



# Hybrid Home Networks Links Quality Supervision

Oscar Olvera Irigoyen

## ► To cite this version:

Oscar Olvera Irigoyen. Hybrid Home Networks Links Quality Supervision. Networking and Internet Architecture [cs.NI]. Télécom Bretagne, Université de Rennes 1, 2012. English. NNT: . tel-00808194v1

**HAL Id: tel-00808194**

**<https://theses.hal.science/tel-00808194v1>**

Submitted on 5 Apr 2013 (v1), last revised 22 Apr 2013 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Sous le sceau de l'Université Européenne de Bretagne**

## **Télécom Bretagne**

**En habilitation conjointe avec l'Université de Rennes 1**

École doctorale - MATISSE

---

### **Hybrid Home Networks Links Quality Supervision**

---

#### **Thèse de Doctorat**

Mention : « Informatique »

**Présentée par Oscar Olvera Irigoyen**

Département Réseaux, Sécurité et Multimédia (RSM)

**Directeur de thèse : Laurent Toutain**

Soutenue le 20 Décembre 2012

#### **Jury :**

**M.** Gerardo Rubino

**Mme.** Catherine Rosenberg

**M.** Martin Heusse

**M.** Lauren Toutain

**M.** Abdesselem Kortebi

- Professeur à l'IRISA, Université de Rennes 1 (Président)

- Professeur à Waterloo University (Rapporteur)

- Professeur à l'Institut Polytechnique de Grenoble (Rapporteur)

- Professeur à Télécom Bretagne (Directeur de thèse)

- Orange Labs R&D (Co-encadrant)



To my wife Zélie and to my daughters Constance and Elinor

To my Parents Maria-Adelita and Jorge

To the memory of my brother Eduardo



# Acknowledgments

A number of persons have greatly assisted my work to finish this dissertation. A major place and sincere appreciations are owed to my advisors, Dr. Laurent Toutain, Associate Professor at Telecom Bretagne and Dr. Abdesselem Kortebi, R&D Engineer at Orange Labs for their constant support, guidance and encouragement. I have benefited from many valuable discussions throughout this work. Abdesselem Kortebi's expert insights have significantly contributed to enrich my thesis.

I gratefully thank Professor Catherine Rosenberg, Professor at Waterloo University and Martin Heusse, Professor at Institut Polytechnique de Grenoble for accepting to review my dissertation and for their valuable remarks. I would like to thank Professor Gerardo Rubino at IRISA Université de Rennes 1 and Dr. David Ros, Associate Professor at Telecom Bretagne for accepting to be members of the jury for this dissertation.

I am very grateful to Orange Labs, for the sponsorship of my PhD study. I also acknowledge and thank Telecom Bretagne and my Ecole Doctorale MATISSE for having me as a member of their PhD students.

I am thankful to all my colleagues, academic and support staff at Orange Labs and Télécom Bretagne for the support rendered to me. I specially thank Dr. Emanuelle Cressan responsible of the "Multiservices Architecture of Network and Gateways" team and Dr. Philippe Chanclou responsible of the "Advanced Studies on Home and Access Networks" team, both at Orange Labs, for their continuous managerial support. I very thank to Dr. Martial Bellec, Dr. Jean-Philippe Javaudin and Dr. Pierre Jaffré for their valuable insights and comments throughout all the period of my research at Orange Labs. I would also like to thank Oumar Diarra and Hakim Senhadji, two master students, for their invaluable help with the experimental part of this dissertation

I would like to extend my gratitude to my parents and family for their unconditional support and teaching. They taught me the values that have allowed me to become successful in life.

Finally, I am most grateful to Zélie, my wife and Constance and Elinor, my daughters for standing with me throughout both difficult and good times, and for their love, patience and sacrifices.

# Abstract

Home networks are evolving vigorously and are also becoming more and more heterogeneous and complex. Home networks contain new services and numerous devices with wired and wireless links. At the same time, users demand high levels of quality of service for many new applications. Hence, the supervision of links quality tends to be mandatory in emergent home networks, to trigger QoS mechanisms. In this context, this dissertation proposes the utilization of the available bandwidth as a performance indicator in hybrid home networks (e.g. for remote service-evaluation, diagnosis and fault detection) and as link-state metric for a number of QoS mechanisms (e.g., admission control, path selection and load balancing).

In this dissertation, we firstly explain why available bandwidth probing is a fundamental tool in the new QoS architectures for hybrid home networks and we explore different use cases. We investigate the networking constraints that affect bandwidth probing in hybrid home networks. We explain why most tools to measure available bandwidth on Internet paths are not very efficient on home networks. We make a taxonomic study of a number of state-of-the-art probing techniques and probing tools for available bandwidth. At that aim, we propose a functional framework, called Metrics Profile Capture Filter and Feedback (MPCFF). Based on these studies, we choose Iperf in TCP mode as an attractive tool to estimate available bandwidth. Then, we conduct a performance evaluation on an experimental test bed, to compare Iperf in lightweight TCP-mode vs. various state-of-the-art tools (Wbest, Pathchirp, Pathload and IGI/PTR). We show that Iperf in lightweight TCP mode is superior in terms of accuracy and speed of convergence. Finally, we integrate Iperf in lightweight TCP mode into a path selection protocol using the Inter-MAC software (developed in the Omega European project). We show, on a test bed, how Iperf in lightweight TCP mode enhances the Inter-MAC path selection by eliminating performance deterioration due to time-varying capacity links. Our results show that Iperf in lightweight TCP mode, accurately, triggers path selection and load balancing mechanisms, to adapt the network resources to the strong degradation of WiFi and PLC links.

Available bandwidth probing detects and helps to control real-time deviations of the quality of experience, associated of sensitive applications. At the same time, available-bandwidth probing helps the service provider to isolate failures and to monitor the end-to-end residential service, on demand or on long-term periodicity basis.



**Keywords:** Home Network, Network Performance, Available Bandwidth, Link Capacity, Probing, QoE, QoS, Link Quality, Iperf, Path Selection.

# Résumé

Les réseaux domestiques connaissent une évolution importante et deviennent de plus en plus hétérogènes et complexes. Ils intègrent de nouveaux services, des équipements variés et des technologies de connectivité filaires et sans fil. De plus, les utilisateurs exigent une qualité de service d'un niveau très élevé pour de nombreuses nouvelles applications. Par conséquent, le contrôle de la qualité des liens tend à être obligatoire dans les réseaux domestiques émergents, pour déclencher les mécanismes de QoS. Dans ce contexte, cette thèse propose l'utilisation de la bande passante disponible en tant qu'indicateur de la performance, dans les réseaux domestiques hybrides, par exemple, pour la détection de défaillances, l'évaluation de service, et le diagnostic. La bande passante disponible peut-être aussi une métrique d'état de lien pour un certain nombre de mécanismes de qualité de service, par exemple, le contrôle d'admission, la sélection de chemins et l'équilibrage de charges.

Dans cette thèse, nous expliquons pourquoi la mesure active (« probing ») de la bande passante disponible est un outil fondamental dans les architectures de QoS pour les réseaux domestiques hybrides. Nous explorons aussi différents cas d'utilisation. Nous étudions les contraintes de réseau qui affectent le « probing » de la bande passante dans les réseaux domestiques hybrides. Nous expliquons pourquoi la plupart des outils de mesure de la bande passante disponible pour mesurer les chemins d'Internet ne sont pas très efficaces sur les réseaux domestiques. Nous faisons une étude taxonomique d'un certain nombre de techniques et d'outils état de l'art pour le « probing » de la bande passante disponible. À cette fin, nous proposons un Framework appelé « Metrics Profile Capture Filter and Feedback » (MPCFF).

Sur la base de ces études, nous avons choisi « Iperf en mode TCP » comme un outil intéressant pour estimer la bande passante disponible. Ensuite, nous procédons à une évaluation de la performance sur test bed, pour comparer « Iperf en mode TCP léger » vs divers outils de l'état de l'art (Wbest, Pathchirp, Pathload et IGI / PTR). Nous montrons qu'Iperf en mode TCP léger est supérieur en termes de précision et de vitesse de convergence. Enfin, nous intégrons Iperf en mode TCP léger dans un protocole de sélection de chemins, en utilisant le logiciel Inter-MAC (développé dans le projet Européen Omega). Nous montrons, sur un test bed, comment Iperf en mode TCP léger améliore la sélection du chemin de l'inter-MAC en contrôlant la dégradation des performances due aux liens à capacité variable dans le temps. Nos résultats montrent qu'Iperf en mode TCP léger

déclenche la sélection de chemins et les mécanismes d'équilibrage de charges de façon précise. Ça permet d'optimiser les ressources réseau en présence d'une forte dégradation des liens WiFi et des liens PLC.

Le « probing » de la bande passante disponible détecte et aide à contrôler, en temps réel, les dégradations de la qualité d'expérience, associées aux applications sensibles. Dans le même temps, le « probing » de la bande passante disponible assiste le fournisseur de services réseau en isolant les pannes et en surveillant les services résidentiels de bout à bout, à la demande ou sur une périodicité à long terme.

**Mots-clés** : réseau domestique, performance réseau, bande passante disponible, capacité de lien, « probing », QoE, QoS, qualité du lien, Iperf, Sélection de chemin.

# Contents

Acknowledgments .....	i
Abstract.....	iii
Résumé .....	v
List of Figures .....	xi
List of Tables.....	xiii
List of Abbreviations .....	xv
Glossary .....	xix
1. Introduction .....	1
1.1. Overview of Home Networks .....	1
1.2. Problem Statement .....	4
1.3. Research Objectives and Contributions .....	6
1.4. Thesis Outline .....	8
2. Background: Available Bandwidth Monitoring and QoS Mechanisms in Home Networks.....	9
2.1 End-to-end Impairments, Management Processes and QoS Mechanisms: Interdependencies .....	9
2.1.1 QoS/QoE budget.....	9
2.1.2 Key Performance Indicators .....	10
2.1.3 QoS Mechanisms, Metrics and Logical Planes .....	12
2.2. Home Network QoS Architectures .....	16
2.2.1 Brief Historical Background.....	16
2.2.2 QoS Home Network Architectures .....	18
2.3. Available Bandwidth Monitoring in Home Networks .....	23
2.3.1 Layer-N Capacity in Home Networks .....	24
2.3.2 Layer-N Available Bandwidth in Home Networks.....	26
2.3.3 Available Bandwidth Monitoring in the Home Network .....	27
2.4. Use Cases .....	28

2.5. Summary .....	33
3. State of the art of Available Bandwidth Probing: Network Constraints, Techniques and Tools .....	35
3.1. Networking Constraints for Bandwidth Probing .....	36
3.1.1 Deterministic Measurement vs. Probing Measurement .....	37
3.1.2 Time Scales and Traffic Classes .....	37
3.1.3 End-Centric Probing vs. Network-Centric Probing .....	38
3.1.4 Networking Constrains for Bandwidth Probing on Internet Paths .....	38
3.1.5 Networking Constrains for Bandwidth Probing on Home Network Paths .....	38
3.2. Principles of Bandwidth Probing Techniques .....	39
3.2.1 The Variable Packet Size (VPS) Technique for Capacity Probing (by-hop) .....	40
3.2.2 The Variable Pair Dispersion (VPD) & Variable Train Dispersion (VTD) Techniques for End-to-End Capacity Probing .....	41
3.2.3 The Direct-Probing or Probe Gap Model (PGM) for End-to-End Available Bandwidth Probing .....	43
3.2.4 The Iterative-Probing or Probe Rate Model (PRM) for End-to-End Available Bandwidth Probing .....	44
3.2.5 Profiler, Capture, Filtering and Feedback (MPCFF) Framework .....	46
3.3. Characterization of Candidate Available Bandwidth Probing Tools for Home Networks .....	51
3.3.1 Spruce .....	51
3.3.2 IGI/PTR .....	51
3.3.3 Wbest .....	52
3.3.4 Allbest .....	52
3.3.5 Pathload .....	54
3.3.6 Pathchirp .....	54
3.3.7 Iperf .....	55
3.4. Selecting Available Bandwidth Probing Tools for Home Networks .....	57
3.4.1 Iperf in Lightweight TCP Mode, an Attractive Alternative .....	58
3.5. Summary .....	59
4. Iperf in lightweight TCP mode in Hybrid Home Networks .....	61
4.1. Introduction .....	61
4.2. Methodology .....	62
4.2.1 Principles of Iperf in lightweight TCP mode .....	62

4.2.2 Tools Performance and Measurement Process .....	63
4.3. Experimental Test bed.....	64
4.3.1 Capture Automation of Performance Descriptors .....	65
4.3.2 Path Scenarios .....	66
4.4. Results .....	67
4.4.1 Benchmarking.....	67
4.4.2 Global Results .....	74
4.5. Summary .....	75
5. Iperf in Lightweight TCP Mode for Path Selection in Heterogeneous Home Networks	77
5.1. Introduction.....	77
5.2. Methodology .....	78
5.2.1 Inter-MAC Path Selection protocol.....	78
5.2.1 Iperf in lightweight TCP mode .....	79
5.3. Experimental Test bed.....	80
5.3.1 Test Bed Features.....	81
5.3.2 Experimental Tests .....	83
5.4. Results .....	84
5.5. Summary .....	88
6. Conclusions and Perspectives.....	89
Appendix A: Résumé étendu en Français.....	93
A.1. Contexte : les réseaux domestiques.....	93
A.2. Problématique .....	96
A.3. Objectifs de la thèse et contributions .....	99
A.4. Organisation de la thèse.....	101
A.5. Conclusion de la thèse .....	101
Appendix B: The Java Test Automation Interface.....	107
Appendix C: Configuring Iperf in Lightweight Mode for Available Bandwidth Probing	
.....	109
C.1. Test Bed Setup .....	109
C.2. Methodology.....	111
C.3. PCs Configuration .....	112
Configuration of PC <sub>1</sub> .....	112
Configuration of PC <sub>2</sub> .....	114
Configuration of PC <sub>3</sub> .....	114

Configuration of $PC_4$ .....	116
C.4. Results.....	116
Bibliography .....	119

# List of Figures

Fig. 1.1. Home Network Services and Hybrid Connectivity Technologies .....	1
Fig. 1.2. “End-to-end residential path” based on ITU-T/Y.2173 [51].....	1
Fig. 2.1. Logical Planes found in the WAN and LAN segments ITU-T/Fig.13/Y.1001.....	1
Fig. 2.2. Timeline QoS Home Networks Evolution .....	1
Fig. 2.3. Physical Topologies: (a) tree, (b) mesh and (c) arbitrary topology.....	1
Fig. 2.4. Layer-N capacity on a Fast Ethernet link in function of L2 PDU size.....	1
Fig. 2.5. Home Network Control Loop and Performance Monitoring.....	1
Fig. 3.1. Bandwidth Probing .....	1
Fig. 3.2. MPCFF functional framework .....	1
Fig. 3.3. Packet pairs and trains dispersion, based on [92] .....	1
Fig. 4.1. Path Scenarios on Test Bed .....	1
Fig. 4.2. Scenario (a) 100Base-TX: Available Bandwidth, L2 Overhead and Probing Interval versus Cross Traffic.....	1
Fig. 4.3. Scenario (b) 100Base/TX-IEEE802.11a: Available Bandwidth, L2 Overhead and Probing Interval versus Cross Traffic .....	1
Fig. 4.4. Scenario (c) 100Base-TX/HPAV/100base-TX: Available Bandwidth, L2 Overhead and Probing Interval versus Cross Traffic .....	1
Fig. 4.6. Scenario (e) 100Base-TX/HPAV/IEEE802.11a/IEEE802.11a/100Base-TX: Available Bandwidth, L2 Overhead and Probing Interval versus Cross Traffic.....	1
Fig. 4.2. Global Results.....	1
Fig. 5.1. Path scenarios on testbed.....	1
Fig. 5.2. $A_1$ and $A_2$ with default implementation (passive + static) measurement.....	1
Fig. 5.3. Spirent report of $F_1$ and $F_2$ bandwidth with default Inter-MAC path selection (passive + static measurements), scenario 2.....	1
Fig. 5.4. $A_1$ and $A_2$ with Iperf in lightweight TCP mode measurement .....	1
Fig. 5.5. Spirent report of $F_1$ and $F_2$ bandwidth with Inter-MAC path selection + Iperf in lightweight TCP mode measurement, scenario 2 .....	1
Fig. A.1. Services dans le réseau domestique et technologies hybrides de connectivité .....	1



Fig. A.2. Chemin du service résidentiel de bout en bout basé sur ITU-T/Y.2173 [51].....	1
Fig. B.1. Screenshot of the Java test interface .....	1
Fig. C.1. Simple example of using Iperf in lightweight TCP mode.....	1
Fig. C.2. Cross traffic injected from PC <sub>3</sub> using Iperf in UDP mode .....	1
Fig. C.3. Available bandwidth measurements from PC <sub>1</sub> with Iperf in lightweight TCP mode .....	1

# List of Tables

Table 2.1. QoS Mechanisms and Performance Metrics in Packet Networks .....	1
Table 2.2. QoS Architectures for Home Networks.....	1
Table 2.3. Example Home Network Paths for the HDTV Service .....	1
Table 2.4. Transport Layer QoE based end-to-end Targets for a Flow of 12Mbps-HDTV-MPEG-4 AVC (from TR-126) .....	1
Table 2.5. Bit Rate Envelope for the Home Network Example .....	1
Table 3.1. Structural Characteristics of the Candidate Tools to Measure Available Bandwidth in Home Networks .....	1
Table 4.1. Global Performance Results.....	1
Table 4.1. Tools Performance Characteristics.....	1
Table 5.1. Monitored Paths.....	1
Table 5.2. Tested Path Capacity Conditions.....	1
Table 5.3. Default Measurements (Mbps) and Inter-MAC Path Selection.....	1
Table 5.4. Iperf in lightweight TCP mode (Mbps) Measurements and Inter-MAC Path Selection .....	1



# List of Abbreviations

AAL	ATM Adaptation Layer
ABR	Available Bit Rate
AGW	Access Gateway
ARPANET	Advanced Research Projects Agency Network
AS	Autonomous System
ATM	Asynchronous Transfer Mode
BAN	Body Area Network
BDP	Bandwidth Delay Products
CAC	Connection Admission Control
CATV	Cable Television
CBR	Constant Bit Rate
CSMA/CD	Carrier Sense Multiple Access / Collision Detection
Diffserv	Differentiated Services
DLNA	Digital Living Network Alliance
DSCP	Differentiated Service Code Point
E2E	End-to-End
FTV	Flow of TV
FTP	File Transport Protocol
HDTV	High Definition Television
HGI	Home Gateway Initiative
HPAV	Home Plug Audio/Video
HTTP	Hyper Text Transfer Protocol
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IGI	Initial Gap Increasing
IMN	Inter-MAC Node
Intserv	Integrated Services
IP	Internet Protocol

IPTV	IP Television
ISDN	Integrated Service Digital Network
ISP	Internet Service Provider
ITU	International Telecommunications Union
KPI	Key Performance Indicator
FIFO	First-in-First-out
LAN	Local Area Network
L1, L2, L3, L4	Layer 1, Layer 2, Layer 3, Layer 4
MoCA	Multimedia over Coax Alliance
MOS	Mean Opinion Score
MPCFF	Metrics Profiling Capturing Filtering and Feedback
MPLS	Multi Protocol Label Switching
MT	Measurement Time
MTU	Maximal Transmission Unit
NAS	Network Attached Storage
OS	Operating System
OSPF	Open Shortest Path First
OWD	One-Way-Delay
PAN	Personal Area Network
PBM	Packet Bunches Modes
PDU	Protocol Data Unit
PDV	Packet Delay Variation
PGM	Probing Gap Model
PLC	Power Line Communication
POF	Plastic Optical Fiber
PREP	Path Replay Message
PREQ	Path Request Message
PRM	Probing Rate Model
QoE	Quality of Experience
QoS	Quality of Service
RTT	Round-Trip-Time
RSTP	Rapid Spanning Tree
RSVP	Resource Reservation Setup Protocol
SCDR	Sub-Capacity Dispersion Range

SDH	Synchronous Digital Hierarchy
SLA	Service Level Agreement
SLO	Service Level Objective
SLoPS	Self-loading Periodic Streams
STB	Set Top Box
STP	Spanning Tree Protocol
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TOPP	Trains of Pairs of Packets
TTL	Time to Live
TV	Television
UDP	User Datagram Protocol
UPnP	Universal Plug and Play
UTP	Unshielded Twisted Pair
VoIP	Voice over IP
VPD	Variable Packet Pair Dispersion
VTD	Variable Packet Train Dispersion
VPS	Variable Packet Size
VP/TD	Variation of Pair/Train Dispersion
WAN	Wide Area Network
WLAN	Wireless Network
WiFi	Wireless Fidelity



# Glossary

Available bandwidth of a link is given by the unutilized fraction of the capacity of this link in a specific period of time. See the definition of available bandwidth in chapter 2.

Cross-traffic is all the traffic not associated with active dummy probes.

Probing-traffic, is the application traffic (passive) or dummy-traffic (active) that is considered to infer, the bandwidth on a measured path.

Bandwidth probing is an in/out process that infers capacity or available bandwidth from the delay behavior of dispersion or bit rate of probing packets after they interact with the nodes and cross-traffic on the measured path.

Convergence time is the period, needed by a probing tool, to provide a metric estimation. Convergence time includes the probing time and the computation time required by the tool, to obtain a metric estimation after the probing tool receiver processes the probing packets.

End-to-end impairments include latency, packet delay variation (PDV) (also called jitter), capacity fluctuations and packet losses, segment. End-to-end-network-impairments are cumulated through the WAN and through the home network segments and impact the performance of end-user applications.

Narrow link: is the link with the minimum capacity on the path and determines the E2E capacity.

Polling interval is the period between the beginnings of two successive executions of a probing tool.

Probing interval is the period between the beginning and the end of the first and the last probing packets of two successive executions of a probing tool.

QoS targets are a number of performance bounds or metrics, such as max delay, max delay variation, max packet losses, required bit rate, etc. The QoS targets can be used in both cases: a) as a design reference, to choose the required QoS mechanisms or b) as a



monitoring reference, to know if a expected QoE is accomplished in the actual network operation.

Test interval is the whole duration of the test, which is chosen to characterize a metric, using a probing tool and given the presence of a cross-traffic pattern on every different path.

The tight link is the one with the minimum available bandwidth on a path.

# Hybrid Home Networks Links Quality Supervision



# Introduction

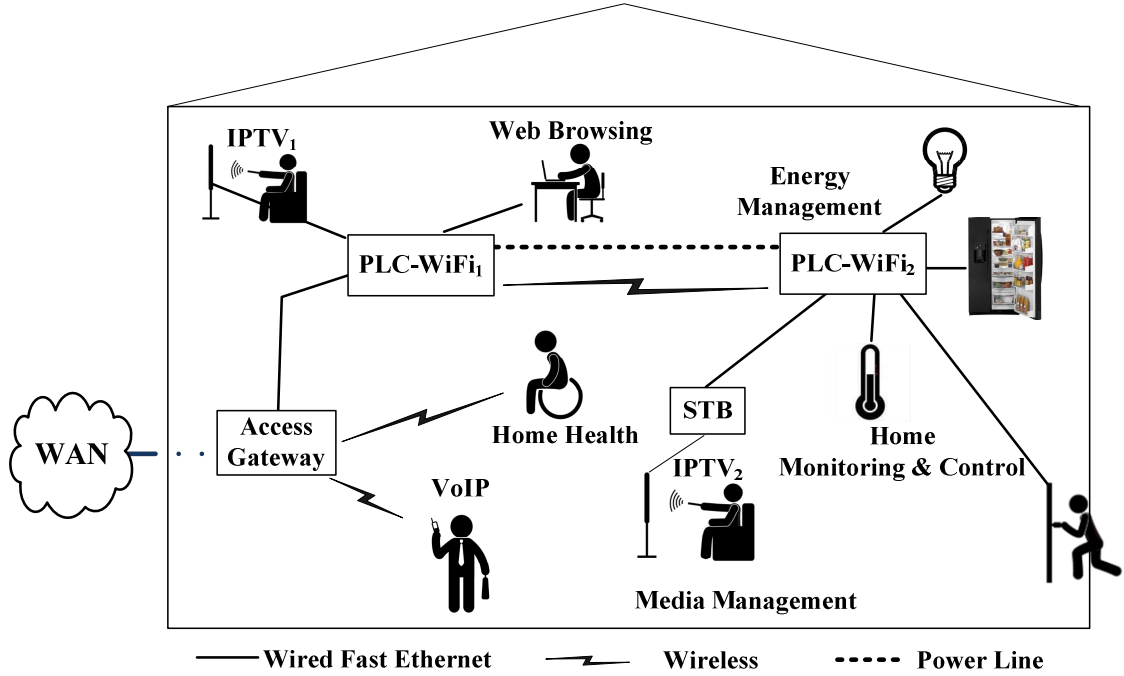
## 1.1. Overview of Home Networks

Home networks show a vigorous evolution and are becoming more and more heterogeneous and complex. Home networks contain new applications, devices with wired and wireless links. At the same time, users demand high levels of quality of service for many new applications.

Let us introduce the context where the home network is placed today, with respect to the end-user perspective and with respect to the service provider perspective.

From the end-user perspective, the home network tends to be increasingly pervasive and heterogeneous due to the constant evolution of its new network applications, new terminals and diverse connectivity technologies. In fact the home network is the interconnection of the Access Gateway (Home Gateway) and the different terminals (including, PCs, tablets, TVs, NAS, etc.). Several infrastructure devices are used in home networks: Ethernet-based switches, wireless interfaces, PLC plugs, etc. Therefore, the home networks can have hybrid paths, which are constituted by a number of serialized links, such as wired links with Ethernet, coaxial, phone wires, power line communication (PLC) links and wireless links within Personal Area Networks (PANs), Body Networks (BANs) and Wireless LANs [116]. Additionally, we can find new home network devices with hybrid links, wireless and PLC, see for instance Fig. 1.1. The home networks also transport several types of flows such as IPTV flows, VoIP flows, bulky flows, streaming, diverse web flows, etc.

Fig. 1.1 shows a home network with the typical central Access Gateway (AGW), which is connected to the WAN and with several serialized links (paths), which interconnect terminals and network nodes.



**Fig. 1.1. Home Network Services and Hybrid Connectivity Technologies**

The evolution of the home network is closely related to the reduction of CPU costs of end-user terminals, the continued increment of interfaces speeds [94] and the strong penetration of the broadband residential access [17]. As we detail in section 2.2, the evolution of the home network begins several decades ago. Even though, the home networks, centered on the access gateway, begin to be popular in the '2000s, different proposals for ISDN/ATM-based home networks can be found in the '80s and '90s. In the '2000s we assist to the Internet boom and we find that the access gateway becomes the convergent point for the Internet access (to substitute the narrow-band/phone-line dialup modems), the CATV and the telephone. The access gateway has been transformed from a layer 2 intermediate nodes (to connect one or a few number of PCs at home, a CATV terminal and a phone) to an evolved router. Today, the access gateway allows multiservice access for a number of IP devices such as electronic tablets, smart phones, game consoles, laptops, household appliances, HD/3D TVs, IP phones, etc.

So, these trends generate a new spectrum of promising services, in the residential market, such as home media management, energy management, health care, home monitoring and control, etc. [82]. However, due to its complexity, the home network is also highly prone to

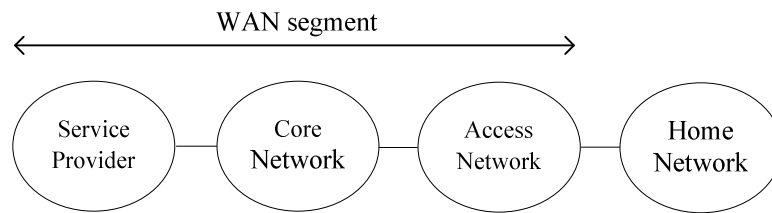
performance variability [13], [115] and becomes a potential source of the end-to-end service quality degradation.

So, these trends generate a new spectrum of promising services, in the residential market, such as home media management, energy management, health care, home monitoring and control, etc. [82]. However, due to its complexity, the home network is also highly prone to performance variability [13], [115] and becomes a potential source of the end-to-end service quality degradation.

That is why there is a great interest, from the service provider perspective, in optimizing the QoS mechanisms, and the remote and local management systems of the home networks. Using, those mechanisms and processes should allow protecting priority flows, maximizing the use of home network paths and facilitating remote failure detection and residential service restoration.

On the other hand, the residential service flows are impacted by the end-to-end network impairments, which are cumulated through the WAN segment and through the home network segment. Those impairments (such as latency, packet delay variation PDV, also called jitter, capacity fluctuations and packet losses) affect the end-user applications, as explained in [52]. For these reasons, the network operators require isolate and control the impairments due to the WAN segment with respect to the home network segment.

So, the WAN segment is basically constituted for the service provider sub-segment, the core network sub-segment and the access sub-segment (e.g., last mile) as shown in Fig. 1.2.



**Fig. 1.2. “End-to-end residential path” based on ITU-T/Y.2173 [51]**

For example, a typical IPTV flow, which is transmitted from the head-end of the IPTV service provider, is impacted by the propagation and queuing latencies, when it traverses all the network nodes through the WAN and home network segments. The packets (of this flow) will arrive to the end-user IPTV terminal with a specific distribution of inter-arrivals (delay and packet delay variation). For instance, if the delay and delay variation are higher than some required (end-to-end) QoS bounds, then the quality of the service will be impacted (e.g., voice/video degradation, high channel switching time, high portal retrieval time, etc.)

At the same time, performance fluctuations on the WAN and on the home network can reduce the capacity of the end-to-end path and make vary the available bandwidth. Those variations can correspond to rerouting changes on the core network, to environmental changes on the access link and perturbations and intensive utilization of bandwidth on the home network links. If the end-to-end available bandwidth reduces to less than the targeted bit rate (required value), when a flow is transmitted, this will have destructive effects over the transmitted flows and will cause packet losses. At the same time, if these packet losses are superior to the permitted packet losses, we will have diverse image degradations and sound interruptions, [40] it will reduce the perceived quality of the service or in more severe cases, it can lead to a complete disruption of the service.

Hence, the selection of correct metrics and performance indicators is critical for QoS control, remote and local monitoring-purposes on links and paths of home networks. Nowadays, we find a number of QoS architectures [108], [14], [44], [45] that allow controlling a number of link/path-state metrics with respect to a series of performance bounds or QoS targets, such as max delay, max delay variation, max packet losses, required bit rate, etc. These architectures should guarantee the performance of end-user applications in the home network. Then, even if there is not a single global metric that can simplify the implementation of these QoS architectures, it is desirable to have a reduced set of metrics, which can have enough performance representativeness.

## 1.2. Problem Statement

As mentioned, the home networks are highly prone to performance variability [13], [115] and become a potential source of the end-to-end service quality degradation. For this reason, there is an increasing requirement of protecting priority flows, maximizing the use of available home network paths, facilitating remote failure detection and assisting the restoration of the residential service. In this context, link/path quality supervision tends to be mandatory in emergent QoS architectures of home networks and network management processes.

As discussed, we find a number of metrics required by reservation or parametric QoS architectures such as max delay, max delay variation, max packet losses, required bit rate, etc. These architectures allow guaranteeing the performance of end-user applications in the home network.

But, what kind of performance metric do we use to characterize link/path quality in the home network?

A representative performance metric with a number of reasonable assumptions in the home network can simplify the deployment of QoS mechanisms and network management processes. This is the case of the available bandwidth that can be used as a link/path-state metric or as part of a key performance indicator (KPI). The reason we focus on available bandwidth is because this metric, differing to other metrics, is a direct measure of the additional load that a link or a path can carry, before it becomes saturated. Other metrics, such as delay, delay variation, packet loss rate and capacity can only determine whether a path is already congested, which reduces its degree of prevention of potential service degradation [56]. Additionally, available bandwidth can simplify the home network service monitoring, can facilitate fault detection and it appears as a recurrent link-state metric in different QoS architectures for home networks.

Then, the question that arises naturally is: how do we measure available bandwidth in hybrid home networks?

There are basically two approaches: the deterministic passive-measurement and the active probing-measurement. Deterministic measurement reads directly the current traffic statistics of physical interfaces without statistical post-treatment for inference purposes. The active-probing measurement requires sending probing packets to sample the behavior of queues, the transmission conditions and processing through the paths. Active probing requires the statistical treatment of the captured samples to infer a specific network metric. Some of the advantages of deterministic passive-measurement are high accuracy on paths with fixed capacity, low or no overhead and relative simplicity of deployment. However, passive measurement suffers of inaccuracy on paths with time-varying capacities, is dependent of lower-layer measurement-mechanisms and do not reflect, the nearest as possible, the behavior of the application layer.

Thus, this dissertation proposes the utilization of transport available-bandwidth as a performance indicator and link/path-state metric tool. We focus on bandwidth measurement through active probing.

The networking constraints that we consider in the processes of available bandwidth probing include:

- a) Space-time-scopes diversity



- b) Strong delay variability in probing flows, induced by multiple-access mechanisms and time-variant link capacities in the home network
- c) The need of a simple calibration and transparency to lower layers
- d) Measurement accuracy, low convergence time, low overhead and stability in different system platforms and home network conditions
- e) The need to facilitate end-centric or network-centric deployments
- f) The need to use badwidth probing with traffic classes

So, the problem is that most of the state-of-the-art tools, that have been developed for Internet paths, are inaccurate, slow and not stable enough (sensitive to clock resolution, timing slips, system interruptions and number, rate and size of probing packets [117], [98], [23], [7], [91]), particularly in the context of hybrid home networks. Therefore, we need to find a way to measure available bandwidth, considering all these networking constraints in home networks.

### 1.3. Research Objectives and Contributions

Taking into consideration the limitations of current probing tools and the constraints of the measurement process, this dissertation proposes using Iperf in a lightweight TCP mode to obtain maximum throughput of non-disruptive TCP flows and estimate the available bandwidth. According to the conducted performance evaluation tests, we support the hypothesis of Iperf as a convenient tool, because:

- a) Iperf is a well-known benchmarking tool and conversely to the other tools is stable
- b) Iperf has been developed for different operating systems including Windows, MacOS, Linux and Android
- c) Iperf, based on the TCP protocol, can be configured to be not disruptive, when real time applications traverse the measured paths
- d) Iperf can be used as an end-to-end available bandwidth estimation tool for residential services
- e) Home networks have short slow-start periods and a predictable stable TCP performance, in opposition to networks with large bandwidth delay products
- f) We can use TCP Iperf in a lightweight TCP mode, with short probing intervals instead its ordinary brute force utilization on WAN paths [103]

Our contributions consist, firstly in benchmarking, on an experimental test bed, Iperf in lightweight TCP mode with respect to several bandwidth probing tools (IGI/PTR, Wbest, Pahtload and Patchirp), in terms of accuracy, convergence time and intrusiveness.

Secondly, we propose and reproduce, with a prototype Java test interface, a probing methodology considering probing, polling and test time-scales to weight the Iperf overhead, depending on home network applications.

These two contributions have been published as: “Available bandwidth probing in hybrid home networks as part of the Local Metropolitan Area Networks” (LANMAN), 2011 18th IEEE Workshop on, 2011.

Thirdly, we applied the proposed available bandwidth measurement tool in an interesting use case of interface/path selection in home networks. More specifically, we integrate Iperf in lightweight TCP mode to two nodes with redundant hybrid links (WiFi/PLC) running an existing path selection protocol (developed during the OMEGA European project)

Fourthly, we compare, on a test bed, the path selection protocol and Iperf with respect to the original implementation. We introduce external interferences on the WiFi/PLC links between two devices, while two IPTV flows transit in the home network. Then, we show how path selection coupled to Iperf in lightweight TCP mode is able to react to the link performance degradation and to avoid the performance blind-spots due to the default way to measure available bandwidth of the original protocol. The “original path selection” biases the available bandwidth measurement, when the capacity (of WiFi or PLC links) changes. Conversely, the path selection protocol coupled to Iperf is capable to detect the implicit variations of capacity and estimate the available bandwidth, accurately, quickly and with low level of overhead.

These last contributions have been submitted and accepted under the title: “Available Bandwidth Probing for Path Selection in Heterogeneous Home Networks”, as part of GC'12 Workshop: The 7th IEEE International Workshop on Heterogeneous, Multi-Hop, Wireless and Mobile Networks (GC'12 Workshop - HeterWMN 2012).

Finally, we have done a survey of the main QoS mechanisms found in state-of-the-art architectures for home networks and their relationships with the available bandwidth. We also explore the interdependencies that exist between the definition of QoS/QoE targets,

network management processes and the QoS mechanisms through the residential path (including the home network). We underline the importance of available bandwidth as a link/path-state metric and as a key performance indicator to monitor and optimize home network resources. We also show different use cases where we use (transport-layer) available bandwidth to trigger QoS mechanisms such as admission control, best path selection and load balancing in the home networks.

## 1.4. Thesis Outline

The rest of this dissertation is composed as follows. In the second chapter, we explain why available bandwidth probing is a fundamental tool in the new QoS architectures for hybrid home networks and we explore different use cases, where available bandwidth probing can be implemented.

In the third chapter, we investigate the networking constraints that affect bandwidth probing in hybrid home networks. We make a taxonomic study of a number of state-of-the-art bandwidth probing techniques and available bandwidth probing tools. At that aim, we propose a functional framework called MPCFF to highlight the structural decomposition of the different techniques. Based on these studies, we choose Iperf in TCP mode as an attractive tool to estimate available bandwidth.

In the fourth chapter, we compare, on an experimental test bed, Iperf in lightweight TCP mode vs. Wbest, Pathchirp, Pathload and IGI/PTR available probing tools. We develop a homogeneous Java platform of tests and a methodology of available bandwidth probing in home networks. We show how Iperf in lightweight TCP mode is superior in terms of accuracy and speed of convergence.

Finally, using an experimental test bed, we apply Iperf in lightweight TCP mode to provide available bandwidth to an existing path selection protocol in hybrid home networks. The obtained results show that the measurements Iperf in lightweight TCP mode enhance the QoS and eliminate the performance blind-spots.

# Background: Available Bandwidth Monitoring and QoS Mechanisms in Home Networks

In this chapter, we study the main QoS mechanisms found in state-of-the-art architectures for home networks and their relationships with the available bandwidth. We also study the interdependencies that exist between the definition of QoS/QoE targets, network management processes and the QoS mechanisms through both the WAN and the home network. We aim to show the weight of available bandwidth as a link/path-state metric and as a key performance indicator to monitor and optimize home network resources. To illustrate that, we show some examples where we can use the transport-layer available bandwidth (obtained by available bandwidth probing) to trigger QoS mechanisms, such as admission control, best path selection or load balancing in the home networks.

## 2.1 End-to-end Impairments, Management Processes and QoS Mechanisms: Interdependencies

### 2.1.1 *QoS/QoE budget*

The definition of Quality of Service in ITU-T E.800 [42] refers to "the collective effect of service performance which determines the degree of satisfaction of a user of the service." As we have mentioned in the introduction, these collective effects reflect many of the

interdependencies among network impairments (delay, delay variation and packet losses) due to the serialized inter-nodal-connectivity-technologies and QoS mechanisms through the WAN and the home network.

As mentioned, the end-to-end impairments targets, along the WAN and the home network must be respected to avoid residential service quality degradation. Accordingly, understanding the QoS experienced by the end-user (called Quality of Experience (QoE) [46]) helps to define such set of maximal impairments supported by the network and the bit rate requirements for specific applications [43]. Ultimately, the QoS targets can be used in both cases: a) as a design reference, to choose the required QoS mechanisms or b) as a monitoring reference, to know if the expected QoE is accomplished in the actual network operation.

Even though the relationships between QoE and QoS parameters are complex and not linear, they can be derived empirically [105]. For example, the Mean Opinion Score (MOS) of a representative group of test persons can range from 1 (worst) to 5 (best) with respect to the perceived quality of a network application. In this sense, we can associate a combination of QoS impairment targets that can be supported for an IPTV application without visible degradation and which requires a specific standard coding and a specific encoding bit rate.

In recommendation TR-126 [105], we find some tables with a set of end-to-end targets to ensure the QoE for specific applications (e.g., IPTV, voice, Web-browsing, bulk data transfer/retrieval, interactive games, etc.)

In this way, a given set of end-to-end targets of QoS is associated to a specific QoE behavior for a given set of applications and a representative group of test users. At the same time, these targets can be split for the WAN and the home network segments. The methodologies of mapping QoE levels to QoS parameters are out of the scope of this work. Our aim is to use a series of “standardized” performance targets as a performance reference, for network management processes and QoS mechanisms.

### *2.1.2 Key Performance Indicators*

The need of defining a set of QoS targets for network management purposes can be related with the concept of Key Performance Indicators (KPIs) or called simply performance indicators.

Key Performance Indicators help to specify service performance targets in harmony with the business objectives of an organization [112]. In this sense, we can have a number of

low level QoS targets associated to an expected QoE for a set of specific applications. On the other hand, we choose a subset of those low level parameters to compose performance indicators, in terms of what simplifies the management processes and is important to the organization, what is accountable and impacts the contractual relationships between service providers and end costumers.

One example of this is the concept of Service Level Objective (SLO) of a Service Level Agreement (SLA), which can include a set of performance indicators. Those performance indicators must not be violated by any of the contract owners, in order to maintain the quality of a specific service and avoid contractual penalization [112]. For instance, if the network operator does not offer a sustained level of bandwidth and a bounded latency, for a videoconference service or if the costumer exceeds its contractual volume of transmitted data, then the network operator or the end-user will be obliged to pay the correspondent economic adjustments.

Another example of key performance indicators is found in the TR-160 recommendation [106]. A number of performance indicators are proposed to monitor the IPTV flows for a residential service. Some of those performance indicators are related with the user interface performance, such as the “IPTV Portal Information Retrieval Time” or the “IPTV Channel Switching Time.” Other performance indicators are related to the video and audio quality such as “IPTV Media Quality.” The IPTV Media Quality is a performance indicator that is associated to the MOS, according to the ITU-T P.911 recommendation and is expressed as follows in [106]:

$$TvQI [MOS\_estimate] = Function \{audio\ codec\ type\ and\ bit-rate,\ video\ codec\ type\ and\ bit-rate,\ transmission\ bit-rate,\ IP\ packet\ loss,\ video\ size,\ audio\ format,\ video\ format,\ audio\ and\ video\ delay\ difference\ and\ delay\ variation\} \quad (2.1)$$

Therefore, these performance indicators can be obtained using specific set top box (IPTV decoder) reports and end-to-end network measurements.

In this dissertation, we use the term of performance indicator to identify a network metric or a set of network metrics that can have enough representativeness of the home network paths performance. These metrics can impact the perception of the end-user and can possibly simplify the control and monitoring processes, related with the service provider operations. Therefore, we focus on the transport-layer available bandwidth in home

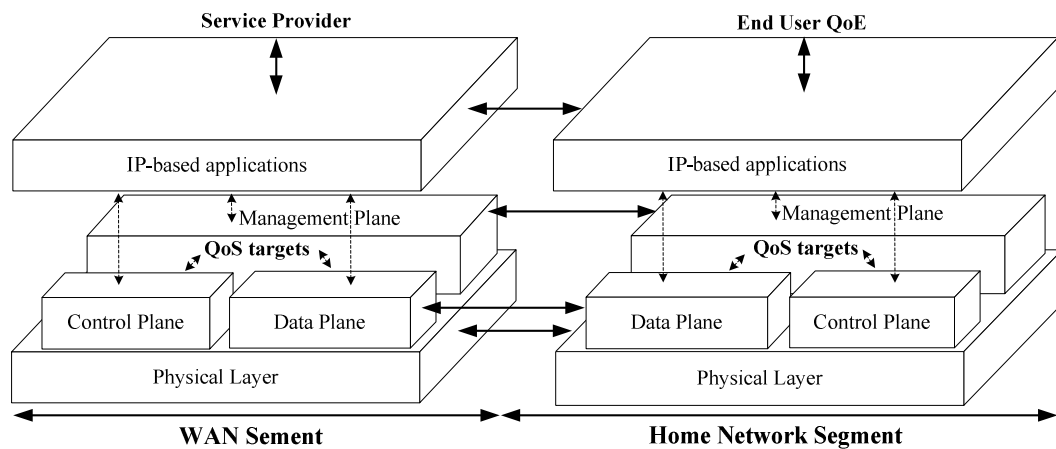
network paths, because it can simplify the service monitoring in the home network, it can facilitate fault detection and it appears as a recurrent link state metric in different QoS architectures for home networks. We detail this below.

### 2.1.3 QoS Mechanisms, Metrics and Logical Planes

From the early '80s innumerable architectures and QoS mechanisms have been proposed to support convergent services in large-scale data networks [114], [113]. We find, for instance, the following architectures: B-ISDN/ATM [79], Intserv/RSVP [6], Diffserv [4], MPLS [111] and other flow-based architectures such as Flow-Aware Networking [61] and Flow-State-Aware Transport [41], etc. Nowadays, some of the basic building QoS mechanisms, used in these architectures, tend to be integrated in the novel QoS home networks architectures.

In this context, the "separation principle" of logical planes [68] can be very useful to classify QoS mechanisms and management processes [9], [50], [74], [2]. Those logical planes include the Data/Forwarding Plane, the Control Plane and the Management Plane. For example, in the case of IP, it was designed as a datagram service that represents the "minimum network service assumption"[10], which coupled to a "large enough" address space, enables "horizontal" scalable connectivity, on a large number of independent administrative domains. Accordingly, IP was not defined considering these convergence planes, so it is necessary to map them on associated sub-protocols [48], [50].

Fig. 2.1 is based on the ITU-T/Fig.13/Y.1001 recommendation [48] and shows graphically, the relationships among logical planes along the WAN and the home network segments.



**Fig. 2.1.** Logical Planes found in the WAN and LAN segments ITU-T/Fig.13/Y.1001

The separation principle allows categorizing the QoS mechanisms and processes, depending on the time-scales of the network metrics, which are required to accomplish such management and control processes.

In this sense, the QoS mechanisms of large-scale networks architectures, cited previously, can be associated to the logical planes and correspondent performance metrics. However, for the moment, we are interested in the relationships between the building QoS mechanisms and management processes, separated from their architectures. Our aim is to show the “time-scope” of typical performance metrics.

For instance, the management processes, included in the Management Plane, require metrics that have a macroscopic scope, with respect to the tasks and mechanisms, found in the Control Plane and Data Plane. So, we have separated the metrics in three categories: the key performance indicator (KPI) category, which have a global or macroscopic scope, the “path-state” category and the “link/interface-state” category, both with a microscopic scope. These categories can give more or less weight to some metrics than others, depending on the studied scenarios.

We synthesize the correspondence between logical planes, management processes, QoS mechanisms and some performance metrics in Table 2.1. Specific technical details of the relationships between performance metrics, QoS mechanism and management processes (showed in Table I) are out of the scope of this dissertation. As mentioned, our aim is to show the “time-scope” of these performance metrics.



TABLE 2.1. QoS MECHANISMS AND PERFORMANCE METRICS IN PACKET NETWORKS

Logical Plane	QoS Mechanisms & Management Processes	Metric Time-Scope	Delay	Delay Variation	Packet losses	Available bandwidth	Hop count	Reference (mechanism processes)
Management Plane	Service Level Evaluation	KPI	X	X	X	X		[74], [114]
	Event Application Monitoring	KPI	X	X	X	X		[80]
	Event Network Monitoring	KPI	X	X	X	X		[80]
	Traffic Restoration	KPI	X	X	X	X	X	[74], [68]
Control Plane	Routing Path selection Traffic Engineering	Path-State	X		X	X	X	[74], [2], [114]
	Admission Control	Path-State	X	X	X	X		[2], [114]
	Resource Allocation	Path-State	X	X	X			[68], [74]
Data Plane	Traffic-policing and shaping	Link-Interface-State	X	X	X	X		[74], [2]
	Traffic- queuing and scheduling	Link-Interface-State	X	X	X	X		[74], [2]

In Table 2.1, we have included five performance metrics: delay, delay variation, packet losses, hop counts and available bandwidth that are typically used to characterize network and application performance [80]. Those metrics can have a large time granularity (except hop count) and can exhibit relevant changes in the range of fractions of micro seconds to months or even years.

The Management Plane is related to the service evaluation, the event monitoring and the traffic restoration procedures, etc. The time-scales of the related management processes, at this plane, can range from seconds to days or months. For instance, each several minutes, a surge pattern of packet losses can affect the WAN-access segment (Fig. 1.2). It can cause, for instance, the IPTV service degradation in the home network terminals. On the same way, the available bandwidth (unutilized capacity) on PLC links or WiFi links (of the home network) can affect the speed of file sharing applications due to the RF interference of household appliances, or neighbor transmitters.

On the other hand, the involved metrics on the Management-Plane processes may also be oriented to the business objectives of the service provider, for service-accounting purposes. The service provider can be interested in knowing the availability, the usage-periodicity or the audience of a number of services during periods of days, weeks or months. Therefore,

these metrics can be part of key performance indicators (KPI) that reflect the MOS as showing in equation (1) or QoS targets in SLOs. For these reasons, the Management-Plane metrics have a macroscopic time-scope regarding the perception of the network services.

In contrast, the QoS mechanisms used typically on the Control Plane must react more rapidly than the processes found in the Management Plane. For instance the QoS routing, or path selection must track unpredictable outages of network paths, to allow rerouting the network flows on time. Then, the time-scope of the involved performance metrics will be of the order of one or many round-trip times of the transmitted flows (e.g., some milliseconds to several seconds, depending on the degree of tolerance of the associated network applications).

In backbone networks, such as SDH rings, the reaction time to reroute, in presence of path failures, is about of 50 ms [88]. However, in routed networks, using for example OSPF, the reaction time can largely increase from 30 to 40 seconds [20]. In local networks, the reaction time is of some units of second, when Rapid Spanning Tree (RSPT) [93] is used and about 30 seconds, when Spanning Tree Protocol (SPT) is active. In home networks, the reaction time to reroute can be of the order of several milliseconds, using for instance the Inter-MAC prototype software [62].

The Data-Plane Mechanisms deal with the forwarding treatment and flow control (sometimes after the classification and marking of flow packets). The Data-Plane requires time-scales of reaction that can be virtually instantaneous, when the packets are forwarded from an input interface to an output interface or from a source to a specific destination on a path. The packets can be classified depending on diverse traffic classes and applying packet priorities or drop-packets rules. Implicit flow control can be used to adapt the load of data applications to the available bandwidth on the network, which is the case of using TCP. Some video/audio applications can also react to the conditions of the path and adapt its coding bit-rate. In such flow control scenarios, the time of reaction can be of some round-trip-times of the flow trajectory, typically several milliseconds.

As we can see, the choice of QoS metrics for the home network is a key task. For instance, one may wish to have the lower number of QoS metrics that accurately represents the performance of a home network. This is very important for implementing the correct QoS mechanisms, monitoring and fault-detection tools. It is easier to deploy, routing, best path selection or admission control algorithms with only one link-state metric than with several metrics [114]. On the other hand, the fault diagnosis and network management processes may require a good compromise between the number of metrics, accuracy and complexity

to deploy measurements and alarms. The latter is crucial, since the end-user and the service provider require simplicity for self-care and remote support in the home networks.

## 2.2. Home Network QoS Architectures

The evolution of home networks and their services have accompanied the development of telecommunications. As we have discussed, home networks tend to assimilate many of the QoS mechanisms found in large-scale networks. In this section, we illustrate a brief historical-background of the services and home network technologies that has determined the emergency of QoS standards in the home networks. We also discuss the most relevant architectures for QoS in home networks. We aim to underline the main QoS-mechanisms, found in these architectures, and associate them to their performance metrics.

### 2.2.1 *Brief Historical Background*

In the last five decades, we have seen the transformation of home network services and technologies, from the visionary stage to markets, technologies and standards in a significant period of maturation.

The '60s mark the revolution of digital communications [73] and the emergency of packet networks [95]. At the same time, several potential residential services are projected for taking advantage of the new digital networks. For example in [90], the author mentions the home-to-home video-telephony as a desirable service, but clearly some obstacles to its development, at that time, were the cost and reliability of the terminals.

The '70s is the beginning of the ARPANET and the TDM-networks for telephone services. At that time, naturally, many digital services for the residential market begin to be imagined. For instance in [72], the authors mention some of these services, which include “the home healthcare monitoring,” “the home instruction,” “the gas/weather leakage monitoring” and “the electronic babysitting.”

In the '80s the definition of traffic classes and performance bounds for specific services becomes central for network dimensioning. At that moment, the home networks begin to be studied in terms of their potential analogical/digital signals (broadcast TV, broadcast radio, home automation, etc) and the required infrastructure for the home connections (wired and wireless links). In [24], a project called “Homenet” studies the control of the traffic in the home network, over heterogeneous transmission-media (such as power lines and infrared wireless links). In [28] and [18], “the home bus” is proposed as a uniform

cabling infrastructure. Also in [28], the authors contemplate several applications for the home network that only today begin to be deployed. These applications include digital HDTV, home-to-home videoconference, rich digital text, home automation and home security. In [107] and [81], a broadband home network is defined using ISDN-based service classes. Additionally in [27], the Power Line Bus (PLbus) with CSMA/CD spread spectrum is considered as a connectivity alternative.

In the second half of the '90s, the Internet is commercially opened [69], and it is required the characterization of the applications running on it for QoS purposes. From the point of view of the home network applications, in [63], the author lists a number of residential applications and associates them to three classes: streaming and block-transfer (both real-time), and applications of non-real-time. The author also associates these traffic-classes to some QoS requirements that include maximal supported delay, minimal bandwidth and the packet size distribution. In [21], a prototype of a home network in mesh, called "Warren" is described. Warren is based on an ATM switch of 25 Mbps ports, as a central home network node. Different ATM traffic classes are used, for instance Constant Bit Rate (CBR) on ATM adaptation layer 1 (AAL1) and Available Bit Rate (ABR) on AAL5. Warren connected home network terminals, including HiFi sound systems, TV, PCs and "baby monitors" (video cameras). Warren was compatible, to interconnect devices, with unshielded twisted pair (UTP), Plastic Optical Fiber (POF) or infrared links.

It was only, until the '2000s that the development of QoS standards got importance [96]. We describe some of the most relevant QoS standards for home networks in next section. Fig. 2.2 shows a time-line of the evolution of the home networks that we discussed before.

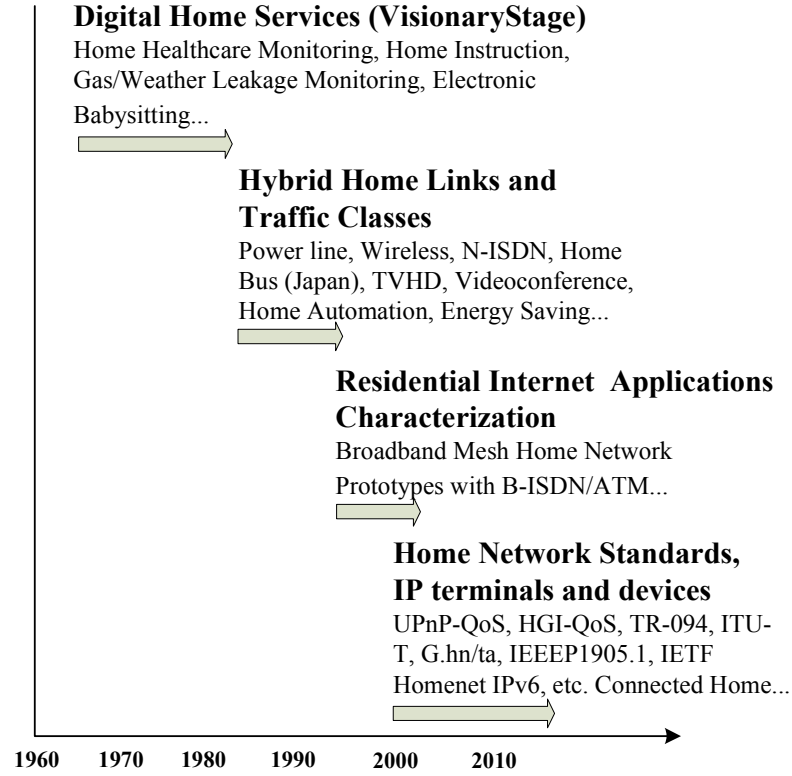


Fig. 2.2. Timeline QoS Home Networks Evolution

### 2.2.2 QoS Home Network Architectures

The home network QoS architectures tend to integrate characteristics of large-scale networks to prioritize and protect traffic classes. In this section, we compare most relevant QoS home network standards that we have found in literature. The detailed description of each standard is out of the scope of this work. Our aim is to identify some patterns in terms of management processes, QoS mechanisms and metrics, which are associated to these QoS architectures.

The considered standards are the Universal Plug and Play (UPnP) QoS [108], The Digital Living Network Alliance (DLNA) [14], the Home Gateway Initiative (HGI) QoS [30], the TR-094/143 [104], the ITU-T G.hn/ta [44], [45], the IEEE P1905.1 [39] and the IETF Homenet IPv6 [1].

Now, let us compare the logical Planes of Management, Control and Data of the studied architectures of the home network, see Table 2.2.

On the Management-Plane, the HGI-QoS, the TR-094/143, and the ITU-T G.hn/ta consider some monitoring processes. These standards are more related with the telecom operator perspective than the rest of studied standards. For instance, the TR-143

recommendation is a guide to measure throughput and other metrics that can help to diagnose the network performance from side of the network operator. The HGI-QoS, and the ITU-T G.hn/ta recommend local monitoring. The involved measurements (such as throughput or one-way-delay) can be done by active or passive probes. TR-143 and ITU-T G.hn/ta standards propose, for instance, using active probing.

On the Control-Plane, the UPnP-QoS, the HGI and the ITU-T G.hn/ta standards recommend the utilization of admission control and resource allocation, considering different metrics (e.g., delay, delay variation, packet-losses, and bandwidth). The three standards recommend the utilization of a single subnet, on tree-based-topologies. The IEEE P1905.1 [draft] [39] proposes the utilization of best-interface-selection and load balancing. The ITU-T G.hn/ta suggests the utilization of L2-optimal-routing and support mesh topologies. The IETF Homenet IPv6 [draft] recommends the reuse of existing protocols and the dual-stack IPv4/IPv6 operation. The IETF Homenet IPv6 [draft] is the unique standard that suggests using arbitrary topologies and multiple subnets. For example, it recommends the use of routing such as OSPFv3 based on delay, delay variation, packet-losses, or bandwidth. Finally, UPnP-QoS, HGI and TR-094/143 do not suggest the use of routing, nor L2-interface-selection, nor L2-path-selection.

On the Data Plane, the UPnP-QoS, the TR-094/143 and the ITU-T G.hn/ta standards suggest the utilization of IEEE 802.1D (Annex G)-based prioritization. On the other hand, the IEEE P1905.1 [draft] considers 802.1Q-based prioritization. The HGI-QoS does not define a specific prioritization protocol; however, it describes traffic classification and packet scheduling. Even though the DLNA standard has not a complete QoS-architecture, the “Home Networked Device Interoperability Guidelines” [14] considers 802.1Q-based prioritization. In [116], the authors mention the use of Differentiated Service Code Point (DSCP) for wireless and Ethernet, for DLNA devices. Finally, the IETF Homenet IPv6 [draft] recommends the use of Diffserv-based prioritization.

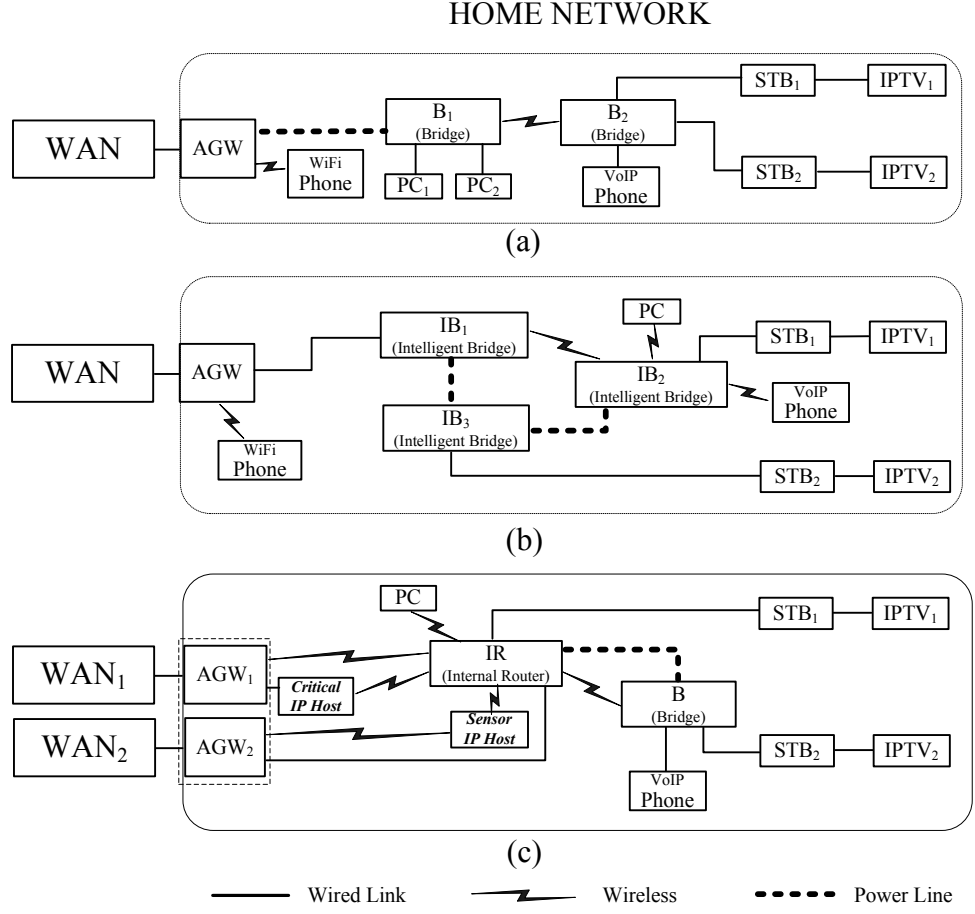
Table 2.2 summarizes this standard comparison.

TABLE 2.2. QoS ARCHITECTURES FOR HOME NETWORKS

QoS Mechanisms and Management Processes	UPnP-QoS	DLNA-based	HGI-QoS	TR-094/143	ITU-T G.hn/ta	IEEE P1905.1	IETF Homenet IPv6
<b>Management Plane</b>							
Access router diagnostics				X			
Access and home network segments measurement. (e.g. using UDP Echo Plus, probes some metrics Round Trip Time, Throughput)				X			
Channel Estimation Protocol (end-to-end active probes)					X		
Links Performance Monitoring (Including WiFi and PLC links)			X				
<b>Control Plane</b>							
Admission Control	X		X		X		
Resource allocation (with delay, jitter, packet-losses, bandwidth as flow and link-state metrics)	X						
Bandwidth Allocation			X		X		
L2 interface selection & load balancing (using packet errors or/and MAC-throughput as link-state metrics)						X	
L2 optimal routing (shortest path tree using hop counts)					X		
L3 routing (using delay, jitter, packet-losses, bandwidth as link-state metrics, e.g. OSPFv3)							X
Multiple subnets							X
Support of tree topologies	X		X	X			
Support of mesh topologies					X		
Support of arbitrary topologies							X
<b>Data Plane</b>							
Classification/Marking/Queuing/Dropping (e.g. WRR)			X				
IEEE 802.1D (Annex G)-based prioritization	X			X	X		
IEEE 802.1Q-based prioritization		X				X	
Diffserv prioritization							X

It is worth of mentioning, that all those architectures consider a number of heterogeneous-links such as: Ethernet, WiFi, PLC, and Coaxial cabling. On the other hand, G.hn considers combinations of Telephone Wiring, PLC and coaxial. The IETF Homenet IPv6 [draft] consider sensor networks.

Fig. 2.3 shows three kinds of home networks defined over different topologies: the tree topology, the mesh topology and the arbitrary topology.



**Fig. 2.3. Physical Topologies: (a) tree, (b) mesh and (c) arbitrary topology**

In Fig. 2.3 (a), we have the simplest of the three scenarios. It depicts the WAN link, which is connected to the access gateway (AGW) and to the two bridges ( $B_1$  and  $B_2$ ). The AGW and the  $B_1$  are connected with a PLC link, while  $B_1$  and  $B_2$  are connected with a WiFi link. A WiFi phone is also directly connected to the AGW.  $B_1$  has connected two PCs ( $PC_1$  and  $PC_2$ ) using wired links. Additionally,  $STB_1$ -IPTV $_1$  and  $STB_2$ -IPTV $_2$  are connected with wired links to  $B_2$ . As we can see, the topology is a simple tree topology without redundant links.

In scenario 2.3 (a), different management procedures for fault detection or service evaluation can be required, such as end-to-end performance monitoring, supported by the HGI-QoS, TR-143 or ITU-T G.hn/ta standards. Some home network performance indicators that can be included are the delay, delay variation, throughput, capacity or available bandwidth. These performance indicators can be measured from the WAN (TR-143) to the AGW, or to the STBs or to the PCs. The metrics can also be measured locally (HGI-QoS and ITU-T G.hn/ta), between the AGW and the home network devices.



In scenario 2.3 (a), different mechanisms of the Control Plane can also be implemented. These mechanisms include admission control, bandwidth allocation and resource allocation (UPnP-QoS, HGI-QoS and ITU-T G.hn/ta). The Control Plane mechanisms can be implemented for instance, in a centralized manner. The AGW can be a control point of some devices, including the STB and storage devices. Since the tree topology has not redundant links, the interface/path selection and routing are not necessary.

In scenario 2.3 (a), any of the schemas of L2/L3-prioritization (of Table 2.) can be applied. Different metrics such as delay-packet-variation, or bit rate per traffic class can be used.

The scenario 2.3 (b) introduces bridges with enriched functionalities for the Control Plane. These functionalities include interface selection, routing and path selection on the layer 2. The bridges can load balance or chose the best path from redundant paths. For this reason, we call them, intelligent bridges ( $IB_1$ ,  $IB_2$  and  $IB_3$ ).  $IB_1$ ,  $IB_2$  and  $IB_3$  are connected in a mesh physical topology (each node is connected with every other node). These kind of intelligent bridges are only considered in the G.hn/ta standard and the IEEE P1905.1 [draft].

The particular characteristic of scenario 2.3 (b) is that the redundant links of  $IB_1$ ,  $IB_2$  and  $IB_3$  allow increasing the home network reliability for forwarding packets. If  $IB_2$  and  $IB_3$  are distributed in different rooms of the house and some kind of interference affects one or two of their links, then the intelligent bridge should allocate the flows in function of the number of hops or in function of the performance and utilization of their available paths. In this sense, the G.hn/ta can be implemented to manage optimal routing, with the shortest path on the tree. Although the IEEE P1905.1 [draft] does not consider mesh topologies, this standard can be used for interface selection by using the state of the MAC-throughput or packet losses on the network interfaces. For example the IEEE P1905.1 [draft] can be used, if we have direct redundant links between  $IB_3$  and  $IB_2$ . To do that, we can eliminate  $IB_1$  and connect  $IB_3$  with both AWG and  $IB_2$ , see Fig. 1.1.

On the Data Plane, in scenario 2.3 (b), it is possible to use prioritization. For instance, the IUT-T G. hn/ta considers prioritization based on 802.1D (Annex G) traffic classes, while the IEEE P1905.1 considers 802.1Q, using the PCP field. Different link/interface-state metrics such as delay packet variation, or the bit-rate of links and interfaces per traffic class can be applied.

The scenario 2.3 (c) depicts the more complex of the three scenarios. It has the largest number of types of devices and services (including, for instance, critical and premium services). This scenario is representative of an arbitrary topology with WAN multi-homing, which is only considered by the IETF Homenet IPv6 [draft]. As we can see, two WAN

connections (multi-homing from two network service providers) are present. Then, we have two access routers ( $AGW_1$  and  $AGW_2$ ) that can be embedded into the same box, for instance. These redundant access links can be associated to critical services such as remote security or health sensors that require high access-availability. In Fig. 2.3 (c), we have a critical IP-Host and (e.g., a video surveillance device) a sensor IP-Host (e.g., a health monitoring device) that have links connected directly to the  $AGW_1$  and  $AGW_2$  and wireless links that go to the central router. The internal router (IR) can run, for instance, OSPF v3 with available bandwidth as link-state metric, and be compatible with the dual IPv4/v6 stack. IR has multiple interfaces and can have different subnets. For example, one subnet can be used for the IPTV and VoIP services, another one for the critical and sensor terminals, and a last one for web-based services. Best path selection and load balancing are not recommended by the IETF Homenet IPv6 [draft]. Therefore, only one of the available links is used, while the others are on standby.

Finally, on the Data Plane, in the scenario 2.3 (c) it is possible to define a priority schema based on Diffserv, with L3/L2 traffic classes, as recommended in IETF Homenet IPv6 [draft]. In scenarios 2.3 (a) and 2.3 (b), different link/interface-state metrics such as delay packet variation or % of link/interface utilization (bit rate) per traffic class can be applied. These metrics should be measured and used by the routing and signaling protocols, used by the IPv6/IPv4 routers.

## 2.3. Available Bandwidth Monitoring in Home Networks

Table 2.1 shows a number of performance metrics that can be used for QoS and QoE in the home network. However as argued in chapter one and in the next sections, we focus on available bandwidth. The available bandwidth of layer-N of a link (here, we use link in the sense of the direct connection between two nodes and not only as a layer-2 link) relates to the unused or “residual” capacity of this link on a specific layer-N, during a specific period of time. So, before detailing the concept of layer-N-available-bandwidth, let us explain the layer-N capacity.

### 2.3.1 Layer-N Capacity in Home Networks

The layer-N-capacity  $C_N$  (above the physical layer of a link) can be expressed as a fraction of the nominal physical capacity  $C_n$  and is proportional to the protocol data unit (PDU) size (at this layer)  $PDU_N$  and inversely proportional to the layer-2  $Frame\_Size_{(L2)}$ , as follows:

$$C_N = C_n * \frac{PDU_N}{Frame\_Size_{(L2)}} \quad (2)$$

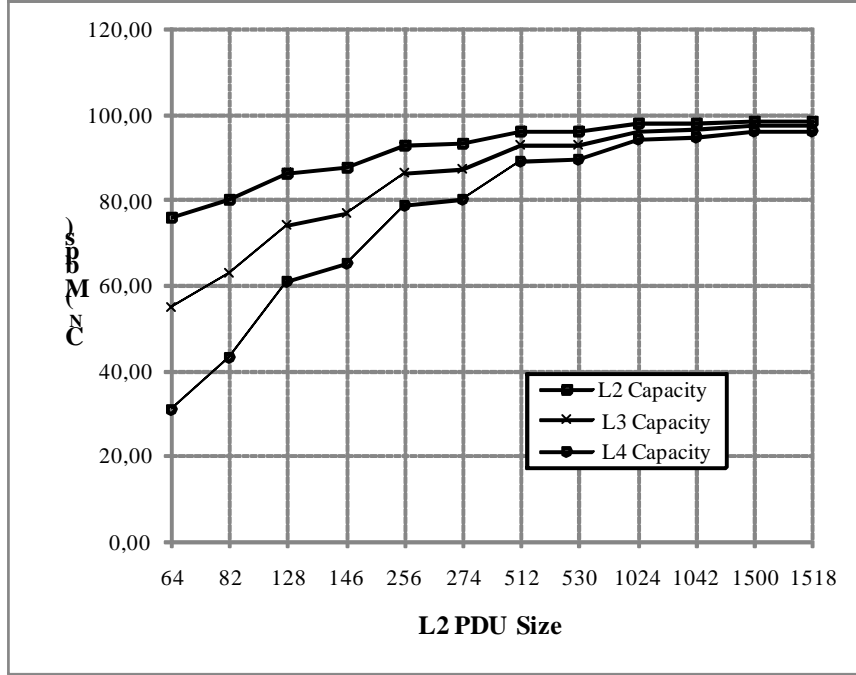
For instance, for a 1500B IP packet ( $PDU_{L2} = 1518B$ ), a Fast Ethernet link typically supports frames with a  $Frame\_Size_{(L2)} = 1538B$ , (8 bytes of frame preamble + 12 bytes of inter-frame gap + 14 bytes of MAC header + 4 bytes of CRC trailer). Then the IP capacity on a Fast Ethernet link can be calculated as:

$$- 100 \text{ Mbps} * (1500B / 1538B) = 97.5 \text{ Mbps}$$

And, the IP capacity for 46B IP packets ( $PDU_{L2} = 64B$ ) with  $Frame\_Size_{(L2)} = 84B$ , (8 bytes of frame preamble + 12 bytes of inter-frame gap + 14 bytes of MAC header + 4 bytes of CRC trailer), can be calculated as:

$$- 100 \text{ Mbps} * (46B / 84B) = 54.76 \text{ Mbps}$$

See in Fig 2.4, the layer-N capacity  $C_N$  for layer-2 to layer-4 in function of different PDU sizes. As we see on Fig. 2.4, the capacity of a link clearly varies with respect to the packet size of different applications and the header on each communication layer (e.g., a typical VoIP packet size can be about 200B, and a typical IPTV packet size can be about 1300B). See, for instance, the reduction of the capacity at layer-4, considering 64 Byte packets. Additionally every layer can exhibit different performance on an end-to-end path, because a path can be constituted by intermediate nodes and terminals, each with different layer processing and different transmission and networking mechanisms.



**Fig. 2.4. Layer-N capacity on a Fast Ethernet link in function of L2 PDU**

Additionally, the nominal physical capacity is a time-varying metric on wireless and PLC channels. This variability is associated, on one hand, to the electromagnetic interference and fading caused by different sources of noise, and on the other hand, to the corresponding transmission-mechanisms (e.g., using different PHY encoding, adaptation rate) to alleviate these transmission perturbations. Those sources, on PLC links, include narrow-band RF sources, commuted power supplies, house appliances and physical conditions of specific electrical installations [83]. On wireless links, the interference can be caused by neighbor devices, which transmit on the same range of frequencies and the attenuation due to obstacles in the house.

The layer-N capacity of home-network links is impacted by the external interference, the packet sizes of the transported flows and especially by the overhead of the PHY layer. However at the end, these perturbations will impact the transport and application layers (upper-layers). So, we assume that the capacity measured on upper layers has enough accuracy to be used on the Control Plane and on the Management Plane. Therefore, the analysis of the overhead, introduced by different PHY technologies in different transmission conditions, and the exact computation of PHY capacity, are not covered on this dissertation.

### 2.3.2 Layer-N Available Bandwidth in Home Networks

The reason we focus on available bandwidth is because this metric is a direct measure of the additional load that a link or a path can carry before it becomes saturated. Other metrics, such as delay, delay variation, packet loss rate and capacity can only determine whether a path is already congested [56]. For instance, losses can be random and/or happen after a link has been congested. The delay may be governed by propagation latencies. The links capacity, by definition, does not take traffic load into account [92]. The available bandwidth, on the other hand, allows determining whether a link or path has sufficient capacity to carry a flow before triggering a reactive QoS-mechanism or network-management process.

Let us define the available bandwidth of a link  $i$ , at the layer  $N$ , for a specific average packet-size and over a certain time interval. If  $C_{Ni}$  is the capacity of link  $i$  and  $u_{Ni}$  is the average utilization of that link, in the given time interval, then the average available bandwidth  $A_{Ni}$  of this link is given by the unutilized fraction of the capacity [92], as follows:

$$A_{Ni} = (1 - u_{Ni})C_{Ni} \quad (2.3)$$

Extending the previous definition to a K-link path, the end-to-end available bandwidth on this path is the minimum available bandwidth on the path. Where, the link with the minimum available bandwidth is often called the “tight link” [16].

For instance, in Fig 1.1, we have a path, between the access gateway AGW and the IPTV<sub>2</sub>, composed by three serialized links. The first Fast Ethernet link connects the access gateway and the first PLC-WiFi<sub>1</sub> device (AGW  $\rightarrow$  PLC-WiFi<sub>1</sub>). The second link connects the two PLC-WiFi devices (PLC-WiFi<sub>1</sub>  $\rightarrow$  PLC-WiFi<sub>2</sub>) and the third link connects the PLC-WiFi<sub>2</sub> to the STB (PLC-WiFi<sub>2</sub>  $\rightarrow$  STB).

Considering this path (AGW  $\rightarrow$  PLC-WiFi<sub>1</sub>  $\rightarrow$  PLC-WiFi<sub>2</sub>  $\rightarrow$  STB) and assuming 1500B IP packets, then the layer-3 nominal capacities for the PLC and WiFi (e.g., 802.11a) links are typically about 80 Mbps and 27 Mbps, respectively. These layer-3 capacities are inferior to the 97.5 Mbps of the Fast-Ethernet links, at the edges. If 50 Mbps of traffic traverses this path, only in the sense of the AGW  $\rightarrow$  PLC<sub>1</sub>  $\rightarrow$  PLC<sub>2</sub>, then the minimal available bandwidth will be about 30 Mbps and will be clearly located on the PLC<sub>1</sub>  $\rightarrow$  PLC<sub>2</sub> link, which becomes the “tight link” of the path.

As we have mentioned, in scenarios like this or in the ones of Fig. 2.3, the PLC and WiFi links are prone to be degraded due to the electromagnetic interference. In that case, the

PLC and WiFi links are not only the ones with the lower nominal capacity, but the more vulnerable ones, because of their time-varying capacity nature. Consequently, they can exhibit the strongest variability of available bandwidth on the paths of the home network.

For instance if the capacity on the  $PLC_1 \rightarrow PLC_2$  link reduces to 50 Mbps, the available bandwidth, in the previous case, will be 0 Mbps and no other flow will be able to traverse the path, without impacting the performance of the involved user-applications.

Hence, if we assume that the regimes of latency and packet delay variation on the home network are relatively low; thanks to the short propagation-distances and the relative limited number of intermediary nodes on the home network paths, then the transport-layer available bandwidth can be very representative of those paths. Available bandwidth can be used for routing, path selection, load balancing, bandwidth allocation, fault detection and performance monitoring. Also available bandwidth can be correlated with the bit-rate, which is required to guarantee a KPI (for instance, as the shown in equation 1).

### 2.3.3 Available Bandwidth Monitoring in the Home Network

In the ITU-T X.641 recommendation [47], the QoS monitoring is defined as: “the use of QoS measures to estimate the values of a set of QoS characteristics actually achieved for some systems activity.” Also QoS monitoring can be seen as a mechanism for the collection and analysis of information, regarding the relationships and internal states [15] of the home network and its interactions with their applications. To illustrate the home network monitoring, we have depicted a control loop on the home network, in Fig. 2.5, which is based on [102]. This control loop includes three main tasks a) the Measurement of QoS Metrics, b) the evaluation (commitments and violations) of QoS-metrics against performance targets and c) informing and warning about the state of network resources.

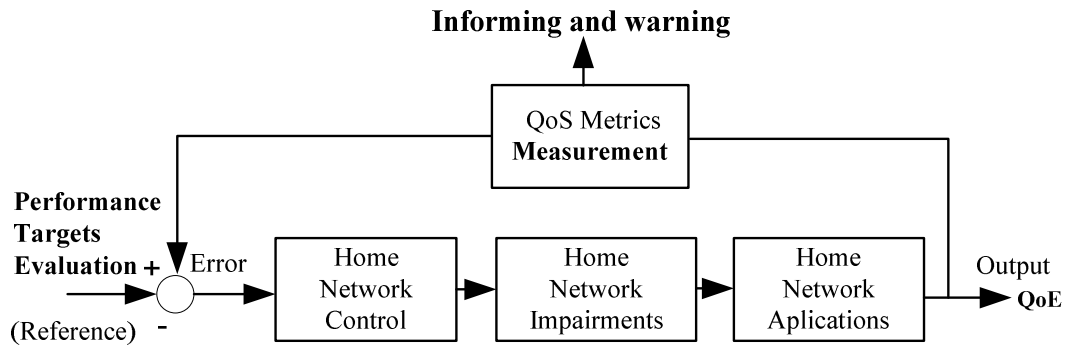


Fig. 2.5. Home Network Control Loop and Performance Monitoring

As we show in Fig. 2.5, the QoS measurement is a key element for performance monitoring. Basically, the techniques to measure QoS can be classified into two families: passive and active [49]. Passive techniques give metrics from actual packet statistics, which are read directly from the network interfaces (e.g., monitoring of queues depth) [30]. On the other hand, active techniques send dummy packets, which emulate the application flows.

A number of tools to measure available bandwidth are publicly available. For instance, a description of a list of active tools, which are developed in open source, can be found in [78].

Although, in the literature, we find numerous active tools to measure available bandwidth, the selection of a convenient tool for hybrid home-networks is challenging. This is because the constraints of home networks that we have cited in chapter 1 and which, we explain in chapter 3. In fact, in chapter 3, we study the principles of available bandwidth probing and a number of state-of-the-art techniques and tools. For the moment, we assume that available-bandwidth probing can be used on the Control Plane and on the Management Plane of the home network. In next section, we illustrate some use cases of probing the available bandwidth in home networks.

## 2.4. Use Cases

Figures 2.3 (a, b and c) show three scenarios to monitor available bandwidth in home networks. In these scenarios we consider two typical IPTV-flows ( $F_{TV_1}$  and  $F_{TV_2}$ ) of 12 Mbps each. Then,  $F_{TV_1}$  and  $F_{TV_2}$  traverse the access gateway and the intermediary network-devices to arrive to the  $STB_1$  and  $STB_2$  and finally reach  $IPTV_1$  and  $IPTV_2$ . For simplicity, we assume that there is no other traffic than  $F_{TV_1}$  and  $F_{TV_2}$  on the studied networks. The paths subject of analysis in Fig. 2.3 (that are traversed by  $F_{TV_1}$  and  $F_{TV_2}$ ), are described in Table 2.3.

TABLE 2.3. EXAMPLE HOME NETWORK PATHS FOR THE HDTV SERVICE

Figure	Path 1 ( $P_1$ )	Path 2 ( $P_2$ )
5 (a)	$AGW \rightarrow B_1 \rightarrow B_2$	-
5 (b)	$AGW \rightarrow IB_1 \rightarrow IB_2$	$AGW \rightarrow IB_1 \rightarrow IB_3$
5 (c)	$AGW_1 \rightarrow IR$	$AGW_1 \rightarrow IR \rightarrow B$

As illustrated in section 2.3, the performance of the paths on the home network is prone to degrade due to their time-varying nature. For instance, the total available bandwidth in the home network can down just to the required bit rate for a single IPTV flow or even less.

Therefore, we need to monitor the paths and flows and protect them by choosing one or more network management procedures and QoS control mechanisms. To do that, we can use (for instance) the QoE-based recommendations that we found in the standard TR-126, [105].

So, we proceed as follows. Firstly, we define a group of QoS metrics that characterize an acceptable transfer of  $F'_{TV_1}$  and  $F'_{TV_2}$ , through the end-to-end path. We call this group of metrics, the “end-to-end targets of performance”. Secondly, given our group of QoS metrics, we split the end-to-end impairments (transfer delay, jitter and packet loss) from the bit-rate, which is required by each TV flow. Thirdly, we define the fraction of the end-to-end impairments, which the home network should be able to manage. Finally we define the Control Plane and Management Plane mechanisms that will be required in case of performance degradation of  $F'_{TV_1}$  and  $F'_{TV_2}$ .

Table 2.4 shows the end-to-end targets of performance for an IPTV flow, which requires a bit-rate of 12 Mbps. These values are suggested on the standard TR-126.

TABLE 2.4. TRANSPORT LAYER QOE BASED END-TO-END TARGETS FOR A FLOW OF 12MBPS-HDTV-MPEG-4 AVC (FROM TR-126)

bit rate (Mbps)	latency (ms)	jitter (ms)	loss rate
12	< 200	< 50	1.22E-6 (5 pkts/h)

Now, we need to split the home network impairments from the WAN impairments of a hypothetical path such the one shown in Fig. 1.2. Clearly, the impairments of a particular path can change depending on the geographical separation between the service provider and the home network. Then, we consider the empirical examples of paths, which are found in the appendix I of ITU-T Rec.Y.1542 [52], as follows.

To show a hypothetical example, we use the approach of [52]. We take the targets of the home network (HN) as about 1% of the end-to-end impairments:

- $latency_{HN\_target} = 2 \text{ ms}$
- $loss\_rate_{HN\_target} = 0.01 * loss\_rate_{end-to-end\_target}$
- $jitter_{HN\_target} = 0.01 * jitter_{end-to-end\_target}$



Now as discussed in section 2.3, we aim to use available bandwidth as a link/path metric and as a performance indicator for service evaluation. So, we make the following assumptions:

- a) The WAN measurements can be considered independent of the local measurements in the home network. The WAN segment can include the following sub-segments: access segment, metropolitan segment and regional segment [52]. Therefore, the WAN impairments depend on the technologies, on the transmission quality and on extension of the WAN sub segments.
- b) The path capacity of the home network is time-varying (due to wireless and PLC links)
- c) The diameter of the home network (longest path) contains a relatively low number of hops (less than 5)
- d) The IPTV traffic is prioritized in the WAN, using a suitable class of service (e.g., Class 6 or 7, recommended in Table 3 of ITU-T Y.1541 [53], which can be associated to the DSCP value of  $0x28 = 40_{\text{dec}}$  [3]) and which can be mapped to a class in the home network (e.g., 802.1D with a priority value of  $100_{\text{bin}}$  [104])
- e) There are no systematic total interruptions of short duration in the available paths. If a series of total interruptions appear at regular intervals, due to hardware damage of any network element, misconnections, or external perturbations then, the home network is not considered as in nominal operation. Then, the user should call the technical support, for fault detection.

Consequently, to the previous assumptions, we have:

- f) The latency/jitter in the home network can be considered globally low with respect to the WAN latency/jitter. The average one-way latency of WiFi and HomePlug AV transmissions is of the order of several ms with a reasonable limited number of hosts [83], [71].

Considering previous assumptions, the QoS/QoE targets on the home network paths can be completely characterized by the bit rate of a HD/IPTV flow. Additionally, we define a bit rate envelope, as a simple preventive measure. In this way, we protect the IPTV flows of any transient congestion with an additional 10% of bit rate target, as shown in Table 2.5.

TABLE 2.5. BIT RATE ENVELOPE FOR THE HOME NETWORK EXAMPLE

Parameter	Target
Bit rate (Mbps)	$\geq 1.1 \cdot 12 = 13.2$

We have defined the bit rate of an IPTV Flow as a part of the considered set of QoS-metrics. Hence, the available bandwidth on the considered paths must be at least equal to the  $bit\_rate\_target = 13.2$  Mbps. It allows preventing the degradation of the FTV<sub>1</sub> and FTV<sub>2</sub>. Then, available bandwidth  $\mathcal{A}$  becomes the link/path-state metric and the performance indicator to monitor the home network paths.

Now, we need to define the measurement mechanisms we wish to implement in the home network. In this context, the estimated-available-bandwidth  $\mathcal{A}_e$  can be evaluated against the available-bandwidth-target  $\mathcal{A}_r$ . So, we need to define the possible states of the monitored paths (described in Table 2.3) with respect to FTV<sub>1</sub> and FTV<sub>2</sub>. Finally, we will define the control actions and support procedures, which are required when  $\mathcal{A}_e$  exceeds  $\mathcal{A}_r$  on the home network paths.

So, it is possible implementing the probing of available bandwidth at the transport-layer (as mentioned in section 2.3). It is extremely useful when passive methods (by physical interface) do not provide consistent measurements, due to the time-varying capacities of wireless or PLC links.

For example, we can probe  $\mathcal{A}_e$  in manual or automatic ways. In fact, it is possible to send periodic probing-flows, through the paths of the home network. We can situate a probing server of available bandwidth, embedded on intelligent nodes, or at home management-devices, at the access gateway or on the service-provider side. The probing clients can be placed on the PC terminals, on the STBs, or on intelligent network peers.

The next examples show the performance states of the paths of the home network and the control actions that can be triggered, when the estimated-available-bandwidth  $\mathcal{A}_e$  does not meet the available-bandwidth-target  $\mathcal{A}_r$ .

In Fig. 2.3 (a), we can probe available bandwidth in a continuous and non-intrusive way. We monitor the available bandwidth  $\mathcal{A}_{e-p1}$  of the single path  $P_1$  (AGW $\rightarrow$ B<sub>1</sub> $\rightarrow$ B<sub>2</sub>). Some possible conditions in this scenario could be:

- ( $\mathcal{A}_{e-p1} \geq \mathcal{A}_r$ ), the estimated-available-bandwidth in  $P_1$  is superior or equal to the available-bandwidth-target and consequently  $\mathcal{A}_{e-p1}$  respects the target or
- ( $\mathcal{A}_{e-p1} < \mathcal{A}_r$ ) and FTV<sub>1</sub> and FTV<sub>2</sub> are transmitted simultaneously and the estimated-available-bandwidth in  $P_1$  is inferior to the available-bandwidth-target and consequently  $\mathcal{A}_{e-p1}$  violates the target.

These conditions can trigger the following QoS mechanisms and management processes:

- Use Admission Control, if  $(A_{e-p1} \geq A_t)$ , then accept a new flow; otherwise refuses the connection and send a warning to the end-user. In this case,  $A_{e-p1}$  can be seen as link/path-state metric, associated to the reactive mechanisms of QoS.
- Use Remote Fault Detection, if  $(A_{e-p1} < A_t)$  and FTV<sub>1</sub> and FTV<sub>2</sub> are not being transmitted, then look for root cause on home network paths; otherwise look for root cause on terminals or on the WAN segment. In this case  $A_{e-p1}$  can be seen as a simple performance indicator, associated to procedures of trouble shooting.
- Warn the management systems of the service provider if  $(A_{e-p1} < A_t)$  and FTV<sub>1</sub> and FTV<sub>2</sub> are not being transmitted and inform about the service restitution-time, after  $(A_{e-p1} \geq A_t)$  and FTV<sub>1</sub> and FTV<sub>2</sub> are recovered. In this case  $A_{e-p1}$  can be seen as a simple performance indicator, associated to contractual guaranties.

Now we want to monitor available bandwidth  $A_{e-p1}$  and  $A_{e-p2}$  on the two paths of Fig. 2.3 (b) and Fig. 2.3 (c):

Fig. 2.3 (b):  $P_1$  (AGW→IB<sub>1</sub>→IB<sub>2</sub>) and  $P_2$  (AGW→IB<sub>1</sub>→IB<sub>3</sub>)

Fig. 2.3 (c):  $P_1$  (AGW<sub>1</sub>→IR) and  $P_2$  (AGW<sub>1</sub>→IR→B)

Then, we can apply continuous available bandwidth probing on each home network path. Some possible conditions in these scenarios could be:

- $(A_{e-p1} \geq A_t \ \& \ A_{e-p2} \geq A_t)$  the estimated-available-bandwidth in  $P_1$  and  $P_2$  are superior to the available-bandwidth-target and consequently  $A_{e-p1}$  and  $A_{e-p2}$  respect the target or
- $(A_{e-p1} < A_t)$  and/or  $(A_{e-p2} < A_t)$  and FTV<sub>1</sub> and FTV<sub>2</sub> are transmitted simultaneously and the estimated-available-bandwidth in  $P_1$  and/or  $P_2$  is inferior to the available-bandwidth-target and consequently  $A_{e-p1}$  and/or  $A_{e-p2}$  violate(s) the target.

These conditions can trigger the following QoS mechanisms and management processes:

- Use Path/Route Selection & Load Balancing for one or the two flows, if  $(A_{e-p1} > A_{e-p2} \geq A_t)$  and then select  $P_1$  or if  $(A_{e-p2} > A_{e-p1} \geq A_t)$ , select  $P_2$ . In this case,  $A_{e-p1}$  and  $A_{e-p2}$  can be seen as link/path-state metrics, associated to reactive mechanisms of QoS (e.g. path selection, admission control, load balancing).
- Use Remote Fault Detection, if  $(A_{e-p1} < A_t) \ \& \ (A_{e-p2} < A_t)$  and FTV<sub>1</sub> and FTV<sub>2</sub> are not being transmitted, look for root cause on home network paths; otherwise look for root

cause on terminals or on the WAN segment. In this case,  $A_{e-p1}$  and  $A_{e-p2}$  can be seen as performance indicators, for trouble shooting procedures.

- Warn, about the service restitution time, to the management systems of the service provider and the care systems of the end-user if  $(A_{e-p1} < A_l)$  and/or  $(A_{e-p2} < A_l)$  (and FTV<sub>1</sub> and FTV<sub>2</sub> are not being transmitted) after we have that:  $(A_{e-p1} \geq A_l \ \& \ A_{e-p2} \geq A_l)$  and FTV<sub>1</sub> and FTV<sub>2</sub> are recovered. In this case,  $A_{e-p1}$  and  $A_{e-p2}$  can be seen as simple performance indicators, associated to contractual guaranties.

These actions will be conditioned by the specific implemented algorithms. For the sake of brevity, we only give some examples here. Our aim is to highlight the importance of available bandwidth monitoring for network management processes and QoS mechanisms (Management Plane and Control Plane) in home networks.

## 2.5. Summary

In this chapter, we studied the interdependencies between network management processes, QoS mechanisms and performance metrics on the WAN and on the home network. We have explored the available bandwidth monitoring and specifically, the probing of available bandwidth in home networks. Finally, we showed some use cases, where we underline available bandwidth on the Control Plane and Management Plane of the home network.



# State of the art of Available Bandwidth Probing: Network Constraints, Techniques and Tools

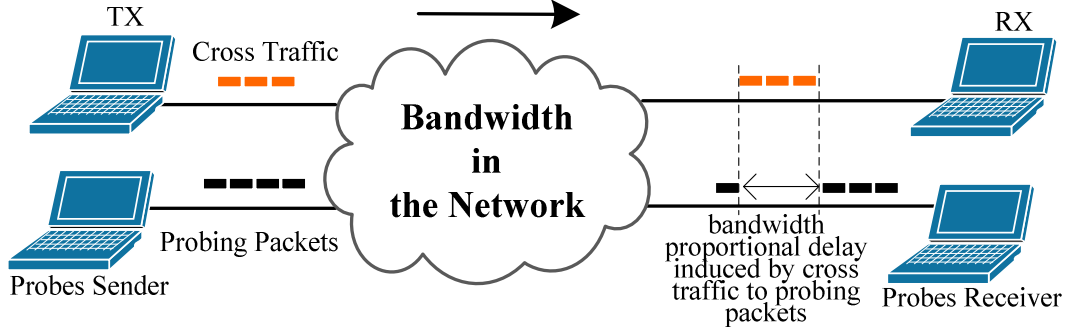
In this chapter we describe a brief state-of-the-art of bandwidth probing. More specifically, after some useful definitions, we highlight network constraints that need to be taken into account. We also present the principles of bandwidth probing. Finally, we introduce some existing tools to estimate available bandwidth.

It is known that bandwidth measurement can be performed mainly through two approaches: active or passive. Active probing is extremely important in presence of the limitations of deterministic measurements (often called passive measurements). Some of these limitations include: the prohibitive volume of captured data on high-capacity links [78], heterogeneous-administrative-domains that do not provide local information and the capacities that vary in the time, such as the ones that are found in home networks [87].

Active probing can be done using UDP or TCP transport layer packets to capture the nearest behavior of the network that impacts the application layer. Active probing can be used to estimate different metrics, including delay, delay variation, packet losses, throughput, capacity and available bandwidth. As argued in chapter 2, we are interested in bandwidth probing.

Bandwidth probing is an in/out process that infers capacity or available bandwidth from the packet dispersion or bit rate of dummy packets. The dummy packets interact with the nodes and the user traffic (called cross traffic), on the measured path, see Fig 3.1.

Bandwidth probing can be deployed only in the ends of the paths (end-centric basis) or on the intermediary nodes (network-centric). We can also test the bandwidth of paths using probing flows per traffic classes.



**Fig. 3.1. Bandwidth Probing**

One of the reasons that make interesting the study of bandwidth probing in home networks is the fact that most existing probing tools were thought for the Internet scale. Consequently, the adaptation of the probing tools to the scale of home networks is challenging. In fact, the majority of those tools were proposed for point to point links with constant capacity. As we have suggested in section 2.3, we are interested in choosing an available bandwidth tool that can be accurate, fast and non-intrusive, in home network paths.

Before establishing what kind of available bandwidth probing tool can be convenient for home networks, we explore the principles of bandwidth probing. These principles include, firstly, the networking constraints that impact the measurement process at the Internet scale and especially at the home network scale. Secondly, we explore not only the techniques for available bandwidth probing but also the techniques for capacity probing, because, by definition, available bandwidth is directly dependent on the capacity.

### 3.1. Networking Constraints for Bandwidth Probing

The networking constraints that make challenging the bandwidth probing in Internet paths and home network paths are related, on one hand, to the way of sampling the bandwidth on different scales of space and time (along the network) and on the other hand, to the nature of transmission links and the cross traffic that traverses the measured paths.

### *3.1.1 Deterministic Measurement vs. Probing Measurement*

In general, deterministic measurement for capacity, delay, throughput or available bandwidth is only possible in single administrative systems (ASs). As mentioned in section two, the deterministic measurement consolidates the counters of physical or virtual interfaces on the network (e.g., queues depth monitoring [30]). At the same time, the deterministic measurement supposes the knowledge of the state of all the interfaces of the nodes on the network. Therefore, the deterministic measurement is only accurate on paths with links of constant capacity. For instance, deterministic available bandwidth can be computed as the subtraction between the capacity and the link/path utilization, where capacity is considered constant. However, if the capacity varies, the available bandwidth measurement will be biased.

Conversely, probing measurement (synonym of inferential measurement) obtains performance metrics by the correlation of timing descriptors of probing packets. Probing measurement requires enough samples to be accurate and it maybe tends to be intrusive and less granular than deterministic measurement. Then, network probing is accurate and extremely useful, when the traversed links have time-varying capacities or do not provide any state information. For the rest of the document, we use the term capacity or available bandwidth probing to refer to the active probing using dummy profiled packets.

### *3.1.2 Time Scales and Traffic Classes*

As we have mentioned, active probing is based on finding dispersion or rate patterns, when the probing packets have traversed the measured links or paths. The dispersion and rate patterns are the result of the interaction between probing packets, cross-traffic and all the interface capacities along the path. Thus, timing on probing-flows interactions has different statistical characteristics if we sample it in the scale of microseconds or in the scale of minutes or days. So, available bandwidth can be used as a state metric if a short time-scope is chosen. Otherwise, available bandwidth probing can be used as a performance indicator, if a macroscopic time-scope is defined. In section 2.4, we have discussed some examples that illustrate the differences between time-scopes in home networks.

The home network can allocate different traffic classes. So, the home network must guarantee specific values of available bandwidth, according to these traffic classes. For instance, FTP is elastic in terms of bandwidth requirements. But, multimedia and real time



applications can require fixed amounts of bandwidth. For this reason, measuring the available bandwidth by traffic classes can be convenient.

### *3.1.3 End-Centric Probing vs. Network-Centric Probing*

The end-centric probing has not access to the state of the intermediary nodes in the path, but instead, it sees the network as a black-box. It relies on end points to perform the measurement. On the other hand, network-centric probing integrates the probing patterns of intermediary nodes. The network-centric probing supposes a “node-state” monitoring and a unique administration authority.

Network-centric or end-centric probing should be chosen, depending on the measurement requirements by use case. For instance, if we want to measure available bandwidth for end-to-end evaluation or for remote-fault-detection, end-centric measurement can be more suitable, because we want to test both, the WAN and home network. Since, traversing an Internet path supposes the absence of explicit state of intermediary links, most probing tools have been thought for end-centric measurement. However, if we have access to the intermediary nodes of the paths (as is often the case of home networks) the network-centric approach can be better.

### *3.1.4 Networking Constrains for Bandwidth Probing on Internet Paths*

In summary, some important networking constraints that are considered for bandwidth probing on Internet paths are:

- a) Bandwidth probing should consider diverse time-scopes
- b) Bandwidth probing in end-centric basis is preferred, when different ASs are traversed
- c) In general, bandwidth probing does not consider traffic classes. The Store-and-Forward discipline, in routers, is the origin of packet queuing effects
- d) Probing flows can be seen as possible threats

### *3.1.5 Networking Constrains for Bandwidth Probing on Home Network Paths*

We consider that the home network (itself) can be managed as a single administrative system (AS). Then, some important networking constraints are:

- a) Bandwidth probing should consider diverse time-scopes

- b) Bandwidth probing should consider strong delay variability in probing flows, which can be induced by multiple-access mechanisms and time-variant link capacities
- c) Bandwidth probing calibration should be simple and transparent to lower layers
- d) Bandwidth probing should have enough accuracy, low convergence time, low overhead and stability in different system platforms and network conditions
- e) Bandwidth probing should be done in end-centric or network-centric basis depending on the use case, accuracy, and simplicity of measurement implementations
- f) Bandwidth probing should consider traffic classes.

### 3.2. Principles of Bandwidth Probing Techniques

In literature, we find a number of seminal techniques to probe capacity and available bandwidth. At the same time, we find publically available tools (software) applying these techniques. In this context, we can associate techniques and tools. To define a homogeneous taxonomic comparison, we have assembled most popular techniques into four categories [56], [92], as follows:

For Capacity Probing, we have:

- The Variable Packet Size (VPS) technique for capacity probing (by-hop)
- The Variable Packet-Pair Dispersion (VPD) & Variable Train Dispersion (VTD) technique for capacity-probing (end-to-end)

For Available Bandwidth Probing, we have:

- The Direct-Probing or Probe Gap Model (PGM) for end-to-end probing of available bandwidth
- The Iterative-Probing or Probe Rate Model (PRM) for end-to-end probing of available bandwidth

In general, the four bandwidth probing techniques can be described as an in/out-process with input and output variables. The input variables characterize probing flows before they traverse the monitored links or paths, while the output variables characterize probing flows after they arrive to their destination. For instance, the input variables referred to the probing techniques can be: packet sizes, packet pair dispersion (also called packet gaps) or

input probing rates. The output variables can be associated to Round Trip Times (RTTs), One Way Delays (OWDs), output-pair/train-gaps and output-probing-rates.

### *3.2.1 The Variable Packet Size (VPS) Technique for Capacity Probing (by-hop)*

#### **3.2.1.1 Early Implementations**

Variable Packet Size (VPS) probing for by-hop capacity is proposed by Bellovin, in the early '1990s [99]. VPS injects a set of input ICMP-ECHO requests to all the IP addresses of intermediary routers along the path. The ICMP-ECHO packets change in size. The output RTT samples are captured and their minimum is filtered by linear interpolation. Considering perfect symmetry in the transmission and reception channels, for each hop, the half of the slope (of the straight-line that fits the intersection of the packet sizes  $L$  and the RTTs) gives the capacity for each hop.

Jacobson, in [55], proposes a variation of Belovin's VPS but that does not require the knowledge of intermediary routers. This variation of VPS is the most popular version and we explain it, below.

Another alternative to measure by-hop capacity is found in the work of Lai, in [65]. Lai uses a pair of back-to-back TCP packets, the first packet is the largest as possible and the second is the smallest as possible. Lai calls this technique Tailgating. The RTTs are captured and their output dispersion is filtered by linear interpolation, similar as in [58] and [99]. In contrast with [99], TTLs (Time To Live) are not necessary.

#### **3.2.1.2 The General Model**

The probing technique Variable Packet Size [92] estimates the capacity  $C_b$ , hop-by-hop based on the minimum of sampled RTTs. Multiple probing packets of a given size are transmitted from the sending host to each hope along the path. The RTT to each hop include three delay components in the forward and reverse paths: serialization delays, propagation delays, and queueing delays. The technique assumes that at least one of the probing packets, together with the ICMP reply that it generates, will not meet practically any queueing delays. Then, the minimum of sampled RTTs is obtained.

To attain every intermediary hop, the Time-To-Live (TTL) field of the probing packets, in the IP header, is configured to  $i$  successive values (e.g. between 1 and 255). Then, the fixed TTL forces the probing packets to expire at a particular  $i$ -th hop. The router at that hop

discards the probing packets and returns an ICMP “Time-exceeded” error to the host source, which includes the IP address of the hop.

Specifically, the minimum RTT for a given packet size  $L$ , up to hop  $i$  is expected to be:

$$T_i(L) = \alpha + \sum_{k=1}^i \frac{L}{C_k} = \alpha + \beta_i L \quad (3.1)$$

where,  $C_k$  is the capacity of  $K^{th}$  hop,  $\alpha$  represents the delays up to hop  $i$  (does not depend on the probing packet size  $L$ ), and  $\beta_i$  is the slope of minimum RTT up to hop  $i$  against probing packet size  $L$ . Repeating the minimum RTT measurement for each hop  $i = 1, \dots, H$ , the capacity estimated at each hop  $i$ , along the measured path, is:

$$C_i = \frac{1}{\beta_i - \beta_{i-1}} \quad (3.2)$$

### *3.2.2 The Variable Pair Dispersion (VPD) & Variable Train Dispersion (VTD) Techniques for End-to-End Capacity Probing*

#### **3.2.2.1 Early Implementations**

The Variable Pair Dispersion (VPD) is proposed in the work of Bolot [5], Carter and Corvella [8] and Lai [65]. The Variation of Train Dispersion (VTD) technique is proposed by Paxson [89]. Both probing techniques estimate the end-to-end path capacity in function of the packet pair/train dispersion that the probing pairs/trains suffer after traversing the measured paths.

Bolot, in [5] uses NetDyn and makes the dispersion of pairs of UDP packets vary. NetDyn is a tool that allows generating UDP packets and can be used to measure end-to-end capacity. The RTT of packet  $n$  is captured and plotted against packet  $n + 1$ , while their output dispersion is filtered manually by linear interpolation. The interpolated line on the curve represents the transmission interval and the inverse of its slope, the end-to-end path capacity.

In [8], Carter and Corvella make vary the input dispersion of successive probing groups of ICMP-ECHO pairs of packets. The RTTs are captured and their output dispersion is filtered by using union and intersection of rate histograms. In this process, Carter and Corvella use heuristics to approximate the capacity mode. This technique resolves the issue

raised by Bolot in [5], about how to build a tool that automates the measurement process and it provides the key insight about how to do the filtering.

Using TCP packets, Lai, in [65], observe the input dispersion of a set of packet pairs vary. The OWDs or RTTs are captured and their output dispersion is filtered using adaptive histograms. The end-to-end capacity is derived from the strongest dispersion mode. If cross-traffic samples are used, Potential Bandwidth Filtering selects the correct size of samples. In contrast with the heuristic filtering of [89], Lai uses Adaptive Kernel Estimation, which is a deterministic technique that defines automatically the bin width (granularity) of histograms.

To probe multi-channel<sup>1</sup> links, Paxson in [89], makes the input dispersion of selected TCP trains vary. Paxson calls this technique PBM (Packet Bunches Modes). The RTTs are captured and their output dispersion is filtered heuristically to obtain the end-to-end capacity mode from histograms

### 3.2.2.2 The General Model

The Variable Pair/Train Dispersion works as follows:

If on a link of capacity  $C_0$  (that does not transport other traffic) is injected a back-to-back pair of packets or a train of  $N$  packets of size  $L$ , the dispersion of the packet pair is expressed in equation 3.3 and the dispersion of the packet train is expressed in equation 3.4.

$$\Delta_0 = \frac{L}{C_0} \quad (3.3)$$

$$\Delta_0 = \frac{(N-1)L}{C_0} \quad (3.4)$$

In general, considering a link  $i$  of capacity  $C_i$ , on a path of  $H$ -links, if the dispersion of a packet pair, prior to enter to that link, is  $\Delta_{in}$  and assuming that there is no other traffic on that link, the output dispersion, after the probing packet traverses the link, can be expressed as:

$$\Delta_{out} = \max\left(\Delta_{in}, \frac{L}{C_i}\right) \quad (3.5)$$

After a packet pair goes through the path, the packet pair dispersion that the receiver will measure is:

---

<sup>1</sup> Each channel of a multi-channel link has separated bandwidth (e.g. the BRI ISDN links) and the aggregated bandwidth of the link is the sum of the bandwidths of all channels.

$$\Delta_R = \max_{i=0,\dots,H} \left( \frac{L}{C_i} \right) = \frac{L}{\min_{i=0,\dots,H} (C_i)} = \frac{L}{C} \quad (3.6)$$

Where,  $C$  is the end-to-end capacity of the path, in other words the smallest capacity along the path. Thus, the receiver can estimate the path capacity from:

$$C = \frac{L}{\Delta_R} \quad (3.7)$$

$C$  can be also derived from train dispersion, considering  $N$  packets, as explained in [92]. The choice of pairs or trains of packets depends on different criteria. These criteria include the variability of the measurement process and the degree of interaction of the probing flows with the cross traffic. When,  $N$  increases, the measurement variability tends to decrease, however the degree of the interaction with the cross traffic also tends to increase, which can bias the capacity measurements with cross traffic. Often, capacity is measured using pairs of packets, while available bandwidth measurement (that needs to reflect the cross traffic behavior) is implemented with trains of packets.

### 3.2.3 The Direct-Probing or Probe Gap Model (PGM) for End-to-End Available Bandwidth Probing

#### 3.2.3.1 Early Implementations

The direct-probing technique is also called the Probing Gap Model (PGM) [67] or simply Gap Model [84]. Keshav, in the early s'1990 [60], defined for the first time the “packet pair principle”. He injects back-to-back TCP pairs of packets, where the rate of the pairs is considered larger than the narrow link capacity (is the link with minimum capacity on the path). The RTTs are captured and the output dispersion rate is filtered by exponential averaging. This dispersion rate is the available bandwidth in the measured paths used to admit or reject new Available Bit Rate (ABR) connections for admission control mechanisms. This technique requires that the routers uses min-max fair scheduling, that always exists data to send, and that narrow link capacity is computed by sending initial back-to-back pairs.

Cprobe is one of the earliest software tools that was developed by Carter and Corvella, in [8] and measures the end-to-end available bandwidth. The Cprobe tool injects groups of 8 ICMP back-to-back pair of packets. The RTTs are captured and the average dispersion rate is filtered. The inverse of this average is multiplied by  $(N-1)L$  to obtain the available

bandwidth. However, as showed in [16], the dispersion rate obtained in this way differs from the available bandwidth.

### 3.2.3.2 The General Model

We discuss the principles of direct-probing/PGM following the explanation of Jain Manish in [56].

In direct-probing/PGM, each probing stream saturates the path during short probing periods, which results in a sample of the available bandwidth. The sender transmits a periodic packet stream of rate  $R_i$  and the receiver measures the output rate  $R_o$ . The basic idea is that, if  $R_i$  is larger than the available bandwidth  $A$ , then  $A$  is expressed as:

$$A = C_t - R_i \left( \frac{C_t}{R_o} - 1 \right) \quad (3.8)$$

Notice that direct-probing/PGM samples the available bandwidth with each packet train, as long as the input rate is sufficiently high, for instance equal to the capacity of the link, where the sender is located. The main assumption in the direct probing approach, however, is that the tight link capacity  $C_t$  is known and that the cross-traffic is path persistent (cross traffic follows exactly the same path as probing traffic without addition or losses at intermediate nodes). Additionally, it is assumed that  $C_t$  can be estimated with end-to-end capacity probing tools. On this way, however, we estimate the capacity  $C_n$  of the narrow link (the link with lower capacity), which may be different than  $C_t$ . We recall that the tight link is the link with lower available bandwidth. For this reason, direct-probing/PGM technique can be inaccurate when the tight link does not coincide with the narrow link [67]. However, the relative simplicity of the direct-probing/PGM algorithm can simplify the development of software tools.

## 3.2.4 The Iterative-Probing or Probe Rate Model (PRM) for End-to-End Available Bandwidth Probing

### 3.2.4.1 Early Implementations

Jacobson, in the latest s'1980 [54], proposed the famous congestion control mechanism for TCP connections by varying the input window size of TCP traffic. The output packet losses and the RTT increments of TCP flows are captured and filtered by linear interpolation. Specific thresholds of packet losses and RTTs are the feedback signals that

represent the changes in the available bandwidth and consequently, the iterative adaptation of the TCP window size to the congestion conditions. This iterative technique is a seminal work that has permitted to elucidate the notions of packet dispersion, congestion avoidance and available bandwidth over Internet paths.

Trains of Packet Pairs (TOPP) is a technique developed by Melander et al., in [77]. The authors make vary the input dispersion between successive trains, until the input rate  $R_i$  becomes larger than the output rate ( $R_i > R_o$ ). Then, a feedback signal must be sent to adjust the increments of  $R_i$ . The output dispersion is filtered by linear interpolation. When  $R_i$  tends to  $R_o$ , it indicates that  $R_i$  approaches the available bandwidth  $\mathcal{A}$ . This condition allows detecting the “turning point”, when  $R_i \approx \mathcal{A}$ . This technique is considered one of the first iterative techniques to probe available bandwidth.

### 3.2.4.2 The General Model

Jain Manish, also in [56], explains the iterative-probing technique. This technique is also called the Probing Rate Model in [67]. In iterative-probing/PRM, we do not need to know the capacity of the tight link  $C_r$ . The sender transmits a periodic probing stream  $k$  with rate  $R_i(k)$ . The rate  $R_i(k)$  varies either linearly, or adaptively as a function of the output-variables, e.g. OWDs, output rates. The probing process is as follows:

- If the  $k$ -th packet train arrives to the receiver with an output rate that is smaller than the input rate in the transmitter ( $R_o(k) < R_i(k)$ ) and therefore, the OWDs of that stream will show an increasing trend, then the input rate will be superior to the available bandwidth  $R_i(k) > \mathcal{A}$ .
- Otherwise, the input rate will be smaller or equal to the available bandwidth  $R_i(k) \leq \mathcal{A}$ .

The basic idea is that, through a sequence of trains with different rates, iterative probing can converge to  $\mathcal{A}$ . Then, summarizing the principles of direct-probing/PRM, we have:

$$R_o(k) < R_i(k) \text{ if } R_i(k) > \mathcal{A}; \text{ else } R_i(k) \leq \mathcal{A} \quad (3.9)$$

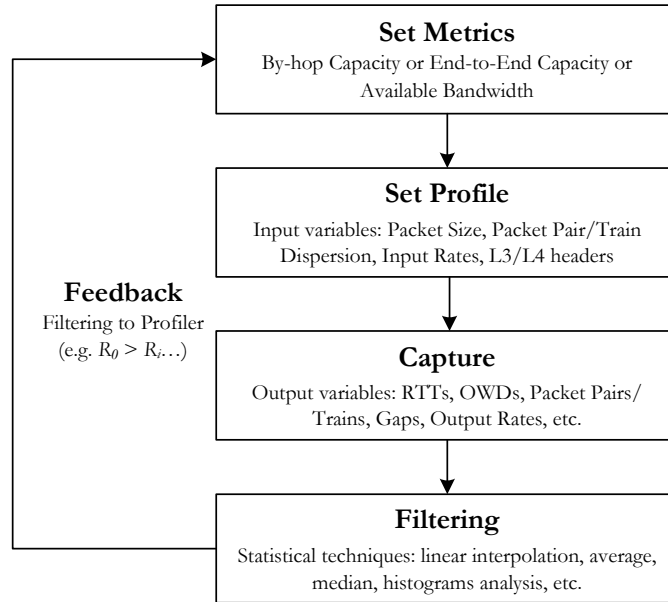
A key point about iterative-probing/PRM is that it does not sample the available bandwidth process; instead, it only samples whether a rate is larger than the available bandwidth or not by introducing a feedback mechanism, to adapt the input rate in function of the output rate and finds when  $R_i(k)$  begins to be close to  $\mathcal{A}$ .



### 3.2.5 Profiler, Capture, Filtering and Feedback (MPCFF) Framework

In order to facilitate the comparison among probing techniques and respective tools, we have defined a functional framework called MPCFF. It helps us to decompose the mentioned probing techniques in five main functions. So, MPCFF intends clearly isolating structural characteristics of the four seminal techniques and facilitate the taxonomic description of probing tools.

In Fig. 3.2, we illustrate the five probing functions of the MPCFF framework: Set metrics, Set Profile, Capture, Filtering and Feedback. In the following sections, we mention the precursor references that have been involved to define these functions. We do not focus on an exhaustive description of each reference, but on pointing out the relationships with the five functions.



**Fig. 3.2. MPCFF functional framework**

#### 3.2.5.1 Set Metrics

Before defining any probing technique or tool it is necessary to clearly establish the kind of bandwidth metric, which we aim to measure. As we mentioned at the beginning of this chapter, the metrics measured by the four techniques are derived from capacity and available bandwidth. Specifically, the measured metrics can be:

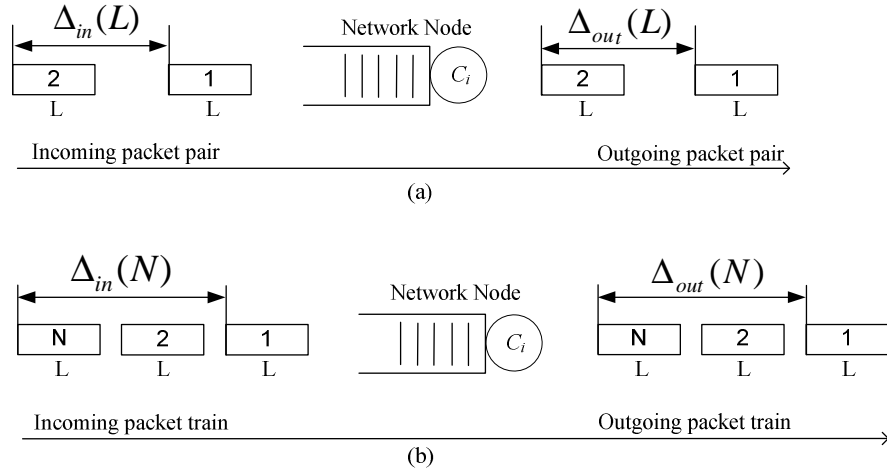
- a) By-hop capacity
- b) End-to-end capacity and
- c) End-to-end available bandwidth

A number of influential references have considered these metrics for Internet paths. For instance, for the by-hop capacity, we find the work of Bellovin [99]. For probing of end-to-end capacity (on transatlantic Internet paths), we find the work of Bolot in [5]. For end-to-end available bandwidth, we find the TCP congestion avoidance mechanism developed by Jacobson [54] and the Packet Pair Flow Control Protocol developed by Keshav [60].

### 3.2.5.2 Set the Profile of Probing Flows Function

The input dispersion between packets or trains, the rate of the probing flows and the forwarding treatment (suffered by the probing flows) will affect the bandwidth conditions of the measured paths.

Since the dispersion of probing packets or trains is a recurrent concept, we explain it, next. The dispersion ( $\Delta$ ) of a packet pair or train at a specific link of the path is the time distance between the last bit of the first and last packets. Fig. 3.3 shows the dispersion of a packet pair or train before and after the packet pair or train goes through a link of capacity  $C_i$  (assuming that the link does not carry other traffic). Then, the dispersion after traversing a node with capacity  $C_i$  is inversely proportional to the bandwidth of this capacity, see equations 3.3 and 3.4.



**Fig. 3.3. Packet pairs and trains dispersion, based on [92]**

Another important input variable is the input rate  $R_p$ , which is associated to a probing flow just before traversing the measured path. Additionally, the size of packets and trains, of the probing flows, will impact the noise resistance and accuracy of the bandwidth measuring process, as explained in [16].

On the other hand, the IP addresses (source and destination, included in the probing packets headers) that correspond to the sender and receiver probing terminals, the

transport protocol and the marking for specific traffic classes will determine the way in how the probing flows are forwarded and the interactions (with the cross traffic), through the measured paths.

Then, we have called the function of “profile of probing flows” the one that sets a combination of the next input-variables:

- a) Packet Sizes
- b) Packets Dispersion
- c) Trains Dispersion
- d) Input Rate
- e) IP address source/destination
- f) Transport protocol and
- g) L2/L3 Marking (if used)

We find different usage of input-variables through the precursor techniques. For instance, the dispersion and size among TCP segments were considered in the earlier work of Jacobson [54]. Packet sizes variation to probe by-hop capacity is proposed in Bellovin [99]. Packet pair dispersion is used in the work of Keshav [60], Bolot [5], Carter and Corvella [8] and Lai [65]. Paxon in [89] uses, for the first time<sup>2</sup>, packet train dispersion instead of the packet pair dispersion.

As mentioned, the profiles of the probes are also dependent of the TCP/IP stack. We find the measuring of TCP traffic flows in passive manner, in the work of Keshav [60], Paxon [89] and Lai [65]. ICMP-ECHO packets are implemented, in the work of Bellovin [99], Carter and Corvella [8], while UDP probes are used by Bolot [5].

No specific traffic classes and respective marking have been found for probing techniques in the seminal references, since they were developed for Internet paths (best-effort disciplines). However, traffic classes can be important in complex QoS WAN/LAN/Home-Networks scenarios.

### 3.2.5.3 Capture Function

The capture function allows obtaining and storing the output variables of the probing process. The output variables represent the behavior of probing flows, when they interact with cross-traffic and the network nodes along the paths. Then, the output variables depend on the number of nodes, the end-to-end network impairments (latency, packet

---

<sup>2</sup> We will find later the use of trains in [11]

delay variation, etc.) and the cross traffic load, along the path. The considered output variables are:

- a) RTTs
- b) OWDs
- c) Output Gaps and
- d) Output Rates

In a symmetric TX/RX channel with no traffic, we can consider that  $RTT = 2 * OWD$ . The OWD can be comprised of three delay components: a) the transmission time due to the transmission interfaces, b) propagation time due to the electromagnetic transfer of signals in a link and c) the queuing time due to the serialization of forwarded packets.

Since most popular techniques were used to measure capacity and available bandwidth over Internet paths, it was not always possible to install software at both sender and receiver nodes. In general, if only RTTs are required, the sender and the receiver can be embedded in a single node, such as the popular, tool to test reachability: “ping”. Conversely, if we need to avoid the reverse path interference, due to acknowledgements and round-trip asymmetry, we may require OWDs. Using OWDs increases accuracy and reduces the number of probing packets. On the other hand, since the OWDs include end-to-end timing components between two nodes, timing synchronization between the nodes can be required.

The output gaps, as a synonym of output packet/train dispersion, are characterized as the difference between the arrival times: the time space between the last bit of the first and last packets of a pair or a train. Since these space times are relative with respect to the packet pairs or trains of the same probing flow, we do not need strict end-to-end timing synchronization. Output rates of the probing flows can be derived from the time stamps of output probing packets, for instance.

The RTTs are used in the work of Jacobson [54], Keshav [60], Bellovin [99], Bolot [5], Paxson [89] and Carter and Corvella [8]. OWDs are used in the work of Lai [65], Melander [77] and Jain et al. [75]. The use of packet gaps, as the difference between Inter Arrival Time Stamps, is proposed by Hu et al. [84]. In this way, we can simplify the capture of relative time stamps, without considering end-to-end synchronization. The input rates of probing flows are considered in the Jain’s work [75].

#### 3.2.5.4 Filtering Function

The filtering function allows separating the representative samples from the noise. The filtering function permits to find trends in output variables samples (RTTs, OWDs or gaps, etc.)

For instance, linear interpolation is used (to find the slope of a straight line resulted from interpolated samples), in the work of Jacobson [54], Keshav [60], Bolot [5], Lai [65] and Melander [77]. Fussy exponential averaging is implemented in the work of Bellovin [99]. It is based on the assumption that a system can be thought of as belonging to a spectrum of conditional behavior that ranges from “steady” to “noisy”. Multimodal<sup>3</sup> discrimination filtering is used in the Paxson’s work [89]. Kernel Density Estimation, obtaining adaptive histograms, is proposed in Lai’s work [65]. Heuristic mechanisms are implemented using union/intersection filtering in [8], of Carter and Corvella.

Each filtering technique separates different sources of noise from the representative samples. In the work of Jacobson [54], Keshav [60], Bellovin [99], Bolot [5], Carter and Corvella [8] and Melander [77], the considered sources of noise include probing drops and cross-traffic interference. The noise derived from clock skews is taken into account, in the work of Paxson [89] and Lai [65].

#### 3.2.5.5 Adaptive Feedback Function

The adaptive feedback function is an algorithmic characteristic that controls the input rate or dispersion of packet pairs/trains to adapt them to the load traversing the measured paths. For instance, we can progressively increase the input rate of probing packets until the OWDs reach a certain threshold that indicates the convergence to the available bandwidth. In this way, we are able to control, in real time, the interactions of probing flows with the cross-traffic. Therefore, if we implement an iterative-probing/PRM algorithm, we can track available bandwidth without the need of knowing in advance the capacity of the tight link, as the direct-probing/PGM requires.

One of the first references that introduce feedback is the work of Jacobson [54] with the TCP mechanism for window-size adaptation. It controls the transmitted load in function of packet losses and RTTs. In this way, the TCP protocol is self-clocked, tracks the available bandwidth and allows avoiding congestion. Another adaptive feedback mechanism, that explicitly measures available bandwidth, is found in the work of Melander et al. [77] for the

---

<sup>3</sup> He explains that link diversity is one of the causes that packet/train dispersion can exhibit multimodal behaviors

Trains of Packet-Pairs (TOPP) methodology and in the work of Jain et al. [75] for the Self Loading Periodic Streams (SLoPS) methodology.

In the next section, we describe some existing probing-tools for available bandwidth.

### 3.3. Characterization of Candidate Available Bandwidth Probing Tools for Home Networks

After analyzing the four seminal techniques for capacity and available bandwidth probing, now we aim to choose suitable candidates of software tools. We aim to test these tools in hybrid home networks. Therefore, we compare seven state-of-the-art tools. The chosen tools are the tools with best performance and some of the ones that have been tested on LAN hybrid paths. These tools are Spruce [100], Wbest [71], Allbest [12], Pathload [75], Pathchirp [109], IGI/PTR [84] and Iperf [103], [78]. We describe these tools, below, using our MPCFF framework.

#### 3.3.1 Spruce

Strauss et al., in [100], proposes “Spruce” which uses direct-probing/PGM by injecting UDP pairs. The dispersion of individual input pairs (intra packet pair gaps) is expressed as  $\Delta_{in}$  and the narrow link (which is assumed to be the tight link) has a known capacity  $C_n$ , by means of an external capacity tool. The dispersion between different pair of packets (inter packet pair gaps) follows an exponential distribution with a mean of 5% of  $C_n$  to reduce the volume of probes and produce a Poisson sampling process. The intra packet pair gaps are captured and the output dispersion  $\Delta_{out}$  is filtered as the average of the samples. Then available bandwidth is:

$$A = C_n \times \left( 1 - \frac{\Delta_{out} - \Delta_{in}}{\Delta_{in}} \right) \quad (3.10)$$

Even if Spruce’s accuracy decreases on paths with multiple bottleneck links and with capacities beyond the 100 Mbps, Spruce seems to be very accurate on Internet paths with a single bottleneck [97].

#### 3.3.2 IGI/PTR

Hu et al., in [84], propose “IGI” (Initial Gap Increasing) to infer available bandwidth using the direct-probing/PGM but choosing the  $R_i$  in a iterative manner (PRM). First,  $R_i$  is the largest as possible and it is reduced until getting the “turning point”, when  $R_{in}$  (input) is equal to the  $R_o$  (output). IGI/PTR infers the available bandwidth by injecting input UDP

trains of 60 packets with the rate  $R_t$ . The train gaps are captured and the output  $R_o$  samples are filtered. Using basically equation (3.8), the available bandwidth is obtained.

In [101], Sundaram et al. show that IGI/PTR is relatively accurate with respect to Pathload and Pathchirp on LAN hybrid paths. However since IGI/PTR uses direct-probing/PGM to obtain the available bandwidth, its accuracy is affected when the tight link does not coincide with the narrow link and with heavy load regimes on the measured path.

### 3.3.3 Wbest

Li et al., in [71], propose a two phase algorithm called “Wbest” to probe available bandwidth in paths with a wireless node at one of the ends and where the narrow link should be the same as the tight link. In the first phase, Wbest obtains the path capacity (effective capacity  $C_e$ ) by sending UDP pairs of 1500 bytes with a rate of 500 kbps and inter packet pair gaps of 10 ms (VPD probing). The samples of the output intra packet pair gaps are captured, the median is filtered and its inverse is multiplied by  $L=1500B$  to estimate  $C_e$ . In the second phase, Li et al. use direct-probing/PGM by injecting input UDP trains of  $N$  packets of  $L$  bytes (default values:  $N = 30$ ,  $L=1500B$ ), with a rate  $C_e$ . The samples of the output trains dispersion  $\Delta_o$  are captured and its inverse is multiplied by  $(N-1)L$  to estimate  $R$  (called achievable throughput). Using the values of  $C_e$  and  $R$ , the available bandwidth is estimated as:

$$A = C_e \times \left( 2 - \frac{C_e}{R} \right) \quad (3.11)$$

WBest detects packet loss in packet pairs and packet trains and finally filters these losses. For a packet train, loss rate  $p$  is recorded and the available bandwidth estimate reduced. If  $p > 0$  then  $A (A \times (1 - p))$ .

WBest does not use a search algorithm to determine the probing rate (which helps to reduce the probing duration). Instead WBest computes the ratio of the effective capacity, which is measured with packet pairs. For this reason, WBest can converge quickly and yield low estimation errors over single link paths or on paths that respect the direct-probing/PGM assumptions. WBest can suffer inaccuracy if the last hop on the wireless path does not coincide with both the tight link and the narrow link.

### 3.3.4 Allbest

Delphinanto et al. propose a prototype tool called Allbest [12]. The sender and the receiver are implemented in the same node Allbest. Allbest uses ICMP ECHO pairs (or trains) of

packets to probe available bandwidth. Allbest, similarly as Wbest, is based on the direct-probing/PGM but Allbest uses back-to-back ICMP ECHO pairs of packets. Allbest copes with the effect of the extra overhead of the wireless medium due to the random contention between the ICMP ECHO requests and their replies.

Allbest estimates capacity by sending a single request with 2 times the size of a MTU, instead of sending two packets of a MTU size each. In this way, rather than waiting that the first reply leaves the bottleneck (and delays both, the second request and the two replays), the single request is fragmented, in the network. Then, two request packets and a single reply will be sent, allowing the measuring of the correct  $RTT_2$  (where  $RTT_2$  corresponds to the second packet of the pair, and  $RTT_1$  corresponds to the first packet). The  $RTT_1$  of the first request is obtained separately with another independent request using a single packet, with a size of MTU. With the corresponding values of  $RTT_1$ ,  $RTT_2$ , and  $L$  (the MTU of the probe packets), Allbest calculates the capacity using next equation [12]:

$$C = \frac{L}{\min_{i=1\dots n}[RTT_2(i)] - \min_{i=1\dots n}[RTT_1(i)]} \quad (3.12)$$

Then, to estimate the available bandwidth, the authors of Allbest consider that the difference between the average and the minimum of the  $RTT_1$ s of probing packets reflects the dispersion mainly due to cross traffic and then, they infer the available bandwidth using:

$$A = \frac{L}{\min_{i=1\dots n}[RTT_2(i)] + \text{avg}_{i=1\dots n}[RTT_1(i)] - 2 \min_{i=1\dots n}[RTT_1(i)]} \quad (3.13)$$

The main drawback of Allbest is that it has been tested as a prototype on Ethernet and wireless links, but the authors do not mention if there is a software version to be tested in different scenarios that include complex hybrid paths (e.g., adding PLC links). Allbest basically consists of a configurable UDP packet generator (to measure the real capacity as a reference value) and a configurable ICMP Ping packet generator, combined with Wireshark.

Allbest as Wbest can also suffer inaccuracy if the last hop on the wireless path is not located at the end of the measured path (typically a Wifi Access Point). This constraint could make Allbest not suitable for complex paths such as the one shown in Fig 4.1 (e). In this scenario, the narrow capacity link can be located whether on the PLC or on the WiFi link in function of the transmission conditions.



### 3.3.5 Pathload

Jain et al. in [75], propose “Pathload” using iterative-probing/PRM. Pathload basically injects two streams with rates  $R_i^{min}$ ,  $R_i^{max}$  respectively, that vary according to a binary searching manner. It is done successively, until one or the two output rates  $R_o^{min}$ ,  $R_o^{max}$  become inferior than the respective inputs  $R_i^{min}$ ,  $R_i^{max}$  (because  $R_i^{min}$ ,  $R_i^{max}$  or both exceed the available bandwidth on the measured path). At this time, a feedback signal is sent to adjust the variation of  $R_i^{min}$ ,  $R_i^{max}$  until a specific trend of OWDs increments is detected. The median of OWDs samples is filtered by Pairwise Comparison and Difference Tests<sup>4</sup> to obtain the OWDs increment trends.  $R_o^{min}$ ,  $R_o^{max}$  are computed with the filtered values of the OWDs and the equation (3.9) is applied to obtain the available bandwidth.

Pathload is able to report a minimum and a maximum value of available bandwidth, instead of the average values. It can be used in some QoS scenarios where the limits min and max of available bandwidth are required to anticipate network conditions. Additionally, Pathload is suitable to make end-to-end probing on complex paths because Pathload does not assume that the narrow link must be the tight link and its estimations present resistance to multiple bottlenecks. Pathload has also been tested with gigabit interfaces and preliminary tested on LAN hybrid links in [101]. As reported in [97], some of the drawbacks of Pathload include that the applicability of Pathload may be limited to non-real-time applications, or applications that do not need to have a bounded response time. Additionally, Pathload can show high variability of estimation convergence time.

### 3.3.6 Pathchirp

Ribeiro et al., in [109], propose “Pathchirp” based on iterative probing. Pathchirp varies the dispersion between successive input pairs (inter pair gaps) in a chirp (a group of pairs) by a “spread factor”  $\gamma$  that follows an exponential distribution. Then,  $N$  packet pairs of a single chirp test the paths to  $(N - 1)$  different rates. Pathchirp sends chirps until detecting queuing excursions when the input dispersion rate  $R_i$  becomes larger than the output dispersion rate  $R_o$ . The output gaps are captured and the correspondent queuing increments thresholds are filtered to give the available bandwidth. As in Pathload, as long as  $R_o = R_i$ , the available bandwidth on the measured path is larger than or equal to the input rate ( $A \geq R_i$ ). Conversely when the output dispersion rate begins to be lower than the

---

<sup>4</sup> Pairwise comparison generally refers to any process of comparing entities in pairs to judge which of each pair is preferred, or has a greater amount of some quantitative property. Pairwise difference test is a type of location test that is used when comparing two sets of measurements to assess whether their population means differ.

input dispersion rate ( $R_o < R_i$ ) then, the available bandwidth is lower than the input dispersion rate ( $A < R_i$ ).

Pathchirp has a command line that permits configuring a number of parameters to tune it on different network scenarios. As IGI/PTR and Pathload, Pathchirp has also been tested on paths with serialized wireless and PLC links [101]. One of the drawbacks of Pathchirp is that it can give erroneous measurements with heavy load on Internet paths. Conversely, Pathchirp, using Poisson sampling, can reduce the number of probes and as a consequence its intrusiveness in the end-to-end estimation.

### 3.3.7 Iperf

In contrast with most available bandwidth tools, which are based on UDP probes, Iperf [33] measures UDP and TCP throughput. The UDP throughput is proportional to the L4 capacity (considering the UDP/IP/MAC overhead) when no cross traffic is present. However if one aims to estimate available bandwidth when cross traffic is present, it can be really difficult. This is because, the way to probe the UDP throughput is by sending UDP flows (with a specific UDP payload size) with a rate equals to the UDP capacity of the measured path. It should be done during a convenient period of time to obtain enough accuracy. As we can see, the UDP probing flows will completely saturate the path. When no cross traffic is present, the capacity can be obtained accurately. However, when cross traffic is present, if for instance, the UDP flows are of the same traffic class that the cross traffic, then both the probing traffic and the cross traffic will tend to share the path capacity. In this case, the possible estimations of capacity or available bandwidth will be strongly biased. Furthermore, if the measured path has links with time-varying-capacity, setting accurately the probing rate of UDP flows could be impossible, without an automatic capacity tracking procedure or an external capacity probing tool.

On the other hand, the TCP throughput is the total data that can be transmitted in function of at least the following parameters: available bandwidth, RTTs, loss rate, efficiency of the TCP protocol and upper layer interactions [29], [16]. By definition, a TCP flow adjusts itself to the available bandwidth in the paths it traverses. It tends to maximize the occupancy of these paths by adapting the number of transmitted segments in the congestion window. In coordination with the slow start and congestion avoidance mechanisms, TCP determines congestion, based on packet losses (e.g. “3 dupacks”) and RTTs. Then, when a given congestion window begins to overwhelm available bandwidth or when the transmission path is saturated, packet losses and timeouts can occur and the

congestion window must be reduced. Contrary to the UDP behavior, the slow start and congestion avoidance mechanisms of TCP need a feedback control loop. This control loop is based on ACK packets, which adds overhead and algorithmic complexity. Fortunately, for a large number of applications, operating systems, and network scales, TCP has been optimized and is stable in different operating systems, which is not the case of most open-source available bandwidth probing tools.

The source code of the tools we presented in this section is publicly available and their structural characteristics are summarized in Table 3.1 based on the proposed MPCFF framework.

TABLE 3.1. STRUCTURAL CHARACTERISTICS OF THE CANDIDATE TOOLS TO MEASURE AVAILABLE BANDWIDTH IN HOME NETWORKS

Probing Functions		Probing Tools						
		Spruce	Wbest	Allbest	Pathload	Pathchirp	IGI/PTR	Iperf
<b>METRIC REPORT FUNCTION</b>								
Path Capacity			X	X				
Available Bandwidth Direct-probing/PGM		X	X	X				
Available Bandwidth Iterative-Probing/PRM					X	X	X	
UDP/TCP throughput								X
<b>PROBE-PROFILER FUNCTION</b>								
Input variables	UDP Packets Pairs/Dispersion	X	X					
	UDP Trains/Input Rate		X			X	X	
	UDP Streams /Input Rate				X			X
	TCP Segment Size (variable window size)							X
	ICMP packets (fixed rate)			X				
<b>CAPTURE FUNCTION</b>								
Output variables	RTTs			X				X
	OWDs/output rate				X			
	Gaps	X	X			X	X	
	Packet losses							X
<b>FILTERING FUNCTION</b>								
Filter Techniques	Average		X	X			X	
	Median		X					
	Average of output gaps from a Poisson sampling	X				X		
	Pairwise Comparison and Difference Tests of median OWDs				X			
	TCP algorithm based							X
<b>FEEDBACK FUNCTION</b>								
Adaptative feedback for input variables					X		X	X

### 3.4. Selecting Available Bandwidth Probing Tools for Home Networks

Now, we aim to select a candidate tool that can be tested on home networks. However, the selection of the correct tool is a complex task, because, with the exception of Wbest and Allbest, most of the tools have not been developed for hybrid LAN environments.

Additionally, the estimation of the available bandwidth in hybrid home networks is challenging because of the network measurement constraints as explained in section 3.2. These constraints include: strong delay variability in probing flows induced by multiple-access mechanisms and link with capacities that vary in the time (e.g. in wireless and PLC links). Furthermore, the selected tool should be: easy to calibrate, transparent to lower layers, accurate, fast (low convergence time), not intrusive (low overhead) and stable in different system and network conditions.

As explained, Spruce, Wbest, Allbest and IGI/PTR are part of direct-probing/PGM. One of the advantages of direct-probing/PGM tools is the possibility of simple algorithmic implementations due to their single non-adaptive incoming rate  $R_{in}$ . However this simplicity limits their accuracy over multiple bottleneck paths. Additionally, direct-probing/PGM tools assume that the tight link capacity corresponds to the narrow link capacity and is accurate when the cross-traffic is “path persistent” (follows, strictly, the same path that the probing flows), which is not always the case [67], [76].

On the other hand, Pathload and Pathchirp are iterative-probing/PRM tools. Unlike direct-probing/PGM tools, iterative-probing/PRM tools do not require the a priori knowledge of the tight link capacity because they try to adapt the probing flows to the load on the path. The input rate variations are obtained from fine-grained changes of delay indicators (e.g. One Way Delays, Round Trip Times) of a relatively small amount of probes. For this reason, complex filtering techniques are required to clean noise effects on the paths. Usually PRM tools seem to better deal with multiple bottlenecks than the PGM tools [101], [76].

In fact, preliminary benchmarking of the probing tools has been done over hybrid local networks [12], [101], [7]. Those results confirm the limitations of the analyzed probing tools on hybrid local networks. They show that these tools are inaccurate, slow and not stable enough (sensitive to clock resolution, timing slips, system interruptions and number, rate and size of probing packets [117], [98], [7], [91]). Only IGI/PTR [101] [84] seems to be relatively accurate but slow on paths that include 100Base-TX, HPAV (PLC plugs), and IEEE802.11. Wbest [71] and Allbest [12] seem to be accurate and fast only on paths that include 100Base-TX and IEEE802.11 at the last hop.

### *3.4.1 Iperf in Lightweight TCP Mode, an Attractive Alternative*

Considering the limitations of current probing tools and the constraints of the measurement process, using Iperf [33] in home networks is attractive, assuming that:

- a) Iperf is a well-known benchmarking tool and conversely to the other tools is a stable one
- b) Home networks have short slow-start periods and a predictable stable TCP performance, in opposition to networks with large bandwidth delay products (BPD), where to reach maximum TCP throughput can require several seconds. In WAN scenarios, if RTTs and path capacities grow, the time needed to adjust the window size of TCP segments, to reach the maximal throughput, can grow dramatically, until tens of seconds [66], [103], [25].
- c) Iperf has been developed for different operating systems (Windows, MacOS, Linux and Android)
- d) Iperf, based on the TCP protocol, can be configured to be not disruptive, when real time applications traverse the measured paths
- e) Iperf can be used as an end-to-end available bandwidth estimation tool for residential services
- f) We can use ***Iperf in lightweight TCP Mode***, with short probing intervals instead of its common brute force utilization on WAN paths [103]. The crucial aspect, we propose, to reduce the inherent intrusiveness of Iperf in TCP mode, is to utilize it in a *lightweight mode*:
  - by using probing intervals of less than one second (duration of the Iperf TCP probing connection). See the details in section 4.2.1.

### 3.5. Summary

In this chapter, we have explored the networking constraints, which impact bandwidth probing at the Internet scale and especially at the home network scale. We presented a state-of-the-art of the techniques and tools for available bandwidth probing. We proposed a framework that helps to analyse the structure of existing bandwidth probing techniques. Then, we established a taxonomical comparison of the existing techniques and tools based on our MPCFF functional-framework. Finally, we showed that Iperf can be an attractive tool to be tested in home networks.

In next chapter, we show through an experimental study the convenience of the utilization of TCP-Iperf with periods of less than one second (lightweight mode), to estimate the available bandwidth, on home networks. We also compare Iperf with respect to IGI/PTR,

Wbest, Pahtload and Patchirp, in terms of accuracy, convergence time and intrusiveness. This comparison obeys to the results of preliminary tests on home networks, cited above.

# Iperf in lightweight TCP mode in Hybrid Home Networks

## 4.1. Introduction

In section 3.3, we have used our MCPFF functional framework to do a taxonomic comparison of the characteristics of a number of state-of-the-art probing tools, which are based on PGM/direct probing, on PRM/iterative probing and on UDP/TCP throughput measurement. As discussed, a number of studies show that most of current probing-tools, for available bandwidth, exhibit low accuracy, and slow convergence in Home Networks. Some reasons of this behavior include the high queuing variability of probing flows, which is induced by multiple-access mechanisms, time-varying capacity links of hybrid paths, and the instability of beta probing tools. Considering these constraints, in section 3.4, we have listed a number of properties that make Iperf a very attractive tool to be tested in home networks.

In this section, we present a performance evaluation of different bandwidth probing tools by means of an experimental test bed. We test Iperf in a lightweight TCP mode to obtain maximum throughput of non-disruptive TCP flows and estimate the available bandwidth.

Our contributions include, firstly, the benchmarking of Iperf in lightweight TCP mode with respect to IGI/PTR, PathChirp and Wbest, on representative hybrid paths, and considering accuracy, overhead and convergence time. Secondly, we propose and reproduce, with a prototype Java test interface (that we developed with the help of an intern), the utilization of a measurement process based on separated measurement time



scales of probing, polling and test intervals to weigh the Iperf overhead, depending on potential client applications.

These two contributions have been published in the “18th IEEE (LANMAN) Workshop, 2011” [87].

Our experimental results show that Iperf doubles the accuracy of the estimates compared with IGI/PTR, Pathchirp and Wbest, on representative hybrid paths. We also show that due to the low bandwidth delay products (BDP), of home network paths, we can keep low overhead using TCP flows, with convenient probing/polling intervals, which reduces the intrusiveness associated to the default Iperf usage.

The rest of the sections are organized as follows. The section 4.2 shows our test performance methodology. The section 4.3 explains our experimental test bed. The section 4.4 gives the performance evaluation results. Finally, we conclude.

## 4.2. Methodology

### 4.2.1 Principles of Iperf in lightweight TCP mode

Iperf, based on the TCP protocol, can be configured to be not disruptive when real time applications cross the measured paths. We explained the general mechanisms of the TCP protocol in section 3.3. The key aspect, we propose, to reduce the inherent intrusiveness of Iperf in TCP mode, is to utilize it in a *lightweight mode*, by using probing intervals of less than one second (duration of the Iperf TCP probing connection). In fact, in WAN environments, transmissions can experiment high round trip times (RTTs), tens or hundreds of milliseconds, basically because long distance propagation times, consequently, paths with large capacities can have high bandwidth delay products (BDP). When RTTs and path capacities grow, the time, needed to adjust the window size of TCP segments to reach the maximal capacity, can also grow dramatically [103], [25]. So, TCP may require several seconds to converge to its maximal throughput and longer probing intervals must be required if we use Iperf. However, in LAN environments, we assume that the RTTs regime is globally low, of few milliseconds, and TCP flows can converge rapidly to the throughput. Consequently, we can use Iperf in TCP mode, with probing intervals of less of a second (less than 1/10 times of its default 10s), much less than its usual brute force utilization on WAN paths [103].

As a result of these considerations, Iperf in lightweight TCP mode may better deal with the multiple-access queuing of hybrid paths, since it can react faster by sending adaptive and

more aggressive probing streams (of short TCP flows) being more accurate, fast, and reducing the overhead associated to the default Iperf usage with long probing intervals.

#### 4.2.2 Tools Performance and Measurement Process

To verify our hypothesis, we are interested in benchmarking Iperf in TCP lightweight mode with respect to PathChirp, Wbest, and IGI/PTR over hybrid paths in home networks. For this purpose, we report three main performance parameters: accuracy, probing interval, and correspondent overhead of each tool. Then, we divide the measurement process into two phases: the *platform setup phase* and the *measurement phase*. We briefly describe these phases.

The platform setup phase is critical and consists in the setup of a homogenous, stable and calibrated test platform. In our case, we have used low cost commercial hardware and open-source software. Some of the experimental criteria that must be carefully set to perform unbiased measurements are:

- a) Stability and compatibility of PC hardware and NICs (100Base-TX, IEEE802.11) with their corresponding software controllers over a specific operating system,
- b) The calibration of the probing tools by optimizing different configuration parameters in terms of accuracy, probing intervals, and overhead.

The measurement phase allows generating regular measurement tests of available bandwidth and related accuracy, overhead, and probing intervals of each benchmarked tool. To this aim, we developed a prototype Java test interface, which is described in section 4.3. It automates the measurement process using the following measurement scales:

- Test interval  $T$  is the whole duration of the test defined to characterize the behavior of a probing tool, given the presence of a cross traffic pattern on every different path. In our study, we define cross traffic following equation (4.3) of section 4.4.
- Cross traffic step interval  $T_s$  is a constant fraction of  $T$ , during which, a single cross traffic rate  $R_i$  is injected. In a network, in operation,  $T_s$  should characterize the traffic duration in specific home network conditions.
- Polling interval  $T_{po}$  is the period between the beginnings of two successive executions of the probing tool, which defines the periodicity of available bandwidth ( $\mathcal{A}$ ) reporting. We chose the  $T_{po}$  sufficiently large in order to reduce the polling weighted overhead  $O_{po}$  (defined in inequality (4.2)) while preserving the accuracy of  $\mathcal{A}$  measurements.

- Probing interval  $T_{pr}$  is the period between the beginning and the end of the first and the last probing packets of a single execution to obtain an available bandwidth measurement  $\mathcal{A}_i$ .  $T_{pr}$  is used to compute the probing overhead in inequality (4.2).
- Convergence time  $T_c$  is the period, needed by a probing tool, to provide an  $\mathcal{A}_i$  value. This time is defined as:  $T_c = T_{pr} + T_{co}$ , where  $T_{co}$  is the computation time required by the tool to obtain  $\mathcal{A}_i$  after it processes the probing packets in the receiver. In the case of our tests with Wbest, PathChirp, and Iperf, we have seen that  $T_{co}$  is negligible in most cases and  $T_{co} \approx T_{pr}$ . However, IGI/PTR requires a considerable  $T_{co}$  (between several milliseconds to several seconds), for this reason, we take  $T_{pr}$  separately, from the output called “Probing uses”, which is given by the tool after each execution.
- The probing bits  $b_p$  are the total amount of bits which are involved in the  $\mathcal{A}$  probing process. These bits correspond to the probing packets plus the required signaling packets interchanged between the sender and the receiver probing terminals to process and report a single  $\mathcal{A}_i$  measurement.  $b_p$  can be obtained from the total sent bytes  $B$  of layer one or upper layers kernel statistics, for instance.

From the previous definitions we have derived the following inequalities:

$$T \geq T_s \geq T_{po} \geq T_{pr} \quad (4.1)$$

$$O_{po} = \frac{b_p}{C_n T_{po}} \leq O_{pr} = \frac{b_p}{C_n T_{pr}} \quad (4.2)$$

where  $C_n$  is the narrow link capacity, which is the bottleneck capacity (the lowest) of the path.

### 4.3. Experimental Test bed

We have defined five path scenarios to test the selected tools, see Fig 4.1. In each scenario, two laptop-pairs are connected at the ends of the paths. The laptops have 2 Ghz Core Duo processors, 2 GB of RAM, Broadcom Nextream Gigabit Ethernet controllers, Intel PRO/wireless 3945AG controllers, and Linux Kernel 2.6.26-2-686. One pair of laptops (cross traffic TX & RX terminals) transmits and receives the cross-traffic (generated by Iperf in the UDP mode). The other pair (SND & RCV probing terminals) sends and receives the bandwidth probes generated by the installed probing tools, as indicated in

Fig. 4.1. We have used 3Com Wireless 7760 11a/b/g PoE Access Points and Devolo dLAN 200 HPAVs PLC plugs.

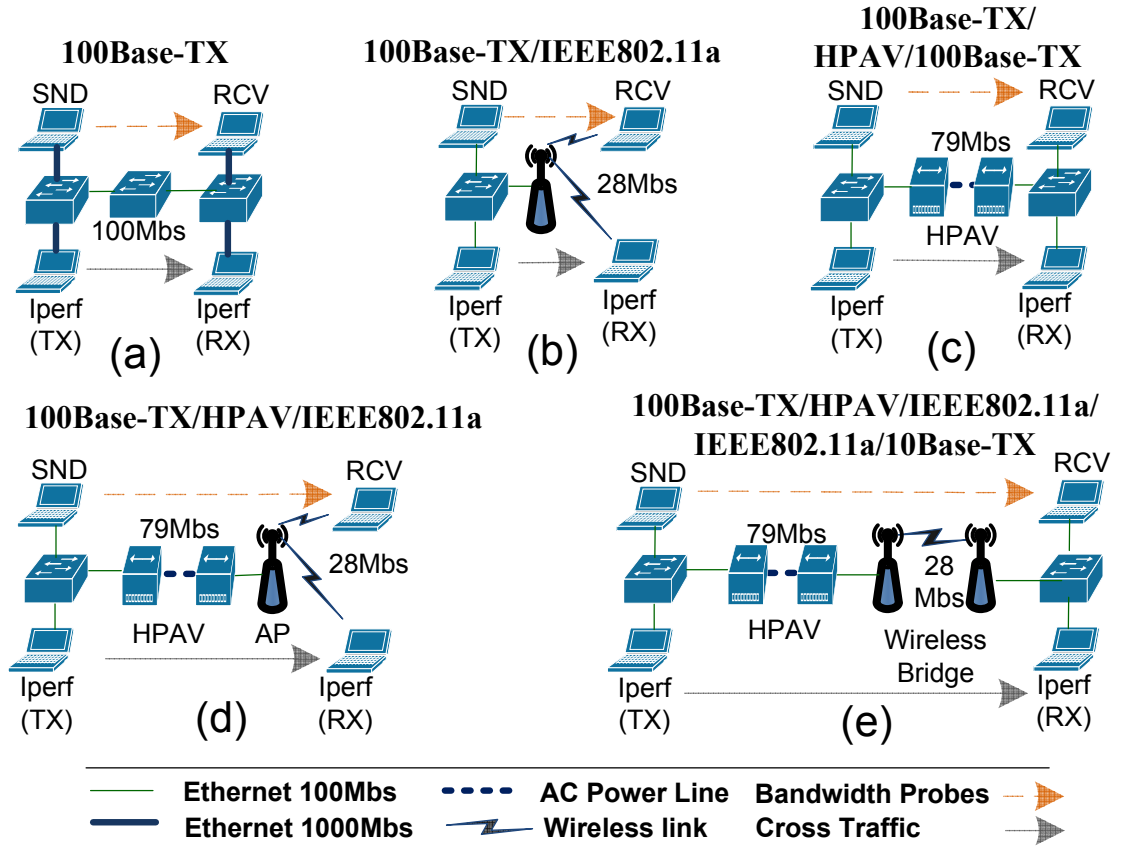


Fig. 4.1. Path Scenarios on Test Bed

#### 4.3.1 Capture Automation of Performance Descriptors

The four tools were interfaced to our prototype Java tests interface (screen captures are shown in Appendix B). The interface consists of a Java application that resides in the SND probing terminal. The overall tasks, accomplished by the interface are:

- Set the configuration parameters required by each probing tool to get an  $A_i$  measurement
- Set a number  $E$  of executions, dividing  $T$  by  $T_{po}$
- Activate, simultaneously, the TX cross traffic terminal using a Linux remote session, and capturing time stamps of the beginning and the end of each cross traffic step  $T_s$
- Capture, for each probing tool execution,  $A_p$ ,  $T_{pr}$ , and  $b_p$  (using `proc/net/dev` Linux statistics)

- e) Plot graphs, in real time, of  $A_p$ ,  $T_{pr}$ , and  $O_{pr}$ , and f) concatenate all values of  $A_p$ ,  $T_{pr}$  (including time stamps), and  $B$ , as a report for further analysis.

To generate the cross traffic from the TX terminal, we have defined  $K = 10$  monotonic cumulative cross traffic rates  $R_p$ , called cross traffic steps as:

$$R_i = \frac{iC_n}{K} \text{ where } i = 0, 1, 2, \dots, K-1 \quad (4.3)$$

Each step  $R_p$  has a duration  $T_s$ . For PathChirp, Wbest, and Iperf, we have:  $T = 600s$ ,  $T_s = 60s$ , and  $T_{po} = 5s$ . Whereas, for IGI/PTR (due to its longer convergence times):  $T = 1200s$ ,  $T_s = 120s$ , and  $T_{po} = 10s$ . This results in 120 available bandwidth measurements (trials) correspondent to 120 individual executions per tool for each path configuration. To obtain the accuracy, the probing intervals, and the overhead of each cross traffic step, we compute the average of the corresponding 12 executions. For the global results, we compute the average of the 120 trials per tool, given a confidence interval of 95%.

To measure  $C_n$  (the bottleneck capacity of the path), we use successive UDP Iperf flows with an increasing rate, near to the nominal narrow link capacity, similar as in [12]. The L4 throughput, when no losses occur or are very low, indicates the  $C_n$  capacity. We also used IPTraf [32] and Wireshark [37] to verify the Iperf UDP measurements.

As capacity on hybrid paths tends to change, in order to obtain a steady capacity reference, we have minimized possible perturbations on Wi-Fi and HPAV links. We have set the Wi-Fi interfaces with the IEEE802.11a standard, using the channel 48, and we have fixed TX/RX rates to 54Mbps and disabled rate adaptation and the RTS/CTS option as in [12]. In the case of the HPAV, we connected the HPAV devices on a separated extension cord. We have done the experiments late in the night in a country side house, far from other IEEE802.11a devices and far from strong noise sources over the electrical outlets. We avoided connecting, at AC circuits, cheap power adapters such as cell phones chargers or domestic motor based appliances.

### 4.3.2 Path Scenarios

As we can see in Fig. 4.1, the path configuration complexity increases from scenario (a) to scenario (e). In each scenario, a bottleneck is placed over the measured path. For example, at scenario (a) there is a 100Base-TX bottleneck, at scenario (c) a HPAV bottleneck and finally in scenarios (b), (d), (e) an IEEE802.11a bottleneck. To the best of our knowledge,

it is the first time that such a benchmarking over paths composed of 100Base-TX, HPAV, and IEEE802.11 links has been reported. In regards to the probing tools calibration (as mentioned in section 4.2), we looked at setting a minimum of parameters but trying to maximize accuracy and minimize probing intervals and overhead. We have tested Wbest using its default parameters of 30 packet pairs to measure the capacity and 30 packets per train as in [71], [12] to obtain available bandwidth. We set the UDP probing payload to 1472B instead of 1460B. We have calibrated PathChirp [109] using the scenario (c). Thus, we have set the probing interval to 2s (instead of default 600s), the size of probing payload to 1472B (instead of 1000B), the average probing rate to 3Mbps (instead of 0.3Mbps), the inferior limit of sampling rate to 1Mbps (instead of 10Mbps), the decrease factor  $F$  to 1.4 (instead of 1.5), the busy period length  $L$  to 8, and the number of estimates per execution to 15. Decreasing  $F$  or increasing  $L$  tends to overestimate the available bandwidth. In opposition, increasing  $F$  or decreasing  $L$  will lead to underestimation of available bandwidth. For IGI/PTR, we used the PTR outputs of available bandwidth as in [101] and we have set the packet size of the trains to 1472B as in the precedent cases. Finally, Iperf was set to use 0.8s of probing interval and fixed the TCP window size at the SND to 85.3KB to equalize the RCV window size. We have tested Iperf in TCP mode (for the sake of brevity, we do not include the details here) with 0.2s, 0.3s, 0.5s, 0.8s and 1s over the five scenarios. We found that from 0.8s, accuracy begins to be more stable. It is straightforward, that sending TCP flows (of relative short duration) of 0.8s is 12.5 times less disruptive than sending flows of 10s as is the default utilization of Iperf in TCP mode. Although in [103], it is suggested to use Iperf within periods of a second over the Internet. No report in the literature uses this scale in home networks.

## 4.4. Results

### 4.4.1 Benchmarking

Figures 4.2 to 4.6 and Table 4.1 show the results of the available bandwidth, the overhead, and the probing interval in function of cross traffic variations for the five tested scenarios. In figures 4.2 to 4.6, we trace the actual available bandwidth as the difference between the narrow link capacity and the cumulative cross traffic.

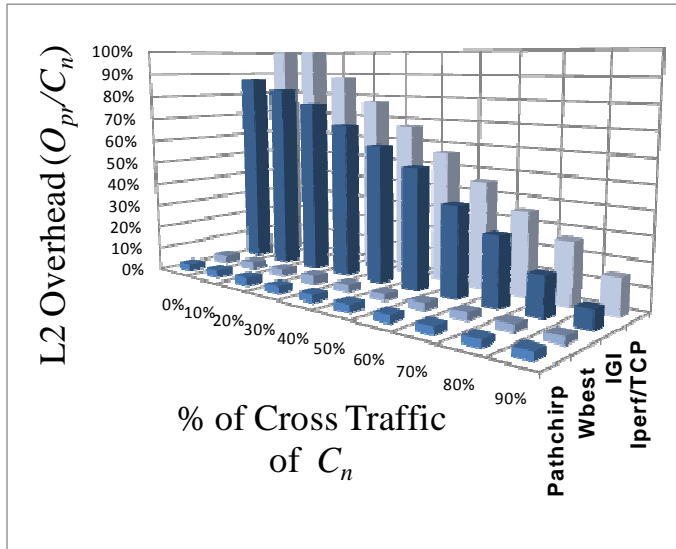
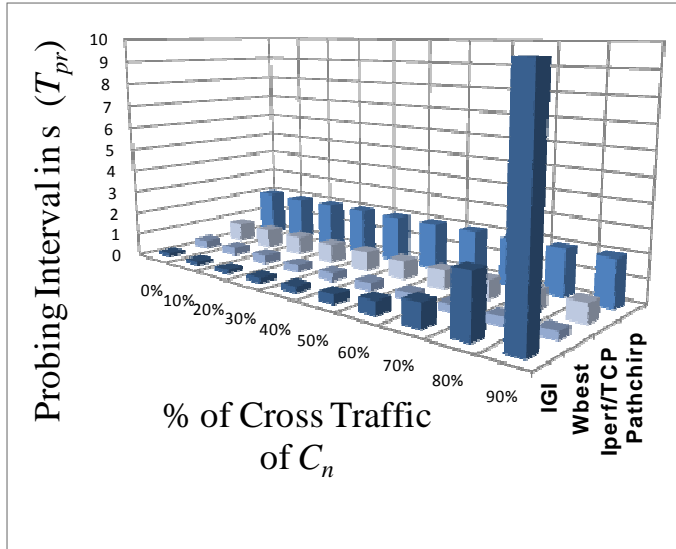
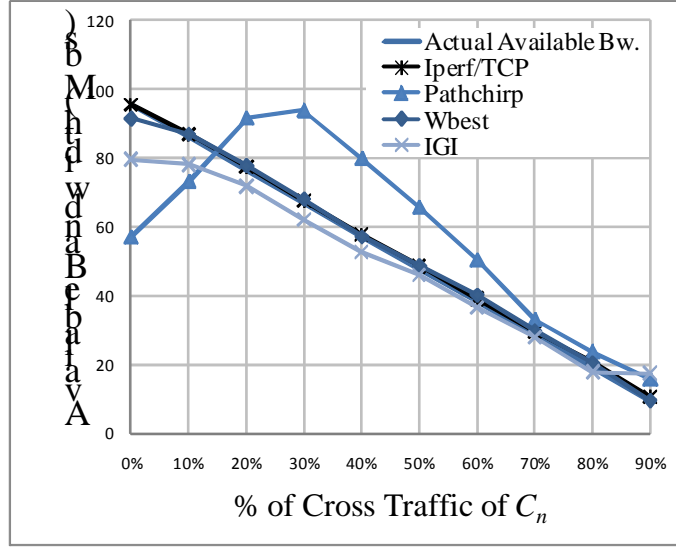


Fig. 4.2. Scenario (a) 100Base-TX: Available Bandwidth, L2 Overhead and Probing Interval versus Cross Traffic

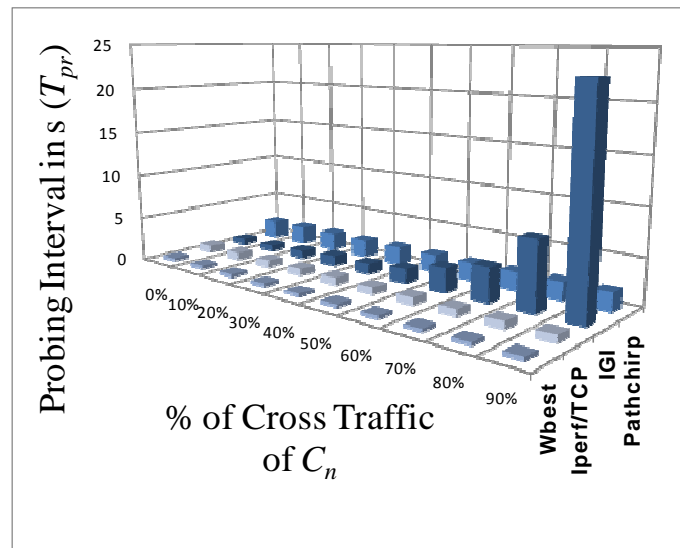
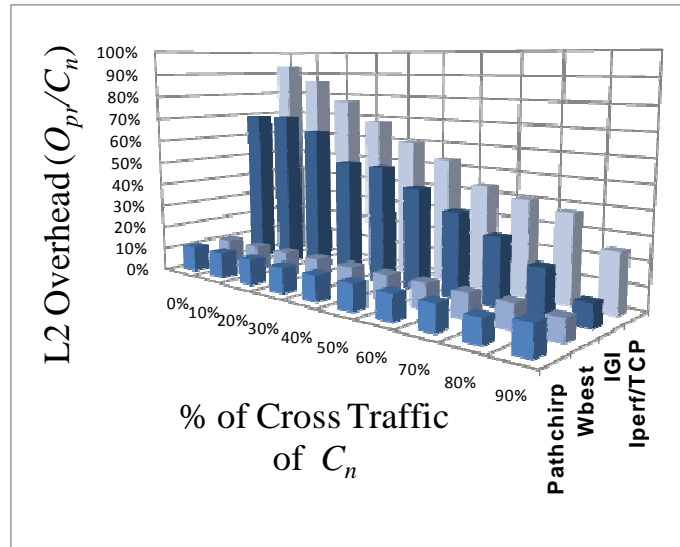
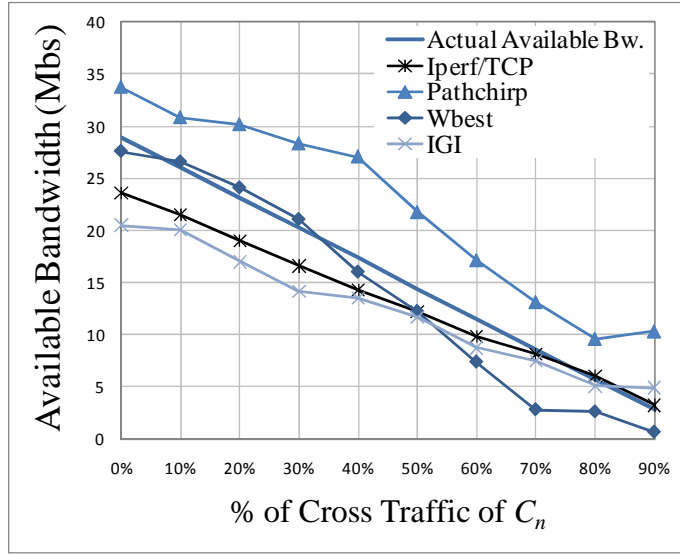


Fig. 4.3. Scenario (b) 100Base/TX-IEEE802.11a: Available Bandwidth, L2 Overhead and Probing Interval versus Cross Traffic



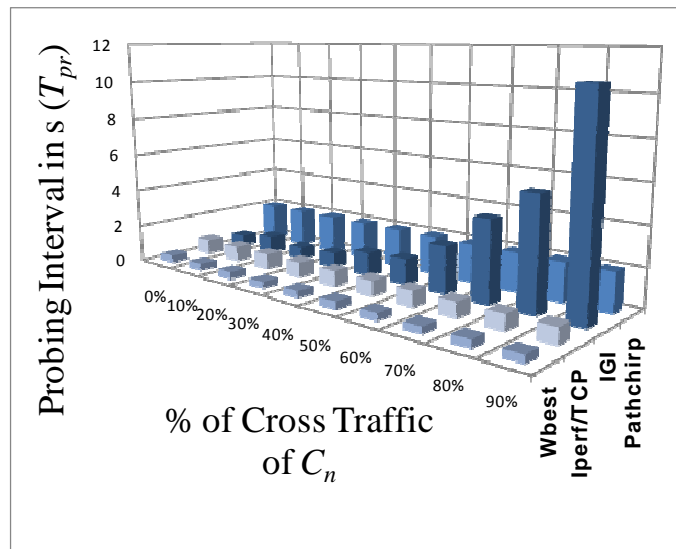
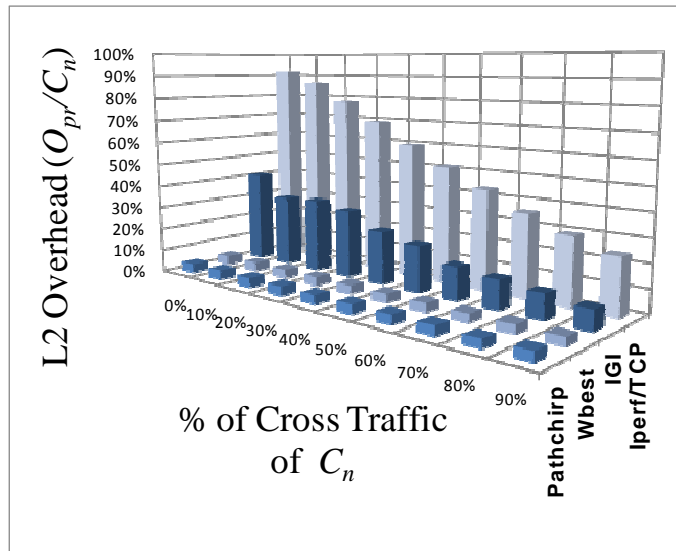
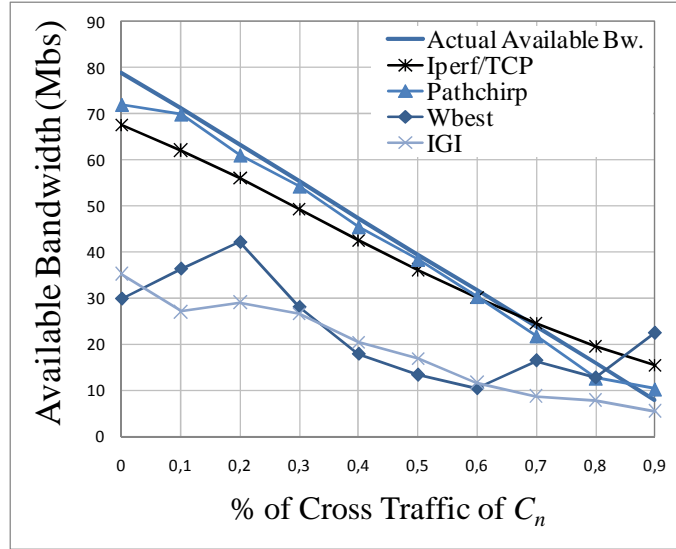


Fig. 4.4. Scenario (c) 100Base-TX/HPAV/100base-TX: Available Bandwidth, L2 Overhead and Probing Interval versus Cross Traffic

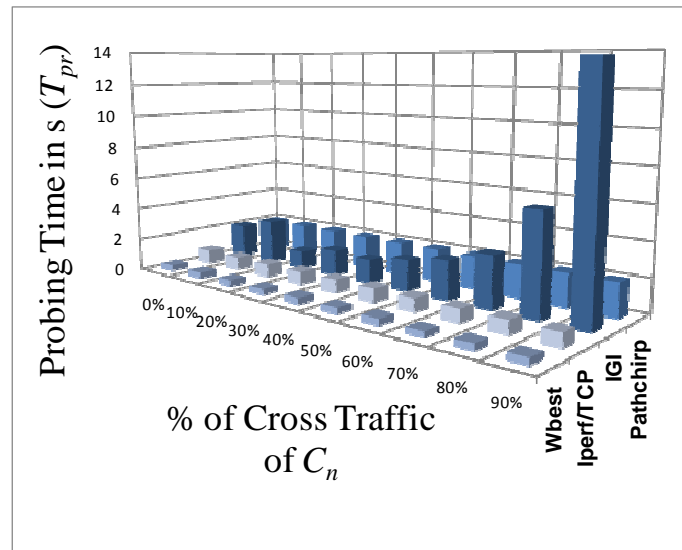
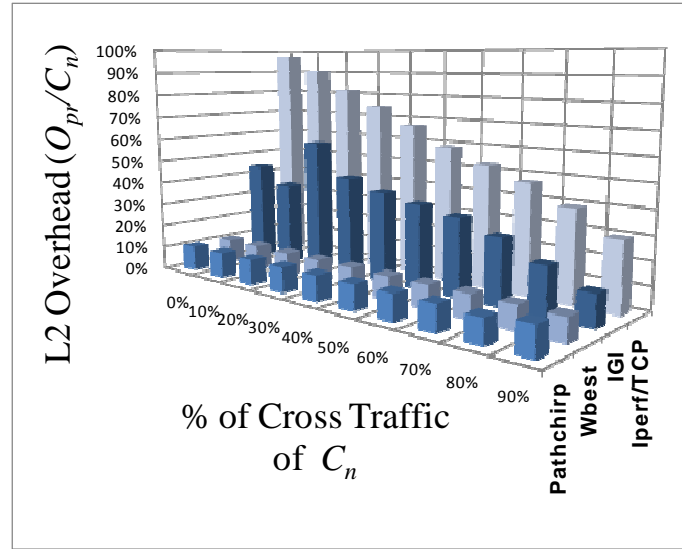
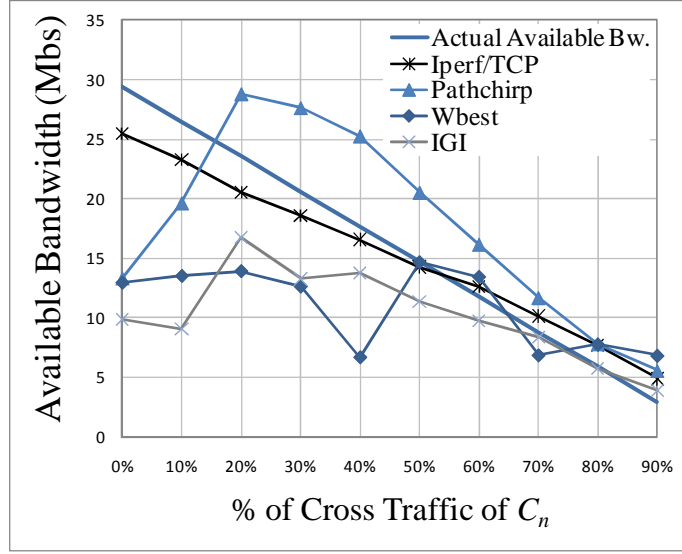


Fig. 4.6. Scenario (e) 100Base-TX/HPAV/IEEE802.11a/IEEE802.11a/100Base-TX: Available Bandwidth, L2 Overhead and Probing Interval versus Cross Traffic

TABLE 4.1. GLOBAL PERFORMANCE RESULTS

PERFORMANCE METRICS	PATH SCENARIO	<i>Pathchirp</i>		<i>Wbest</i>		<i>IGI</i>		<i>Iperf</i>	
<b>Average Error</b> (95% of Confidence Interval)	<i>a</i>	28%	+3%	4%	+1%	10%	+3%	3%	+1%
	<i>b</i>	63%	+22%	38%	+6%	22%	+5%	15%	+1%
	<i>c</i>	14%	+3%	66%	+10%	51%	+4%	20%	+5%
	<i>d</i>	18%	+3%	35%	+6%	36%	+5%	20%	+2%
	<i>e</i>	42%	+6%	58%	+10%	28%	+5%	17%	+3%
	<b>Global</b>	<b>33%</b>	+7%	<b>40%</b>	+7%	<b>29%</b>	+4%	<b>15%</b>	+2%
<b>Average Probing Interval</b> (s)	<i>a</i>	2.0		0.3		1.1		0.8	
	<i>b</i>	2.0		0.4		2.8		0.8	
	<i>c</i>	2.0		0.4		2.3		0.8	
	<i>d</i>	2.0		0.4		3.9		0.8	
	<i>e</i>	2.0		0.4		2.9		0.8	
	<b>Global</b>	<b>2.0</b>		<b>0.4</b>		<b>2.6</b>		<b>0.8</b>	
<b>L2 Average Overhead</b> in Tight Link over Total Probing Interval (%)	<i>a</i>	3%		3%		19%		60%	
	<i>b</i>	12%		10%		24%		55%	
	<i>c</i>	4%		4%		14%		55%	
	<i>d</i>	11%		10%		20%		49%	
	<i>e</i>	11%		10%		25%		60%	
	<b>Global</b>	<b>8%</b>		<b>7%</b>		<b>20%</b>		<b>56%</b>	
<b>L2 Average Overhead</b> in Tight Link over Total Polling Interval (%)	<i>a</i>	1%		0.2%		2%		10%	
	<i>b</i>	4%		0.7%		8%		9%	
	<i>c</i>	2%		0.3%		3%		9%	
	<i>d</i>	4%		0.7%		9%		8%	
	<i>e</i>	4%		0.7%		8%		10%	
	<b>Global</b>	<b>3%</b>		<b>0.5%</b>		<b>6%</b>		<b>9%</b>	

In terms of accuracy, PathChirp shows a strong underestimation in scenarios (a) and (d) with 0% and 10% of cross traffic. In fact, in most scenarios, we found a trend of initial underestimation under light load. We do not have a detailed explanation of this effect, but we believe that it is related to the exponential increments of probing packet interspaces within a chirp. In scenarios (a), (b), and (e), PathChirp exhibits strong overestimation when cross traffic is between 20% and 60% of the narrow link capacity. The explanation of this is that we have calibrated PathChirp to perform accurately over scenario (c), used as a reference for paths with HPAV. It is also possible to calibrate PathChirp to perform accurately over the scenario (a), if we change the spread factor  $s$  to 1.1, the decrease factor  $F$  to 3.5, and the rate superior limit  $\mu$  to 100Mbps, for instance.

Since, PathChirp underestimates bandwidth over paths with HPAV links, using the default parameters; we have used a low value for  $F$  (1.4) and a large value for the busy period  $L$  (8), which, unfortunately, induces the overestimation in scenarios with no HPAV links. PathChirp is also accurate over the scenario (d).

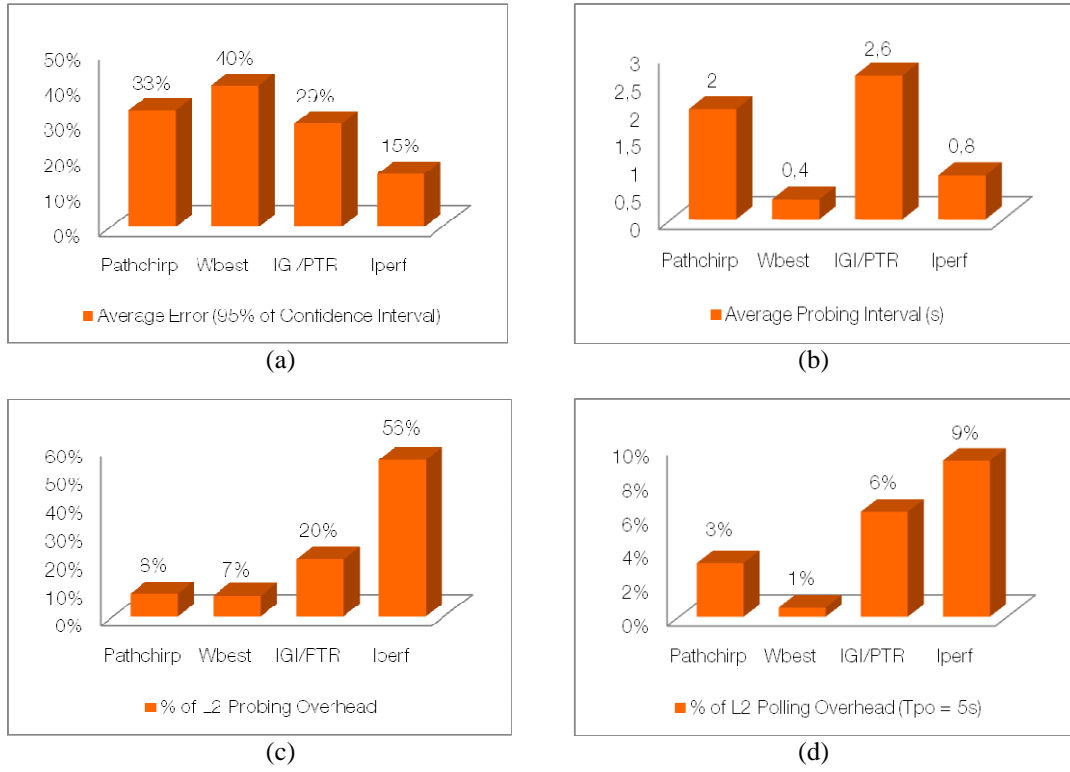
Wbest exhibits high accuracy over scenario (a). However, in scenarios (b) and (d), it lightly underestimates available bandwidth, when cross traffic increases more than 50%. A plausible explanation of this is that cross traffic, at high loads, strongly interacts with the Wbest probing trains, that traverse the wireless link, and then it tends to accelerate queuing increments that are not correctly filtered by Wbest.

In scenarios (c) and (e), Wbest leads to a strong underestimation even without cross traffic. This is due to the fact that these scenarios present a configuration that does not correspond to Wbest assumption, which is to have the Wi-Fi bottleneck in the last link on the path. Then, Wbest is the least accurate of the four tested tools given 40% of relative error from actual available bandwidth.

Among the four tested tools, IGI/PTR is the second most accurate, after Iperf, having a 29% of global relative error from the actual available bandwidth. However, IGI/PTR shows light underestimations in scenarios (a) and (b) and strong underestimations in scenarios (c), (d), and (e). In all cases IGI/PTR performs better under high load than low load, as reported in [101], section 4.4.

Iperf in lightweight TCP mode outperforms all available bandwidth tools. It exhibits only 3% of relative error at scenario (a) and in the other scenarios shows conservative estimations given a global relative error of 15%, which is almost 50% less than the IGI/PTR relative error and almost 75% less than the Wbest relative error.

#### 4.4.2 Global Results



**Fig. 4.2. Global Results**

Highlighting the global results of our performance comparison, we can see, in Fig. 4.2 (a) (from Table B.1) that Iperf outperforms globally the other tools being 50% more accurate than IGI/PTR, while Wbest deviates as far as 40%.

In Fig. 4.2 (b), we see that Iperf is the second fastest tool requiring only 0.8 s to converge to accurate available bandwidth estimations while the fastest one is Wbest and the slowest is IGI/PTR with 2.6 s.

In Fig. 4.2 (c), we see that Iperf is the most intrusive of the compared tools giving more than 50% of probing overhead with respect to the narrow link capacity on the measured paths, but we can reduce its inherent intrusiveness using convenient polling intervals. For example if we use polling intervals ( $T_{po}$ ) of 5 s, see Fig. 4.2 (d), we can reduce the Iperf polling overhead more than five times than the probing overhead, because we reduce the periodicity over the test time of the probes. Obviously, the use of a suited polling interval depends on the target application requiring bandwidth measurement in the home network.

In Table 4.1, we summarize the performance characteristics of the tools. We see that Wbest is the fastest and less intrusive one, but it is extremely inaccurate when the bottleneck is not the last hop, or when high load prevails.

TABLE 4.1. TOOLS PERFORMANCE CHARACTERISTICS

Tool Performance Criteria	<i>Pathchirp</i>	<i>Wbest</i>	<i>IGI/PTR</i>	<i>Iperf</i>
Accurate over Ethernet Single bottleneck	YES	YES	YES	YES
Support multiple bottlenecks	YES	NO	YES	YES
Adaptative probing to the load	NO	NO	YES	YES
Good resolution with Low load	NO	NO	NO	YES
Good resolution with high load	NO	NO	YES	YES
Resilience to multiple access mechanisms on hybrid paths	YES	NO	NO	YES
Easy configuration	NO	YES	YES	YES
Fast sampling on hybrid paths	NO	YES	NO	YES
Low probing intervals variability	YES	YES	NO	YES
Low probing overhead	YES	YES	NO	NO
Global error (95% confidence interval)	33%	40%	29%	15%
Average probing interval (s)	2.0	0.4	2.6	0.8
Probing overhead in the narrow link	8%	7%	20%	56%
Polling overhead using 5s of Tpo	3%	1%	6%	9%

Even though IGI/PTR is the second in terms of accuracy, it presents a relative high error of 29%. Furthermore, it has high convergence times under high load regimes, which could make it unusable for many applications where the reactivity is crucial. Only Iperf in lightweight TCP mode gives very good accuracy over all tested scenarios, firstly due to the adaptive and more aggressive way to sample available bandwidth, when multiple access require longer probing packets and sequences to give maximum throughput, and secondly due to the stability of current TCP versions used by Iperf. In addition, it requires a few number of configuration parameters, it is quick, and its probing overhead is less than ten times the overhead of the default configuration in TCP mode (10s of probing interval). Clearly, Iperf requires to be weighted to optimize its polling overhead over polling intervals associated to potential client applications.

## 4.5. Summary

Considering the limitations of current probing tools and the constraints of the measurement process, in this chapter we proposed to estimate available bandwidth, in home networks, using Iperf in a lightweight TCP mode. The use of probing intervals of less than 1s is much less than its usual brute force utilization on WAN paths, which drastically reduces the overhead, associated to the default Iperf usage with long probing intervals. We conducted an experimental performance evaluation to compare Iperf in a lightweight TCP mode with the tools PathChirp, Wbest, and IGI/PTR. We have setup a test bed of five representative hybrid path scenarios. We have used our prototype Java test interface to capture, in a homogeneous way, available bandwidth, total probing bytes,

probing intervals, and cross traffic. To the best of our knowledge, it is the first time that all these path configurations have been included to benchmark open-source available bandwidth probing tools, and that Iperf is used in lightweight TCP mode for home networks.

Results show that, despite the difference between TCP throughput and available bandwidth, Iperf in lightweight TCP mode can estimate available bandwidth in home networks, where the overall regime of RTT's (few milliseconds) and bandwidth delay products (BDP) are low. In fact, Iperf in lightweight TCP mode gives very good accuracy over all the tested scenarios, firstly due to the adaptive and more aggressive way of sampling when multiple access require longer probing packets and sequences, and secondly due to the stability of current TCP versions used by Iperf. Besides, Iperf requires a few number of configuration parameters, converges rapidly and gives an acceptable overhead that can be weighted over the polling intervals associated to specific applications.

In the next chapter, we show the use of Iperf in lightweight TCP mode for a realistic use case for path/interface selection in the home network. Indeed, after benchmarking these different bandwidth probing tools, the intent has been testing the most appropriate one for potential applications.

# Iperf in Lightweight TCP Mode for Path Selection in Heterogeneous Home Networks

## 5.1. Introduction

As stated in the first and second chapters, two attractive applications for bandwidth measurement in home networks is path selection and load balancing. In this chapter, we present an experimental study, to evaluate a path selection protocol in heterogeneous home networks based on available-bandwidth metric. At this aim, we enhance an existing implementation based on the Inter-MAC concept introduced in the project OMEGA<sup>5</sup> and described in [62], [85]. The Inter-MAC is an abstraction layer between IP and MAC layers to handle heterogeneous home networks technologies. Inter-MAC has been used as the starting point for the IEEE P1905 working group about convergent home-networks. The new capability, we aim to show, is the introduction of the Iperf in lightweight TCP mode probing technique, studied in chapter 4, to estimate links available-bandwidth, combined with a path selection protocol. This way of estimating available bandwidth supports time-varying capacity links. Note that, path selection in hybrid-mesh home networks has been

---

<sup>5</sup> <http://www.ict-omega.eu/>



studied by the authors of [112]. However, the used protocol applied a metric based on control packets loss rate. The use of available bandwidth aims to improve the overall performance.

Our contributions include, the benchmarking, on a test bed, of Inter-MAC path selection protocol with Iperf with respect to the original Inter-MAC. We introduce external interference on the WiFi/PLC links between to Inter-MAC devices, while two IPTV flows transit the home network. Then, we show how Inter-MAC path selection coupled to Iperf in lightweight TCP mode is able to react to the link performance degradation and to avoid the QoS deterioration due to the default way to measure available bandwidth of the original Inter-MAC. The original Inter-MAC biases the available bandwidth measurement, when the capacity of WiFi or PLC changes. Instead, Inter-MAC coupled to Iperf is capable to detect the implicit variations of capacity and estimate accurately, fast and with low level of overhead, the available bandwidth.

These contributions led to an article accepted in the “7th IEEE International Workshop on Heterogeneous, Multi-Hop, Wireless and Mobile Networks HeterWMN, in Globecom 2012 conference” [86].

The next sections are organized as follows. Section 5.2 discusses our methodology used to benchmark Inter-MAC path selection with lightweight TCP probing. Section 5.3 describes our experimental test bed. Section 5.4 discusses the experimental results and we conclude in section 5.5.

## 5.2. Methodology

In this section, we discuss how we associate Inter-MAC path selection and Iperf in lightweight TCP mode.

### 5.2.1 *Inter-MAC Path Selection protocol*

The inter-MAC path selection protocol has been implemented in the OMEGA project as a software module under Linux OS. The path selection is based on a reactive approach, which means that routes are computed on demand. It is basically executed as follows:

In the beginning of the procedure, when a new flow arrives, a path request message (PREQ) is transmitted on each interface of the Inter-MAC source node. Intermediate nodes receiving a PREQ message rebroadcast it to their remaining adjacent nodes and the available bandwidth, correspondent to the links within the PREQ, is updated (we will

indicate later how the available bandwidth is calculated). Notice that, a sequence number prevents from loops caused by the flooding process. When the destination node receives a PREQ, it responds with a path reply message (PREP). The PREP message contains the information of end-to-end path including the intermediate nodes and the corresponding available bandwidth. The source node collects all incoming PREPs and selects the most suitable path: the one having the highest available bandwidth. Then, this path is activated by sending a path confirmation message (PCNF) along the selected path.

Thus, path selection relies on the node-by-node current available bandwidth. This information is collected and maintained by a functional block, which is called the monitoring engine. The monitoring engine computes the available bandwidth as the subtraction of the current link occupation (counting length and number of L2 frames during a measurement period) from the L2 capacity of the Inter-MAC node interfaces (which is a static value indicated in a configuration file).

After the selection of a suitable path is done, the path selection protocol verifies periodically (every second in our case) the state of the available bandwidth through the path. If an alternate path offers more available bandwidth than the active one, the flows traversing the active path can be reallocated thanks to path check messages. Furthermore, path verification takes into account LinkUp and LinkDown events. To avoid the “ping pong” effect of the path verification procedure, when the available bandwidth of two paths are quite near, we have introduced a rule in the verification procedure code. A flow is reallocated to the alternate path only if the available bandwidth in the alternate path is superior to the available bandwidth of the current path with a margin higher than the flow is actually requiring. This rule is easy to apply with CBR flows, which is the case for IPTV flows mentioned below.

### *5.2.1 Iperf in lightweight TCP mode*

To introduce our active probing technique (Iperf in lightweight TCP mode) into the Inter-MAC code, we have also modified the original Inter-MAC code. Instead of counting the frames on the Inter-MAC node interfaces, we probe the L4 available bandwidth by sending short TCP probing flows. To do this, we periodically execute the Iperf client in the transmission Inter-MAC interface and we activate the Iperf server in the reception Inter-MAC interface of half-duplex links. In full-duplex links, it is required to activate two Iperf clients and two Iperf servers in the correspondent transmission and reception ends of each

link (typically Ethernet full-duplex interfaces). In this section, we focus on monitoring half-duplex PLC and WiFi media links.

The probing process we used to estimate the L4 available bandwidth in the two-node Inter-MAC link (see Fig. 5.1) can be split into the following measurement time scales:

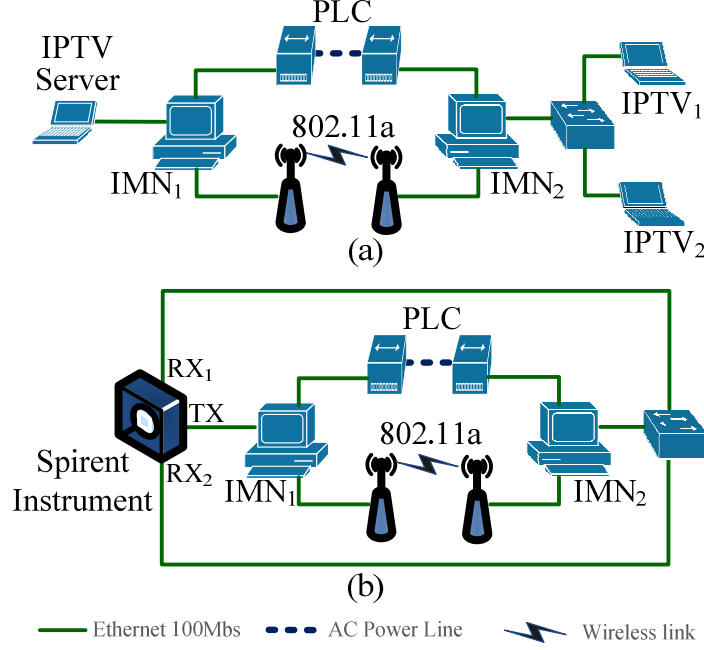
- We call the test interval  $T$  the whole duration of a test defined to characterize the behavior of the Inter-MAC path selection, given the presence of a cross traffic pattern and possible performance variability on every different path. In our experimental study,  $T$  will be in the order of several minutes. Our cross traffic of interest will be two IPTV flows, i.e.,  $F_1$  and  $F_2$  and the two considered paths  $P_1$  and  $P_2$  are listed in Table 5.1 and depicted in Fig. 5.1.
- The polling interval  $T_{po}$  is the period between the beginnings of two successive executions of the Iperf client, which defines the periodicity of available bandwidth ( $\mathcal{A}$ ) reporting values. We chose  $T_{po}$  sufficiently large in order to reduce the polling weighted overhead while preserving the accuracy of  $\mathcal{A}$  measurements.
- And finally, the probing interval  $T_{pr}$  is the period between the beginning and the end of the first and the last probing packets of a single Iperf client execution to obtain an available bandwidth measurement  $\mathcal{A}_i$ . We chose  $T_{pr}$  sufficiently short in order to reduce the probing weighted overhead while also preserving the accuracy of  $\mathcal{A}$  measurements.

From the measurement time scales, the inequality (4.1) of chapter 4 is applied.

### 5.3. Experimental Test bed

First, we present our test bed with Inter-MAC path selection coupled with Iperf. Then, we describe the tests we have done.

### 5.3.1 Test Bed Features



**Fig. 5.1. Path scenarios on testbed**

In Fig. 5.1 (a), we depict the topology of this test bed. Three laptops, called IPTV Server, and two clients, called IPTV<sub>1</sub> and IPTV<sub>2</sub>, are connected at the ends of the paths. Another pair of laptops is used as Inter-MAC Nodes, which we call IMN<sub>1</sub> and IMN<sub>2</sub>. All laptops have 2 Ghz Core Duo processors, 2 GB of RAM, Broadcom Nextream Gigabit Ethernet controllers and Linux Kernel 2.6.26-2-686. The laptop server also represents the home access gateway, which is supposed to guarantee the flow forwarding from the operator network. It sends two IPTV unicast flows,  $F_1$  and  $F_2$  to the IPTV<sub>1</sub> and IPTV<sub>2</sub> clients (note that the flows were recorded using Orange TV on a real access). To transmit and play the IPTV flows, we have used VLC media player in a server client mode. We have used “3Com Wireless 7760 11a/b/g PoE” Access Points and “dLAN 200 AV Wireless N Starter Kit+” PLC plugs. The topology of this use case represents a typical configuration to increase the overall throughput when using no new wires (WiFi and PLC links) in the home network.

The IMN<sub>1</sub> and IMN<sub>2</sub> have three 10/100/1000 Ethernet Interfaces and run the Inter-MAC modified software (including the interaction with Iperf measurements mentioned in section 5.2). As explained in section 5.2, for example, IMN<sub>1</sub> is capable to forward the IPTV flows from the IPTV server to the path with the largest available bandwidth, to reach the clients. To avoid the reroute “ping-pong effect”, we have configured the periodic path verification procedure, to reroute a flow, with a bandwidth envelope. Only if the alternate path has at least 8 Mbps ( $F_i$  rate + 1 Mbps) more than the current path, which is the bandwidth for

the largest flow that we will inject in our tests plus a bandwidth tolerance value, see section 5.2. For our experiments, this envelope was convenient; however more research is required to define a mechanism that can work correctly with any bandwidth requirement.

From equation (4.1), we have configured the Iperf probing and polling times for both paths as follows:  $T_{pr1} = 0.9s$  and  $T_{po1} = 7s$  for the PLC path and the  $T_{pr2} = 0.8s$  and  $T_{po2} = 7s$  for the WiFi path. The test time is  $T = 180s$ . Again, these values have been obtained empirically, because they worked fine to show the reactivity of the Inter-MAC path selection and Iperf in lightweight TCP mode, in this scenario. These values can be increased or reduced, depending on the accuracy, and intrusiveness of the Iperf measurements and the desired speed of reaction for the path selection and path verification procedures.

TABLE 5.1. MONITORED PATHS

Path 1 ( $P_1$ )	Path 2 ( $P_2$ )
IMN <sub>1</sub> →PLC→IMN <sub>2</sub>	IMN <sub>2</sub> →802.11a→IMN <sub>2</sub>

Table 5.1 shows the paths  $P_1$  and  $P_2$  that we are interested to monitor when the path selection procedure is activated. In fact  $P_1$  and  $P_2$  are seen by IMN<sub>1</sub> and IMN<sub>2</sub> as paths with a single logical link. This is because the three intermediate physical links (e.g., Ethernet-PLC-Ethernet) are bridged and no other node running the Inter-MAC software is present through  $P_1$  or  $P_2$ . Then, IMN<sub>1</sub> and IMN<sub>2</sub> will select one of the two links, for each video flow in function of the available bandwidth in  $P_1$  and  $P_2$ .

In Fig. 5.1 (b), we have replaced the IPTV server and clients with a Spirent TestCenter platform V.2.20 and chassis SPT-9000A. It permits us to simulate  $F_1$  and  $F_2$ , and analyze their statistics of bandwidth reception and packet loss in real time and in a synchronized manner.

Before making the experimental trials, we measured the UDP capacity as a reference, with no traffic. Actually when we test the UDP capacity, only the Inter-MAC signaling messages transit in the paths, which can represent some kbps. Therefore, we consider this signaling traffic relatively negligible in our scenarios. We call the UDP capacity  $Cr$  the bottleneck capacity of each path. At this aim, we use successive flows of UDP Iperf with an increasing rate, until the flows begin to have losses. The UDP throughput, when no losses occur or are very low, indicates the UDP  $Cr$ . As capacity on hybrid paths tends to change, in order to obtain a steady reference, we have minimized possible perturbations on Wi-Fi (we have disabled the adaptation rate feature) and PLC links. The nominal UDP capacities of  $P_1$  and

$P_2$  are respectively  $Cr_1 = 42$  Mbps and  $Cr_2 = 28$  Mbps. Notice that  $Cr_1$  (on the PLC link) is lower than its nominal capacity (near to 80Mbps) because we have connected in separated extension cords each of the two PLC Plugs and due to the interference of the electrical loads, found on the AC circuit of the lab, where we have done the measurements (servers, PCs, office devices and network devices). However, the obtained value (42 Mbps) was stable. During the tests, the load before the PLC extension cords remained constant (no other PCs or electronic devices were connected) to permit  $Cr_1$  being bounded.

For all the tests, we assumed that available bandwidth  $A_1$  and  $A_2$  of each respective path are related with their respective time-varying capacities and with the total traffic traversing each path. We call  $C_1$  and  $C_2$  the time-varying capacities of  $P_1$  and  $P_2$  respectively, where:  $C_1 \leq Cr_1$  and  $C_2 \leq Cr_2$ .  $C_1$  and  $C_2$  change when the paths are physically perturbed.

If no traffic traverses  $P_1$  and  $P_2$ , then,  $A_1 = C_1$  and  $A_2 = C_2$ . If both IPTV flows are simultaneously transmitted, as in our tests, we have:

$$A_1 = C_1 - (F_1 \text{ or } F_2 \text{ or } (F_1 + F_2)) \quad (5.1)$$

$$A_2 = C_2 - (F_1 \text{ or } F_2 \text{ or } (F_1 + F_2)) \quad (5.2)$$

### 5.3.2 Experimental Tests

In all experimental tests,  $F_1$  and  $F_2$  are always simultaneously transmitted.

Table 5.2 shows the two scenarios depending on whether or not we have introduced external perturbations on the WiFi and PLC links. These perturbations generate fluctuations in the capacity of the links to force path selection adaptability. To introduce perturbations in the PLC link, we plugged in the extension cords some additional commuted power supplies that interfere with the PLC plugs. To generate perturbations on the WiFi link, we manually reduced the data rate transmission.

We have defined two scenarios of tests. In scenario #1, we have not introduced perturbations on  $P_1$  and  $P_2$ . So, the capacities remain equals to the reference capacities, i.e.,  $C_1 = Cr_1$  and  $C_2 = Cr_2$ . In scenario #2, we have perturbed  $P_1$  and  $P_2$  reducing both,  $C_1$  and  $C_2$ , see Table 5.2. As mentioned above, to reduce  $C_1$  we have connected different commuted power supplies on each extension cords. It was done progressively and measuring the UDP capacity, without traffic, until attaining condition #2.

TABLE 5.2. TESTED PATH CAPACITY CONDITIONS

Scenario #	$C_1$ (Mbps)	$C_2$ (Mbps)
1	42	28
2	9	11

To generate  $F_1$  and  $F_2$  we have recorded two HD/IPTV flows in the IPTV Server laptop, obtained with an FTTH access and using the VLC media player. Both flows are MPEG4 CBR flows with L2 frames of constant 1356 bytes. The rates of the flows are:  $F_1 = 7$  Mbps and  $F_2 = 5$  Mbps. We send  $F_1$  and  $F_2$  from the IPTV Server using UDP protocol. For the bandwidth reception and packet loss analysis, we have configured the Spirent TestCenter to transmit two UDP flows with the same characteristics as the IPTV flows.

## 5.4. Results

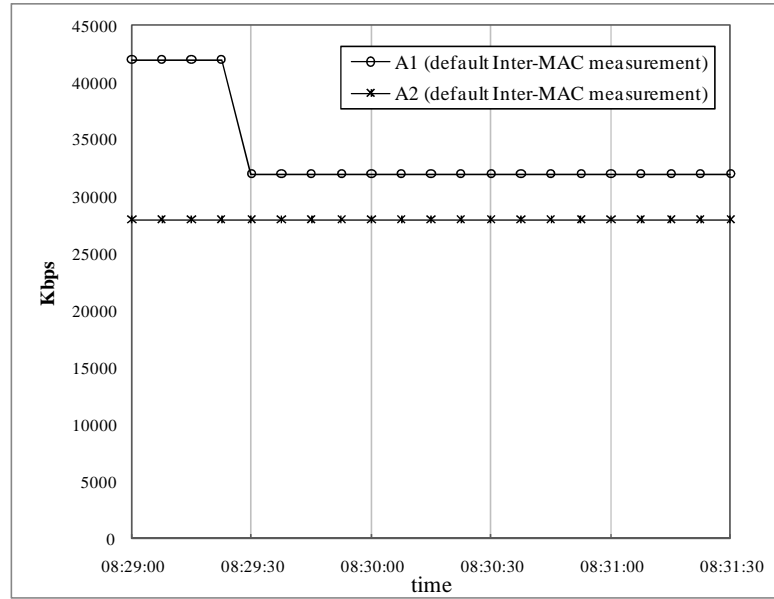
In this section, we firstly discuss the obtained results using the original Inter-MAC path selection that calculates available bandwidth with a static capacity reference as explained in section 5.3. Secondly, we discuss the results obtained with Inter-MAC path selection and Iperf in lightweight TCP mode measurements.

Using the original Inter-MAC path selection we have the following results. In the tested scenario #1 (without path perturbations, see Table 5.2), before transmitting any flow, we have  $A_1 = C_1 > A_2 = C_2$ . Then the path selection procedure will route  $F_1$  and  $F_2$  through  $P_1$ . After the flows begin to be transmitted, the path verification will compute available bandwidth as follows.  $F_1 + F_2 = 12$  Mbps and  $C_1 = 42$  Mbps, then from equation (5.1),  $A_1 = C_1 - (F_1 + F_2) = 30$  Mbps. Before transmitting any flow,  $A_2 = 28$  Mbps, then after transmitting the flows, as we have enough available bandwidth in  $P_1$ ,  $A_1$  remains superior to  $A_2$  and no rerouting is done. So,  $P_2$  does not transmit any flow, and  $A_2$  remains constant, i.e.,  $A_2 = 28$  Mbps. Then,  $F_1$  and  $F_2$  will show satisfactory and steady quality, i.e., no packet losses appear (if no other perturbations are introduced), see scenario #1 in Table 5.3.

TABLE 5.3. DEFAULT MEASUREMENTS (MBPS) AND INTER-MAC PATH SELECTION

#	$A_1$ (actual)	$A_1$ (default)	$A_2$ (actual)	$A_2$ (default)	$P_1$	$P_2$	RX State $F_1$	RX State $F_2$
1	30	30	28	28	$F_1 + F_2$	-	OK	OK
2	0	33	11	28	$F_1 + F_2$	-	BAD	BAD

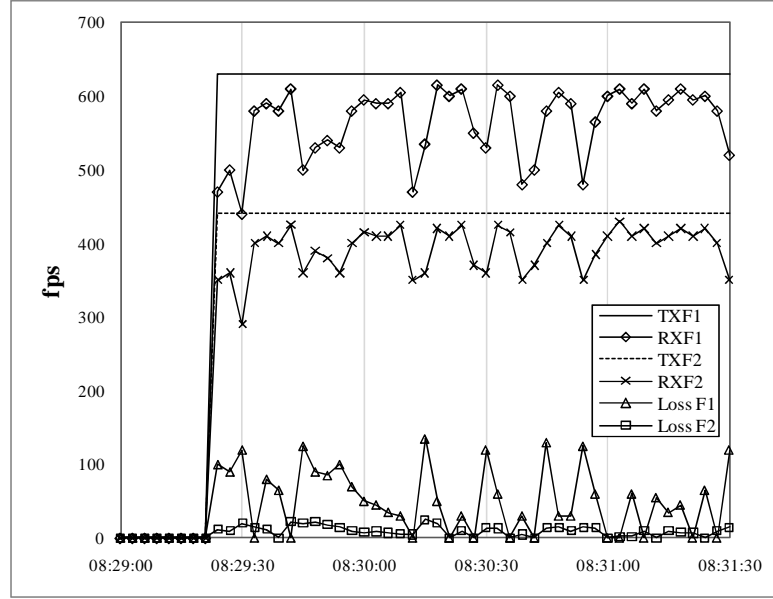
In scenario #2 of Table 5.2, we have introduced external physical perturbations on  $P_1$  and  $P_2$  when the flow is transmitted. Before transmitting any flow, we have  $A_1 = C_1 = 9$  Mbps  $< A_2 = C_2 = 11$  Mbps. It means that actually, neither  $P_1$  nor  $P_2$  supports two flows simultaneously (they would require at least 12 Mbps). However, the path selection and verification procedure cannot detect implicit capacity changes on PLC and WiFi links and will consider that  $A_1 = 33$  Mbps (42 Mbps – 9 Mbps)  $> A_2 = 28$  Mbps (the total rate in  $P_1$  is 9 Mbps instead of 12 Mbps, because the actual  $C_1$  limits the maximum throughput), Fig. 5.2. Then, both flows are transmitted through  $P_1$ .



**Fig. 5.2.  $A_1$  and  $A_2$  with default implementation (passive + static) measurement**

Thus, the path selection procedure will not react to reroute  $F_1$  or  $F_2$  through  $P_2$ . So,  $F_1$  and  $F_2$  will suffer quality degradation. It means that packet losses will strongly increase, because both flows will share the same path, see Fig. 5.3. In this case, since the default measurements (based on a static capacity and passive interfaces statistics) implemented in the original Inter-MAC software cannot reflect the real available bandwidth reduction, a capacity blind spot condition appears.





**Fig. 5.3. Spirent report of  $F_1$  and  $F_2$  bandwidth with default Inter-MAC path selection (passive + static measurements), scenario 2**

On the other hand, using the Inter-MAC path selection and Iperf in lightweight TCP mode measurements, we have the following results. In scenario #1 (Table 5.2), the active measurements are equivalent to the default measurements. The values are not exactly the same as the passive measurements (about 5 %). Iperf tends to lightly underestimate the available bandwidth as reported in [87], see scenario #1 of Table 5.4.

**TABLE 5.4. IPERF IN LIGHTWEIGHT TCP MODE (MBPS) MEASUREMENTS AND INTER-MAC PATH SELECTION**

#	$A_1$ (actual)	$A_1$ (active)	$A_2$ (actual)	$A_2$ (active)	$P_1$	$P_2$	RX State $F_1$	RX State $F_2$
1	30	28.5	28	27	$F_1 + F_2$	-	OK	OK
2	4 or 2	$\sim 4$ or $\sim 2$	6 or 4	$\sim 6$ or $\sim 4$	$F_1$ or $F_2$	$F_1$ or $F_2$	OK	OK

In spite of this relative small difference between the actual and TCP Iperf measurements and before transmitting any flow, the path selection procedure will select  $P_1$  to forward  $F_1$  and  $F_2$ . After the flows begin to be transmitted, the path verification  $A_1$  remains superior to  $A_2$  and no rerouting is required.  $F_1$  and  $F_2$  exhibit a good quality, see scenario #1 in Table 5.4.

In scenario #2, the Inter-MAC path selection and Iperf in lightweight TCP mode outperforms the default Inter-MAC software. When we perturb the PLC link, adding power supplies to the PLC cords, the PLC link capacity experiments a transitory

destructive effect and  $C_i \rightarrow 0$  and consequently also  $A_i \rightarrow 0$ , see Fig. 5.4 between 8:24:00 and 8:24:30. Since  $F_1$  and  $F_2$  were present in  $P_1$ , before  $P_1$  was perturbed, both flows are impacted and reallocated on  $P_2$ , see the decreasing pick of the curve A2 after 8:24:30. In this period, the packet losses have the largest rate, see Fig. 5.5

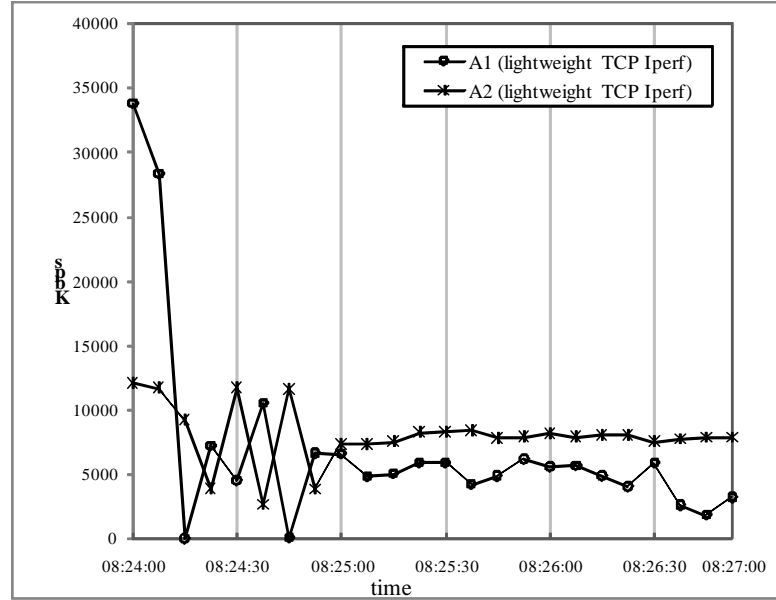


Fig. 5.4.  $A_1$  and  $A_2$  with Iperf in lightweight TCP mode

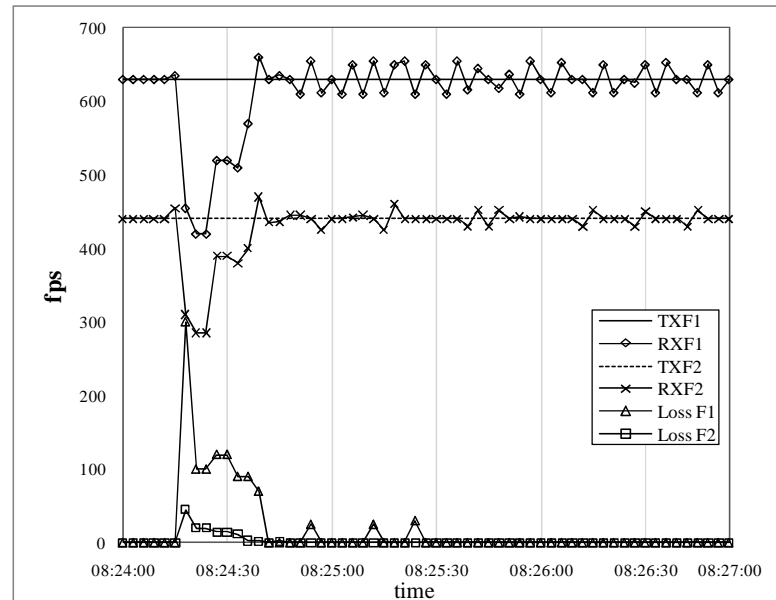


Fig. 5.5. Spirent report of  $F_1$  and  $F_2$  bandwidth with Inter-MAC path selection + Iperf in lightweight TCP mode measurement, scenario 2

After that, the PLC link recovers itself giving  $\mathcal{A}_1 \approx 9 \text{ Mbps} > \mathcal{A}_2$  (growing pick, in curve A1 after 8:24:30) and the two flows are again reallocated on  $P_1$ , second decreasing pick of curve A1 in Fig. 5.4. Finally, the system stabilizes and balances the load,  $F_1$  is placed on  $P_1$  and  $F_2$  in  $P_2$ , the packet losses are controlled and the QoS is satisfactory, as can be seen by comparing Fig. 5.3 and Fig. 5.5. As  $P_1$  was strongly perturbed (it works at about 1/10 of its nominal capacity), it exhibits more variability than  $P_2$ . See the evolution of the curve A1 with respect to the curve A2 in Fig. 5.4.

It is worth mentioning that the optimization of the algorithm to place a specific flow (for instance, the one with the largest bandwidth requirement or the one with most priority on the path with largest available bandwidth) has been left for further work.

## 5.5. Summary

In this chapter, we have shown that coupling Inter-MAC path selection with Iperf in lightweight TCP mode is really convenient to avoid performance blind spots in home network paths with time-varying-capacity-links (including WiFi and PLC links) and optimizing path utilization. We have showed how Iperf in lightweight TCP mode technique developed in chapter 4 and published in [37] can be integrated to monitor efficiently available bandwidth. The scenarios we have tested represent emergent use cases in home networks. For instance, commercial hybrid devices, such as WiFi extenders and “Qualcomm Atheros Hy-Fi,” that combine WiFi, PLC and Ethernet links. Notice that Iperf has been included in UPnP DM V2 bandwidth tests specification (without specifying the details of implementation), for instance.

Our test bed has focus at dual links with two Inter-MAC nodes, however it is possible to extend this path selection process to three or more nodes. We see a promising field of work for active probing and path selection. It can include a test bed that deals with more than two nodes, different types of traffic more than two flows and the optimization of the path selection and path verifications algorithms used in this chapter.

## Conclusions and Perspectives

The home network is the last and probably most fragile segment in the supply chain of network residential services. Moreover, home networks are becoming more and more complex, including various devices, different services and several connectivity-technologies (e.g. Ethernet, WiFi, PLC, etc.). For these reasons, it is very important integrating the correct QoS-mechanisms on the Control Plane and Management Plane of the home network. The ultimate goals are, on one hand, controlling the QoE commitments, which affect the end-user side, and on the other hand, optimizing profits and resources, on the network-provider side.

In this dissertation, we studied the interdependencies between network management processes, QoS mechanisms and performance metrics in the home network. More specifically, we have explored the available bandwidth monitoring in home networks and particularly, the available-bandwidth probing on the transport-layer. Firstly, we have showed different use cases. These use cases are related with the utilization of monitoring of available-bandwidth on the Management/Control/Data Planes, in state-of-the-art architectures.

Then, we have investigated the networking constraints that affect bandwidth probing in hybrid home networks. We demonstrate that bandwidth measurement on hybrid links is a tricky task. In fact, most existing probing-tools are not adapted to home networks. For instance, PLC and wireless-links are subject of interferences and thus have capacities that vary in the time. We have conducted a taxonomic study of a number of bandwidth-probing techniques and probing tools of the state-of-the-art. We have proposed a functional

framework called MPCFF. Based on these studies, we have chosen Iperf in TCP mode as an attractive tool to estimate available bandwidth.

Therefore, we have compared, on test-bed, Iperf in lightweight TCP mode with respect to other probing tools (Wbest, Pathchirp, Pathload and IGI/PTR). This comparison has included accuracy, convergence time and overhead. We have developed a homogeneous platform of tests and a methodology of available bandwidth probing in home networks. To the best of our knowledge, it is the first time that all the tested paths have been included to compare open-source probing-tools, and that Iperf is used in lightweight TCP mode for home networks.

The obtained results show that, despite the difference between TCP throughput and available bandwidth, Iperf in lightweight TCP mode can conveniently estimate the available bandwidth in home networks. Indeed, Iperf in lightweight TCP mode provides very good accuracy over all the tested scenarios, firstly due to the adaptive and more aggressive way of sampling bandwidth when multiple access require longer probing packets and sequences, and secondly due to the stability of current TCP versions used by Iperf. Besides, Iperf requires a few number of configuration parameters, converges rapidly and gives an acceptable overhead that can be weighted over the polling intervals associated to specific applications.

Finally, we have applied our proposed bandwidth measurement method to a path-selection protocol as a typical use case for future home networks. More precisely, we integrated Iperf in lightweight TCP mode into two nodes with redundant hybrid links (WiFi/PLC), running an existing path selection protocol (using the Inter-MAC software). We have showed that coupling Inter-MAC path selection with Iperf in lightweight TCP mode is really convenient to avoid performance blind-spots (in home network paths with WiFi/PLC links) and to optimize link utilization. We have showed how the Iperf in lightweight TCP mode efficiently estimates the available bandwidth for path selection purposes.

The scenarios we have tested represent emergent use-cases in home networks. For instance, nowadays, we find new commercial hybrid devices, such as “Qualcomm Atheros Hy-Fi” that combine WiFi, PLC and Ethernet links and that will support the new convergent digital home network standard IEEE P1905.1. Our test bed has focused on dual-links with two nodes; however it is possible to extend this path selection process to three or more nodes.

The evolution of available bandwidth probing is very important, because in the home networks, the traffic tends to extend itself in finer and larger scales of volume and duration, due to the constant emergence of new connectivity technologies and applications. So, we need more efficient, faster and less intrusive monitoring techniques. The ultimate goal is maintaining and increasing the end-to-end quality of experience of sensitive and critical applications.

We see a promising field of work for available bandwidth probing in hybrid home networks. Future work can include different studies, for example:

- Testing the Iperf in lightweight TCP mode including different types and volumes of traffic classes such as voice, IPTV and web-based traffic
- The adaptation of the server/client Iperf to the access gateway and the STB to verify the quality of the connectivity between both devices (transparently to the transmission technologies)
- The adaptation of Iperf in lightweight TCP mode to a monitoring server on the operator access premises (as illustrated in the TR-143 recommendation) for remote periodic monitoring. This can complement the usage of UDP echo plus (found in TR-143), for instance
- The adaptation of Iperf in lightweight TCP mode with an optimized version of TCP for sensor networks or where the TCP overhead represents a critical concern
- The study of the footprint associated to Iperf in lightweight TCP mode in terms of CPU utilization and memory occupation. In fact, it constitutes a necessary step, for deploying such a probing tool on real home network devices, including access gateway and STB, where resources (CPU, memory) are scarce with respect to PC environment
- The study of the interactions of Iperf in lightweight TCP mode with respect to security concerns (e.g., as a kind of flooding threat on the network)
- Analyze if the processes of Iperf in lightweight TCP mode (e.g. congestion avoidance/slow start) can be optimized for single links and if it is pertinent to adapt such a process to the MAC layer. This could simplify the utilization of the Iperf in lightweight TCP mode in a pure network centric manner.

Furthermore, other perspectives could include:

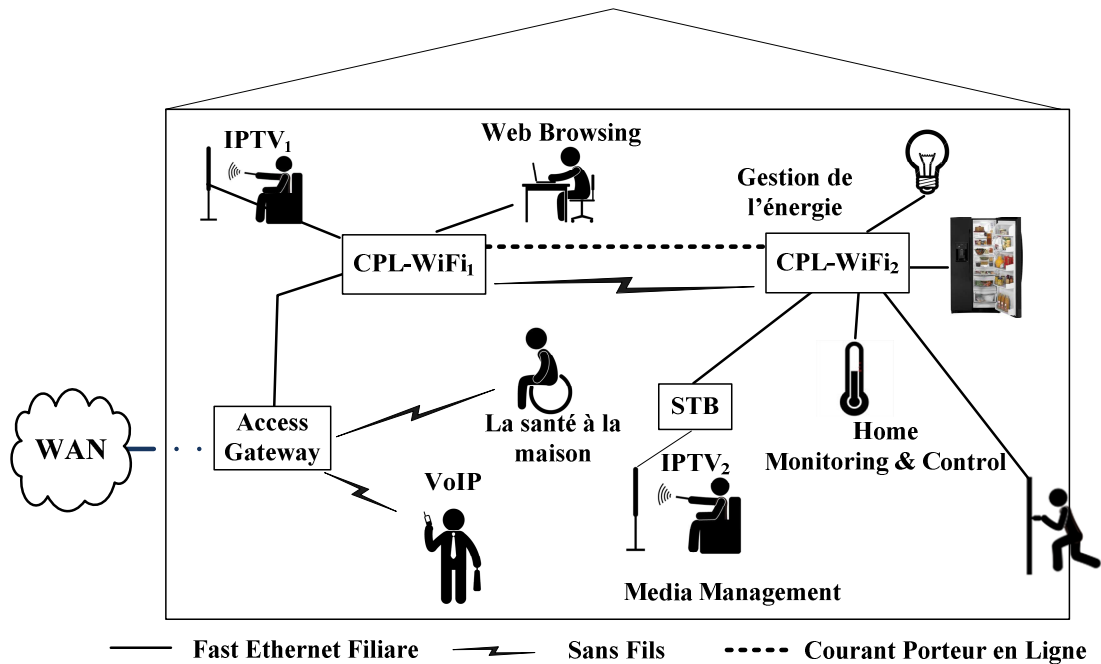
- The analytical study of the queuing phenomena associated with the process of probing considering separated or serialized PLC, WiFi, Ethernet and other wireless links, such as Bluetooth
- The study of the bandwidth probing impact on energy consumption in the home network and its interactions with sleeping mechanisms (on network interfaces in idle state)
- The study of the combination of active probing techniques and passive measurements to improve the overhead while maintaining the accuracy
- The extension of the conducted tests to other applications that exploit available bandwidth information such as service level checking (to verify SLA conformance), home network topology enrichment (by indicating the available bandwidth on each link of the network map), etc., in addition to path selection and load balancing that have been described in chapter 5.

# Appendix A: Résumé étendu en Français

## A.1. Contexte : les réseaux domestiques

Les réseaux domestiques connaissent une évolution importante et deviennent de plus en plus hétérogènes et complexes. Ils intègrent de nouveaux services, des équipements variés et des technologies de connectivité filaires et sans fil. De plus, les utilisateurs exigent une qualité de service d'un niveau très élevé pour de nombreuses nouvelles applications. Ainsi, les réseaux domestiques peuvent contenir des chemins hybrides, qui sont constitués par différents types de technologies de connectivité, tels que des liaisons filaires Ethernet, coaxiales, des câbles téléphoniques, des liens sur lignes électriques (CPL) et des liens sans fil à l'intérieur des PANs, BANs et WLANs [45]. Plusieurs équipements d'infrastructure réseau peuvent être utilisés pour connecter les terminaux (par exemple, les switch Ethernet, les plugs CPL, les WiFi extenders). En outre, nous pouvons trouver de nouveaux équipements réseau résilients tels que les «Qualcomm Atheros Hy-Fi », avec des liens WiFi/PLC en double attachement, lesquels supportent le draft IEEE 1905.1 (voir Fig. A.1). Les réseaux domestiques aussi transportent plusieurs types de flux tels que les flux IPTV, flux VoIP, les flux de transfert de fichiers, le streaming, les flux web divers, etc.





**Fig. A.1. Services dans le réseau domestique et technologies hybrides de connectivité**

La Fig. A.1 montre un réseau domestique avec la passerelle d'accès centrale typique (AGW) connectée au réseau WAN et avec plusieurs liens sérialisés (chemins) qui relient les terminaux et les nœuds du réseau.

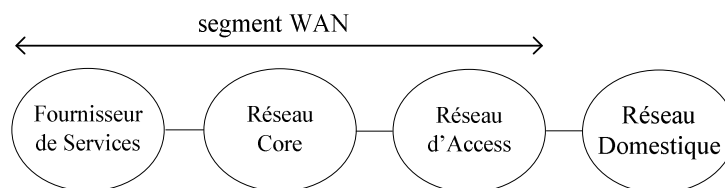
L'évolution des réseaux domestiques est étroitement liée à la réduction des coûts CPU des terminaux utilisateurs, l'accroissement continu de la vitesse des interfaces réseau [42] et la forte pénétration de l'accès résidentiel à large échelle [43].

Comme nous le détaillerons dans la section 2.2, l'évolution du réseau domestique remonte à plusieurs décennies. Même si, les réseaux domestiques centrées sur « l'accès gateway » commencent à être populaire dans les '2000s, on trouve différentes propositions des réseaux domestiques complexes basées sur ISDN / ATM dans les années 80 et 90. Dans les années 2000s, on a assisté à l'explosion d'Internet et « l'accès gateway » était le point de convergence pour l'accès Internet (pour remplacer les modems commutés « narrow-band/phone-line »), le service de télévision par câble et téléphone. L'accès gateway a été transformée d'un nœud intermédiaire couche 2 (pour connecter un ou un petit nombre de PCs à la maison, un terminal de télévision par câble et d'un téléphone) à un routeur très évolué. Aujourd'hui, l'accès gateway multiservice permet d'accéder à un certain nombre de périphériques IP tels que les tablettes électroniques, téléphones, consoles de jeux, ordinateurs portables, appareils ménagers, de téléviseurs HD/3D, NAS, etc.

Ainsi, ces tendances donnent un nouvel éventail de services, dans le marché résidentiel, tels que la gestion des médias de stockage, la gestion de l'énergie, la santé à domicile, le suivi et

le contrôle de la maison, etc. [1]. Toutefois, en raison de sa complexité, le réseau domestique est également très enclin à [2] la variabilité des performances, [53] et devient une source potentielle de dégradation de la qualité du service de bout en bout. C'est pourquoi, il y a un grand intérêt, du point de vue de l'opérateur, pour l'optimisation des mécanismes de QoS et des systèmes de gestion à distance et en local des réseaux domestiques. Cela devrait permettre de protéger les flux prioritaires, en maximisant l'utilisation des liens du réseau domestique et en facilitant la détection de pannes à distance et le rétablissement du service résidentiel.

D'autre part, les flux de services résidentiels sont touchés par les dégradations du réseau de bout en bout, qui sont cumulées à travers le segment WAN et à travers le réseau domestique. Ces contraintes de transport réseau, comme le délai, la variation du délai des paquets (Packet Delay Variation PDV, également appelée gigue), les fluctuations de la capacité et les pertes de paquets impactent les applications des utilisateurs finaux, comme s'expliqué en [38]. Pour ces raisons, l'opérateur du réseau est concerné par l'isolation et le contrôle des dégradations dues au réseau domestique mais aussi celles qui concernent le segment WAN. Par conséquent, la désignation des limites de performance, appelée dans le présent document, « les objectifs de qualité de service », (« targets de QoS ») nécessite la prise en compte des contraintes de transport réseau, à travers le segment WAN également. Ainsi, le segment WAN est constitué essentiellement de trois sous-segments qui incluent : le sous-segment du fournisseur de services (TV, Internet), le sous-segment de l'opérateur réseau et le sous-segment d'accès (par exemple, last mile), comme s'indiqué dans la Fig. A.2.



**Fig. A.2. Chemin du service résidentiel de bout en bout basé sur ITU-T/Y.2173 [51]**

Par exemple, un flux IPTV typique, qui est transmis à partir de la tête de réseau (liée au service IPTV), est impacté par la latence de propagation et les phénomènes des files d'attente, quand il traverse tous les nœuds et les technologies réseau à travers les segments WAN et du réseau domestique. Les paquets qui composent ce flux, arrivent au terminal IPTV de l'utilisateur final avec une distribution d'inter-arrivées particulière (délai et

variation du délai des paquets). Si la variation du délai des paquets devient plus large que les « targets QoS de but en but », la qualité d'expérience, perçue par l'utilisateur final, sera impactée (mauvaise voix / synchronisation de vidéo, temps de zapping élevé, etc).

Dans le même temps, les variations de performances sur le segment WAN et le réseau d'accès peuvent conduire à une réduction de la bande passante disponible de bout en bout. Ces variations peuvent correspondre à des changements de routage dans le réseau du fournisseur de service, à des changements externes environnementaux sur le lien d'accès et à des perturbations sur le réseau domestique telles que la dégradation sur les liens sans fil et CPL. Si le débit disponible est inférieur au débit requis (target), quand un flux est transmis, cela aura des effets destructeurs sur le flux transmis et cela provoquera des pertes de paquets. Dans le même temps, si ces pertes de paquets sont supérieures aux pertes de paquets cibles (target), elles seront traduites comme des artefacts vidéo divers et des interruptions sonores [44]. Cela produira une réduction de la qualité du service perçue ou dans des cas plus graves, cela conduira à une interruption complète du service.

Par conséquent, la sélection des métriques et des indicateurs de performance appropriés est critique pour le contrôle de la qualité de service, pour la supervision à distance et locale des liens et des chemins du réseau domestique. Aujourd'hui, nous trouvons un certain nombre d'architectures de QoS, paramétriques et de réservation de ressources [23], [51], [27], [28] qui permettent de contrôler des métriques d'état de lien/chemin en fonction de différents critères de performances ou de qualité de service cible, tels que le délai maximum, la variation de délai maximum, les pertes de paquets maximum, le débit requis...etc.

Dans cette thèse, on se focalise sur la mesure de la bande disponible sur les liens hybrides du réseau domestique. Elle peut être utilisée comme une métrique :

- D'état de lien/chemin pour déclencher des mécanismes QoS, par exemple, l'équilibrage de charge, la sélection de chemin et l'allocation des ressources
- Dans le cadre de Key Performance Indicators (KPIs) pour le diagnostic, la détection de défaillances de liens et l'évaluation des services.

## A.2. Problématique

Les réseaux domestiques sont sujets à une forte variabilité des performances [2], [53] et deviennent une source potentielle de dégradation de la qualité du service de bout en bout. Pour cette raison, il est nécessaire de protéger les flux prioritaires, en maximisant l'utilisation des chemins disponibles du réseau domestique, et en facilitant la détection à

distance des défaillances et le rétablissement du service résidentiel. Dans ce contexte, le contrôle de la qualité de liens/chemin est primordial dans les architectures de QoS et la gestion des réseaux domestiques émergents.

Plusieurs paramètres tels que le délai max, la variation de délai max, les pertes de paquets max, le débit requis peuvent intervenir dans les architectures de QoS. Ces architectures permettent de garantir la performance des applications des utilisateurs finaux dans le réseau domestique.

Toutefois, la question qui se pose est : quel type de métrique de performance devons-nous utiliser pour caractériser la qualité de lien/chemin dans le réseau domestique?

Même s'il n'y a pas une seule mesure globale qui permette de simplifier la mise en œuvre de ces architectures de QoS, il est souhaitable de disposer d'un ensemble réduit de paramètres (une seule métrique éventuellement) qui, compte tenu d'un certain nombre d'hypothèses pratiques dans le réseau domestique, puisse avoir une représentativité des performances suffisante. C'est le cas de la bande passante disponible au niveau de couche de transport qui peut être utilisée comme une métrique d'état de lien/chemin de déclenchement de mécanismes de QoS, par exemple, l'équilibrage de charge, la sélection de chemin et l'allocation des ressources ou dans le cadre d'un Key Performance Indicator (KPI) pour le diagnostic, la détection de défaillances de liens et l'évaluation des services.

La raison pour laquelle nous mettons l'accent sur la bande passante disponible est que cette mesure, par rapport à d'autres mesures, est une mesure directe de la charge supplémentaire qu'un lien ou un chemin peut supporter, avant que ce lien ne devienne saturé. D'autres paramètres, tels que le délai, la variation de délai, le taux de perte de paquets et la capacité peuvent seulement déterminer si un chemin est déjà saturé, ce qui réduit le degré de prévention pour éviter la potentielle dégradation du service [52]. D'autre part, la bande passante disponible peut simplifier la surveillance (monitoring) des services du réseau domestique, faciliter la détection des défaillances et apparaît comme une métrique d'état de lien récurrente dans différentes architectures de QoS des réseaux domestiques.

Par conséquent, on doit apporter une réponse à la question suivante : comment mesurer la bande passante disponible dans les réseaux domestiques hybrides?

Il existe essentiellement deux approches : la mesure déterministe (réalisée de façon passive) et la mesure à base de sondes actives également appelée « probing actif ». La mesure passive déterministe obtient directement les états courants du trafic à partir des interfaces physiques, avec un niveau faible ou sans aucun traitement statistique. Tandis que le probing actif exige l'envoi des flux de sonde pour estimer le comportement des files

d'attente, les conditions de transmission et le traitement des paquets sur les liens ou chemins mesurés. Le probing actif nécessite le traitement statistique des échantillons capturés pour en déduire une métrique réseau spécifique.

Les mesures déterministes ont une précision élevée sur des liens de capacité constante, ils ont un degré d'intrusivité faible ou nul. Néanmoins, les mesures déterministes souffrent d'imprécision sur les liens/chemins avec des capacités qui varient dans le temps. Elles sont tributaires des mécanismes de couches basses et ne reflètent pas le comportement de la couche applicative d'une manière assez précise.

Ainsi, cette thèse propose l'utilisation du probing de la bande passante disponible au niveau de la couche transport en tant qu'indicateur de performance et en tant qu'outil de mesure de l'état de liens/chemin. Nous mettons l'accent sur la mesure de la bande passante par probing actif.

Les principales contraintes réseau qui sont à prendre en compte pour la problématique de probing de la bande passante disponible dans le réseau domestique sont :

- a) La diversité des échelles de temps
- b) La particularité des liens de type WiFi et PLC qui sont de nature partagée et dont la capacité peut varier dans le temps rend la tâche de mesure de bande passante complexe
- c) L'outil de probing à proposer doit être simple à calibrer et transparent vis à vis des couches inférieures
- d) L'outil de probing à proposer doit être suffisamment précis, rapide (faible temps de convergence), peu intrusif et stable sur divers systèmes d'exploitation et conditions réseau
- e) Le besoin de faciliter les déploiements de bout en bout et repartis
- f) La possibilité d'appliquer le probing avec des classes de trafic

Donc, le problème est que la plupart des outils d'état de l'art ont été développés pour des chemins d'Internet. Ils sont inexacts, lents et pas assez stables (sensibles à la résolution d'horloge, aux glissements de synchronisation, aux interruptions du système, au nombre, débit et à la taille des paquets de probing) [101], [106], [96], [100], [102], en particulier dans le contexte des réseaux domestiques hybrides. Par conséquent, nous devons trouver un moyen adapté de mesurer la bande passante disponible compte tenu de toutes ces contraintes réseau.

## A.3. Objectifs de la thèse et contributions

Compte tenu des limites des outils actuels de probing et les contraintes du processus de mesure dans les réseaux domestiques, cette thèse propose d'utiliser Iperf dans un mode allégé pour obtenir le throughput de flux TCP non-disruptifs et estimer la bande passante disponible. Selon les tests d'évaluation de la performance effectués, nous soutenons l'hypothèse qu'Iperf est un outil très efficace pour être utilisé dans les réseaux domestiques car :

- a) Iperf est un outil bien connu et beaucoup plus stable que les autres outils
- b) Iperf a été développé pour différents systèmes d'exploitation tels que Windows, MacOS, Linux et Adroid
- c) Iperf, basé sur le protocole TCP, peut être configuré pour ne pas être disruptif, lorsque les applications temps réel traversent le chemin mesuré
- d) Iperf peut être utilisé en tant qu'outil pour l'estimation de la bande passante disponible de bout à bout des services résidentiels, eu égard à sa précision
- e) Les réseaux domestiques ont de courtes périodes « slow-start » et une performance TCP prévisible et stable, par opposition aux réseaux WANs dont le produit délai bande passante est plus grand (delay bandwidth products)
- f) On peut utiliser le protocole TCP Iperf en mode TCP léger, avec de courts intervalles de probing, plutôt que son utilisation ordinaire à force brute sur les chemins WAN [94]

Notre première contribution consiste en l'analyse comparative, sur test bed, d'Iperf en mode TCP par rapport à différents outils de probing, l'IGI/PTR, Wbest, Pahtload et Patchirp, en termes de temps de convergence, de précision et d'intrusivité (overhead) associés.

Deuxièmement, nous proposons et reproduisons, avec notre interface prototype de test automatique en Java (que nous avons développé avec la contribution d'un stagiaire), une méthode de mesure tenant compte des échelles de temps de probing, polling et test pour pondérer l'overhead d'Iperf, en fonction des applications du réseau domestique.

Ces deux contributions ont été publiées dans un article intitulé: "Available bandwidth probing in hybrid home networks as part of the Local Metropolitan Area Networks" (LANMAN), 2011 18th IEEE Workshop on, 2011.

Troisièmement, nous intégrons notre technique proposée pour mesurer la bande passante disponible dans un cas d'utilisation intéressant relatif à la sélection de chemin et équilibrage de charge dans le réseau domestique. Plus précisément, nous intégrons (avec l'aide d'un stagiaire) Iperf en mode TCP sur deux nœuds connectés avec des liens redondants hybrides (WiFi / CPL) exécutant un protocole de sélection chemin existant (développé avec le concept Inter MAC au cours du projet Européen OMEGA)

Quatrièmement, nous réalisons un benchmark, sur test bed expérimental, du protocole de sélection de chemin Inter-MAC et Iperf par rapport à la version originale d'Inter-MAC. Nous induisons des perturbations extérieures sur les liens WiFi / CPL, tandis que deux flux IPTV transitent sur le réseau domestique. Ensuite, nous montrons comment la sélection de chemin et Iperf en mode TCP sont capables de réagir à la dégradation du lien et d'éviter les chutes de performance, à cause de la méthode par défaut de la version Inter-MAC originale. La version originale d'Inter-MAC donne des mesures inexactes de la bande passante disponible, lorsque des changements de la capacité sur les liens WiFi ou CPL apparaissent. Au lieu de cela, Inter-MAC et Iperf sont capables de détecter les variations implicites de la capacité et donnent une estimation précise et rapide de la bande passante disponible au niveau de la couche transport, avec un faible taux d'overhead.

Ces dernières contributions ont été présentées dans le « The 7th IEEE International Workshop on Heterogeneous, Multi-Hop, Wireless and Mobile Networks - MENS 2012 » dans un article intitulé : “Available Bandwidth Probing for Path Selection in Heterogeneous Home Networks”

Enfin, nous avons fait une étude sur les principaux mécanismes de QoS trouvés dans des architectures pour les réseaux domestiques et de leurs relations avec la bande passante disponible. Nous explorons également les interdépendances qui existent entre la définition de QoS / QoE « targets », les processus de gestion réseau et les mécanismes de qualité de service à travers la connexion d'accès résidentielle (y compris le réseau domestique). Nous soulignons l'importance de la bande passante disponible, comme une métrique « d'état de lien » ou « état de chemin » et comme un indicateur clé de performance pour surveiller et optimiser les ressources du réseau domestique. Nous montrons également différents cas d'utilisation où la bande passante disponible pourrait être exploitée pour déclencher les mécanismes de QoS tels que le contrôle d'admission, la sélection de chemin et l'équilibrage de charge dans les réseaux domestiques.

Ces dernières contributions ont été soumises dans un article intitulé: « Available Bandwidth Monitoring and QoS Mechanisms in Home Networks: a Survey » dans le journal « IEEE Communications Surveys & Tutorials »

## A.4. Organisation de la thèse

Le reste de cette thèse est composé comme suit. Dans le deuxième chapitre, nous expliquons pourquoi le probing de la bande passante disponible est un outil fondamental dans les architectures de QoS pour les nouveaux réseaux domestiques hybrides. Nous explorons également différents cas d'usage, où le probing de la bande passante disponible peut être mis en œuvre.

Dans le troisième chapitre, nous étudions les contraintes réseau qui affectent le probing de la bande passante dans les réseaux domestiques hybrides. Nous faisons une étude taxonomique d'un certain nombre de techniques et d'outils cités dans l'état de l'art pour le probing de la bande passante disponible. Nous appliquons un framework fonctionnel appelé « MPCFF » qu'on propose afin de décortiquer les différentes techniques et de comparer leur structure. Sur la base de ces études, nous avons choisi Iperf en mode TCP comme un outil très intéressant pour estimer la bande passante disponible.

Dans le quatrième chapitre, nous effectuons une évaluation de performance sur un test bed expérimental, d'Iperf en mode TCP par rapport aux outils existants : Wbest, Pathchirp, Pathload et IGI / PTR. Nous développons une plate-forme de tests automatique en Java homogène et une méthodologie de probing pour la bande passante disponible dans les réseaux domestiques. Nous montrons qu'Iperf en mode TCP est meilleur en termes de précision et de vitesse de convergence.

Enfin, nous intégrons Iperf en mode TCP dans un protocole de sélection de chemin existant. On effectue des tests sur deux nœuds avec des liens redondants hybrides (WiFi / CPL) exécutant le logiciel Inter-MAC (issu du projet européen Omega). Nous montrons comment Iperf en mode TCP améliore la sélection de chemin Inter MAC en éliminant les dégradations de performance induites par la variation de capacité des liens que la mesure passive déterministe de la bande passante ne détecte pas.

## A.5. Conclusion de la thèse

Le réseau domestique est le dernier segment et probablement le plus fragile de la chaîne d'approvisionnement des services résidentiels. De plus, il se caractérise par une complexité



grandissante avec une augmentation du nombre d'équipements et l'utilisation de liens hétérogènes pour les connecter. C'est pourquoi, il est très important d'intégrer des processus de gestion réseau et des mécanismes de QoS dans les architectures du réseau domestique, afin de protéger les classes de trafic. L'objectif final étant de satisfaire des engagements QoE, du côté de l'utilisateur et d'optimiser les profits et les ressources, du côté de l'opérateur (fournisseur de services).

Dans cette thèse, nous avons étudié les interdépendances entre les processus de gestion réseau, les mécanismes de qualité de service et les mesures de performance dans le réseau domestique. Plus précisément, nous avons exploré la supervision de la bande passante disponible et plus particulièrement, le probing de la bande passante disponible au niveau de la couche transport dans les réseaux domestiques hybrides (incluant différents types de liens, par exemple : Ethernet, WiFi, PLC). Nous avons montré différents cas d'usage où la supervision de la bande passante disponible est présente dans les mécanismes de QoS et dans les processus de gestion réseau, en relation avec le management plane, control plane et data plane dans des architectures de QoS cités dans l'état de l'art.

Nous avons étudié les contraintes réseau qui affectent le probing de la bande passante dans les réseaux domestiques hybrides. Nous avons fait une étude taxonomique d'un certain nombre de techniques et d'outils cités dans l'état de l'art pour le probing de la bande passante disponible avec notre framework fonctionnel appelé MPCFF. Sur la base de ces études, nous avons choisi Iperf en mode TCP comme un outil intéressant pour estimer la bande passante disponible. Par conséquent, nous avons comparé, sur un test bed, Iperf en mode TCP avec plusieurs outils de mesure existant : Wbest, Pathchirp, Pathload et IGI / PTR.

Nous avons développé une plate-forme homogène de tests automatique en Java et une méthodologie de probing de la bande passante disponible dans les réseaux domestiques. A notre connaissance, c'est la première fois que l'ensemble des configurations de chemin testés ont été inclus afin d'évaluer tous ces outils open-source de probing de la bande passante disponible, et qu'Iperf est utilisé en mode TCP léger (courte connexion de moins d'une seconde) pour les réseaux domestiques. Les résultats obtenus montrent que, en dépit de la petite différence entre le throughput TCP et la bande passante disponible, Iperf en mode TCP léger peut estimer efficacement la bande passante disponible dans les réseaux domestiques. En effet, Iperf en mode TCP léger donne une très bonne précision sur tous les scénarios testés, d'abord en raison de sa manière adaptative et plus agressive de l'échantillonnage de la bande passante sur les différents liens et d'autre part en raison de la

stabilité des versions TCP actuelles utilisées par Iperf. Par ailleurs, Iperf nécessite un faible nombre de paramètres de configuration, converge rapidement et donne un overhead acceptable qui peut être pondéré sur les intervalles « de polling » (exécutions successives de plusieurs mesures) associés à des applications spécifiques.

Enfin, nous avons utilisé Iperf en mode TCP léger afin de fournir une métrique basée sur la bande passante disponible dans le cadre d'un protocole de sélection de chemin dans le réseau domestique. L'idée étant de montrer, par le biais d'un use case pertinent, comment la mesure de la bande passante, avec la méthode que nous proposons, peut être exploitée. Plus concrètement, nous avons conduit des tests avec deux nœuds connectés par des liens redondants hybrides (Wifi / CPL, en double attachement) exécutant le logiciel Inter-MAC (issu d'un projet collaboratif appelé Omega). Nous avons montré que le couplage du mécanisme de sélection de chemin d'inter-MAC avec le probing Iperf TCP est vraiment efficace pour éviter les dégradations de performance (dans les liens et chemins des réseaux domestiques avec des capacités variables dans le temps comme le WiFi et PLC) et optimiser l'utilisation des ces liens. Nous avons montré comment la méthode d'Iperf en mode TCP léger permet de déterminer la bande passante disponible à des fins de sélection de chemin.

Les scénarios que nous avons testés représentent des cas d'utilisations émergents dans les réseaux domestiques. Par exemple, on trouve de nouveaux équipements commerciaux hybrides, tels que «Qualcomm Atheros Hy-Fi» qui combinent WiFi, CPL et Ethernet (et qui devraient supporter prochainement le nouveau standard sur le 'convergent digital Home network' dénommé IEEE P1905.1).

Ainsi, de façon concise, la démarche suivie au cours de la thèse peut se résumer avec les principales étapes que sont :

- Etude bibliographique du contexte du réseau domestique, des mécanismes de QoS et des processus de gestion réseau associés.
- Etude bibliographique et analyse comparative des méthodes de mesure de la bande passante disponible et des outils associés.
- Evaluation expérimentale de plusieurs outils de mesure de la bande passante disponible.
- Proposition d'utilisation de l'outil Iperf en mode TCP léger (courte connexion) afin de réaliser la mesure de bande passante dans le réseau domestique hybride.

- Application d'Iperf en mode TCP léger dans le cadre d'un use case de sélection de chemin sur une plateforme expérimentale.

L'évolution du probing actif de la bande passante disponible est très importante, parce que dans les réseaux domestiques, le trafic a une tendance à occuper des échelles plus fines et plus importants de volume et de durée, en raison de l'apparition constante de nouvelles technologies de connectivité et des applications. Donc, nous avons besoin de la redéfinition continue des architectures QoS, des processus de gestion réseau et des techniques de supervision réseau, qui doivent être plus efficaces, plus rapides, moins intrusives et plus économiques en termes de leur déploiement et leur consommation d'énergie. Où, le but ultime est de maintenir et d'améliorer la qualité d'expérience des applications sensibles et critiques du réseau domestique.

Nous voyons un nombre de perspectives pour le probing actif de la bande passante disponible, qui peuvent inclure :

- Tester Iperf TCP avec différents types et flux et de classes de trafic telles que la voix, IPTV et trafic Web diverse
- L'adaptation d'Iperf (serveur / client) sur la passerelle d'accès et la STB pour vérifier (de manière transparente aux technologies de transmission) la qualité de la connectivité entre les deux « devices »
- L'adaptation d'Iperf TCP à un serveur de supervision de qualité de liens (dans les locaux de l'opérateur réseau comme illustré dans la recommandation TR-143) pour la surveillance périodique à distance. Ce qui pourrait permettre de comparer ou de compléter l'utilisation de UDP echo plus (recommandé par TR-143), par exemple
- L'adaptation du protocole Iperf TCP léger avec une version optimisée de TCP pour les réseaux de capteurs ou lorsque l'overhead TCP représente une contrainte majeure
- L'étude de la possibilité de déploiement de l'outil Iperf TCP sur de vrais équipements du réseau domestique (notamment sur les passerelles résidentielles et la STB). Pour ce faire, il serait important d'évaluer la consommation CPU et l'occupation mémoire associées
- L'étude des interactions d'Iperf TCP avec les problèmes de sécurité (par exemple, vu comme une sorte de menace de « flooding » réseau)
- Analyser si les processus de probing Iperf TCP (e.g. congestion avoidance /slow start) peuvent être optimisés pour des liens point à point et s'il est pertinent

d'adapter de tels processus à la couche MAC. Cela pourrait simplifier l'utilisation des principes d'Iperf TCP léger, pour le probing d'une manière centrée sur le réseau. Comme les principes de Iperf TCP léger sont agnostiques aux différentes technologies de transmission filaires et sans fil, ces principes pourraient être adaptés à tout appareil du réseau domestique en tant que une fonctionnalité généralisé de surveillance réseau.

Par ailleurs, d'autres perspectives pourraient inclure:

- L'étude de l'impact du probing de la bande passante sur la consommation d'énergie dans le réseau domestique et de ses interactions avec les mécanismes de mise en veille
- L'étude de la combinaison des techniques actives de probing et des mesures passives de la bande passante pour diminuer l'intrusivité tout en conservant la précision
- L'extension des tests effectués vers d'autres applications qui exploitent les informations de bande passante disponible (pour vérifier la conformité vis-à-vis des SLA), pour enrichir la topologie réseau (en indiquant la bande passante disponible sur chaque lien), etc, outre que la sélection de chemin et d'équilibrage de charge qui ont été décrits dans le chapitre 5.



# Appendix B: The Java Test Automation Interface

In this Appendix, we describe the Java test automated interface mentioned in chapter 4.

In order to optimize the experimental performance evaluation, it was necessary to conduct automatic tests. The idea was to launch multiple tests using different cross traffic loads for a given path configuration using a suited probing tool. It was also important to accelerate the collect of results phase and the display of pertinent parameters (such as available bandwidth and intrusiveness) in real time. At this aim, a test automation interface was developed using Java language, thanks to the contribution of an internee. Next figure shows a screenshot of the test interface using Iperf in TCP mode, to estimate the available bandwidth on an Ethernet path with a fixed capacity of 100Mbps. In the figure, we have three graphs corresponding respectively to the available bandwidth (3), the probing interval (4) and probing overhead (5). These values are updated instantly by the Livegraph tool.

In addition to the representation of the results in graphical form, the application displays the instantaneous values in (1) and average values in (2). The test interval corresponding to the cross traffic is 600 seconds. The test duration was set to 605 seconds to ensure that all cross traffic flows have been sent. We notice a periodic decrease in available bandwidth. This decrease is due to the increase of the load of cross traffic. Indeed, the available bandwidth varies inversely according to cross traffic load, see details in [33].

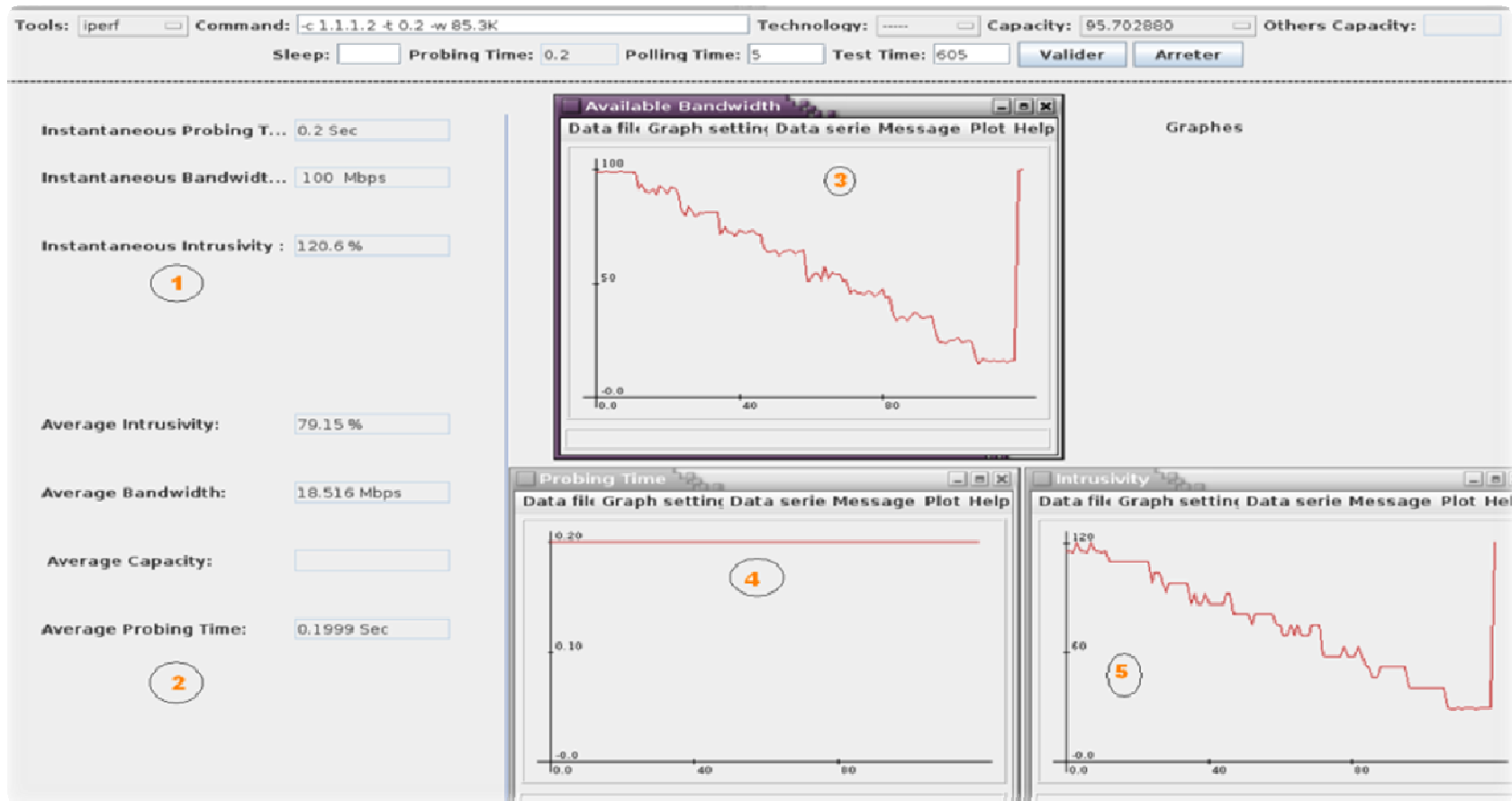


Fig. B.1. Screenshot of the Java test interface

# Appendix C: Configuring Iperf in Lightweight Mode for Available Bandwidth Probing

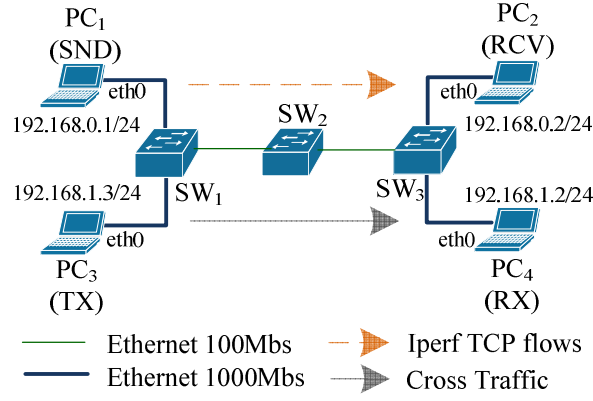
To illustrate how we can configure Iperf in lightweight TCP mode we use a simple test bed to be setup in an easy way. Our aim is to show, how the Iperf in Lightweight TCP mode can be relatively rapidly deployed and configured. The key to configure TCP Iperf in lightweight TCP mode is running Iperf during short intervals of time. These periods are called probing, intervals that have to be executed, leaving time spaces (called polling intervals). This allows reducing the Iperf intrusiveness: the shorter the probing intervals and the longer the polling intervals, the lower the intrusiveness. Then, to set Iperf, we show a script that permits setup Iperf in our test bed.

In next subsections we firstly explain the test bed setup, secondly we recall the probing methodology, thirdly we show the configuration on the PCs and finally we show some results of this procedure on the test bed.

## C.1. Test Bed Setup

To show how we configure Iperf in lightweight TCP mode, we use the example in Fig.C.1.





**Fig. C.1. Simple example of using Iperf in lightweight TCP**

In Fig. C.1., all laptops have Linux Kernel 2.6.26-2-686 and Ethernet interfaces (10/100/1000). We use Iperf 2.0.4 in the four PCs. It is worth mentioning that Iperf is also stable on different Linux distributions and other operating systems including Unix, Windows, MacOS and Android [33], [32], [38]. We have also installed IPTraf on the four PCs, to monitor the bit rate on each PC interface. Using the IPTraf option “Detailed Interfaces Statistics” and choosing a specific interface (e.g. eth0), we can see the L2 incoming and outgoing data rates on the chosen interface.

PC<sub>1</sub> (SND) and PC<sub>2</sub> (RCV) are the TCP-Iperf client and the TCP-Iperf server, respectively. To estimate the available bandwidth, SND sends TCP-Iperf flows that adapt themselves to the available bandwidth on the path (PC<sub>1</sub>→SW<sub>1</sub>→SW<sub>2</sub>→SW<sub>3</sub>→PC<sub>2</sub>), until the flow reaches the RCV. TX and RX are the cross-traffic transmitter and the cross-traffic receiver, respectively.

TX and RX are the cross-traffic terminals and can be a server/client pair or a point-to-point pair. TX and RX are connected through the network path (PC<sub>1</sub>→SW<sub>1</sub>→SW<sub>2</sub>→SW<sub>3</sub>→PC<sub>2</sub>) to transmit and receive application packets. To simplify the example, we emulate cross-traffic flows by using a UDP-Iperf client on PC<sub>3</sub> (TX) and a UDP-Iperf server on PC<sub>4</sub> (RX). So, TX can send flows of constant bit rate (CBR) of different speeds on the home network.

In our example, we physically isolate the probing terminals from the cross-traffic terminals. However, it is possible to execute both, the probing client and cross-traffic client on a single PC, and the probing server and cross-traffic server on another PC.

SW<sub>1</sub> and SW<sub>3</sub> are Ethernet switches of 10/100/1000 Mbps (Netgear GS108T v2) while SW<sub>2</sub> is a 10/100 Mbps switch (3Com 3CFSU08-ME). As the SW<sub>2</sub> interfaces are limited to 100 Mbps, they are the bottleneck.

To depict in real time the measured available-bandwidth, we use the Java LiveGraph 1.1.4 application [36].

In next sections, we assume that we have IP connectivity between  $PC_1$  and  $PC_2$  and between  $PC_3$  and  $PC_4$ . Then, we can do ping between  $PC_1$  and  $PC_2$  and also between  $PC_3$  and  $PC_4$ . We consider that all other traffic between the PCs is negligible, before and during the test.

## C.2. Methodology

As explained in [87], [86] (in sections 4.2. and 5.2. of this dissertation), we probe the L4 available-bandwidth by sending short TCP-flows. To do this, we periodically execute the Iperf client in  $PC_1$  and we activate the Iperf server in  $PC_2$  on the sense ( $PC_1 \rightarrow PC_2$ ). If we want to measure available bandwidth in the opposite direction, it is required to activate another TCP-Iperf client and another-TCP Iperf server, but in the other sense ( $PC_1 \leftarrow PC_2$ ). In this example, we focus on one direction measurement. The probing process that we used to estimate the L4 available bandwidth considers three time scales:

The probing interval  $T_{pr}$  is the period between the beginning and the end of the first and the last probing-packets of a single execution of iperf, to obtain an available-bandwidth measurement  $A_r$ . The probing interval is the shortest Iperf-execution time that allows getting accurate measurements. If the probing interval is too short, the variability of the measurements increases (for repetitive trials, in the same system conditions and traffic load). If the probing interval is too large it will be more intrusive and impact the behavior of the cross traffic. As explained in [87] and section 4, we use empirical probing intervals between 0.8s and 1s.

The polling interval  $T_{po}$  is the period between the beginnings of two successive executions of the Iperf client, which defines the periodicity of available bandwidth of the reporting values ( $A$ ). Since we aim reducing the presence of the probing flows in the network (intrusiveness), we must shift each Iperf execution on the time. Then, we need to choose longer spaces of time “s” between Iperf executions. The polling interval is the addition of the probing interval  $T_{pr}$  plus the time space  $s$ :

$$T_{po} = T_{pr} + s \quad (C.1)$$

As explained in [87], [86] and chapters 4 and 5 we can use polling intervals of several seconds, e.g. 5s to 7s.

The test interval  $T$  is the whole duration of a test defined to characterize the available bandwidth, given the presence of cross traffic.  $T$  will be in the order of several minutes. As mentioned, the probing intervals are chosen empirically. The summation of all polling intervals will give the whole test interval. The inequality (4.1) is applied here. Then, the number of executions  $N$  multiplied by the polling interval gives the test interval:

$$T = N(T_{pr} + s) \quad (C.2)$$

### C.3. PCs Configuration

#### *Configuration of PC<sub>1</sub>*

Software requirements: Iperf, IPTraf, LiveGraph version 1.1.3(4).

The following script *tcp\_iperf\_clnt.sh* is executed on the PC<sub>1</sub>:

```

1      > iperf_clnt.txt
2      read -p "How many executions?" N
3      java -jar LiveGraph.1.14.Complete.jar -dfs Demo-DataFileSettings.lgdfs -gs &
4      for((i=0; i<N; i++)); do
5          iperf -c 192.168.0.2 -t 0.8 -w 85.3K -y C -x C >> iperf_clnt.txt
6          sleep 4
7      done

```

The script *tcp\_iperf\_clnt.sh* allows executing  $N$  times the TCP Iperf client as follows:

- Line 1 cleans the measurement to the file "*iperf\_clnt.txt*" which is the file where the Iperf measurements are written down
- Line 2 asks and reads the number of executions required to probe the available bandwidth
- Line 3 executes the LiveGraph java application. LiveGraph requires some configuration files to be set. In this example we only use the file *Demo-DataFileSettings.lgdfs* [36]. Then, we edit the line `<entry key="DataFile">iperf_clnt.txt</entry>` of *DataFileSettings.lgdfs*.
- Line 4 executes a loop of  $N$  executions.  $N$  multiplied by the addition of the probing interval and the time spaces (sleep in this example) give the test interval, as defined in equation C.2.
- Line 5, executes Iperf. The TCP Iperf client (PC<sub>1</sub>) connects to the server with the IP address *192.168.0.2* (PC<sub>2</sub>). *-t 0.8* is the probing interval  $T_{pr}$  in seconds. *-w 85.3K* allows fixing the window size, which is the default value used by the Iperf server. *-y*

$C -x C$  allows filtering only the specific available bandwidth values without another informational text. Since, we do not specify the number of the used TCP port; the client will utilize the default Iperf port, 5001. Finally, each available bandwidth value is copied to the text file *iperf\_clnt.txt*, which is read by the LiveGraph application, see Fig. C.3 and an example of *iperf\_clnt.txt* at section C.4.

- Line 6 defines the time space  $s$  between two Iperf executions and delays  $s = 4s$  the beginning of a new Iperf execution to define the polling interval:

$$T_{po}: 0.8s + s = 4.8s$$

It means that every 4.8s we will have an available bandwidth measurement.

- Line 7 finishes the loop execution

The script *tcp\_iperf\_clnt.sh* can require different permissions to be executed, for instance:

```
sudo chmod 755 tcp_iperf_clnt.sh
```

To have a deterministic measurement reference of the eth0 interface, before executing *tcp\_iperf\_clnt.sh*, we execute IPTraf with root permissions. Using the IPTraf option “Detailed Interfaces Statistics” and choosing a specific interface (in this case eth0), we can see the L2 incoming and outgoing data-rates.

```
sudo iptraf
```

We can select and see the traffic activity on “eth0” before and after Iperf is activated. When Iperf is executed, the eth0 counters will exhibit the traffic activity, see Fig. C.1.

To execute *tcp\_iperf\_clnt.sh*, we activate the TCP Iperf server on PC<sub>2</sub>. After executing the TCP-Iperf server on PC<sub>2</sub>, we execute *tcp\_iperf\_clnt.sh* as follows:

```
./tcp_iperf_clnt.sh
How many executions?
100
```

If *tcp\_iperf\_clnt.sh* script is executed, without cross traffic going from PC<sub>3</sub> to PC<sub>4</sub>, the available bandwidth (showed in the live graph during 8 minutes in this example) will be near to 95 Mbps, which is the L4 available-bandwidth on the bottleneck links of 100Mbps. When the script is executed, if we choose  $N = 100$ ,  $T_{pr} = 0.8s$  and  $s = 4s$ , from the equation (C.2), the test interval is:

$$T = 100(0.8 + 4) = 480s = 8 \text{ min}$$

### *Configuration of PC<sub>2</sub>*

Software requirements: Iperf, IPTraf.

The next commands have to be run on PC<sub>2</sub>:

```
sudo iptraf
iperf -s
```

In the first line, IPTraf is executed to see the “eth0” traffic activity. In the second line Iperf is executed as a server in the TCP mode. Since, we do not specify the used TCP-Iperf port, the server chooses the port 5001 that will match with the port used by the Iperf client.

### *Configuration of PC<sub>3</sub>*

Software requirements: Iperf, IPTraf, Linux System Monitoring.

In PC<sub>3</sub>, we activate the IPTraf application to see the evolution of the counters on eth0.

```
sudo iptraf
```

PC<sub>3</sub> is the client that emulates the application flows (UDP cross-traffic) that arrive to the server (PC<sub>4</sub>). Clearly, to execute the client, we need before activating the UDP server on PC<sub>4</sub>.

Firstly, we intend to fill the bottleneck link with UDP flows. As the bottleneck link has a capacity of 100 Mbps (as see in the Fig. C.1), we send a UDP flow of 100 Mbps with duration of some seconds from PC<sub>3</sub> to PC<sub>4</sub>, as follows:

```
iperf -u -c 192.168.1.2 -t 10 -b 100M
```

As we can see, the IPTraf counters on output interface of PC<sub>3</sub> and on the input interface of PC<sub>4</sub> will increase near to 96Mbps, which corresponds to the Layer-2 capacity. After 10 seconds (that the UDP flow has traversed the path), Iperf gives the UDP throughput between PC<sub>3</sub> and PC<sub>4</sub>, which is about 95Mbps (due to the encapsulations, L3 and L4).

To emulate a cross traffic pattern that increases and decreases in the time we can use the equation (4.3). We translate this equation to the following script that we call *udp\_iperf\_clnt.sh*. This script will control the increments of UDP flows.

```

1      read -p "How many cycles ? " N
2      for((i=0; i<N; i++)); do
3          iperf -u -c 192.168.1.2 -t 20 -b 0
4          iperf -u -c 192.168.1.2 -t 20 -b 10M
5          iperf -u -c 192.168.1.2 -t 20 -b 20M
6          iperf -u -c 192.168.1.2 -t 20 -b 30M
7          iperf -u -c 192.168.1.2 -t 20 -b 40M
8          iperf -u -c 192.168.1.2 -t 20 -b 50M
9          iperf -u -c 192.168.1.2 -t 20 -b 60M
10         iperf -u -c 192.168.1.2 -t 20 -b 70M
11         iperf -u -c 192.168.1.2 -t 20 -b 80M
12         iperf -u -c 192.168.1.2 -t 20 -b 90M
13         iperf -u -c 192.168.1.2 -t 20 -b 80M
14         iperf -u -c 192.168.1.2 -t 20 -b 70M
15         iperf -u -c 192.168.1.2 -t 20 -b 60M
16         iperf -u -c 192.168.1.2 -t 20 -b 50M
17         iperf -u -c 192.168.1.2 -t 20 -b 40M
18         iperf -u -c 192.168.1.2 -t 20 -b 30M
19         iperf -u -c 192.168.1.2 -t 20 -b 20M
20         iperf -u -c 192.168.1.2 -t 20 -b 10M
21         iperf -u -c 192.168.1.2 -t 20 -b 0
22     done

```

The script *udp\_iperf\_clnt.sh* allows executing *C* times the UDP Iperf client as follows:

- Line 1 asks and reads the number of cycles to generate increments and decrements of cross traffic
- Line 2 executes a loop of *C* executions
- Line 1 to 21, execute Iperf. The UDP Iperf client (PC<sub>3</sub>) connects to the server with the IP address 192.168.1.2 (PC<sub>4</sub>). *-u* indicates the use of Iperf in UDP mode. *-t 20* is the step duration in seconds of a UDP-flow with a rate *-b* growing and decreasing, in steps of 10 Mbps. Since, we do not specify the number of the used TCP-port; the client will utilize the Iperf default port 5001.
- Line 22 finishes the loop execution.

Also, the *udp\_iperf\_clnt.sh* file can require different permission before execution, for instance:

```
sudo chmod 755 tcp_iperf_clnt.sh
```

Finally, to see a graphical evolution of the emulated cross-traffic we can activate the Linux “System Monitoring” from Linux “System Tools”. See the network activity graphs, as in Fig. C.2 (“Historique du trafic réseau”).

## Configuration of PC<sub>4</sub>

Software requirements: Iperf, IPTraf.

The next commands have to be run on PC<sub>4</sub>:

```
sudo iptraf
iperf -u -s
```

In the first line, IPTraf is executed to see the traffic activity on eth0. In the second line, Iperf is executed as a server in UDP mode (`-u -s` arguments). Since, we do not specify the used UDP-Iperf port the server chooses the port 5001 that will match with the port used by the UDP-Iperf client.

## C.4. Results

Fig. C.2 shows the cross traffic, which is generated from PC<sub>3</sub>, using Iperf in UDP mode.

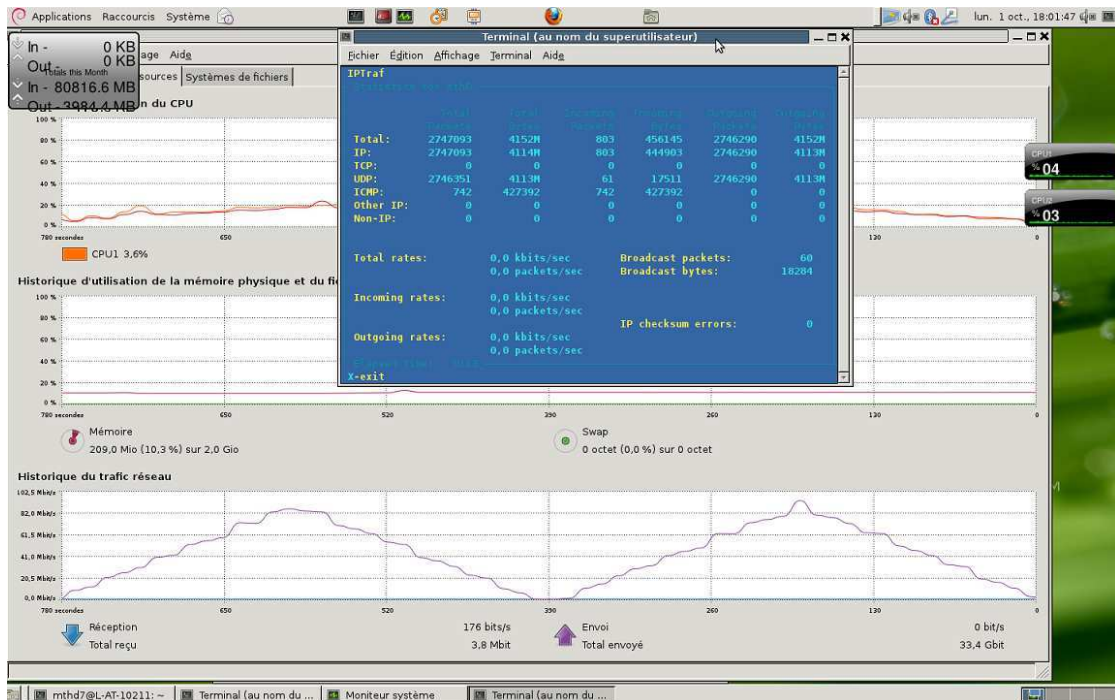


Fig. C.2. Cross traffic injected from PC<sub>3</sub> using Iperf in UDP mode

As plotted in Fig. C.2, using the Linux system monitoring, we can see the step increments of 10 Mbps each 20 sec. We have also used IPTraf to verify the counters in eth0 (blue window).

Fig. C.3 shows the curves of available bandwidth given by LiveGraph plotted from the file *iperf\_clnt.txt* on PC<sub>1</sub>

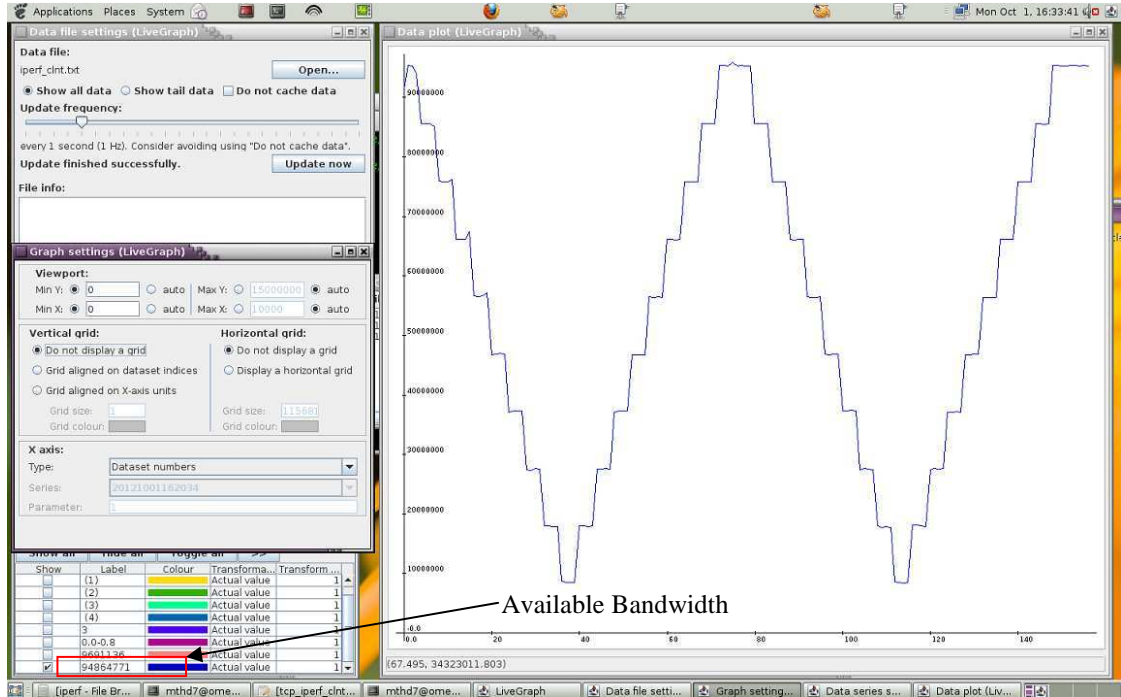


Fig. C.3. Available bandwidth measurements from PC<sub>1</sub> with Iperf in lightweight TCP

As indicated with the red square, we chose this field on the LiveGraph window because it corresponds to the column of available bandwidth in the file *iperf\_clnt.txt*. The file *iperf\_clnt.txt* is:

```
20121001162034,,,,,3,0.0-0.8,9691136,94864771
20121001162039,,,,,3,0.0-0.8,9207808,91465717
20121001162044,,,,,3,0.0-0.8,9601024,95277319
20121001162049,,,,,3,0.0-0.8,9625600,95246859
20121001162054,,,,,3,0.0-0.8,9396224,93723713
20121001162059,,,,,3,0.0-0.8,8593408,85541551
20121001162104,,,,,3,0.0-0.8,8585216,85538447
20121001162109,,,,,3,0.0-0.8,8593408,85533356
20121001162114,,,,,3,0.0-0.8,8536064,84969356
20121001162119,,,,,3,0.0-0.8,7634944,75843563
20121001162124,,,,,3,0.0-0.8,7634944,75791331
20121001162129,,,,,3,0.0-0.8,7643136,75798912
20121001162134,,,,,3,0.0-0.8,7643136,76095588
20121001162139,,,,,3,0.0-0.8,6643712,66118182
20121001162144,,,,,3,0.0-0.8,6627328,66151561
```

Every line gives a time stamp (the first column), the flow ID given by Iperf (3 in this example), the probing interval (0.0 to 0.8 s), the transmitted bandwidth and the measured bandwidth (available bandwidth), in bold fonts.



The curves of figures C2 and C3 are complementary, because when we inject cross traffic, the available bandwidth reduces. As we can see, the available bandwidth decreases before the 40<sup>th</sup> sample in steps of 10 Mbps. The available bandwidth (maximal on layer-4) is about 95Mbps, when no cross-traffic is injected, which corresponds to a TCP segment of 1460 B, from equation 2.2. The minimum available bandwidth is about 8.3 Mbps, when 90 Mbps of cross traffic is sent from PC<sub>3</sub> to PC<sub>4</sub>. From the trial 40<sup>th</sup>, the available bandwidth begins to increase. From the 80 trial to the 160 trial, the curve is symmetric with respect to the first 80 trials.

# Bibliography

- [1] Arkko, J. Home Networking Architecture for IPv6 [draft]-ietf-homenet-arch-01, 2012.
- [2] Armitage, G. Quality of service in IP networks: foundations for a multi-service Internet. Macmillan Publishing Co., Inc., 2000.
- [3] Bhatti, G.; Yige Wang; Jinyun Zhang; Matsubara, F. A. K. M. S. A. S. QoS harmonization for home networks, Broadband Multimedia Systems and Broadcasting, 2010 IEEE International Symposium.
- [4] Blake, S. RFC 2475 An Architecture for Differentiated Services
- [5] Bolot, Jean-Chrysotome “End-to-end packet delay and loss behavior in the internet” SIGCOMM '93: Conference proceedings on Communications architectures, protocols and applications.
- [6] Braden, R. RFC 1633 Integrated Services in the Internet Architecture: an Overview 1994
- [7] Bredel, Michael and Fidler, Markus, “A measurement study of bandwidth estimation in IEEE 802.11g wireless LANs using the DCF”, Proceedings of the 7th international IFIP-TC6 networking conference on, 2008
- [8] Carter, Robert and Crovella, Mark “Measuring Bottleneck Link Speed in Packet-Switched Networks” Boston, MA, USA 1996.
- [9] Clark, D. D.; Partridge, C.; Ramming, J. C. & Wroclawski, J. T. A knowledge plane for the internet, conference on Applications, technologies, architectures, and protocols for computer communications, ACM, 2003, 3-10.
- [10] Clark, D. The design philosophy of the DARPA internet protocols Symposium proceedings on Communications architectures and protocols, ACM, 1988, 106-114.
- [11] D.E. Meddour, A. Kortebi, R. Boutaba. "Mesh Based Broadband Home Network Solution: Setup and Experiments", ICC 2010
- [12] Delphinanto, A.; Koonen, T. & den Hartog, F. “End-to-end available bandwidth probing in heterogeneous IP home networks” Consumer Communications and Networking Conference (CCNC), 2011 IEEE.
- [13] DiCioccio, L.; Teixeira, R. & Rosenberg, C. Impact of home networks on end-to-end performance: controlled experiments. Proceedings of the 2010 ACM SIGCOMM workshop on Home networks.
- [14] DLNA, Digital Living Network Alliance Home Networked Device Interoperability Guidelines Version: 1.0 2004

- [15] Dobson, S.; Sterritt, R.; Nixon, P. & Hinchey, M. Fulfilling the Vision of Autonomic Computing Computer, 2010, 43, 35 -41.
- [16] Dovrolis, C.; Ramanathan, P. & Moore, D. What do packet dispersion techniques measure? INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 2001, 2, 905-914 vol.2
- [17] Edwards, C. Bringing broadband home Engineering Technology, 2011, 6, 74 -75
- [18] Fanshawe, D. Traffic Analysis Yields Efficient Home Network Architecture Consumer Electronics, IEEE Transactions on, 1987, CE-33, 375 -382
- [19] Gorod, A.; Sauser, B. & Boardman, J. System-of-Systems Engineering Management: A Review of Modern History and a Path Forward Systems Journal, IEEE, 2008, 2, 484 -499
- [20] Goyal, M., Ramakrishnan, K. & chi Feng, W. Achieving faster failure detection in OSPF networks Communications, 2003. ICC '03. IEEE International Conference on, 2003, 1, 296 - 300 vol.1
- [21] Greaves, D. & Bradbury, R. Warren: a low-cost ATM home area network Network, IEEE, 1998, 12, 44 -56
- [22] Grinter, R. E.; Edwards, W. K.; Newman, M. W. & Ducheneaut, N. The work to make a home network work European Conference on Computer Supported Coop. Work, Springer-Verlag NY., 2005.
- [23] Guerrero, C. D. & Labrador, M. A. On the applicability of available bandwidth estimation techniques and tools Comput. Commun., Butterworth-Heinemann, 2010, 33, 11-22
- [24] Gutzwiller, F. W.; Francis, J. E.; Howell, E. K. & Kruesi, W. R. Homenet: A Control Network for Consumer Applications Consumer Electronics, IEEE Transactions on, 1983, CE-29, 297 -304
- [25] Ha, S. & Rhee, I. "Taming the elephants: New TCP slow start" Computer Networks, 2011, 55, 2092 – 2110.
- [26] Handel, R. Evolution of ISDN towards broadband ISDN Network, IEEE, 1989, 3, 7 - 13.
- [27] Hanover, G. Networking the intelligent home Spectrum, IEEE, 1989, 26, 48 -49
- [28] Hatori, M.; Mokuno, K.; Lida, K.; Ochiai, R. & Horie, T. Home Informatization and Standardization of Home Bus Consumer Electronics, IEEE Transactions on, 1986, CE-32, 542 -549
- [29] Higginbottom, Gary "Performance evaluation of communication networks" Norwood, MA, 1988 Artech House.
- [30] Home Gateway Initiative HGI-GD013-R2 HGI Guideline Document: QoS white paper 2009.
- [31] [http://grouper.ieee.org/groups/1905/1/Scope\\_Purpose.htm](http://grouper.ieee.org/groups/1905/1/Scope_Purpose.htm)

- [32] <http://iptraf.seul.org/>
- [33] <http://sourceforge.net/projects/Iperf/>.
- [34] [http://www.broadbandhomecentral.com/report/backissues/Report0410\\_pictures.html#DLNA](http://www.broadbandhomecentral.com/report/backissues/Report0410_pictures.html#DLNA)
- [35] <http://www.caida.org/tools/taxonomy/performance.xml>
- [36] <http://www.live-graph.org/>
- [37] <http://www.wireshark.org>
- [38] <https://play.google.com/store/apps/details?id=com.magicandroidapps.iperf&hl=en>
- [39] IEEE P1905.1 IEEE [draft] Standard for a Convergent Digital Home Network for Heterogeneous Technologies IEEE P1905.1/D02, December 2011, 2011, 1 -80.
- [40] ITU-T G.1080 Quality of experience requirements for IPTV services 2008
- [41] ITU-T ITU-T Y.2121 Requirements for the support of flow-state-aware transport technology in NGN Recommendation ITU-T 2008
- [42] ITU-T Rec. E.800 Definitions of terms related to quality of service, 2009.
- [43] ITU-T Rec. G.1011 Reference guide to quality of experience assessment methodologies, 2010.
- [44] ITU-T Rec. G.9960 Unified high-speed wire-line based home networking transceivers – System architecture and physical layer specification, 2010.
- [45] ITU-T Rec. G.9970 Generic home network transport architecture, 2009.
- [46] ITU-T Rec. P.10/G.100 Amendment 1 SERIES P Vocabulary and effects of transmission parameters on customer opinion of transmission quality, 2007.
- [47] ITU-T Rec. X.641 Information technology – Quality of Service: Framework. 1997.
- [48] ITU-T Rec. Y.1001 IP framework - A framework for convergence of telecommunications network and IP network technologies, 2000.
- [49] ITU-T Rec. Y.1543 Measurements in IP networks for inter-domain performance assessment, 2007.
- [50] ITU-T Rec. Y.2011 General principles and general reference model for Next Generation Networks, 2004.
- [51] ITU-T Rec. Y.2173 Management of performance measurement for NGN, 2008.
- [52] ITU-T Rec.Y.1542 Framework for achieving end-to-end IP performance objectives, 2010.
- [53] ITU-T Y.1541 Network performance objectives for IP-based services, 2006.
- [54] Jacobson, V. “Congestion avoidance and control” SIGCOMM '88: Symposium proceedings on Communications architectures and protocols.
- [55] Jacobson, V. pathchar. 1997. URL <ftp://ftp.ee.lbl.gov/pathchar/>.
- [56] Jain, M. End-to-end available bandwidth estimation and its applications Georgia Institute of Technology, 2007

- [57] Johnsson, A. and Bjorkman, M. and Melander, B. "An Analysis of Active End-to-end Bandwidth Measurements in Wireless Networks", End-to-End Monitoring Techniques and Services, 2006 4th IEEE/IFIP Workshop on
- [58] K. Agrawala and D. Sanghi. Network dynamics: An experimental study of the internet. In Proceedings of GLOBECOM, December 1992.
- [59] Kephart, J. & Chess, D. The vision of autonomic computing Computer, 2003, 36, 41 – 50.
- [60] Keshav, Srinivasan "The Packet Pair Flow Control Protocol" University of California, Berkeley, 1991.
- [61] Kortebi, A.; Oueslati, S. & Roberts, J. Cross-protect: implicit service differentiation and admission control High Performance Switching and Routing, 2004. HPSR. 2004 Workshop on, 2004, 56 - 60
- [62] Kraemer, R.; Brzozowski, M. & Nowak, S. Reliable architecture for heterogeneous home-networks: The OMEGA I-MAC approach Telecommunication in Modern Satellite Cable and Broadcasting Services (TELSIKS), 2011 10th International Conference on, 2011, 279 - 284
- [63] Kwok, T. Residential broadband Internet services and applications requirements Communications Magazine, IEEE, 1997, 35, 76 -83.
- [64] Labit, Y. and Owezarski, P. and Larrieu, N. "Evaluation of active measurement tools for bandwidth estimation in real environment", End-to-End Monitoring Techniques and Services, 2005. Workshop on.
- [65] Lai, Kevin and Baker, Mary "Measuring link bandwidths using a deterministic model of packet delay" SIGCOMM '00, Stockholm, Sweden. Copyright 2000 ACM.
- [66] Lakshman, T. & Madhow, U. "The performance of TCP/IP for networks with high bandwidth-delay products and random loss Networking" IEEE/ACM Transactions on, 1997, 5, 336 -350.
- [67] Lao, L.; Dovrolis, C. & Sanadidi, M. "The probe gap model can underestimate the available bandwidth of multihop paths", article 2006.
- [68] Lazar, A. A. & Pacifici, G. Management and Control of Resources in Broadband Networks with Quality of Service Guarantees in Workshop on Distributed Systems: Operations and Management, 1991.
- [69] Leiner, B.; Cerf, V.; Clark, D.; Kahn, R.; Kleinrock, L.; Lynch, D.; Postel, J.; Roberts, L. & S.Wolff Past and Future History of the Internet Communications of the ACM, 1997, 40, 102-108
- [70] Li, M. and Claypool, M. and Kinicki, R., "Packet Dispersion in IEEE 802.11 Wireless Networks", Local Computer Networks, Proceedings 2006 31st IEEE Conference on

- [71] Li, Mingzhe and Claypool, M. and Kinicki, R. “WBest: A bandwidth estimation tool for IEEE 802.11 wireless networks” Proc. 33rd IEEE Conference on Local Computer Networks LCN 2008.
- [72] Licklider, J. & Vezza, A. Applications of information networks Proceedings of the IEEE, 1978, 66, 1330 – 1346.
- [73] Litchman, W. The Future of Digital Communications Communications Systems, IEEE Transactions on, 1963, 11, 149 -158
- [74] Lu, H.-L. & Faynberg, I. An architectural framework for support of quality of service in packet networks Communications Magazine, IEEE, 2003, 41, 98 – 105.
- [75] M. Jain and C. Dovrolis, “Pathload: A measurement tool for end-to-end available bandwidth,” in Proc. Passive Active Measurements, Fort Collins, CO, Mar. 2002.
- [76] M. Jain and C. Dovrolis. “Ten fallacies and pitfalls on end-to-end available bandwidth estimation”. In Proceedings (IMC'03), Oct. 2003.
- [77] Melander, B. and Bjorkman, M. and Gunningberg, P. “A new end-to-end probing and analysis method for estimating bandwidth bottlenecks” Global Telecommunications Conference, 2000. GLOBECOM '00. IEEE.
- [78] Michaut, F. & Lepage, F. Application-oriented network metrology: metrics and active measurement tools Communications Surveys Tutorials, IEEE, 2005, 7, 2 – 24
- [79] Minzer, S. Broadband ISDN and asynchronous transfer mode (ATM) Communications Magazine, IEEE, 1989, 27, 17 -24
- [80] Morton, A. Reporting Metrics: Different Points of View [draft]-ietf-ippm-reporting-metrics-07, 2012
- [81] Mosel, H.-J. Home Communication Systems Selected Areas in Communications, IEEE Journal on, 1986, 4, 633 - 639
- [82] MR-239 Broadband Forum Value Proposition for Connected Home Marketing Rep. 2011.
- [83] Murty, Rohan; Padhye, J. C. R. R. C. A. W. M. Characterizing the End-to-End Performance of Indoor Powerline Networks Harvard University, 2008
- [84] N. Hu and P. Steenkiste. “Evaluation and Characterization of Available Bandwidth Techniques”. IEEE JSAC Special Issue in Internet and WWW Measurement, Mapping, and Modeling, 2003.
- [85] Nowak, S.; Schaefer, F.; Brzozowski, M.; Kraemer, R. & Kays, R. Towards a convergent digital home network infrastructure Consumer Electronics, IEEE Transactions on, 2011
- [86] Olvera-Irigoyen, O.; Kortebi, A. & Toutain, L. Available Bandwidth Probing for Path Selection in Heterogeneous Home Networks, GC'12 Workshop: The 7th IEEE International Workshop on Heterogeneous, Multi-Hop, Wireless and Mobile Networks (GC'12 Workshop - HeterWMN 2012)

- [87] Olvera-Irigoyen, O.; Kortebi, A.; Toutain, L. & Ros, D. Available bandwidth probing in hybrid home networks Local Metropolitan Area Networks (LANMAN), 2011 18th IEEE Workshop on, 2011.
- [88] Ou, C.; Sahasrabudde, L.; Zhu, K.; Martel, C. & Mukherjee, B. Survivable virtual concatenation for data over SONET/SDH in optical transport networks Networking, IEEE/ACM Transactions on, 2006, 14, 218 - 231
- [89] Paxson, Vern E. "Measurements and Analysis of End-to-End Internet Dynamics" EECS Department, University of California, Berkeley, 1997.
- [90] Pierce, J. R. New means of communication Spectrum, IEEE, 1966, 3, 62 -65.
- [91] Portoles-Comeras, Marc and Cabellos-Aparicio, Albert and Mangues-Bafalluy, Josep and Banchs, Albert and Domingo-Pascual, Jordi, "Impact of transient CSMA/CA access delays on active bandwidth measurements", IMC '09: Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference
- [92] Prasad, R. S.; Murray, M.; Dovrolis, C. & Claffy, K. Bandwidth Estimation: Metrics, Measurement Techniques, and Tools IEEE Network, 2003, 17, 27-35
- [93] Prytz, G. Network recovery time measurements of RSTP in an ethernet ring topology Emerging Technologies and Factory Automation, 2007. ETFA. IEEE Conference on, 2007, 1247 -1253
- [94] Roberts, L. G. Beyond Moore's Law: Internet Growth Trends Computer, IEEE Computer Society Press, 2000.
- [95] Roberts, L. The evolution of packet switching Proceedings of the IEEE, 1978, 66, 1307 - 1313
- [96] Rose, B. Home networks: a standards perspective Communications Magazine, IEEE, 2001, 39, 78 -85.
- [97] Shriram, A. and Kaur, J. "Empirical Evaluation of Techniques for Measuring Available Bandwidth" INFOCOM 2007. 26th IEEE International Conference on Computer Communications, IEEE.
- [98] Sommers, J.; Barford, P. & Willinger, W. "A Proposed Framework for Calibration of Available Bandwidth Estimation Tools" Computers and Communications, 2006. ISCC '06. Proceedings.
- [99] Steven M. Bellovin "A Best-Case Network Performance Model" ATT Research, Tech. Rep 1992.
- [100] Strauss, Jacob and Katabi, Dina and Kaashoek, Frans "A measurement study of available bandwidth estimation tools" IMC '03: Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement.

- [101] Sundaram, N.; Conner, W. & Rangarajan, A. Estimation of Bandwidth in Bridged Home Networks Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops, 2007. WiOpt 2007. 5th International Symposium on, 2007, 1 -6
- [102] Tipsuwan, Y. Control methodologies in networked control systems Control Engineering Practice, 2003, 11, 1099-1111.
- [103] Tirumala, A.; Cottrell, L. & Dunigan, T. Measuring end-to-end bandwidth with Iperf using Web100 Web100”, Proc. of Passive and Active Measurement Workshop, 2003, 2003
- [104] TR-094 Multi-Service Delivery Framework for Home Networks, Tech. Rep. 2004
- [105] TR-126 Triple-play Services Quality of Experience (QoE) Requirements Tech. Rep., 2006.
- [106] TR-160 IPTV Performance Monitoring. Tech. Rep., 2010
- [107] Tritton, J. Interactive Home Systems (IHS)-an overview Consumer Electronics, IEEE Transactions on, 1988, 34, 694 - 699
- [108] UPnP-QoS Architecture:3 UPnP Forum, 2008
- [109] V. Ribeiro, R. Riedi, R. Baraniuk, J. Navratil, and L. Cottrell, “pathchirp: Efficient available bandwidth estimation for network paths,” in PAM ’03, La Jolla, CA, USA, Apr. 2003.
- [110] Verma, D. Service level agreements on IP networks Proceedings of the IEEE, 2004, 92, 1382 – 1388
- [111] Viswanathan, A.; Feldman, N.; Wang, Z. & Callon, R. Evolution of multiprotocol label switching Communications Magazine, IEEE, 1998, 36, 165 -173
- [112] Wetzstein, B.; Karastoyanova, D. & Leymann, F. Towards Management of SLA-Aware Business Processes Based on Key Performance Indicators. 9th Workshop on Business Process Modeling, Development and Support (BPMDS’08).
- [113] Wójcik, R. & Jajszczyk, A. Flow Oriented Approaches to QoS Assurance ACM Comput. Surv., ACM, 2012, 44, 5:1-5:37
- [114] Xiao, X. & Ni, L. Internet QoS: a big picture Network, IEEE, 1999, 13, 8 -1
- [115] Yiakoumis, Y.; Yap, K.-K.; Katti, S.; Parulkar, G. & McKeown, N. Slicing home networks Proceedings of the 2nd ACM SIGCOMM workshop on Home networks, ACM, 2011, 1-6
- [116] Zhang, Y.; Yu, R.; Xie, S.; Yao, W.; Xiao, Y. & Guizani, M. Home M2M networks: Architectures, standards, and QoS improvement Communications Magazine, IEEE, 2011, 49, 44 -52
- [117] Zhou, H.; Wang, Y.; Wang, X. & Huai, X. “Difficulties in Estimating Available Bandwidth” Communications, 2006. ICC ’06.