Towards Knowledge-driven QoE Optimization in Home Gateways

Bjørn J. Villa
Department of Telematics
Norwegian Institute of Science and Technology
Trondheim, Norway
bjorn.villa@item.ntnu.no

Poul E. Heegaard
Department of Telematics
Norwegian Institute of Science and Technology
Trondheim, Norway
poul.heegaard@item.ntnu.no

***Abstract* - This paper presents the concept of using distributed knowledge components as basis for a Quality of Experience optimization process. We also present simulation results indicating the potential in using this approach for access and home networks. The main novelty of the paper is the presentation of how specific end user preference information can be combined with specific content provider service information, and used as input to an optimization process in a home gateway device. The results show that the effect of doing this is significant.**

Keywords: QoE, Home Gateway, Adaptive Services

I. INTRODUCTION

The focus on QoE (Quality of Experience) rather than just QoS (Quality of Service) has been growing in strength over the last years. The main reason for this relates to the acknowledgement of that users are not equal. The QoE approach covers not only technical metrics, but also metrics describing the uniqueness of a specific user (cf. Table 1). As a result, it represents a measure of the overall customer satisfaction with a service or vendor. This makes it more suitable for user oriented service delivery architectures [3].

Table 1. EXAMPLE QoS AND QoE METRICS

QoS metrics	**QoE metrics**
Bandwidth	Perception
Delay	Preferences
Packet Loss	Expectations
Jitter	Acceptance
Availability	Price
…	…

The traditional approach of assigning a fixed priority per service or service class, and then implement a QoS design may not be rich enough to support more advanced QoE optimization schemes. Even with the full range of DiffServ values/classes [12], this will be a limiting resource. Further on, the actual QoS implementation with a high number of classes would have significant complexity issues. As an alternative to this, the concept of knowledge based QoE optimization is proposed.

For content providers operating in the Over-The-Top domain it is natural to focus more on the QoE dimension rather than the QoS subset, as the latter would be partly outside of their control. In line with this statement it is easy to understand that this type of content providers would appreciate techniques enabling them to adapt their service delivery according to different users and varying network conditions. Further on, the location of effective optimization processes outside of the network operator domain is beneficial, as this would not require involvement from the network operator.

The structure of this paper is as follows. Section 2 provides an overview of state of the art in the relevant field and also defines the objectives of the research reported in this paper; Section 3 describes the role and components of the Knowledge Plane; Section 4 describes the simulation model; Section 5 presents simulation results; Section 6 presents an analysis of the results; Section 7 provides the conclusions and an outline of future work.

II. STATE OF THE ART

The framework used for QoE optimization in a home network environment is in line with related work as stated in [9][10][11] and illustrated in Figure 1.

The addition of a Knowledge Plane in network architecture as an addition to the well known control and management plane was originally proposed by the authors of [5]. The purpose of this Knowledge Plane was to give a unified view of network aspects, to analyze it – to explain it – and finally also to make suggestions on what to do in order to achieve specified objectives.

The use of a Knowledge Plane in the networking context, and the ideas from autonomic computing [8] was taken further by the MUSE Project "Advanced features for MM enabled access platform" [6]. Their work lead to a proposal of having Monitor Plane (MP) and Action Plane components distributed across a network, including the end systems.

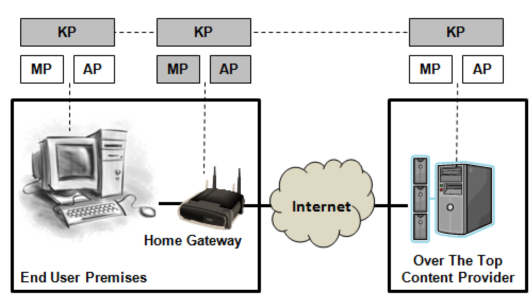

Figure 1. Optimization framework

The main difference between the optimization model used in this paper and earlier work by others is the inclusion of KP/MP/AP components from end user and content provider domains (cf. Figure 1). The KP components from

 ISBN: 978-1-61208-133-5

these domains are used as input to an optimization process in the home gateway which then is studied in this paper.

The content of the Monitor Plane and Action Plane is not the main focus of this paper, as we just assume their presence in the home gateway. More information on this can be found in previous work [9][15][17][18][19]. The type of Action Plane components applied would to a large extent depend on whether the traffic flows subject to control are of a responsive (TCP) or non-responsive (UDP) type. Related work in this area can be found in [1][7][20]. It is also important to note that the location of Action Plane components in the home gateway and not at network edge impose some challenges. Reason being that the congestion point for downstream traffic is at network edge.

The objective of the research documented in this paper is to support the statement that QoE optimization mechanisms for Internet services can be implemented in the home network domain, with the use of appropriate knowledge sources. The chosen method for providing this support is by means of simulation of a defined service usage scenario, with variable input parameters.

III. KNOWLEDGE PLANE

The Knowledge Plane is represented by information objects distributed across the platform components involved.

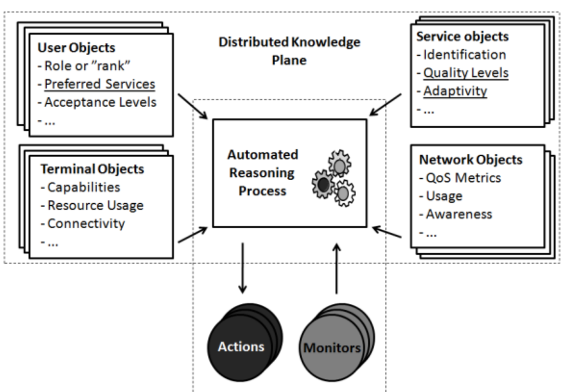

Figure 2. Knowledge Plane input to reasoning process

The use of Knowledge components in an optimization process requires a reasoning process (cf. Figure 2). This reasoning process combines and interprets the different components, allowing them to be used for some actions, and also effects to be monitored and understood.

The information objects used in the work reported in this paper are the user and service objects, and selected parameters from these (user: preferred service, service: quality level and adaptivity)

A. User Objects

The list of user preferences and associated capabilities which, could be used as input to an optimization scheme is potentially long, and depends to a large extent of the type of users being discussed (residential vs. business). What is considered as important by one user may not be of interest at all to another user, and vice versa. The thresholds for what is considered as good or bad quality are also different between users. This dynamic picture of user preferences and profiles are considered important to analyze and structure, in order to use part of it as input to optimization mechanisms.

In addition to the specific preferences of a user, there are also other differences in terms of factors contributing to the per-user QoE. Users are, e.g., different in terms of expectations concerning real quality. This may be directly related to the preferred user terminal capabilities or just basic differences in human perception. User preferences are also influenced by cost factors and assumed user rank in the specific home environment.

B. Service Objects

Many Internet-based services have certain requirements in terms of what is needed in order to be used. These requirements have traditionally been described by QoS parameters (delay, packet loss, jitter and bandwidth). This set of service information is still valid, but should be extended with additional parameters. This is especially important in light of the rapid evolution in content delivery techniques and associated technologies. The concept of adaptive streaming is an example of this. In this scenario the quality levels of a service is able to adjust itself according to end-to-end performance before and during service usage. This makes the bandwidth requirement for a certain service no longer fixed, but rather a variable parameter with some min/max thresholds and granularity. Further on, the concept of multi-source streaming from distributed and shared service platforms is also growing in popularity making it more challenging to recognize and classify services. The distribution of sources also makes the services become less sensitive for high delay, packet loss and jitter as it can pick the best performing streams and compose the service based on this.

C. Reasoning Process

In order to see the effect on using end user and content provider knowledge in the optimization process, three different schemes have been studied. These schemes are to some extent in line with the concept of a DiffServ bandwidth broker [13], but instead of priorities and policies as basis for bandwidth sharing we are using other knowledge components.

The first scheme is the basic FCFS (First Come First Served), where all knowledge use has been disabled and the home gateway operates in a regular best effort mode. The second mode is named STOPINC, where the Action Plane prevents background traffic source from increasing (if attempted) during a period where an end user preferred service is running below its maximum level. The third mode is called STEPDOWN, where the Action Plane in addition to what the STOPINC mode does - also makes a background traffic source decrease its rate according to the end user preferred service granularity. In the latter mode, the purpose is then to make it easier for the preferred service to increase its rate – one step closer to its maximum. For both the

 ISBN: 978-1-61208-133-5

STOPINC and STEPDOWN modes, when a background traffic source is either prevented from increasing or even made decrease its rate – it will be subject to this control for a certain period. This period should be enough so that the adaptive source notices that there is a chance of increasing rate.

IV. SIMULATION MODEL

The user scenario modeled in the simulator is a residential user group present in a typical home environment. The user group is connected to the Internet through a typical broadband connection. The broadband connection represents the resource shared between users and associated service.

A group of 4 users are considered, each of which operate independently of each other. Each user can start a single service at a time. There are no feedback mechanisms implemented in terms of users changing behavior as a result of good or poor performance.

Table 2. Simulation parameters

Parameter	Adaptive Service	Bkgd. Service
Max sessions	1	3
Bitrate (Kbps)	300-900	100-2800
Granularity (Kbps)	300	1
Time to first start (s)	Uniform(3,10)	
Session lifetime (s)	Uniform(10,30)	
Time to next start (s)	Uniform(3,10)	
Conn. capacity (Kbps)	1000-7000	
Control Period (s)	3	
Simulation time per run	7 days	
Number of seeds / run	100	

As can be seen from the simulation parameters, the session lifetimes are very short – much shorter than what could be expected in real life. The purpose of this was to make the simulation scenarios as dynamic as possible.

In order to see the effect of the studied QoE optimization process during different levels of congestion, the access capacity was varied while the service characteristics are kept the same.

The simulator was built using the process oriented Simula [14] programming language and the Discrete Event Modeling On Simula (DEMOS) context class [4].

A. Adaptive Service

The adaptive service (cf. Figure 3) operates between the max/min thresholds of respectively 900Kbps and 300Kbps and the granularity of increase/decrease could be set to values between 50Kbps and 300Kbps in the simulator - corresponding to fine versus course rate granularity. The granularity used in the presented results is 300Kbps and the interval of potential rate change was set to 2 sec. The reason for choosing both these values was that these are common parameter values used by live services [2][16].

The adaptive service will always try to increase its rate if possible, and will remain at max level when reached until it finishes unless if influenced by background traffic bursts. The influence from traffic bursts has been included in the model as it would be difficult to prevent, due to the location of the optimization process in the home gateway after the downstream congestion point.

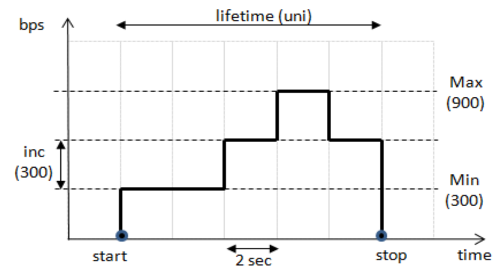

Figure 3. Adaptive service profile

The lifetime of the adaptive service session is taken from a random uniform distribution. A single adaptive service is run at a time, with repeated starts/stops during the simulation period.

B. Background Service

The background services (cf. Figure 4) used in the simulation operates in a rather simple mode, but potentially close to a worst case scenario. The sources are very bursty and pick a new target rate for each interval between a lower (100kbps) and upper threshold (2800Kbps) according to a uniform distribution. The intervals between each rate change is according to a negexp distribution (λ=1).

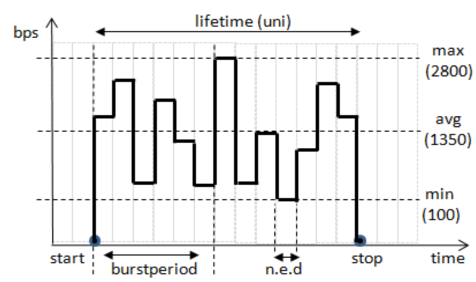

Figure 4. Background service profile

Whenever a background service starts up, it enters a burst period. During this burst period, the background services are allowed to influence the user preferred services, and in the case of congestion – they will make the adaptive service decrease its rate. The reasoning behind this is that the optimization process simulated is placed in the customer home gateway, and therefore after the access congestion point for traffic to the customer.

The duration of the burst period is decided by how fast new background traffic can be put under control by action plane components in the home gateway. Depending on the traffic type (TCP, UDP) and associated application this period will have different values. In the simulation results presented in this paper, the burst period has been varied between zero and 0.6 sec. The value of zero would represent no burst impact (ideal situation).

ISBN: 978-1-61208-133-5

The lifetime of the background service session is taken from a random uniform distribution. Maximum three background services are run at a time, with repeated starts/stops during the simulation period.

V. SIMULATION RESULTS

The parameter studied in the simulations is the average achieved bitrate for the adaptive service as a function of access capacity. Traffic load is kept constant.

A. *FCFS, STOPINC and STEPDOWN results*

In Figure 5 results are presented where the burstperiod is set to 0.2 seconds, the adaptive service has increments of 300Kbps and the background services have n.e.d rate increment intervals with λ=1. The three different models FCFS, STOPINC and STEPDOWN are then compared.

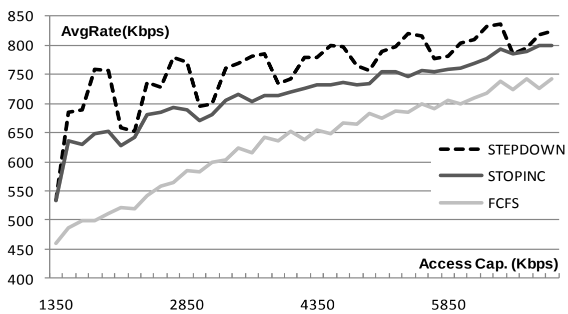

Figure 5. Comparison of optimization models

The 95% confidence intervals for the STEPDOWN model are shown in Figure 6, in order to give see how similar the results from the different simulation runs are.

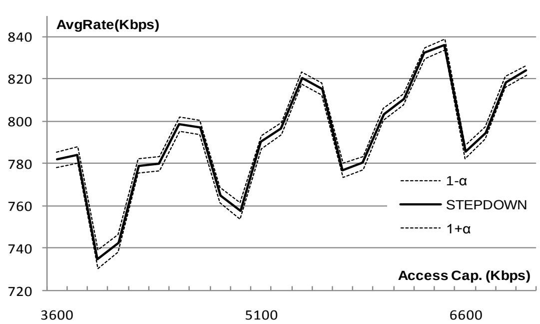

Figure 6. Confidence intervals for STEPDOWN

The confidence intervals are all in the region of +/- 2 to 7 across the studied access capacity range, which is very close to the plotted averages.

B. *Effect of changing burstperiod*

In Figure 7 the effect of changing the burstperiod for the background service is shown for the STEPDOWN optimization model.

The purpose of changing this parameter was to see if it had a major impact on the simulation results, and also to provide an indication on how fast the relevant Action Plane components would have to be in order to support the proposed QoE optimization process.

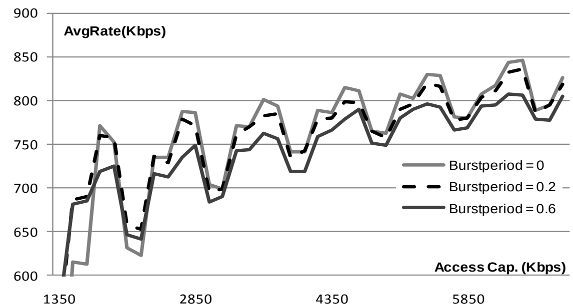

Figure 7. Change of burstperiod for STEPDOWN

The burstperiod values used are 0, 0.2 and 0.6 – whereas the value of 0 corresponds to an ideal scenario where the background services never influences the preferred adaptive service. The higher burstperiod values corresponds to scenarios where the Action Plane require some time interval in order to achieve control on the background services.

VI. ANALYZING THE RESULTS

The results presented in the previous section are considered promising, as they support the statement subject to investigation. The comparison between the FCFS, STOPINC and STEPDOWN modes of operation (cf. Figure 5) shows that for a home gateway the potential improvement in average bitrate for a preferred service is significant, if knowledge about the service granularity is made available. The simulation results show that during high congestion both the STOPINC and STEPDOWN models perform significantly better that FCFS. For the STOPINC mode there is a potential for between 10-30% higher average rate, and for the STEPDOWN mode the same potential is between 10-40%. The STEPDOWN mode performs significantly better that STOPINC for all access capacity levels (cf. Figure 8).

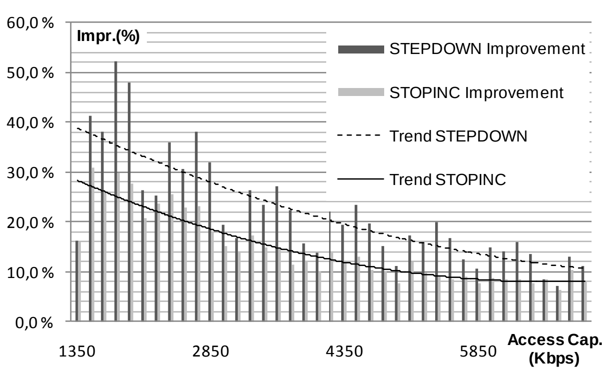

Figure 8. Rate improvements in percentage per model

The results when changing the burst period (cf. Figure 7) illustrate the importance of having an efficient Action Plane supporting the reasoning process. If services are not put

 ISBN: 978-1-61208-133-5

under control as fast as possible, it reduces the potential of STEPDOWN in the order of tens of Kbps.

It should be noted that there is no general 1:1 mapping between an achieved value of a QoS metric such as bitrate and a specific QoE metrics. However, it is a fair assumption that there is a weighted mapping between QoS metrics and related QoE metrics, following the preference and perception levels of a certain user. In line with this, we can say that the achieved increase in bitrate for the preferred adaptive service contributes to an increased QoE level.

VII. CONCLUSION AND FUTURE WORK

Based on the analysis and simulation results presented, the statement of a potential gain in implementing QoE optimization mechanisms in the access and home network domain is strengthened. It is clear that even with just very basic knowledge components available from the user and service objects (cf. Figure 2) a significant improvement in QoE can be achieved.

The presented results may also have value for pure network oriented QoS mechanisms, if this type of stepwise service adaption becomes a success in emerging service delivery architectures. As an example, it is likely that the bandwidth broker concept of DiffServ could benefit from introducing this type of service knowledge in its operation.

As future work in this area, the plan is to investigate more complex user and service scenarios. It is also the intention to make the service models used in the simulator closer to real life traffic. Further on, the logics in the reasoning process together with efficient action plane components will be addressed.

VIII. ACKNOWLEDGEMENTS

The reported work is done as part of the PhD studies for the first author, which is an integrated part of the Road to media-aware user-Dependant self-aDaptive NETWORKS - R2D2 project. This project is funded by The Research Council of Norway.

REFERENCES

[1] A. Aggarwal, S. Savage, and T. Anderson. “Understanding the performance of tcp pacing”. In Proc. IEEE Nineteenth Annual Joint Conf. of the IEEE Computer and Communications Societies INFOCOM 2000, volume 3, pages 1157–1165, 2000.

[2] S. Akhshabi, A. C. Begen, and C. Dovrolis. “An experimental evaluation of rate-adaptation algorithms in adaptive streaming over http”. In ACM Multimedia Systems (MMSys), 2011.

[3] E. Areizaga, L. Perez, C. Verikoukis, N. Zorba, E. Jacob, and P. Odling. “A road to media-aware user-dependent self-adaptive networks”. In Proc. IEEE Int. Symp. BMSB ’09, pages 1–6, 2009.

[4] G. Birtwistle. Demos - A system for Discrete Event Modelling on Simula. School of Computer Science, University of Sheffeld,, July 1997.

[5] D. D. Clark, C. Partridge, J. C. Ramming, and J. T. Wroclawski. “A knowledge plane for the internet”. SIGCOMM’03, August 2003.

[6] H. Dequeker, T. V. Caenegem, K. Struyve, E. Gilon, B. D. Vleeschauwer, P. Simoens, and W. V. de Meerssche. Advanced features for mm enabled access platform (muse project deliverable). Technical report, Alcatel Bell and IBBT - IBCN, December 2006.

[7] H. Jamjoom and K. G. Shin. “Persistent dropping: an efficient control of traffic aggregates”. In Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications, SIGCOMM ’03, pages 287–298, New York, NY, USA, 2003. ACM.

[8] J. O. Kephart and D. M. Chess. “The vision of autonomic computing”. Computer, 36(1):41–50, 2003.

[9] S. Latr´e, P. Simoens, B. De Vleeschauwer, and V. de Meerssche. “An autonomic architecture for optimizing qoe in multimedia access networks”. Computer Networks 53, 10:1587–1602, 2009.

[10] S. Latre, P. Simoens, B. De Vleeschauwer, W. Van de Meerssche, F. De Turck, B. Dhoedt, P. Demeester, S. Van Den Berghe, and E. G. de Lumley. ”Design for a generic knowledge base for autonomic qoe optimization in multimedia access networks”. In Proc. IEEE Network Operations and Management Symp. Workshops NOMS Workshops 2008, pages 335–342, 2008.

[11] S. Latr´e, S. Verstichel, and B. D. Vleeschauwer. “On the design of an architecture for partitioned knowledge management in autonomic multimedia access and aggregation networks”. MACE, pages 105–110, 2009.

[12] K. Nichols, S. Blake, F. Baker, and D. Black. “Definition of the differentiated services field (ds field) in the ipv4 and ipv6 headers”. RFC 2474 (Proposed Standard), Dec. 1998. Updated by RFCs 3168, 3260.

[13] K. Nichols, V. Jacobson, and L. Zhang. “A two-bit differentiated services architecture for the internet”. RFC 2638 (Informational), July 1999.

[14] R. J. Pooley. An Introduction to Programming in Simula. Blackwell Scientific Publications. ISBN: 0632014229, 1987.

[15] P. Simoens, B. D. Vleeschauwer, W. V. de Meerssche, F. D. Turck, B. Dhoedta, and P. Demeester. “RTP connection monitoring for enabling autonomic access network qos management”. In Proceedings of the 12th European Conference on Networks and Optical Communications (NOC), 2007.

[16] TV2. Adaptive webtv streaming solution from TV2 Norway. Available online at www.tv2sumo.no. Last access 01.01.11.

[17] B. J. Villa and P. E. Heegaard. “Monitoring and control of qoe in media streams using the click software router”. In Norsk Informatikk Konferanse (NIK2010), Norway. ISBN 978-82-519-2702-4., volume 1, pages 24–33, Nov 2010.

[18] B. D. Vleeschauwer, W. V. de Meerssche, P. Simoens, F. D. Turck, B. Dhoedta, and P. Demeester. “Enabling autonomic access network qoe management through tcp monitoring”. In Proceedings of the First IEEE Workshop on Autonomic Communications and Management (ACNM), 2007.

[19] B. D. Vleeschauwer, P. Simoens, W. V. de Meerssche, S. Latré, F. D. Turck, and S. V. den Bergheb. “Autonomic qoe optimization on the access node”. In Proceedings of the Broadband Europe, 2007.

[20] H.-Y. Wei, S.-C. Tsao, and Y.-D. Lin. “Assessing and improving tcp rate shaping over edge gateways”. IEEE Trans on Computers, 53(3):259–275, 2004.

