European Commission — FP7

R / T E

REDUCING INTERNET TRANSPORT LATENCY

Project acronym: **RITE**

Project number: 317700

Work package: Dissemination & Exploitation

Deliverable number and name:

Deliverable 4.3: Exploitation Plan

Title: Exploitation Plan
Work Package: WP4

Version: 00A
Date: December 1, 2015
Pages: 14

Author:
Philip Eardley

Co-Author(s):
Achilles Petras, Koen de Schepper, Ing-Jyh Tsang, Bob Briscoe, Christian Tellefsen, Andreas Petlund, Iffat Ahmed

To:
Jorge Carvalho
Project Officer

Status:
[ ] Draft
[ ] To be reviewed
[ ] Proposal
[X] Final / Released to CEC

Confidentiality:
[X] PU — Public
[ ] PP — Restricted to other programme participants
[ ] RE — Restricted to a group
[ ] CO — Confidential

Revision:
(Dates, Reviewers, Comments)

Contents:
This report describes the plans of the partners in the RITE project to exploit its results. Commercial plans of the industrial partners are given only in general terms given the intrinsic confidential nature of such plans. Exploitation for non-commercial purposes, e.g. the development of public educational courses or open source software distribution is also included.

Contents

1 Introduction

This document describes the potential impact and exploitation of RITE's research, within the industrial partners and more widely across the industry as well as in the open source community and academia.

2 Alcatel-Lucent

2.1 Background

Alcatel-Lucent's solutions and product portfolio enable carriers, Internet service providers, enterprises and governments worldwide, to deliver voice, video, multimedia, and data communication services to end-users. Alcatel-Lucent is a leading IP networking, ultra-broadband fixed or wireless access, and cloud technology specialist. Its products have to serve diverse customer segments, which impose a wide range of requirements on its equipment, as it should be capable of performing a variety of different functions under different network conditions. To maintain its competitiveness it is fundamental for Alcatel-Lucent to innovate and address not only the present problems its customers face, but the upcoming challenges of the future telecommunication industry.

Alcatel-Lucent recognizes that the upcoming Interactive Internet applications will run on ultra high bandwidth networks interacting with nearby Cloud systems, and that all ingredients for a very low latency user experience will be in place, except for the current Classic Transport layer. Network latency can be a limiting factor for the deployment of several novel services, from interactive cloud base video application to the deployment of novel network architecture such as promoted by Network Function Virtualization (NFV) and Software-Defined Networking (SDN) paradigm. Moreover, as the access network moves towards infrastructure with higher and higher throughput, at gigabit levels, network latency will be more and more visible and a limiting factor to achieve the required QoS/QoE for possible upcoming services, such as network based virtual reality applications. As a result of the RITE project we know now that without Network support, it will not be possible to drastically evolve the end-systems responsiveness to congestion, required to substantially prevent unnecessary latency.

The results of the RITE project will be extremely useful to decide which features will be required in Alcatel-Lucent's new products and in the evolution of its current products, with which the operators (Alcatel-Lucent's customers) can ensure the end-user's quality of experience while at the same time allowing operators to exploit their network efficiently. The RITE project offered the opportunity for Bell Labs, the research organization of Alcatel-Lucent, to investigate and implement novel techniques to address network latency. In addition to enable the collaboration with leading research centres and researchers in the field.

The plan was to exploit the RITE acquired knowhow in the following ways:

- Enhancement of network products and services to enable distribution of high volumes of interactive low latency content to end users with the required QoE level;
- Contributions to standardization, to accelerate the deployment of cost-effective, open solutions for large scale content distribution by network operators and service providers.

As the technology innovations and the market evolved further during the execution of the project, it became clear that the project results would be beneficial for several exploitation tracks of different nature, some for nearly immediate use and with immediate effect, and some from a more longer term perspective and hence with some more uncertainty about their effectiveness.

2.2 Exploitation – Access and IP Routing Products

At Alcatel-Lucent several topics are under discussion within different departments such as the Fixed Access and IP Routing department. In particular, the implementation and deployment of the DualQ Coupled AQM, as a novel product feature, enhancing their service capabilities. Figures 2.1 and 2.2 shows the possible evolution of the QoS mechanisms at the access network and evolution of the hierarchical QoS at the broadband network gateway (BNG) in

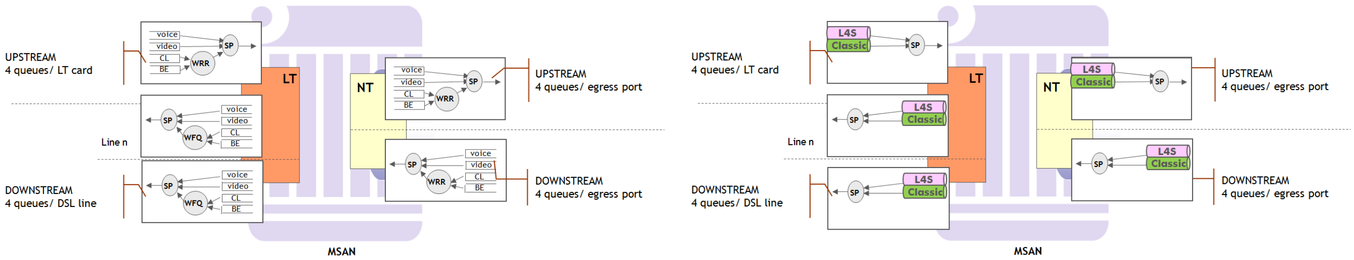

Figure 2.1: Traditional QoS mechanism (left) versus Novel Low-Latency Model (right) at the multiservice access network (MSAN) equipment

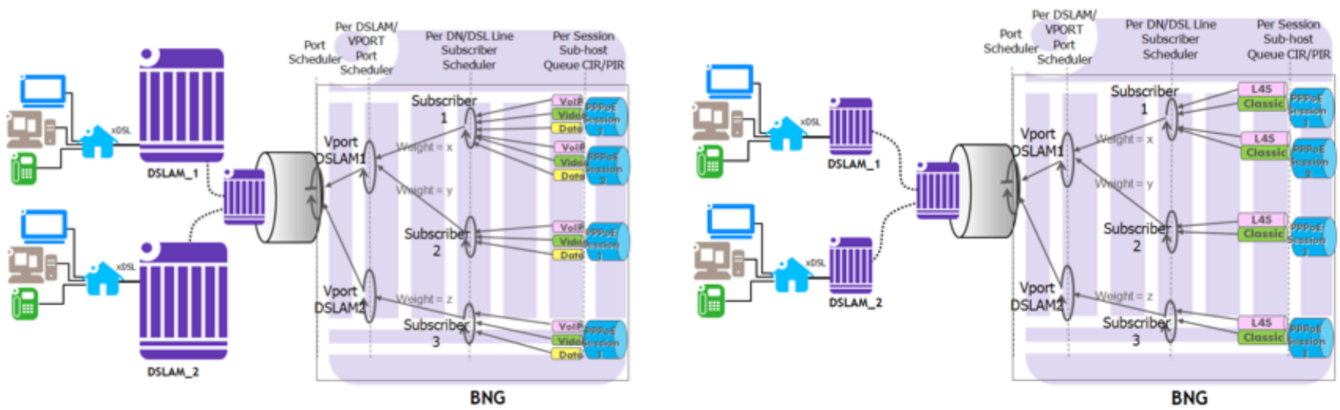

Figure 2.2: Traditional QoS mechanism (left) vs Novel Low-Latency Model (right)at the broadband network gateway (BNG) equipment

the network and similar (not shown) in the residential gateway (RGW) in the home. Instead of multiple queues per service, hierarchical per user, per service, it allows simplification toward classical and scalable (low-latency, low-loss - L4S) services, which would rationalize not only equipment resources, but also reduce configuration and management overhead.

2.3 Exploitation – CDN and Cloud Based Services

Interactive cloud based service, such as virtual reality applications and panoramic interactive video services requires low latency, high throughput which can be enable by the techniques developed at RITE, for example the DualQ AQM offers several advantages over the present QoS models. In addition, server techniques developed in RITE improves services delivered by CDN, such as over-the-top (OTT) HTTP adaptive video (HAS) and generic web content. The Alcatel-Lucent testbed (see Figure 2.2) demonstrated how the mechanisms in RITE could be used for delivering CDN and Cloud based services with the following benefits:

- Diffserv does not apply end-to-end, especially at the home networks, where any service should reach any device. In particular, diffserv is hard to support towards any device at the last hop between the home gateway and end devices. For services that require low latency and high throughput, QoS can be enabled by the DualQ AQM as was demonstrated by the testbed experiments. (See Figure 2.3)
- The testbed experiments also show that the DualQ AQM mechanism in combination with Scalable TCP (DCTCP) delivers a much better quality of experience for video applications using HAS than Classic TCP. The importance is that HAS traffic now dominates the Internet; Netflix accounts for about 35% and YouTube 15% of traffic at peak time. In addition, it was shown to deliver general web traffic more efficiently and with consistent lower completion times. Alcatel-Lucent has vested interest in this mechanism as it can be used as a differentiator for our Velocix CDN products.

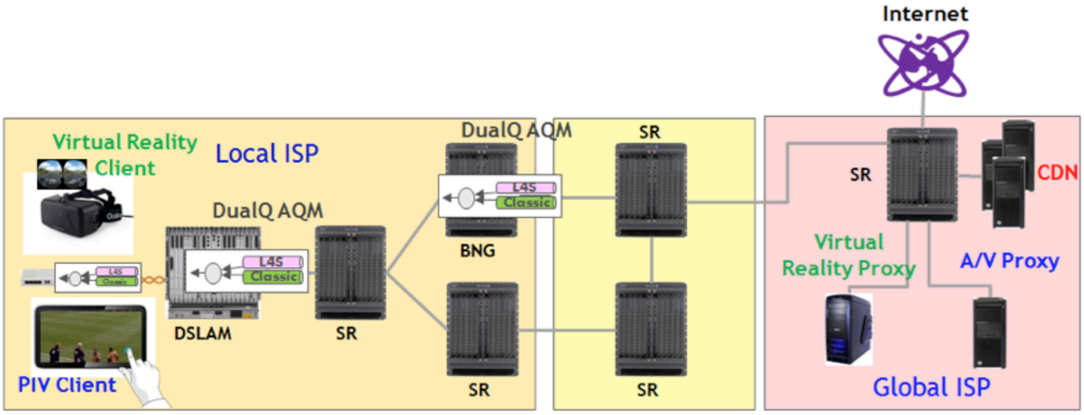

Figure 2.3: Framework for CDN and Cloud Based Services residential interactive, low latency service delivery.

2.4 Exploitation – SDN and NFV Paradigm

Alcatel-Lucent invests, develops, deploys and operates complete networks from the home equipment to metro and core elements, as well as application services, management and orchestration functions. The evolving architectural framework such as promoted by Network Function Virtualization (NFV) and Software-Defined Networking (SDN) paradigm based on cloud nodes requires extreme low latency service network. The work proposed, developed, tested and analysed during the RITE project will be used to promote, enable and validated the network evolution to NFV and SDN (see Figure 2.4).

Alcatel-Lucent develops products, applications and services that enable telecom operator and service provider networks to handle the expected future network architecture and network traffic requirements

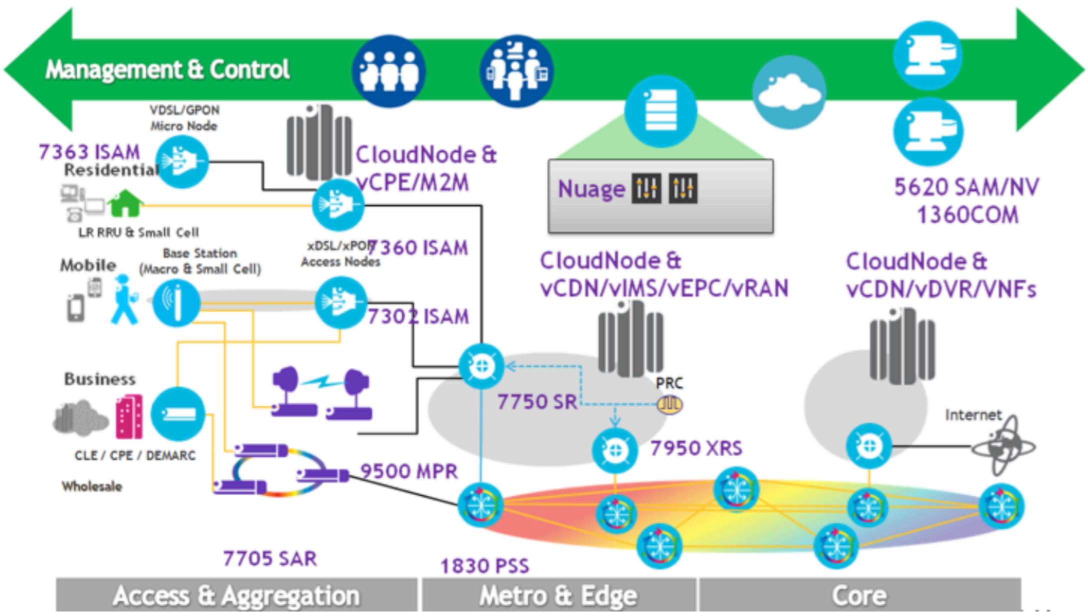

Figure 2.4: The Network Evolution to NFV and SDN

3 BT

Several topics are under discussion for exploitation, in particular on simplifying the operation of Quality of Service (QoS) on the BT broadband network.

3.1 Simplifying QoS on BT network

3.1.1 Purpose

Simplify the operation of QoS scheme at Broadband Network Gateways on BT's broadband network as well as improving the performance of applications. There is also an incentive to reduce cost of BNGs by using standard interface cards as opposed to dedicated and more expensive Media Dependent Adapters (MDAs). These are used by operators with large-scale residential and business service deployments like BT. It requires high scale, high density and fine granularity of queues for assured delivery of multi-service applications including high-speed internet, voice, video, and IPTV to individual subscribers. These cards deliver hierarchical QoS for assured and policy-enforced delivery of all subscriber applications to ensure end users enjoy the highest levels of quality of experience (QoE).

However, the intention is to move from the current Diffserv hierarchical QoS model on Multi Service Edge Routers (BT's distributed BNGs) because:

- Mapping of new services to new queues and rationalising their relative importance becomes unsustainable
- Endeavours to give streams a 'bandwidth' - no regard to delay/loss characteristics
- Partitioning of classes of service with policing to protect lower classes of service assumes a certain traffic mix and could still create partition induced bottlenecks when the traffic mix is different
- The predefined hierarchy of prioritisation may not always address what End Users expect

3.1.2 Summary of relevant Foreground that RITE has developed

RITE has developed a new queue management scheme, called DualQ Coupled AQM, which would operate at the bottleneck link (the broadband network gateway in the downstream direction). This operates in conjunction with ECN (explicit congestion notification marking) and a scalable TCP such as DCTCP. The work is summarised in deliverable D3.3 Section 3.

3.1.3 Potential Impact

Two advantages are foreseen. Firstly, a substantial simplification of QoS and traffic management. This can be seen by comparing the current QoS model with the potential new model:

The proposal has been developed with several requirements in mind: low latency for some (or all) applications; the ability to prioritise some services; and the ability to isolate some traffic within a customer's VLAN. The proposed QoS model has fewer queues than currently, whilst still allowing per customer performance assurances and control over how to distribute 'quality' amongst different applications when the bottleneck is overloaded (whether this is for a short or long period). All this must be within the constraints of regulation, where Openreach operates independently and provides service to all Communication Providers (including BT Consumer) on an equivalent basis.

The second advantage is cost savings due to less expensive interface cards at the BNGs, by using standard interface cards as opposed to dedicated and more expensive Media Dependent Adapters. The latter are currently used by operators with large scale residential and business service deployments like BT. These cards deliver Diffserv hierarchical QoS and achieve high scale, high density and fine granularity of queues for assured and policy-enforced delivery of multi-service applications to individual subscribers. However, there are various downsides that motivate the search for a simpler solution:

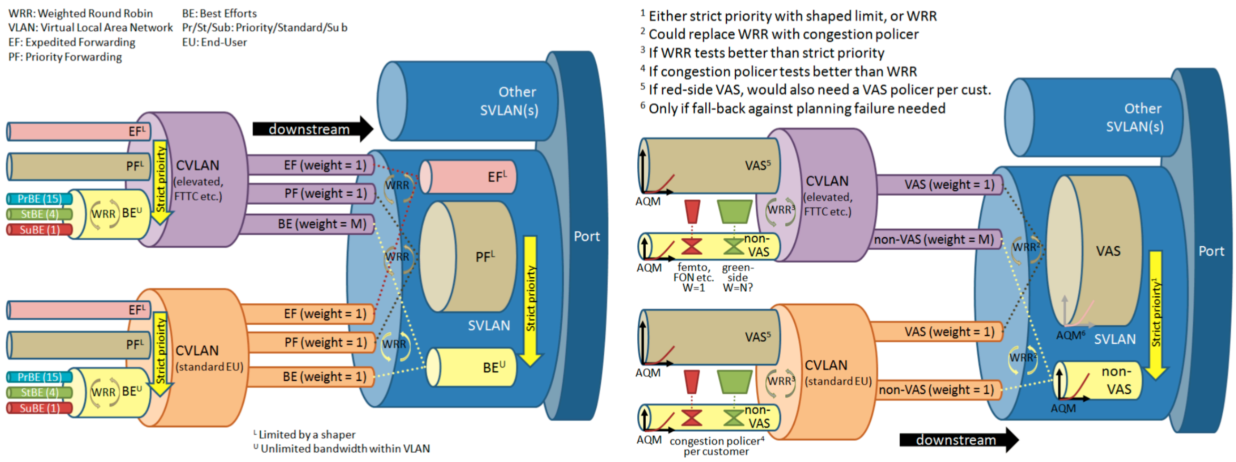

Figure 3.1: BT Existing (left) vs Potential (right)

- Mapping of new services to new queues and rationalising their relative importance becomes unsustainable
- Endeavours to give streams a 'bandwidth' - no regard to delay/loss characteristics
- Partitioning of classes of service with policing to protect lower classes of service assumes a certain traffic mix and could still create partition induced bottlenecks when the traffic mix is different
- The predefined hierarchy of prioritisation may not always address what End Users expect

The third area of potential impact is to give a much better performance for video applications (those that have been upgraded to use scalable TCP), whilst also being fair to non-upgraded applications and in fact improving the latency /download time of other typical applications like web browsing and file downloads.

The final impact is to help dispel the myth that faster access speeds (with widespread deployments of FTTx and upcoming G.Fast) would eradicate any issues with queueing delays. The higher peaks of traffic would transfer the bottleneck from the access network to the backhaul network, create more variable traffic load patterns leading to buffer fill ups that in turn could create increased queueing delays.

3.1.4 Traction

BT's network architects are actively discussing how to simplify the operation of QoS in our network. The above approach grew from iterative discussion with them. However, changing our deployed QoS model is a significant step, so discussions continue.

3.1.5 Further work and research necessary

From a direct BT perspective, the open issues include whether there needs to be a separate elevated CVLAN (customer VLAN), the use of shaping vs Weighted Round Robin and the handling of 'redside' services like WiFi FON. Discussion is also needed with the market facing units, to work through the implications for the different customers (CPs and residential consumers). Other issues include: comparative testing against other upcoming techniques (e.g. Octoshape QUIC) that are based on UDP and higher layer congestion avoidance algorithms; assessment of the impact /risk of UDP flows pushing out classic TCP flows; exploring how DCTCP/ECN can be rolled out to end hosts and the network.

3.2 Other potential exploitation in BT

We have discussed other ways to use RITE foreground within BT:

- Potential collaboration with content providers to provide better interactivity and QoE to innovative services; for example interactive zooming and panning into panoramic video feeds in sports venues, Live Virtual Reality experiences, Reliable unicast IPTV (i.e., provide predictable QoE to adaptive streaming flows that currently suffer from variable TCP throughput)
  - With a panoramic video service the end customer can select their specific view within a wider view captured by the camera – a kind of per-user facility to pan and zoom. This could make watching sport for instance, more personal and exciting. Deployment of such a service on a RITE testbed and its evaluation is described in D3.3 Section 4. For Live VR/immersive experiences, ideas have been explored with BT's external innovation team who have recently been scouting companies like NextVR or 3d-4u in order to explore how live stereoscopic feeds to VR headsets can be supported by the broadband network.
- Offer better performance for BT Global Services' customers in Financial Markets, for example for applications on trading floors. Low latency and reliability is critical, and RITE's DualQ scheme potentially can help.

4 Megapop

Megapop is a game developer and publisher based out of Oslo, Norway. The company was established in 2012, and is founded by some of the most experienced and innovative online game developers in Europe. The focus of Megapop is to create great casual game experiences for players of all ages, on multiple popular gaming platforms. This includes iOS and Android platforms, as well as browser based gaming on the PC. At the moment we do not plan to support any of the traditional game consoles, primarily because they are experiencing a market decline, but they may become of interest again in the next console generation. Platforms which are of great interest though are the Blackberry, the Windows phone and Smart TV. With Megapop's formidable experience in running complex Massively Multiplayer online games, the company seeks to treat their casual games as a service, constantly updating, polishing and honing their live game services. Additionally Megapop wants their games to run on a single server set-up, so while the player may be sitting on some particular platform, can experience the game together with players who own different devices. This is currently cutting edge technology for casual gaming, and can drastically increase the customer experience and fun factor.

4.1 Main technical challenges addressed by RITE mechanisms

Low bandwidth ecosystems – When playing on portable devices "on the go", such as iOS or Android devices, customers will access the services on wildly differentiating speeds. We need to optimise our games, content and servers to cater for this, as well as potentially find solutions that can instantly detect the network and platform, the gamer is accessing our servers on. Reducing the latency for low bandwidth eco-systems is thus vital for the best possible user experience. We also need to detect when players are residing in higher speed areas (e.g. home or office networks), potentially allowing automatic back-end streaming and / or patching of content when on higher speed.

Live vs. a-sync serving of multiplayer content – We need to optimise the serving and experience of multiplayer content. How much data we push through the network, how advanced live multiplayer content we serve (e.g. how many live calculations of movement, projectiles etc.), and optimise both the core and the packages. At the moment there are few companies in the world which are attempting this approach, but the trend is clear; more and more are moving towards live multiplayer content, even on portable devices.

Cross-platform servers – While gamers may access the games on different platforms, our backend server system are "singular" for all the platforms. This requires a keen eye on the specific demands of each platform, and finding common denominators to ensure a smooth service for all players, regardless of where they come from. We would also like to point out that wildly differing platforms may / will demand latency hiding to give all players a feeling of "equal service", but that actually reducing latency will also reduce the urgency and scale of latency hiding techniques.

4.2 Megapop infrastructure

The server infrastructure employed by Megapop consists of services and/or servers deployed in the cloud. Figure 4.1 shows an example of how the infrastructure of the backend systems may be organised. Depending on the needed level of control, we may choose to deploy the backend using Software as a Service (SaaS), Platform as a Service (PaaS) or Infrastructure as a Service (IaaS).

- SaaS allows for little or no control of the mechanisms employed to reduce latency and will be used for services with no strict latency requirements or elements of the backend that is not directly or indireclty influencing the experienced latency of the player.
- PaaS gives us more possibilities for configuration. This option is chosen as a tradeoff between configurability and deployment overhead.
- IaaS will be chosen for the services where we need the larges possible amount of control over kernel mechanism configuration and deployment. This solution costs more to deploy, but allows us to deploy custom patches and solutions to provide the best possible service for latency-sensitive parts of our backend.

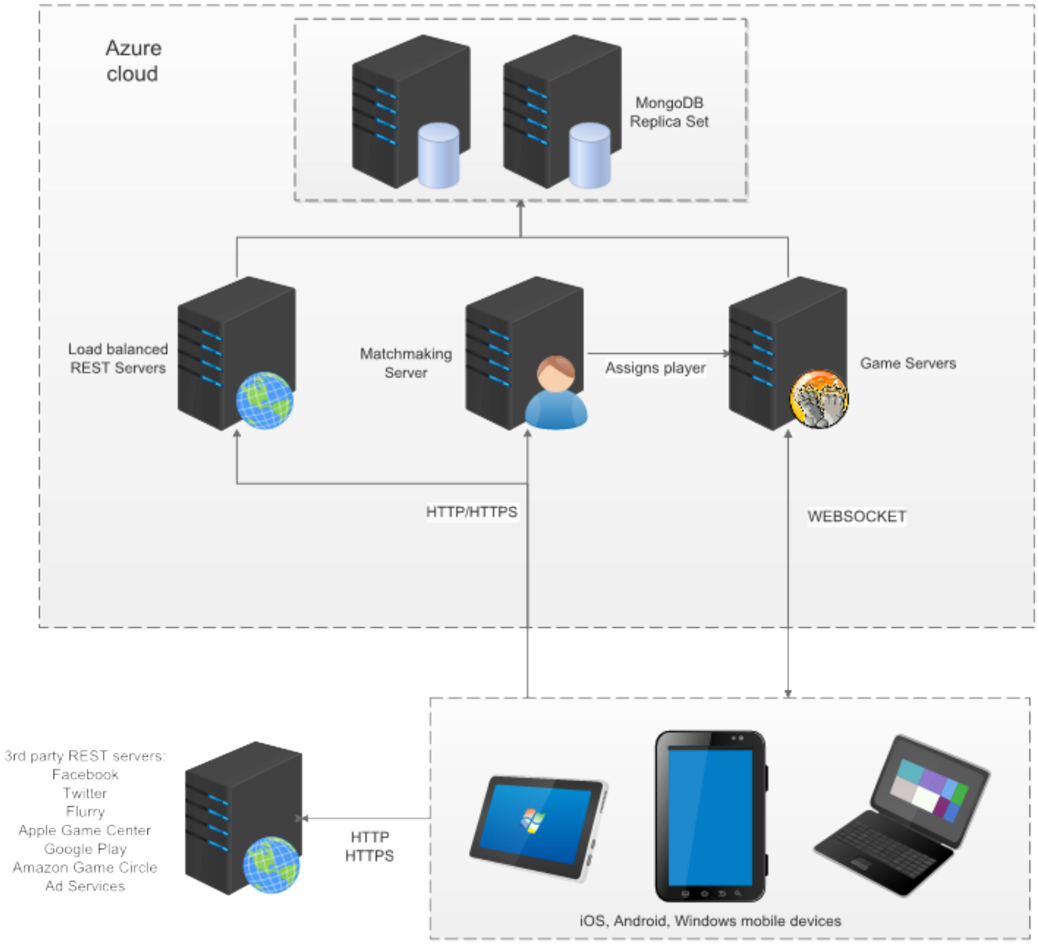

Figure 4.1: Megapop server infrastructure example.

Our targets are the casual gaming market and we aim to deploy them on a wide vareity of platforms to reach as many users as possible. We keep the server infrastructure centralised, allowing us a certain degree of control of mechanisms employed on the server side to the benefit of all connected clients. For deployment of realtime games, we need to deal with the problems that will create noticeable delays and interruptions to gameplay. Since we use TCP, and depend on reliable transfer, retransmissions are the biggest source of delays to be addressed.

Our exploitation plans mainly focus on the PaaS and IaaS solutions as they provide us with the options to reconfigure the systems to a certain degree.

4.3 Exploitation

We can see several places where we can apply RITE mechanisms to improve our service in the near future. Depending on the traffic patterns we see/expect from the different game concepts, we may deploy several of the end-host mechanisms on the servers to reduce the latency.

- New Congestion Window Validation (New CWV) for bursty traffic patterns like live transfer of in-game objects and textures. Reduces transfer time for the objects while being considerate to competing traffic.
- Retransmission TimeOut Restart (RTOR) and Tail Loss Probe Restart (TLPR) for web-like traffic and short flows. Reduces retransmission latency.
- Redundant Data Bundling (RDB) for signal traffic and position updates in real-time gameplay situations. Helps avoiding retransmissions.

For the PaaS and SaaS scenarios, we will work in collaboration with SRL, KaU and ALU to get the patches accepted in the mainline kernel so that they can be globally applied in the cases where we have no control of the kernel deployed in the infrastructure.

When we need greater control, we will use IaaS and apply the patches needed to enable the mechanism helping the game in question.

The work in RITE on Dual Queue Coupled AQM and scalable Low Latency, Low, Loss, Scalable congestion control (L4S) has the potential to be very useful to our scenario as it will allow for reducing both queueing delay and retransmission delays. The deployment, however depends on the network supporting the technology. We will make sure ECN is enabled for our end hosts, awaiting the benefit of more widespread ECN support in the network, as advocated by the RITE project. Assuming an incremental deployment path is found, we will add support for L4S as support becomes available.

Finally, results from the trace analysis and RITE experiments on our testbed server has led us change the parameter configuration of our running services to get an immediate improvement of our service. Our evaluation revealed that Nagle's algorithm was turned on for Node.js by default, even if this platform is commonly used for web 2.0 interactive applications that will suffer from delaying small packets. Further experimentation revealed that the combination of Nagle and Tail Loss Probe (TLP) performed worse than when either one of the two was enabled without the other. The negative interaction has prompted the project to repair the problem for the Linux kernel and a patch is being made ready for Linux kernel mailing list submission.

4.3.1 Potential impact

There is a still untapped market for mobile casual real-time games that has not been successfully exploited by any game company so far. With the advent of lower-latency mobile technology, the realism of deploying such services approaches, and combined with the RITE technology for minimising end-to-end delay, this could lead to breakthroughs in this area and we hope to be in the forefront of this development. Our aim is to be in the lead among the actors that enter this market as it matures, and the RITE project has helped us to get a step further towards this goal.

5 RITE Plans for further Education & Research

Institution	Outcomes
KAU	• Master's course "Topics in Computer Networking", which includes latency understanding gained from RITE • 2 PhD students
UiO/SRL	• RITE influence from bachelor to doctoral research • RITE-related Master theses • 4 PhD students (2 UiO, 2 SRL) • basic computer network education (bufferbloat, RITE latency solution, videos)
UoA	• undergraduate network course (added topics on bufferbloat, latency, RITE videos) • RITE 3 PhD students (WP1 & WP2)
IMT	• RITE topics in Master's level courses • Specific courses on TCP now feature topics on bufferbloat, AQM, latency • Student projects on RITE topics such as MPTCP

Table 5.1: RITE exploitation in Academia

Academic partners will exploit the outcomes of RITE along the following lines:

- **Teaching**: The portfolio of experiences gained in RITE will be used by academic partners in communication-related courses to educate both under-graduate and graduate students.
- **Advanced Education**: The experience in RITE will help the new generation of experts of Internet transport to gain awareness of latency-related problems and possible countermeasures. This has been done by applying knowledge gained from RITE in teaching and supervision at the PhD and MSc level. The expertise maturated in these experiences will constitute the basis for future PhD theses and will add to the research base at each academic partner.
- **Public Engagement**: RITE partners coordinated a number of events (http://riteproject.eu/category/events) to spread the RITE findings to a broader public beyond the academic and standardisation community. Motivated by the success of these events, RITE partners plan to continue fostering the outcomes of RITE (e.g building on other international collaborations) beyond the end of RITE.
- **Research & Innovation**: The visibility/influence acquired by leading standardisation groups and the research community on latency-related problems is a key outcome of RITE. The links between RITE and key players will be used to foster innovation in latency-related solutions and as a starting point for new collaborative research projects.
- **Contribution to open-source projects**: RITE project was strongly development-oriented and a set of contributions have been made publicly available as open-source software (see Section 6). RITE partners will continue to maintain the software within the open-source project (such as Linux and FreeBSD) and use it for future experimentation.

The above objectives are clearly long-term targets and RITE partners have already started to implement some of these during the course of the project. Table 5.1 summarises the steps that academic partners have taken to exploit RITE outcomes in education. To achieve the educational targets RITE partners are capitalising on results and experience developed along the project including the literature produced by the project, the know-how achieved, the educational materials (videos, interactive games, etc) and the open-source software.

Part of the scientific production is being used directly for education and training. In particular, the survey on latency reducing techniques [REF] is a good reference for under-graduate students who wish to expand their understanding of the causes of Internet delay. The survey is an excellent starting point for doctorate students that need to have a big picture of the state-of-the-art before delving into the details of a specific technical problem. Doctorate students will also benefit from reading the articles produced in the project to undertake new lines of research.

The number of high-quality academic contributions and the success of RITE proposals in standardisation groups demonstrate the leading role that RITE partners achieved in Internet transport innovation. This puts RITE partners in a position to influence decision-making and leveraging the technologies developed in RITE. This will be done by means of a continued participation to standardisation groups, contribution to open-source project (such as Linux and FreeBSD), and a renew collaboration with industry partner. This will further increase their involvement in future research projects and reinforce their networks of European research partners. In this respect, it is important to remember the TAPS (Transport Services) working group at the IETF, which has been created by UiO with the support of several RITE participants. TAPS is a successful initiative which has attracted a number of key players from industry and promise to be the right channel to keep alive links with industry. The academic partners have also already initiated several joint research projects with industry that (partly) build on the RITE results, such as NEAT[1], MONROE[2] and MAMI[3].

RITE animated videos are valuable resources that RITE partners are going to use as teaching materials. Although they were initially seen as a way to sensitise a broad audience to the problems of latency and influence decision-makers, they proved to be effective also in education. The videos have already been used by several partners in networking classes as an enjoyable and clear introduction to fundamental networking concepts. The success of the initial trial means that academic partners are planning to reuse the videos in future lectures.

6 Open-Source Software

The RITE project had a stated goal to make tools and mechanisms available to a wider set of interested parties through Open Source Sofware (OSS). This software is all freely available under the Linux (GPL) or FreeBSD licence. This section describes our open source contributions, potential third party exploiters, and how to access each of the OSS mechanisms/tools. Table 6.1 lists the open source software shared through the project. We have divided the table into two sections: **Mechanisms** and **Tools**.

The main exploitation benefits of Open Sourced code is that it is possible for companies and academics to build upon our results and employ the mechanisms and tools in their own businesses and research. If a mechanism is made available as part of the mainline kernel, it will become available to millions of users as updates to the main distributions are rolled out. This includes the majority of web servers and super computer clusters in the world. Currently the RITE project has gotten one of its target mechanisms, the Linux implementation of Caia Delay Gradient, that has been improved in several ways from the FreeBSD version, into the mainline Linux kernel[4].

6.1 Mechanisms

The mechanisms that were developed in the project and prototyped in the Linux kernel are shown in the first part of table 6.1. For the mechanism already accepted into the mainline Linux kernel, the kernel version first integrating the mechanism has been listed. For other mechanisms, links to where the patch can be found are shown. The table also points to which RITE deliverable describes the Mechanism in detail.

[1] https://www.neat-project.org/
[2] https://www.monroe-project.eu/
[3] https://mami-project.eu/
[4] https://lwn.net/Articles/644440/

RITE Mechanisms

Mechanism	Details	How shared
Caia Delay Gradient	D1.3	Mainline Linux kernel (from 4.2)
RTO Restart patch	D1.3	https://github.com/perhurt/rtorestart
TLP restart patch	D1.3	https://github.com/perhurt/rtorestart
RDB patch	D1.3	https://bitbucket.org/mpg_code/rdb
New CWV patch	D1.3	https://github.com/rsecchi/newcwv/
ABE patch	D1.3, D2.3, D3.3	http://heim.ifi.uio.no/naeemk/research/ABE/
Dual Queue AQM	D2.3	http://riteproject.eu/dctth/ when approval process completed

RITE Tools

Tool	Details	How shared
TCP Evaluation Suite	QMR4+	https://bitbucket.org/hayesd/tcp-evaluation-suite-public http://trac.tools.ietf.org/group/irtf/trac/wiki/ICCRG_tcpeval
Pcap latency analysis tool	D3.1	https://bitbucket.org/mpg_code/tstools_analysetcp
Active measurement AQM detection tool (TADA)	D2.3	https://bitbucket.org/mkargar/tada-tool_automatic_detection_aqm
RITE integrated kernel	D1.3	http://riteproject.eu/resources/rite-integrated-linux-kernel/

Table 6.1: Open-source software shared from the RITE project.

Patches not included in the mainline kernel are available for interested parties through patching of the kernel version it was made for. If the patch does not cleanly apply to their preferred kernel version, they have access to the code so they can port it to their kernel of choice. The mechanism can also be inspected by third parties for compliance with their corresponding standard or for bugs. There are several patches that have been refined to a point where kernel mailing list submission is viable, most notably RTO Restart, RDB[5] and New CWV.

New Congestion Window Validation

Lead partner: UoA

The New CWV mechanism attempts to choose a good size for the TCP congestion window after idle periods when the window has not been recently validated. The mechanism is useful for applications with bursty traffic patterns. It has been thoroughly tested and the algorithm has been standardised in the IETF as RFC 7661 (experimental). The New CWV mechanism and patch was presented to the Linux community and Industry at Linux Conference Europe 2015. A patch is being refined for kernel mailing list submission.

RTO Restart

Lead partners: KaU, UiO, SRL

Both TCP and SCTP use a retransmission timeout (RTO) timer as a last resort for data loss recovery. The standardised

[5] https://lkml.org/lkml/2015/10/23/720

way in which this timer is restarted often unnecessarily extends the loss recovery time by at least one round-trip time (RTT). The RTO Restart (RTOR) mechanism can be used to reduce the retransmission delay when recovery is done by RTO. The mechanism reduces the recovery time for tail loss and is applicable for short flows like most web traffic. It has been thoroughly tested and the the mechanism has come far in the IETF process of becoming an experimental RFC. The RTOR mechanism and patch was presented to the Linux community and Industry at Linux Conference Europe 2015. A patch has been developed and prepared for mailing list submission.

TLP Restart

Lead partner: KaU, UiO, SRL

Linux makes use of a non-standardised retransmission scheme called Tail Loss Probe (TLP) that is designed to tackle the slow loss recovery in the end of flows. Like the Retransmission timeout, TLP also makes use of a timer that has the same restart problem. To mitigate this performance problem, RITE has developed a restart strategy for TLP, which we call TLP Restart (TLPR), that is similar to RTOR. The mechanism reduces the recovery time for tail loss and is applicable for short flows like most web traffic. The TLPR mechanism and patch was presented to the Linux community and Industry at Linux Conference Europe 2015. A patch has been developed and prepared for mailing list submission.

Alternative Backoff with ECN (ABE)

Lead partners: UiO, UoA

ABE is a simple sender side TCP modification that allows end-hosts to gain performance benefits from Explicit Congestion Notification (ECN) signals from the network. This provides incentives for latency lowering Active Queue Management (AQM) schemes with ECN to be deployed in the network. ABE was developed by RITE in collaboration with the Centre for Advanced Internet Architectures (CAIA), Swinburne University of Technology (Melbourne, Australia). An experimental patch has been made available for both FreeBSD and Linux.

Redundant Data Bundling

Lead partner: SRL

Redundant data Bundling (RDB) is a sender-side only mechanism that tries to preempt the experience of packet loss by redundantly bundling all unacknowledged data as long as the resulting segment does not surpass 1 Maximum Segment Size (MSS). This allows for effectively avoiding high retransmission delays without sending any more packets into the network. The mechanism will be effective for flows that transmit only a small amount of data with relatively large inter-packet interval times or *thin streams*. The RDB mechanism and patch was presented to the Linux community and Industry at Linux Conference Europe 2015. A patch has been submitted to the *netdev* mailing list and are currently being discussed[6].

Caia Delay Gradient Linux implementation (CDG)

Lead partners: SRL, UiO CAIA Delay Gradient is a delay-based congestion control that takes its backoff decisions from estimates of whether the queue is growing or shrinking based on the gradient of delay measurements within RTT periods. The mechanism includes a heuristic that switches the congestion control to "New Reno" when it detects that it is competing with other loss-based congestion control mechanisms. The aim of this is to get the queueing delay benefits when alone on the bottleneck or competing with delay-based congestion controls, and still be able to achieve a fair throughput when competing with loss-based congestion control. In addition, CDG can be deployed as a less-than-best-effort congestion control by turning the loss competition heuristics off. The CDG Linux implementation was presented to the Linux community and Industry at Linux Conference Europe 2015. Kenneth Klette Jonassen developed the Linux version in collaboration with David Hayes and Andreas Petlund. As of version 4.2, CDG is a part of the mainline Linux kernel.

[6] https://lkml.org/lkml/2015/10/23/720

Dual Queue AQM

Lead partners: ALU, SRL, BT The Dual Queue Coupled AQM is a novel AQM that allows DCTCP (or other scalable / L4S) flows to coexist with Classic TCP flows. It assures that the Classic TCP flows have an approximately equal throughput (TCP-Fair) compared to the DCTCP flows, while the DCTCP flows maintain their ultra low latency capability. Throughput compatibility is achieved by applying a dropping/marking probability to Classic TCP packets that is the square of the marking probability applied to the L4S flows. Low latency for the L4S packets is achieved by scheduling the L4S packets with priority. At the moment of writing, the approval process within ALU for the release of the code is still ongoing. The latest status can be checked at the DCttH webpage http://riteproject.eu/dctth/. When approval is received, also the patches will be published.

6.2 Tools

The bottom section of table 6.1 lists tools developed and used in the RITE project. This is code that can be used for testing, analysis or measurement related to the study of Internet latency. These tools will be useful to any third party that wants to evaluate the mechanisms developed in RITE. There are also tools of more general interest that will be valuable to other communities, like the AQM detection tool that can be used by measurement researchers and Industry trying to survey the deployment of this important latency-reducing class of mechanisms. The list of tools also contains a patch set and build script that allows for building an integrated kernel with the key RITE mechanisms available as Linux patches. The table points to deliverables where more details about the tool can be found.

TCP evaluation suite

Lead partner: UiO

This suite of realistic tests implemented in the ns2 simulator facilitates testing of TCP modifications in a number of standard scenarios. The software has built upon previous work by Gang Wang and Yong Xia (NEC China) and David Hayes at Swinburne University of Technology (Melbourne, Australia). The public bitbucket repository contains test scripts along with a tailored ns2 2.35 simulator. The traffic traces are available from the ICCRG tools page. Elements of this tool have been used in the evaluation of a number of RITE proposals.

Pcap latency analysis tool

Lead partners: SRL, KaU

RITE has developed a tool for analysis of PCAP formatted files with a special focus on latency. The tool can be used to get separate statistics on retransmission latency as well as aggregated latency statistic over a group of different connections. Functionality for calculating one-way delay above the base OWD and compensate for clock drift between sender and receiver computer is implemented. It also has support for analysing the bytewise latency for a TCP stream when the sender bundles old data with new using RDB or if bundles are made upon retransmissions, which occur in the Linux kernel. The tool was developed by Andreas Petlund, Kristian Evensen, Bendik Rønning Opstad and Jonas Markussen. It is available through a public Bitbucket repository.

Tool for Automatic Detection of AQM (TADA)

Lead partners: SRL, KaU

The last years have seen a lot of effort to show the benefits of AQMs over simple tail-drop queuing and to encourage deployment. Yet it is still unknown to what extent AQMs are deployed in the Internet. RITE has developed an end-to-end active measurement method to detect AQMs on the path bottleneck. We have developed an active measurement tool, *TADA*, and evaluated our measurement methodology on a controlled experimental testbed. Experimental results show that the proposed approach provides the basis to identify whether an AQM is deployed on the bottleneck. The latest stable version of the tool is available through a public Bitbucket repository.

RITE Integrated Kernel

Lead partners: SRL, KaU, MEGA

The RITE project has prepared a patch set for the 3.18.5 Linux kernel containing the key end-host mechanisms developed in the project. The patch set also includes a build script that will download the correct kernel source, apply the patches, build the kernel copying the running kernel configuration and install the kernel for testing. The mechanisms provided with the kernel are New CWV, RTO Restart, TLP Restart, Redundant Data Bundling and a tuning option for experimenting with different sizes for the TCP initial window upon restart. The integrated kernel allows for interested parties to have a low-threshold way of experimenting with and build upon RITE mechanisms. It is shared from the Resources page on the RITE website.

7 Summary

We believe that RITE's results will be exploited extensively by industry, the open source community and academia. Examples of the exploitation routes include:

- The reduction in end-to-end delay with RITE mechanisms can have a great benefit for on-line real-time games, interactive video applications and other cloud-based services, and future virtualised network services. There is notable interest from the commercial /operational departments of BT, Megapop and ALU.
- The new queueing mechanism on network equipment is simpler, as well as more effective, which can reduce both costs and operational complexity.
- Various end host mechanisms developed by RITE (some in conjunction with the new queueing mechanism) can achieve lower end-to-end latency.
- We have been very open about our work as we believe this is the best way to achieve widespread exploitation via standardisation and deployment in operation systems.
- We have achieved many successes at the IETF, including helping in the creation of 2 new working groups, 3 RFCs, 2 specifications approved for publication as RFCs and 5 working group documents. These are detailed in Deliverable 4.2.
- One RITE mechanism has already been incorporated into the mainline Linux kernel, and thus is available to millions, whilst several other mechanisms are about to be submitted to the Linux kernel gatekeepers.