even a completely frozen video is observed upon reception when the BE QoS Cube is assigned (for comparison purposes), which highlights the performance of QTA-Mux for achieving QoS differentiation in the RINA network. In the 90% load scenario, results have been similar except for C-2-2, where moderate video stuttering is experienced even when assigning the RINA flow supporting the HD video session the Gold QoS Cube. This has happened because the 90% offered load scenario is setup without accounting for the transmitted HD video streaming. Hence, when transmitting the video, A1 and A2 urgent traffic in the link between IR-2-1 and IR-2-2 was filling 100% of its capacity on average, experiencing the HD video session very high congestion, the same experienced by HD Video Call, Online Gaming and VoIP flows.

A simple QoS Cube reconfiguration at the MF DIF to better isolate HD video streaming session traffic would have solved this degradation, though. For instance, a 3x2 C/U matrix could be defined in the MR DIF with the following SPN to MR DIF QoS Cube mappings: Gold -> A1; Silver -> B1; Sensitive BE -> B2; Bronze -> C1; BE -> C2. In this way, with the same QoS Cube assignment at the SPN DIF, HD video session traffic would only share the highest urgency queue at the MR DIF with HD Video Call traffic, delivering perfect QoE in all cases.

V. CONCLUSIONS

This paper has reported the results of the Fed4FIRE+ medium experiment ERASER. Configuring a 37-Node metro/regional RINA network scenario in the Virtual Wall experimentation facility, QTA-Mux scheduling policy performance has been validated by measuring QoS metrics experienced by synthetic traffic flows reproducing heterogeneous applications. Moreover, a HD (1080p) video streaming demo over RINA has been successfully performed in highly congested network scenarios, with perfect QoE in most cases when assigning flows tunneling the HD video sessions the highest (i.e., Gold) QoS Cube.

Interested readers can find all ERASER experimentation data and detailed descriptions on how to reproduce the conducted experiments in the Fed4FIRE+'s chosen repository Zenodo (DOI: 10.5281/zenodo.1420391).

REFERENCES

[1] J. Day, "Patterns in Network Architecture: A Return to Fundamentals", Prentice Hall, 2008.

[2] PRISTINE FP7-619305, "Final specification and consolidated implementation of scalable techniques to enhance performance and resource utilization in networks," Deliverable 3.3, June 2016.

[3] ARCFIRE H2020-687871, "Converged Service Provider Network Design Report," Deliverable 2.2, May 2017.

[4] N. Davies, "Delivering predictable quality in saturated networks," PNSol Technical Report, September 2003.

[5] ARCFIRE H2020-687871, "Execution of experiments, analysis of results and benchmarking against KPIs," Deliverable 4.4, June 2018.

[6] PRISTINE FP7-619305, "Consolidated software for the use cases and final report on the use cases trials and business impact," Deliverable 6.3, June 2016.

[7] S. Leon et al., "End-user traffic policing for QoS assurance in polyservice RINA networks," Telecommunications Systems, pp. 1-13, July 2018.

[8] S. Leon et al., "Assuring QoS Guarantees for Heterogeneous Services in RINA Networks with ΔQ," NetCloud 2016 Workshop, Luxembourg, December 2016.

[9] Fed4FIRE+ Virtual Wall test-bed, description available at: https://www.fed4fire.eu/testbeds/virtual-wall/.

[10] IRATI RINA Stack, available at: https://github.com/IRATI/stack.

[11] The Rumba repository, available at: https://gitlab.com/arcfire/rumba.

[12] VLC VideoLAN webpage, available at: https://www.videolan.org/.

[13] jFed webpage, available at: https://jfed.ilabt.imec.be/.