ON EXCESS PROCESS POLICING IN B-ISDN NETWORKS*

Finn Arve Aagesen** Bjarne Emil Helvik***

In a B-ISDN network a contract between a subscriber and the network is fundamental for the optimum allocation of resources and the supervision of the utilization. Various aspects of contracts and policing is discussed, and a policing scheme denoted as the Excess Process Policing Scheme (EPPS) is presented. EPPS is based on the policing of excess counts related to various independent time-resolution regions of the source processes.

For image-related services, the frame is an important physical quantity, reflected in the source processes. The inter-frame processes are the average and the smoothed frame processes, and the intra-frame processes are the burst and extreme processes. Inter-frame excess-statistics for some real-time image sources are presented, and the application of EPPS for policing of the inter-frame aspects of these sources is considered. The policing of intra-frame processes is also discussed.

1. INTRODUCTION

The traffic flow management for the individual virtual circuits in a B-ISDN network may be based on flow control or cell policing. Flow control will act as a back-pressure mechanism, while the policing function will mark and possibly discard cells.

The primary objective of the policing function is to prevent the users from violating the agreed contract. The definition and introduction of a policing function is in no way straight-forward. The obstacles are of both conceptual and technical nature. Concerning the conceptual ones, the contract as well as the joint operation with other traffic-related functions has to be defined. The technical problems are primarily related to the fact that the complexity and the cost must be kept very low, because the policing is on a VC basis.

A policing scheme is an implementation-independent definition of the policing function, while a policing mechanism is a description of how the policing function is carried out. The policing-related works have been focusing on policing mechanisms such as LB (Leaky Bucket), EWMA (Exponential Weighted Moving Average) a.o. Mechanism descriptions and some analysis results are found in [HEMM89], [RATH90] and [MONT90].

This paper is concerned about the conceptual aspects of policing, with focus on real-time image sources. In Sec. 2, some general aspects of policing is discussed. In Sec. 3, the Excess Process Policing Scheme (EPPS) is presented. In Sec. 4, the application of EPPS is discussed, based on statistics from variable bit-rate coded real-time image sources.

2. POLICING - SOME GENERAL CONSIDERATIONS

2.1. Policing contracts

The contract comprises within this paper the total set of "policing-related" quantities negotiated between the source and the network. A contract may be based on:

- the source parameters, such as cell intensity during a burst, the duration of a burst, the mean cell intensity etc.
- the policing mechanism parameters, such as Leaky Bucket "leak rate" and "maximum bucket size".
- traffic quantities that are related to the source, but that is "decoupled" from the source parameters as well as the policing mechanism parameters.

We are in favour of the third type of contract, which means that the contract is the given reference for the source as well as the policing function. This allows a contract independent of specific sources as well as policing mechanisms. Hence an uncoloured source must adjust to the contract. This adjustment is denoted as source inconvenience.

2.2. Time resolution aspects of source processes

The applications are here classified as real-time, interactive and computer-computer traffic. An often applied scheme for characterizing the source processes, is to apply models at the connection, dialog, burst and the cell levels [HUI88]. Here a source is described by its extreme, burst, (smoothed) frame, and average process. These processes describe the behaviour of the sources in various time resolution regions as indicated in Fig. 2.1.

---

** The work reported in this paper is done under contracts with Norwegian Telekom, Research Department

** Norwegian Institute of Technology, N-7034 Trondheim Norway. E-mail: aagesen@idt.unit.no

*** ELAB-RUNIT, N-7034 Trondheim Norway. E-mail: helvik@elab-runit.sintef.no

---
The concept "frame" is here a generalization of the concept video-frame, which is typically between 25 and 40 msec. Frame-time is generally interpreted as a time for presenting a quantity of information. For general image sources, the frame-time is defined as the time for presenting one terminal screen. For voice, the frame concept is not so intuitive, but can be related to the time for presenting units of the speech that has meaning for the intelligibility, i.e. phonems. For computer-computer-traffic, the frame-time has no direct physical interpretation.

The (smoothed) frame process is the source behaviour during consecutive frames. The extreme and burst processes are defined as intra-frame processes, and the frame and average processes as the inter-frame processes. For real-time image traffic, the inter-frame processes are determined by the image scene process, the screen and the video codec. For the intra-frame processes, there is a high degree of freedom in how the burst of cells are sent within a frame [ROBE90].

2.3. Flow management alternatives

Table 2.1 illustrates flow management alternatives for the various time resolution aspects of various traffic classes. Real-time traffic needs policing for all its time regions. It is our feeling that flow control should be applied where possible. It is a self-stabilizing feedback mechanism without loss.

<table>
<thead>
<tr>
<th>PROCESS TRAFFIC</th>
<th>Extreme, Burst, Frame</th>
<th>Average:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real-time</td>
<td>Policing</td>
<td>Policing</td>
</tr>
<tr>
<td>Interactive</td>
<td>Policing</td>
<td>Policing/Flow control</td>
</tr>
<tr>
<td>Computer-computer</td>
<td>Policing</td>
<td>Policing/Flow control</td>
</tr>
</tbody>
</table>

Table 2.1. Flow management alternatives

The policing of the average aspects of real time traffic should be done softly, i.e. the contract should be dynamically modified in accordance with the observed traffic. This for the following reasons: 1): The average value is not easy to predict. 2): The coding parameters and the codec operation are on a frame basis. Intra-frame and frame-to-frame operations can be adjusted by short-time quality changes, but the policing of the deviation from a contractual average value can involve a huge amount of cells. 3): The time resolution of the average process may allow such a negotiation.

2.4. The policing function environment

The policing function must cooperate with other traffic-related functions as indicated in Fig. 2.2. In this constellation, the agreement has a vital position. The agreement is established by the call acceptance control function, and comprises agreement elements (deals) related to policing, grade of service, charging and flow control. The agreement element P is the policing contract, G is a set of grade of service parameters, C the agreed charging principle and F is the flow control scheme and parameters.

The quality of the policing scheme is dependent on the ability to prevent speculative misuse of the contract. Soft policing needs cooperation between the policing and the call acceptance control function.

3. THE EXCESS PROCESS POLICING SCHEME (EPPS)

An excess process is the process observed when the traffic process exceeds some predefined limit. EPPS is based on independent policing of excess processes related to the time resolution regions defined in Sec. 2.1. Sec. 3.1 explains the basic principle, and Sec. 3.2 is discussing the contents of the policing contracts.

3.1. The excess process policing principle

Denote the considered time resolution region for region Q, where index $Q \in \{E,B,F,A\}$ indicates the extreme, burst, frame and the average process-regions, respectively. Let $X_Q$ be the excess process duration. An excess process duration example as illustrated in Fig. 3.1 is given below. First, the following quantities related to process-region Q are introduced:
\[ t_0 \]: the life-time variable for process \( \Omega_0 \).
\[ N_0(t_0) \]: the accumulated cell count after an elapsed life-time \( t_0 \).
\[ \lambda_0 \]: the contractual input rate.
\[ \Omega_0(t_0) \]: the excess process, i.e. the accumulated excess count after an elapsed life-time \( t_0 \).
\[ \lambda_0(t_0) = [N_0(t_0) - \Lambda_0(t_0)]^* \].
\[ L_0(t_0) \]: the upper limit for the excess process, i.e. the maximum deviation between the observed count and the contractual count after an elapsed life-time \( t_0 \). See (3.1) below.
\[ \Lambda_0(t) \]: the instantaneous arrival rate at time \( t \).
\[ T_0 \]: the interprocess period.

The definition of the instantaneous arrival rate, \( \Lambda_0(t) \), is related to the specific time regions. All basic observations are related to cells. Let:

\[ T_0 \]: the interarrival time between cells \( j-1 \) and \( j \).
\[ N_0^j(t) \]: the number of cell arrivals during time \( t \).

The arrival rate for the intra-frame processes is defined at the instants of cell arrivals as the inverse of the cell interarrival time, i.e. \( \Lambda_0(t) = 1/T_0^j \), \( Q = E,B \). For the inter-frame processes, the arrival rates are the number of arrivals per smoothing interval \( T_0 \), i.e. \( \Lambda_0(t) = (N_0(t) - N_0(t-T_0^j)) / T_0^j \); \( Q = F,A \), where \( t \) now take discrete values \( 1 \leq t \leq 1, 2, ... \).

The excess process \( \Omega_0(t_0) \) is considered in Fig. 3.1. The cell arrival rate is observed, and the excess period is started \( (t_0=0) \) when the rate is higher than the contractual rate. In this example the excess period is finished when \( \Omega_0(t_0) \) goes to zero.

\[ \Omega_0(t_0) = [N_0(t_0)-\Lambda_0(t_0)]^* \leq L_0(t_0); \ 0<t_0 \leq X_0+T_0 ^* \] (3.1)

\( L_0(t_0) \) is the limiting quantity (or factor). If \( L_0(t_0) \) is a constant \( (L_0^*) \), then (3.1) is denoted as a \( LB \) (Leaky Bucket) scheme for process-region \( Q \) with parameters \( \lambda_0 \) and \( L_0^* \).

3.2. The EPPS generic contract

The relation (3.1) can now be used for all time regions, with the appropriate excess period duration definition. The following quantities related to the smoothing of the inter-frame processes are defined:

\[ t_F \]: the duration of a frame
\[ t_A \]: the average process smoothing interval
\[ \Lambda_0 \]: the frame excess process duration counter
\[ \Lambda_0 \]: the average process smoothing interval counter

The conditions for the excess process duration might be defined differently from the case in Fig. 3.1. In the example, the observations are based on stepping intervals. The observations might be based on jumping intervals as well, i.e. no consideration of the excess duration. In the stepping interval case, there might be different conditions related to the excess period stop condition. Within EPPS, the duration \( X_0 \) is simply the call duration, and \( T_0 \) is the time between calls, i.e. the observation of the average process is based on jumping intervals. For the other processes, stepping intervals are assumed in this conceptual consideration.

The policing contract \( P \) defined in Sec. 2.4 can be structured as follows:

\[ P_1 \]: Traffic source intensity quantities.
\[ P_2 \]: Time and process duration quantities.
\[ P_3 \]: Limiting quantities introduced for policing.
\[ P_4 \]: Observed and policed quantities.

The relationship (3.1) contains the following set of generic contract quantities:

\[ \{ \lambda_0; \ Q = E,B,F,A \} \]
\[ \{ t_F; t_A; X_0; \ Q = E,B,F,A \} \]
\[ \{ N_0(t_0); \ Q = E,B,F,A \} \]
\[ \{ N_0(t_0); \lambda_0(t_0); \ Q = E,B,F,A \} \]

4. EXCESS PROCESS POLICING OF REAL-TIME IMAGE SOURCES

This section will discuss the applicability of EPPS based on observations of real-time image sources, and with emphasis on the inter-frame processes, to which the observations are related.

4.1. Inter-frame excess statistics

The statistics are based on measurements made available by Dr. W. Verbiest at Bell Alcatel. The videocodec as well as some results are presented in [VERB88] and [VERB89]. The statistics are based on four applications: A: Video-phone, B: Video-conference, C: A TV-serie-video and D: A Sport-video.

The results are concerned about excess characteristics. In addition to the excess count \( \Omega_0(t_0) \), the excess rate \( \gamma_0(t) = \Omega_0(t_0)/t_0 \), \( \Lambda_0(t_0) / t_0 \), and excess factor \( B_0(t) = \Omega_0(t_0)/L_0 \), are introduced for a convenient presentation of the statistics.
Let $Z$ be a general stochastic quantity. The operators: $E(Z)$, $\sigma(Z)$, $\text{Max}(Z)$, $c(Z)$ and $b(Z)$ are the mean, the standard deviation, the maximum, the coefficient of variation ($\sigma(Z)/E(Z)$) and burstiness ($\text{Max}(Z)/E(Z)$) of $Z$, respectively.

Arrival intensity statistics $\Lambda_F$ are given in Table 4.1. The frame duration $t_F = 40$ msec.

<table>
<thead>
<tr>
<th></th>
<th>Video-phone</th>
<th>Video-conf.</th>
<th>TV-serie</th>
<th>TV-sport</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(\Lambda_F)$</td>
<td>4.38</td>
<td>4.23</td>
<td>17.12</td>
<td>29.57</td>
</tr>
<tr>
<td>$c(\Lambda_F)$</td>
<td>0.44</td>
<td>0.36</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>$b(\Lambda_F)$</td>
<td>6.27</td>
<td>6.50</td>
<td>1.87</td>
<td>1.34</td>
</tr>
</tbody>
</table>

Table 4.1. Frame process arrival intensity (Intensities in Mb/s/sec)

The figures 4.1-4.2 are related to the frame process excess intensity. As contractual intensity is used $\lambda_F = E(\Lambda_F)$. The excess period stop condition was defined as: $\Lambda_F(t_F^*) - \lambda_F \leq 0$. Figure 4.1 shows the basic curves $E(\beta_F(t_F))$ and $\text{Max}(\beta_F(t_F))$, for application A, which is Video-phone. Figure 4.2 shows $\text{Max}(\beta_F(t_F))$ for all the applications A-D.

Fig. 4.3 shows the average process $\text{Max}(\beta_A(t_A))$. The smoothing intervals $\tau_A = t_F = 40$ msec and the contractual intensities $\lambda_A = E(\Lambda_A) = E(\Lambda_F)$.

For both of the inter-frame processes, the maximum excess factor can be expressed by the burstiness of the frame intensity as follows:

$$\text{Max}(\beta_A(t_A)) = b(\Lambda_F) - 1 ; \text{Q} = \text{F, A} \quad (4.1)$$

The influence of the characteristics of $\Lambda_F$ on $\beta_A$ is explained by the fact that $\tau_A = t_F$, i.e. $\lambda_A = \lambda_F$ in this presentation.

4.2. The average process policing

What concerns the limiting quantity, $L_A(t_A)$, it can be set by various criteria. By determining the quantity based on the maximum value, i.e.

$$L_A(t_A) = \text{Max}(\beta_A(t_A)) \lambda_A \Lambda_A \quad (4.2)$$

we will have an average process policing scheme with no inconvenience related to the present observation material. Fig. 4.4 shows the limiting quantity $L_A(t_A)/\lambda_A$. It seems as $L_A(t_A)$ approaches a limiting constant value. If so, the maximum excess count is independent of time when $t_A$ is of a certain duration. This limiting value $L_A^*$ is seemingly nearly proportional to the contractual rate, i.e. $L_A^* = c_A \lambda_A$. The constant $c_A$ is being quite close for the various applications.
A simplification is to apply one limiting value $L_A(t_A) = L_A^*$ for all $t_A$, i.e.:

$$L_A^* = \text{Max}(L_A(t_A); t_A \leq X_A)$$  \hspace{1cm} (4.3)

This is a LB scheme with parameters $\lambda_A$ and $L_A^*$. What is so the potential power of the average process policing, and what is the consequence of choosing the LB scheme? Concerning the power, the maximum time-dependent deviation from the contractual value is the value of $\text{Max}(\beta_A(t_A))$ illustrated in Fig. 4.3. The average process policing cannot be better, unless limiting values lower than given by (4.2) is chosen. This will cause inconvenience. So the curves in Fig. 4.3 reflect what is attainable without the adjustment of the source patterns. All applications have $\text{Max}(\beta_A(t_A))$ less than 0.1 for $t_A > 180$ sec.

![Figure 4.4](image)

Fig. 4.4. The normalized limiting quantity $L_A(t_A)/\lambda_A$. (In sec)

The suggested scheme has a possibility for misuse, especially for $t_A < 60$ sec. Misuse can be avoided by charging a session with length $X_A$ for the actual cell value $N_A(x_A)$, rather than charging by the count calculated from the contractual rate, i.e. by $\lambda_A x_A$.

The omission of the time dependency in $L_A$ has minimal impact. Let:

$\Psi_b$: time to rise to the limit by burst intensity

$\Psi_c$: time to rise to the limit by contract intensity.

These time durations can generally be expressed as: $\Psi_b = L_A^*/\lambda_A b(\lambda_A)$ and $\Psi_c = L_A^*/\lambda_A$, and for the existing statistics $\Psi_b = c_{\text{A}} b(\lambda_A)$ and $\Psi_c = c_{\text{A}}$. For application A this give $\Psi_b = 3.4$ sec and $\Psi_c = 21.6$ sec. Only one rise and fall period can be accomplished before the time region for the appropriate value of $L_A^*$ is reached. Similar results are obtained for the other applications.

4.3. The smoothed frame process policing

Note that the excess process stop condition was defined by zero excess intensity in this case. Limiting values for the smoothed frame process can be found in the same way as for the average process. As is seen from Fig. 4.1, $E(\beta_f(t_f))$ is increasing with the excess time $t_f$. However, the maximum value of the excess factor has a decreasing tendency. So applying $\text{Max}(\beta_f(t_f))$ for determining limiting values with accordingly small inconvenience, is a possible approach.

Equation (4.2) with indexes "F" can be applied for studying the limiting quantity $L_F(t_f)$. Because the curves in Fig. 4.2 are rather flat, the limiting quantity will be roughly linearly increasing with $t_f$.

A three-level scheme can be applied. These levels reflect the maximum, the typical and the zero excess intensity. The limiting factors for the levels are:

$$L_F(\kappa F t_F) = (b(\lambda_F) - 1) \lambda_F \kappa F t_F ; \text{ for } \kappa_F = 1, 2$$  \hspace{1cm} (4.4.a)

$$L_F(\kappa F t_F) = (b(\lambda_F) - 1) a \lambda_F \kappa F t_F ; \text{ for } 2 < \kappa_F \leq K_F$$  \hspace{1cm} (4.4.b)

$$L_F(\kappa F t_F) = L_F(K_F t_F) ; \text{ for } \kappa_F > K_F$$  \hspace{1cm} (4.4.c)

The values for applications A-D for the parameter "a" are found to be 0.43, 0.24, 0.57 and 0.74 and for $K_F$ to be 12, 12, 44 and 44 sec, respectively. The power of the scheme is seen from Fig. 4.2. For utterly limitation of the frame process, inconvenience must be introduced.

A LB scheme is also possible in this case. A limiting value is defined by the maximum excess count, and assume that the maximum excess intensity during a frame is $\lambda_f(b(\lambda_A) - 1)$. Let the time to reach the limiting value for the three-level scheme be $\Psi_c$, and for the LB scheme $\Psi_b$. The ratio between these can then be approximated by:

$$\Theta = \Psi_b/\Psi_c = a$$  \hspace{1cm} (4.5)

A LB scheme can rise 1/a times faster than the three-level scheme. The variance allowed is significantly in disfavour of the LB scheme.

4.4. On the burst and extreme process policing

The burstiness of the cell arrival intensities are important for determining the limiting quantities. In the discussion of the inter-frame processes, we used $\lambda_f = \lambda_f$ and $t_f = t_f$, and accordingly one burstiness $b(\lambda_f)$ was defined. Let in a similar way the intra-frame processes be represented by one contractual rate $\lambda_f$ and burstiness $b(\lambda_f)$. The intensity and burstiness related to the intra-frame processes, i.e. the burst and the extreme process is then defined as:

$$\{ [\lambda_f, b(\lambda_f)], [\lambda_e, b(\lambda_e)] \} = \{ [\lambda_f, b(\lambda_f)], [\lambda_f b(\lambda_f), 1] \}$$

A tempting choice is $\lambda_f = \lambda_f b(\lambda_f)$. The values of $b(\lambda_f)$ for the presented statistics are between 1 and 7. If the maximum rate of 160 Mbit/sec is used during consecutive cells, the intra-frame process burstiness will be in the same range as the inter-frame process burstiness.

The extreme process policing is now expressed by the following relationship:

$$N_E(t_E) - \lambda_f b(\lambda_f) t_E \leq 1 ; 0 \leq t_E \leq X_E$$  \hspace{1cm} (4.6)
The policing of the maximum number of cells during a frame time is done by:

\[ N_C(t_F) - \lambda_F b(\Lambda_F) t_F \leq 0 \]  

(4.7)

The relations (4.6) and (4.7) represents the "upper" and "lower boundary" on the cell pattern, respectively. If there shall be utterly control of the detailed pattern, there must also be a contract on the pattern. One possibility is of course to agree on a policing mechanism as for instance a Leaky Bucket with defined parameters rather than the pattern, and to give the source the freedom that this mechanism allows. However, two policing schemes denoted as the stretching scheme and regularity scheme are defined below. These schemes will give a simple intra-frame policing with the pattern controlled by the extreme process policing. The extreme process policing relationship (4.6) is replaced by a new policing relationship, taking care of the cell pattern. The relationship (4.7) is unchanged.

When using the stretching scheme, the source must send the total frame content equally stretched within the frame. This means that the extreme rate never will be higher than the maximum frame rate, i.e. the relationship (4.6) turns to:

\[ N_E(t_E) - \lambda_E b(\Lambda_E) t_E \leq 1 ; \quad 0 \leq t_E \leq X_E \]  

(4.8)

Under the regularity scheme, the cells are sent with a rate \(\lambda_E\) independent of the frame rate \(\Lambda_F\), and the extreme process policing is expressed by:

\[ N_E(t_E) - \lambda_E t_E \leq 1 ; \quad 0 \leq t_E \leq X_E \]  

(4.9)

Since the real-time requirements of the sources are related to the frames, imposing the stretching or the regularity scheme does not need to negatively influence the service as seen from the users. At the same time these schemes represent smooth traffic streams offered to the network.

5. CONCLUSIONS

Policing and flow control are the alternative flow management schemes to be applied on individual VCs in B-ISDN networks. For real-time sources, policing must be used for all the time resolution regions. All flow management may give an inconvenience, and any policing scheme must be based on the trade-off between inconvenience for uncoloured sources and insufficient policing.

An Excess Process Policing Scheme (EPPS) has been presented, based on independent policing of various time resolution regions of the source process-ess. While the inter-frame processes reflect the physical applications, there is a seemingly freedom related to the specific generation of the cell pattern of the intra-frame processes. The contract for the inter-frame policing must be based on the physical characteristics of the sources. For the intra-frame process policing, there is seemingly a question of decision.

Based on inter-frame statistics from real-time image sources, EPPS for the inter-frame process region has been considered. The limiting factor can for the average process be independent of time and for the smoothed frame process proportional to the elapsed excess time. EPPS is more flexible and more related to the sources than the Leaky Bucket scheme, and its potential policing power is accordingly higher.

The intra-frame policing can be based on an agreed cell pattern within frames. If either a regular scheme or a stretching scheme is agreed, the intra-frame policing is simple.

REFERENCES