The BETSY project on timeliness and energy aspects of wireless video streaming

M. Sénéclauze, J-D Decotignie (CSEM), P. van der Stok, H. de Groot, M. van Hartskamp, G. van Doren, D. van Heesch, C. Otero Perez (PR), M Joosten (C-lab), C. Blanch, J. Bormans (IMEC), M. Geilen, T. Basten, B. Theelen (TU/e), C. Koulamas, G. Papadopoulos, A. Prayati (ISI), G. Fohler. D. Isović (MDH), G. A. Papadopoulos, P. Cheng, Z Abraham (UCy)

Abstract

The BETSY project focuses on the seamless adaptation of multimedia streams on wireless hand-held devices to fluctuating network conditions and available terminal resources. Consequently, the user can enjoy true multimedia experiences with freedom of movement in a networked home or at any hot-spot. A stream model is developed during the project to optimize the energy consumption and satisfy the timeliness constraints and optimize the energy consumption. The project is in its first 8 months. The initial stream model that is the basis of further work is described and motivated.

1. BETSY project structure and goals

The BETSY (Being on Time Saves energY) project is executed by a consortium of 8 partners: Philips Research (NL), CSEM (CH), IMEC (BE), ISI (Gr), MDH (S), Siemens C-lab (D), TU/e (NL) and Univ of Cyprus (UCy). The total duration is 30 months with a starting date of 1 September 2004.

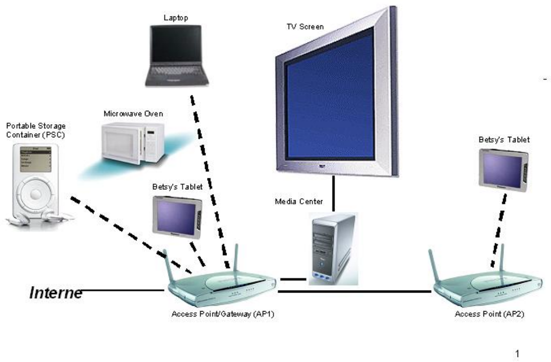

Figure 1 home network example

Figure 1 shows an example network that BETSY addresses. Video is transported over both wired and wireless communication media. Sources of video streams can be a broadcast, the Internet, a camera or a storage device within the home. More than one Access Point are involved, leading to hand-overs and changing bandwidth availability during the displacements of a user who watches video on a hand-held device. The presence of a microwave shows the interference from an outside source on the WiFi connection, similar to the interference coming from Bluetooth devices [1]. Display devices of different qualities are involved. The BETSY project aims at increasing the user satisfaction for mobile video rendering devices by:

1. providing techniques which allow the rendering device to seamlessly adapt to fluctuating network conditions
2. reducing the energy consumption of the rendering device such that a user can enjoy the video without interruption over an extended period.

These high level objectives will be addressed by studying the trade-offs and the timeliness requirements of the video streams. The timeliness requirements can be guaranteed after an analysis of all timeliness aspects over the complete transport chain. This chain is based on components for which trade-offs can be determined.

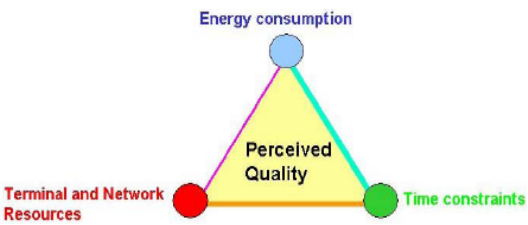

Figure 2 trade-off triangle

Figure 2 shows the optimisation criterion, which is the user-perceived quality. This quality can be improved by trading-off between any two end-points on the side of the triangle (e.i. resources versus energy consumption, time constraints versus energy consumption, and resources versus time constraints). Examples of resources are network bandwidth, CPU cycles, and memory size.

Given the dynamic nature of the utilization of the network utilization, an approach is needed that allows the optimisation device settings every time network conditions changes. Examples of changes are:

- the addition or removal of a video stream,
- video passing from one Access Point to another due to device mobility.
- Environmental changes, such as the level of interference.

It is not sufficient to optimise settings of individual network segments or the individual devices, but an integrated approach is needed that involves all components that collaborate to stream a given video.
The project intends to define a common resource model that maps to the mobile devices as known to-day. The project aims at understanding the trade-offs and quantifying them for the common resources. Once the quantification is done for the individual components, the optimisation of the whole communication chain can be tackled. The final result is an integrated approach for optimising the energy consumption for a given perceived quality over the whole transport chain. This integrated approach constitutes the added value of BETSY.

2. Problem description

A video consists of a sequence of frames, where each frame completely defines a picture. Video frames are transported over a network in a digital format that is constructed according to the MPEG-x standard, with x being 1,2, or 4 [2]. The most popular standard at this moment is MPEG-2. The BETSY project however concentrates on the more recent MPEG-4 standard. The performance of the MPEG-4 encoder is superior to the MPEG-2 encoders (less bits for a given quality) and the consortium expects that the popularity of MPEG-4 will grow with increasing video usage on hand-held devices (e.g. UMTS combined with video on your telephone). The consortium expects that the bit-rate of the encoded video can be adapted to the bandwidth availability on the transmission medium. It is expected that the simple profile of MPEG-4 will give this possibility while at the same time consuming little energy.
At the rendering device, frames are generated with an interval of 20 ms or larger. At the source, the video is either stored on a storage device in MPEG-4 format or created on-line by a camera. Two types of video can be distinguished:

1. *Real-time* in which there is a short delay (<500 ms) between the creation of the video and its rendering at the destination.
2. *Off-line* in which the video can be rendered with any delay and trick play like fast forward, rewind, pause, is possible.

At the receiver side, the video is decoded and the values of the individual pixels are calculated for each frame. The quality of the video is given by its Signal to Noise Ratio (SNR) value defined as the negation of the logarithm of the differences between the encoded picture and the original picture for this frame. The higher the SNR value the better the quality. The quality of a frame with a SNR value lower than 30 is generally considered poor, because artifacts become visible whereas a frame with a SNR value of 40 or higher is generally considered of excellent quality. The relation between SNR value and quality depends on many things such as the type of video (cartoon, movement, etc.). It takes about twice as many bits to go from SNR value 35 to SNR value 40. Above 40 a slight increase in SNR value adds very little perceived quality improvement.
For a given frame, there is a clear relation between video quality and the number of bits that defines the frame. A video can be stored or encoded at a given quality and consequently with an associated bit-rate. However, for a given video at a given quality, the number of bits still changes drastically from frame to frame.
The bandwidth of a transport medium defines the number of bits that can be transported per unit time. Wired networks have a constant bandwidth. However, wireless networks suffer from interference from other sources and transmission degrades with increasing distance between source and destination. Bandwidth fluctuations can go from maximum to half in less than a few ms.
Another aspect is the relation between the capacity of the transmission medium and the calculation capacity of the rendering device. A set of videos to be transported may need more bandwidth than provided by the communication media. At the receiving side a higher quality video may lead to higher demands on the capacity of the decoding device. Perturbations during transport may require additional computation power to remove the effects of these perturbations.
We are confronted with a fluctuating size of the video, a fluctuating bandwidth of the wireless media, and the capacity of the medium and the rendering destination device. Choices need to be made about resource allocation to the video streams.
The energy consumption of a hand-held device is important. The fact that after a few minutes of video rendering the battery needs to be recharged for a few hours is a major drawback with respect to user-friendliness. Although the major consumption currently resides in the display technology, the advent of new techniques like *Polymer*, or *Polymeric Light Emitting Diode (PLED)* eliminates the display as major bottleneck in power consumption. Consequently, the encoder or decoder becomes an important energy consumer. BETSY investigates whether the simple profile defined by MPEG-4 will serve all needs of power reduction and adaptability to fluctuating transmission conditions.

3. Scenarios

The BETSY project has taken two scenarios as starting point, which will also be used to validate the BETSY results. Each scenario consists of three parts, where each successive part constitutes an addition to or a refinement of the former part. Scenario 1 is centered around the home with the network represented in Figure 1. Scenario 2 concentrates on hot-spots distributed in a residential area, typically covering 1-3 streets.

Scenario 1 covers the effects of hand-over between access-points causing a fluctuating image quality of the video shown on a tablet. The owner of the tablet moves from living-room to a room upstairs and later to a pool outside. The creation of a Picture in Picture (PiP) view from the kitchen onto the tablet at the pool leads to quality changes, and the switching on of the microwave leads to a disruption of the video in the kitchen. The scenario is intended to quantify the effects on video quality dependent on network load during every change in network condition and configuration.

Scenario 2 equips children with cameras while playing with their bicycles in the neighborhood. The scenario describes how the lifetime of the batteries in the cameras can be increased by reducing the video quality. Also the effect of all children gathering within the area of one hot-spot is analyzed. The presentation of the image as PiP or as standard television image is envisaged. It is assumed that no hiccups or quality degradations occur when the children cycle from one hot-spot coverage into another one. As in scenario 1, calculations on bandwidth usage and battery lifetime will be presented.

4. Breeze concept

To avoid the confusion coming from the use of overloaded terms like "stream" or "flow", the multimedia content sent from the source to the destination was named a "Breeze". The objective is to create a concept that

1. is abstract enough to hide all design choices to be made later during the project, and
2. is expressive enough to contain all the functionality perceived by a user manipulating the remote control.

Choosing already existing concepts, like a SIP connection [3], was rejected, because the concepts address (for BETSY unnecessary) functionality that might be distractive during the discussions.

A Breeze is a piece of content, on which a sequence of functions is applied for processing, storing and communicating data items in an end-to-end delivery chain, on which only one entity has control.

Its main characteristics are:

- Its content can be displayed on many screens
- Audio and Video are transported along the same path
- It is implementation independent

The chosen abstractions bring many advantages to the description of the end-to-end process involved in BETSY.

1. Since audio and video are transported along the same path and since the same content displayed on two different screens will be coming from the same breeze, the jitter between the two screens can be considered small and acceptable enough not to require a strong synchronization.
2. As the Breeze goes from the acquisition to the display, the format of its content evolves. The Breeze term is used over the whole chain simplifying the abstraction.
3. If its content needs to be stored for a short or a long time, the stored content is either seen as the endpoint of a breeze (end storage) or part of the breeze (intermediate storage). An example of intermediate storage is the delay-buffering concept, which allows skipping commercials.
4. Trade offs made along the path to optimize the complete network (e.g. bandwidth, energy consumption) are done at the Breeze level by adapting the content format or size.

The Breeze concept is not an end-goal of BETSY. It is a way to describe the end-to-end multimedia involved in BETSY. At later point during the project, an existing video chain description (SIP session for example) might be chosen either because it suits our evolving needs, because it may be easier to deploy, or makes it easier to explain the BETSY results. Nevertheless, the Breeze concept has a vital importance because in this early stage all partners speak about the same thing with the same meaning for all partners, and it focuses the partners on the project's issues.

As noted earlier, the Breeze content may evolve along the end-to-end path. This evolution is the main mechanism to the overall network optimisation. For example, the camera used on a child's helmet can run for one additional hour by going from high quality to low quality MPEG-4 thanks to the capability of the receiving display to handle low resolution MPEG-4 content.

All the content adaptations will be made along the Breeze by a set of Functions. During the definition of the set, the goal was to come up with a small, yet sufficient, set of functions without making any implementation or design choices at this early stage. Identified functions which can be manipulated by end-user choices are:

- Breeze Start-points: Capturing, Retrieving
- Breeze End-points: Recording, Rendering
- Other functions: Encoding and decoding, Multiplexing and de-multiplexing Delay buffering and transporting.

The functions are mapped to physical components, which together constitute one or more devices (a box one can buy in the shop). The physical components consume resources, shared between the functions, and the low availability of one (or many) of them will necessitate choices to find a better point in the trade-off triangle. The common resources consumed by the functions are CPU, memory and energy. For example the functions, Retrieving and Capturing, involve permanent storage.

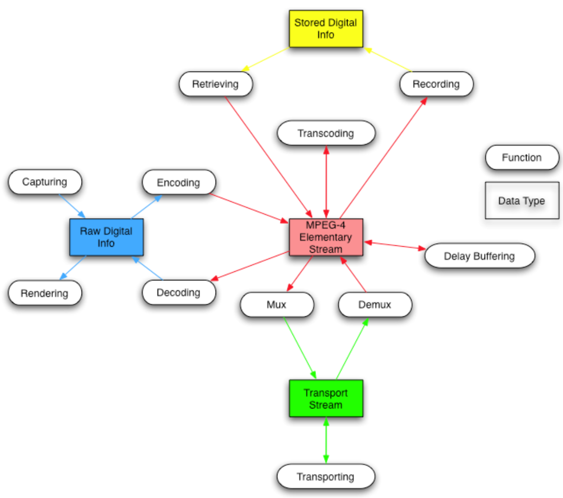

Figure 3 Breeze data path

Figure 3 shows the functions, their input formats, and output formats. The data types, used as input or output of the functions, are mainly MPEG-4 elementary streams but three other types are needed in order to describe the full end-to-end chain:

- Raw Digital Info: only the decoding and the encoding function locally within the start/endpoint devices will use this data type.
- Stored Digital Info: this represents a file on a permanent storage device that stores the content of the stream.
- Transport Stream: this is the elementary stream enhanced for transport purposes and needs not conform to the MPEG-2 Transport Stream.

5. Breeze concept validation

The breeze concept is validated in four stages:

1. verify that currently devices in the market are sufficiently described by the breeze functions
2. decompose current devices in physical components (e.g storage unit, processing unit)
3. map the breeze functions to physical components
4. express the scenarios in breeze functions and physical components.

The last stage actually represents the actual validation of the Breeze concept. It shows that the Breeze concept is complete (no functionality is missing). It also shows that all breeze functions are necessary (all functionality is used in the scenarios).

Quite a lot of effort has gone in validating the Breeze Concept. However, it does not guarantee the absence of surprises during the project. Surprises are bound to occur because the validation does not show whether the Breeze Concept hides all design decisions. For example, by specifying that a multicast of the same video is one breeze, has as consequence that synchrony between the different end-points is fixed and the content at the source is the same. However, during implementation, it may be decided to use two different content formats of the same video to serve the different quality requirements of the two destinations. Given such an implementation decision, the synchrony between destinations cannot be guaranteed any more. The expectation is that the breeze concept definition needs to be tightened to include source equivalence and timing constraints as well. The difficulty lies in only adding those refinements, which do not restrict the design choices afterwards.

The validation of the breeze concept is an on-going process that will teach us the (possibly orthogonal) parameters that are necessary to optimise the network utilisation for a given set of breezes.

6. Conclusions

Future hand-held devices will become more popular and need capabilities to adapt to the fluctuating network conditions. At the same time energy consumption is important to guarantee that the video can be enjoyed during the life-time of the video without battery replacement actions. The BETSY project looks at the optimization of the energy consumption while maintaining the timeliness aspect for a video of a given quality. Trade-offs between energy-consumption and video quality are quantified. The project will provide an integrated approach to optimize the settings of all components involved in transporting and rendering the video.

References

1. A.S. Tanenbaum, *Computer Networks*, 4th edition, Prentice hall PTR, 2003.
2. Rob Koenen *From MPEG-1 to MPEG-21: Creating an Interoperable Multimedia Infrastructure*, ISO/IEC JTC1/SC29/WG11 N4518, www.chiariglione.org/mpeg/, December 2001.
3. A. Johnston, S. Donovan, R. Sparks, C. Cunningham, and K. Summers, *Session Initiation Protocol (SIP) Basic Call Flow Examples,* RFC 3665, December 2003.