Adaptive CAC for SVC Video Traffic in IEEE 802.16 Networks
M. Abdennebi, Yacine Ghamri-Doudane

To cite this version:

HAL Id: hal-00794611
https://hal.archives-ouvertes.fr/hal-00794611
Submitted on 7 Mar 2013

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.
Adaptive CAC for SVC Video Traffic in IEEE 802.16 Networks

Marwen Abdennebi + and Yacine Ghamri-Doudane *

+ Laboratoire deTraitement et de Transport de l’Information (L2TI), Université de Paris 13
marwen.abdennebi@univ-paris13.fr

* École Nationale Supérieure d’Informatique pour l’Industrie et l’Entreprise (ENSIIE) and Laboratoire d’Informatique Gaspard Monge, CNRS LIGM
yacine.ghamri@ensiie.fr

Abstract – Call Admission Control is a key function that guarantees the Quality of Service (QoS) for users. In radio networks, this function is usually based on traffic models and ensures that sessions are admitted only if the estimated available bandwidth is enough for the entire call duration. For video on IEEE 802.16, the CAC function must ensure that the bandwidth to be reserved is compatible with the resource availability. For the enhanced SVC (Scalable Video Coding) systems, the CAC function must take into account all the layers and their characteristics. In this paper, we propose an enhanced CAC function for SVC that adapts the admission according to the statistical behaviour of the video sessions. The main goal is to use measurements in the 802.16 base station (BS) to update the traffic model of SVC video flows, this for the different layers of SVC flows. We then use the variability of the traffics generated to adapt the CAC according to the characteristics of incoming flows. To perform that, we use a Markovian model that adapts for each flow instead of using a generic static one as used in most of the papers. Performance evaluation is given to illustrate the interest of our proposal.

I. INTRODUCTION

IEEE 802.16 defines different services with different bandwidth reservation mechanism. Each defined service should map to the appropriate traffic and satisfy the Quality of Service (QoS) required by users. However, for each of those defined service, a Call Admission Control function, or CAC, must be added. CAC is a key function for IEEE 802.16 [1] systems as it ensures that a new session will be accepted only if the Quality of Service can be guaranteed for the entire session duration. It is essential to ensure that this function is well implemented as it allows the operator to admit the maximum number of users in the system while keeping a good quality. For each service, this function must accept a session if, and only if, the QoS can be guaranteed.

In 802.16, for the particular case of real-time Polling Service (rtPS), the bandwidth reservation scheme must take into account the mean estimation and the probabilities of each bandwidth request and its variability in time. This estimation should be based on analytical traffic models for the sources. System modelling will give the probabilities of each traffic state and hence will allow the operator to guess the probabilities of each bandwidth request related to bitrate variations.

Usually, the corresponding algorithms use single traffic models, one model per traffic source. This approach does not take account of the variations between different kinds of video traffics and hence, will not be accurate for an efficient CAC. In our previous work [2], we proposed an approach that takes into account those differences between videos flows by classifying them into different categories, using 3 different kinds of traffic models. Measurements were done to classify flows in the appropriate category. However, this approach was applicable only for the simple case of AVC (Advanced Video Coding) video where the traffic characterising was done for a single layer.

We then propose in this paper to extend the adaptive CAC solution to SVC video flows by using models for each considered enhancement layer. In this case, we have to distinct between the layers of admitted video sources and have a traffic model for each category of enhanced layer. We then obtain an adaptive CAC that takes account of the number of video sessions in each category and for each layer, and hence adapts the admission decision as well as the rate through enhancement layers.

Hence, in the following, we first present a background study on IEEE 802.16 services, on SVC coding and on the CAC function for video streams. Then, section III describes our proposed adaptive CAC system and the measurements and processes it involves at the BS level. In section IV we show the benefits of our proposal in term of performances. Finally, in section V, we conclude the paper and propose some future research directions to improve and develop this work.

II. BACKGROUND

A. IEEE 802.16 System

Unlike for UGS (Unsolicited Grant Service), where the bandwidth request/grant mechanism is fixed for the entire session and hence the CAC function simply consists in
admitting only sessions that fit the mean available bandwidth, for rtPS (and also for extend rtPS or ertPS), the bandwidth is not fixed and the request/grant vary during the session. Those variations must be taken into account before admitting a session to ensure that peaks of bitrates will not cause packet losses.

Hence, the CAC function [3] must be aware of the traffic models of requesting sources. Indeed, those models will give the bitrate variations and hence the bandwidth requests for the IEEE 802.16 frames. As the traffic models used will finally give the distribution of the bandwidth requests (in term of probabilities), it is very important that those models should be as accurate as possible, for both single and aggregated traffic sources. That consists in first computing, for each time interval, the probability of the bandwidth request related to each flow, this using the simple traffic model for the corresponding source [4]. Then, it should be compared with the available bandwidth for this service in the IEEE 802.16 frames that will be given using the aggregated traffic model for all admitted sources in rtPS. If the grant is possible for the defined number of frames (given from the polling interval), the bandwidth will be allocated, otherwise, the packets are delayed or even dropped.

B. SVC: Scalable Video Coding

SVC is an extension proposed to the H264/MPEG-4 compression standard to enhance the video quality using layered images [6]. Some works have been proposed to enhance performances of SVC flow transmissions on broadband radio networks, but most of them like [8] [9] in deals especially with the scheduling or the coding and modulation schemes.

This extension is based on the addition of enhancement layers that, when added, improve the quality of the video. The enhancement process is based on the addition of layer frames, in time and/or space. The obtained quality improvement is related to the number of enhancement layer (figure 1 above). Each enhancement layer decoding is dependency based, meaning that it needs the base and intermediate layers to be decoded.

In the following section, we consider the particular case of Coarse Grain Scalability, CGS, which consists in enhancing the quality through the increase of prediction adding the corresponding precision in each enhancement layer.

C. CAC for video traffic

Usual CAC algorithms often use the same model considering a single video stream. Video streaming are especially challenging for as the traffic flow rate and its variations are related to the motion level of the videos. We then demonstrate that the combined models will be more accurate than a single one. The most commonly used CAC computes the probability states of the modelled system and then determine the maximum number of sources that should be accepted [5]. The main goal is then to ensure that the QoS condition is respected for the traffic sources in term of bandwidth/rate, delay, jitter and drops. The system can also accept or reject sessions “on the fly”, according to bandwidth availability, but, this time, the system will not be able to accurately predict variations for this bandwidth and hence the guarantee would not be respected.

In fact, in rtPS, the regular polling of admitted sources implies various bandwidth requests in time and space related to the throughput variations of the corresponding streams. It means that almost all requests must be granted in order to satisfy the QoS in terms of loss rates and/or delays. Actually, for each time interval, the bandwidth requested should be compared to the available one (in terms of time slots and/or symbols in frame) to ensure the grant.

The figure 2 below illustrates the differences between the CAC methods, (a), (b), (c) we consider. The first is the simple one that takes into account the available bandwidth and static models to make the decision of accepting an incoming session. The second is the adaptive CAC we proposed in [2] and takes into account the measurements of flows to update the decision algorithm and make it more efficient. The third is the one we propose in this paper and which improve to the adaptive CAC the rate adaptation through adding enhancement layers.

III. ADAPTIVE CAC FOR SVC

The proposed CAC adapts the video session admission
decision for SVC traffic taking advantage of the rate adaptation through enhancement layers. This adaptation uses flow measurements to classify them into 3 categories according to bandwidth consumption, this for each considered SVC layers. This classification gives an idea about flows partition among the models and is used by CAC for the admission decision and rate adaptation.

In the sub-sections below, we present the traffic model we use, then the streams measurements process and classification/mapping with the proposed models, and finally the adaptive CAC scheme. The whole process is summarized in the figure 3 of sub-section B.

A. Layer based Traffic Model

In this part, we propose to model each traffic layer of individual source and for base and enhanced layers. We use in this purpose as in [2] rate-based single dimension rate based Markovian chain. This model is appropriate for aggregated frames when the time scale of sampling/quantization is about 100 ms. This means that we can use this model in for the rtPS 802.16 as the bandwidth request concerns more than one MAC frames.

Actually, for video traffics, the rate levels and the rate variations differ from one flow to another according to the motion included in each video and is not related to the enhancement layer itself in CGS especially. As we use a Markovian process to model the rate, we will have, for each SVC video flow, one model for each layer.

The quantized video flows are modelled for each layer, base and enhanced. Each obtained model describes the bit rate evolution from one state to another. These states are the different value that the bitrate of each layer can take. Hence, the number of states will be the number of quantized values defined by the measurement process, with a maximum value defined for each model.

As described in figure below, each layer \( L \) of each flow is described by the process parameters \( \lambda^L \) and \( \mu^L \) which are the mean arrival and service rates from one quantized bitrate to another.

Those parameters also determine the statistical properties of the video bitrate \( D \) which is characterized by the mean rate \( E[D] \), the variance of the rate \( Var[D] \), and the correlation function \( corr(t) = e^{-\alpha t} \).

\[
E[D^{0.1.2}] = nD^{0.1.2} \frac{\lambda^{0.1.2}}{\lambda^{0.1.2} + \mu^{0.1.2}}.
\]

\[
Var[D^{0.1.2}] = n.(D^{0.1.2})^2 \frac{\lambda^{0.1.2} \mu^{0.1.2}}{(\lambda^{0.1.2} + \mu^{0.1.2})^2}
\]

\[
corr(\tau) = e^{-\alpha \tau} \quad (\tau, \mu, \lambda)
\]

The state probabilities are obtained by solving the following linear equations system:

\[
\pi_i^L = \left( \frac{N}{i} \right) \frac{(\rho^L)^i}{(1 + \rho^L)^N}
\]

\[
with \quad \rho^L = \frac{\lambda^L}{\mu^L} \quad \pi_0^L = \frac{1}{(1 + \rho^L)^N}
\]

\[
L \in \{0,1,2\} \ (in \ our \ case)
\]

The discrete values of modelled bitrate can have \( M+1 \) quantized levels. The durations \( T_i^L \) in each bitrate state \( i \), for each layer \( L \), are exponentially distributed:

\[
T_i^L = \frac{1}{(n-i)\lambda^L + i\mu^L}
\]

B. Measurements and Streams Classification

The goal of this step is to classify the video session into the appropriate Markov chain model from the 3 defined categories by mapping it with the parameters one of the Markov chains corresponding to the category proposed.

With the adaptive proposed CAC, we aim to take account of the bitrate variability for the session admissions. It means that, according to each classification and flows repartition, we have a maximum number of sessions that will also vary. This classification/repartition is according to the bitrate, i.e., the motion levels, but also according to the enhancement layers (eLayers). We use for the traffic characterising 3 levels of motion as recommended in [12] [13], which means 3 simple models, one for each flow category. We also consider 3 layers including the base layer combined to 3 categories of traffic models for each layer.

According to the traffic repartition the CAC must ensure a...
maximum number of admitted sources \(N\) if considering the base layer, \(N_1\) if considering the layer 1 and \(N_2\) for the layer 2.

For the CAC, we need to compute through measure the repartition of streams in each layer. We must determine \(\eta_{iL}\) that denotes the relative number of flow in the traffic class \(i\) and the layer \(L\) flow in each traffic category for each layer.

If the repartition in each layer \(L\) is \(\eta_{i1}, \eta_{i2}, \eta_{i3}\), we then have:

\[
\eta_{i1}, \eta_{i2}, \eta_{i3}\text{ for the base layer,}
\eta_{i1}, \eta_{i2}, \eta_{i3}\text{ for layer 1,}
\text{and } \eta_{i1}, \eta_{i2}, \eta_{i3}\text{ for the layer 2.}
\]

This means that \(\eta_{i1}\%\) of the flows have the characteristics of \(\eta_{i1}\) model in the base layer, \(\eta_{i2}\%\) of the flows have the characteristics of \(\eta_{i2}\) model in the layer 1, etc. This leads us to 3 models for each layer, all bitrate based, to map well with bandwidth request/grant mechanism of IEEE 802.16.

Each traffic model, for each layer, is characterised by the rates levels and the Markov chain as explained in the subsection above. Moreover, those models are also characterised by the mean rates states (levels) and the mean durations in those states. Those statistics give the PDF function (probability density function) which is used for mapping the measure with the model.

The figure 4 below summarizes this process, from the measurements step to the CAC update.

The measurements are handled by the BS and should be able to give information on the statistical behaviour of video traffics for the classification. Then, those measurements are done for each admitted flow and this for each layer \(L\) that was allowed. The measurements are done for the bitrate, value and duration, as the traffic model is also bitrate based.

The rate measured is first quantized with the same level as for the models, which means that we have discrete values for the measures according the Markov chain states and hence the PDF functions.

Then we map the quantized streams measures (discrete rate levels) with the models to find the probability density function (rate levels for example) of the flow and finally classify it. The statistical behaviour is mapped to the histogram through the appropriate PDF and hence according to the parameters of the appropriate model.

Once identifying the parameters of the incoming flow, for all its enhancement layers, the classification is done into the appropriate traffic category \(i\) for each layer \(L\).

After the classification, we are able to update the flow repartition \(\eta_{iL}\), then the global system model parameters and hence update the CAC function.

As the measurements are done for bitrate, the measurement duration must be enough to obtain the convergence of the parameters (the fact that the measured parameters reach a fixed value and does not diverge for more than \(n\) frames). For the considered case of video flows, it means that this duration is in the scene time scale, which means at least seconds/minutes. It is important to note that for each flow measurements should be performed repeatedly to update the system.

Once the repartition determined, the CAC process takes place by computing the maximum number of sources that can be admitted in the system. This number can be different if we can cancel some layers in case of congestion. Hence, if \(N_0, N_1\) and \(N_2\) is the maximum number of sources respectively in base layer, enhancement layer 1 and layer 2, we will have the CAC function, \(\{N_0, N_1, N_2\} = \eta_{iL}, i\) is the traffic class and \(L\) the layer.

The proposed adaptive CAC will then, for each admitted flow, regularly update those parameters using measurements, and hence compute the repartitions parameters \(\eta_{iL}\). Once updated, the CAC function should compute, thanks to offered bandwidth and the Quality of Service parameter (delay or drop rate threshold), the new corresponding maximum number of sessions that the system can admit, i.e., \(N_0, N_1,\) or \(N_2\) depending on the layer.

C. CAC Adaptation

For each accepted flow, the measurements are done in order to guess its statistical characteristics after measures convergence. For the average value for example, it means that we can update the CAC after the measurement of one flow, until reaching a threshold value \(l\) defined according the measured rate for each layer \(L\) at each measure frame \(k\) (time).

\[
\left| \sum_{L=0}^{\text{layer}} \sum_{\text{frame } k=1}^{\text{frame max}} \frac{1}{k} \text{rate}^L(k) - \sum_{\text{frame } j=1}^{\text{frame max}} \frac{1}{j} \text{rate}^L(j) \right| < l \tag{6}
\]

Then, the \(\eta_{iL}\) could be determined and the corresponding rates probabilities in each state of the 3 Markov chains are
In the case of rtPS scheduling, the number of frames related to the bandwidth request/grant should be also taken into account as this number is related to the minimum time granularity for the quantization step. It means that this value is directly related to the number of bytes that the request concerns, and hence also the states probabilities of each considered flow.

Once the measurements convergence step is finished, we obtain the probabilities of each bitrate state through a histogram of the levels. In addition to this PDF function we also use state duration values, in order to guess the value of $\lambda^L$ and $\mu^L$. Equations (4) and (5) will be used to guess the individual flow characteristics, using the Markov behaviour of the quantized rate levels. The mapping is then deduced after the measurement of the mean durations in each state $i$ and the probabilities distributions. As we have several parameters to be mapped, we use LMS (Least Mean Square) to classify flows in the appropriate category.

Once the classification obtained and the layer/flow mapped to the appropriate Markov chain, we should be able to find the appropriate CAC. After each update interval, we compute the updated states probabilities $\pi \_{i, j, k}$ and this for each layers 0, 1, 2. The values of these states probabilities $\pi \_{i, j, k}$ are obtained by multiplying single states by each other ($\pi^L_{i, j, k}$ is $\pi^L_i \pi^L_j \pi^L_k$) and hence we find all the probabilities for all combined states that belong to the same layer.

For the CAC, the decision is done according to the mean bandwidth consumed (rate) and the bandwidth available. Depending on the layer rate this consumed bandwidth in this considered layer $L$ (from 0 to 2) will be:

$$BW^L = \sum_{i=0}^{2}\sum_{j=1}^{2}\sum_{k=0}^{2}\pi^L_{i, j, k} \cdot [rate^L(i) + rate^L(j) + rate^L(k)] \quad (7)$$

This value must not exceed the available bandwidth allowed in the system (for all rtPS connections). This limitation is usually set according to objective parameters like delay, jitter and drop rates and/or to PSNR (Peak Signal to Noise Ratio).

**IV. PERFORMANCE ANALYSIS**

**A. Performance Evaluation**

In this part, we validate the proposed adaptive CAC and we evaluate its performances compared to a non-adaptive CAC. The goal is to show that the proposed scheme allows to admit more video sessions using the same allocated bandwidth and with no QoS degradation, i.e., with a mean QoS that stays under a defined threshold. This means that we can either admit more flows or even transmit more layers for video quality enhancement. At the same time, we show that we can also limit the admission in to avoid degradation while a non-adaptive CAC will not be able to detect this.

We added to the NS2 simulator a IEEE 802.16 module [15]. We considered only the 802.16 rtPS service in which we are able to configure all the parameters of its reservation mechanism. The number of 802.16 MAC frames per bandwidth request was set to 5 and the MAC frame itself was 20 ms duration. This leads to set the quantization scale to 100ms (5 x 20 ms). Moreover, we focus only on the rtPS service with 10 Mbps allocated bandwidth in the MAC frame.

In this purpose, we use SVC video trace files downloaded from [14]. Each file describe for each layer, the data frame type (I, B or P), its size and time the frame is generated. As in [2], we first analyse the statistical behaviour using Matlab. Those videos are various in type, from very high bitrates traffic (soccer games) to medium and low bitrates (with talks/news...)

Then, for each layer of each video session, we are able to find the appropriate parameters and hence classify it in the appropriate flow. Note that we considered only CGS (Coarse Grain Scalability) SVC videos traffic files for our simulations.

An example of a “NBC News” video [14] flow in base layer is given in the figure 5.

Figure 5. Bitrate (blue) and quantized rate (red) for base layer

The measurement process should update the system model parameters, and hence the CAC in term of possible admissions. The measurements show that the convergence corresponds to about 3 minutes of measurements (illustrated by the green line) which is small enough compared to the session duration. After measurements mapping to the states probability functions, the classification is performed.

Below, the measurements for enhancement layer 2.

Figure 6. Bitrate (blue) and quantized rate (red) for enhancement layer 2
Like for the base layer, the measurements mapping to the states probability functions and then classification are performed for enhancement layers. The appropriate adaptive CAC is obtained. Actually, the measurement process continues to update the system in case of strong variability.

The obtained models flow characteristics for “News” video category are given in the table 1 below.

<table>
<thead>
<tr>
<th>Layer</th>
<th>λ</th>
<th>µ</th>
<th>Mean rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Layer</td>
<td>0.48</td>
<td>0.8</td>
<td>0.30 Mbps</td>
</tr>
<tr>
<td>Enhanced Layer 1</td>
<td>0.56</td>
<td>1.4</td>
<td>0.67 Mbps</td>
</tr>
<tr>
<td>Enhanced Layer 2</td>
<td>1.5</td>
<td>2.8</td>
<td>1.94 Mbps</td>
</tr>
</tbody>
</table>

Table 1. Detected parameters of each layer for “News” video

The same tables are obtained for the high rate traffic model that map for example to soccer games and for the medium rates model that is be used to characterise intermediate rates video streams like movies.

B. Adaptive CAC Performances and Results

To evaluate the performance of the proposed adaptive CAC, we used different combinations of traffic flows, with very different statistical behaviours and bitrates, for all the considered layers (from base to eLayer 2).

For each simulation time reference described in the plots, we increase the number of high data rate videos flows and we adapt the CAC in consequence.

We first adapt the CAC in term of sources admission and then in term of enhancement layer admissions. It means that in the first case, more sessions are accepted with the same quality, while in the second case, more layers are added, then better QoS is expected.

Figure 7 illustrates the first case and we clearly see that the CAC average number of session is increased thanks to the adaptation. Compared to the non-adaptive case, the number of admitted sources is increased in all cases, except in the beginning of the simulation where the number of flow is reduced to avoid losses. This is due to the fact that, as the convergence of measures is not completed yet, and to preserve QoS, the adaptive CAC should be restrictive in term of admission.

![Figure 7. Admission for adaptive and non adaptive CAC](image)

However, except the beginning of the simulation, we are able to improve the CAC in term of flow admission. Unlike the static CAC that does not change the admission according the incoming flow traffic characteristics, the proposed adaptive CAC is able to use measurements to admit more session if needed or to restrict the admission to avoid the negative effect on QoS.

In figures 8 and 9 below, we compare the observed performances between adaptive and non adaptive, in case of adaptation in term of enhancement layers.

As explained above, the enhancement is obtained through two mechanisms which are highlighted in the obtained plots.

First, the CAC avoids having admitted flows that would degrade the quality of all admitted sessions. That will, in the other hand increases the number of blocked sessions like explained above, but it will keep the QoS of admitted streams acceptable for users.

Second (starting from the time reference 3.0), as the adaptation concerns the number of enhancement layers, we are able to increase the number of admitted sources. But since also adapting the rate through eLayers admissions, we are also able to reduce packet losses by reducing congestion and hence improve the mean quality of the admitted flows.

![Figure 8. Comparison between packets loss rates](image)

This adaptation is performed thanks to the ability of the proposed system to react to the changes of the statistical behaviour of the admitted flows. Unlike the classical static CAC, the proposed one is able to reduce the session admission if possible and/or the enhancement layers for rate limitation and hence reducing losses. This reduction is significant, from about 2% to 1% in some cases (for some flow repartition configurations).

For the non adaptive case, the packet losses are not reduced significantly as the increased bandwidth is not always able to allow all enhancement layers to be admitted and hence causing packet losses.

The figure 9 gives the video quality results in term of PSNR. We clearly see that the system was able to add more layers in reaction to the bandwidth load. However, thanks to the system measurements and adaptations, we are able to increase the number of eLayers only if this does not affect the...
whole system performance. As expected from the packet losses results, we clearly see an enhancement for the PSNR and hence the quality perceived by the end user.

Like for the losses, the enhancement is particularly observed for some traffic repartitions (up to 3 dB for the PSNR) in where our system is able to reduce losses while maintaining a good CAC at the same time.

V. CONCLUSION AND FUTURE WORK

In this paper we propose an adaptive CAC for enhanced SVC coding based on measurements and classification. This classification gives categories of traffic of similar statistical behaviour and can be done in addition to model parameters refinement. As illustrated by performance results, operator can enhance the QoS by admitting more enhancement layers for example and/or the Grade of Service (GoS) by admitting more flows without changing the mean QoS for users. The improvements are significant (up to 3 dB for PSNR) in term of QoS for users.

In future work, we are working on the impact of the SINR on CAC in IEEE 802.16 rtPS systems. We will then adapt the bandwidth allocation and the CAC, by flow and by layer, according to the radio conditions of each mobile node. Moreover, we are also working on an extension of our work to the 3GPP LTE (Long Term Evolution) and LTE Advanced air interface [16].

REFERENCES