Predictive Bandwidth Control for GEO Satellite Networks

L. Chisci\textsuperscript{1}  R. Fantacci\textsuperscript{2}  T. Pecorella\textsuperscript{3}

\textsuperscript{1}Dpt. di Sistemi e Informatica
Universit\`a di Firenze, Italy.

\textsuperscript{2}Dpt. di Elettronica e Telecomunicazioni
Universit\`a di Firenze, Italy.

\textsuperscript{3}CNIT - Unit\`a di Firenze, Italy.

e-mail: chisci@dsi.unifi.it, \{fantacci,pecos\}@lenst.det.unifi.it

The IEEE Int. Conf. on Communications, Paris, 20 - 24 June 2004
1. Satellite Systems
2. Quality of Service (QoS)
3. Dynamic Bandwidth Allocation (DBA)
   - DBA Scheme
   - Satellite Gateway (SG)
4. Results
5. Conclusions
Satellite Networks

Pros
- Easy deployment
- Cost-effective
- Mobility
- Broadcasting
- Geographical coverage

Cons
- In GEO links propagation delay is very high (RTT \(\simeq 500\) ms).
- High delay*bandwidth systems are critical for TCP flows.
- The bandwidth is scarce, so a careful management is needed.

Chisci, Fantacci, Pecorella
Predictive Bandwidth Control for GEO Satellite Networks
Quality of Service (QoS)

QoS metrics
- Bandwidth
- Delay
- Jitter
- Packet Loss

<table>
<thead>
<tr>
<th>Application</th>
<th>Required Band</th>
<th>Sensitivity to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Delay</td>
<td>Jitter</td>
</tr>
<tr>
<td>VoIP</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Web</td>
<td>Avg.</td>
<td>Avg.</td>
</tr>
<tr>
<td>E-mail</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

Internet Engineering Task Force (IETF) proposals
- Integrated Services (IntServ)
- Differentiated Services (DiffServ)

Chisci, Fantacci, Pecorella
Quality of Service (QoS)

**QoS metrics**
- Bandwidth
- Delay
- Jitter
- Packet Loss

**Application**

<table>
<thead>
<tr>
<th>Application</th>
<th>Required Band</th>
<th>Sensitivity to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Delay</td>
</tr>
<tr>
<td>VoIP</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Web</td>
<td>Avg.</td>
<td>Low</td>
</tr>
<tr>
<td>E-mail</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

Internet Engineering Task Force (IETF) proposals

- Integrated Services (IntServ)
- Differentiated Services (DiffServ)
In order to compute the bandwidth requests, each SGs use 2 blocks:

- Traffic Predictor
- Bandwidth Controller
Buffers (FIFO queues)

Input packets are divided into $n$ buffers (FIFO queues), according to the DiffServ policy. (e.g. EF, AF, BE classes $\Rightarrow n = 3$)

$\mathbf{y}(t) \in \mathbb{R}^n$ vector of queue sizes
$\mathbf{w}(t) \in \mathbb{R}^n$ vector of input flows
$\mathbf{v}(t) \in \mathbb{R}^n$ vector of output flows

\[
\mathbf{y}(t + 1) = \mathbf{y}(t) + \mathbf{w}(t) - \mathbf{v}(t)
\]
Scheduler

\( v(t) \): bandwidth assigned at time \( t \) to the SG from the NCC

\( v(t) \in \mathbb{R}^n \): transmission flows at time \( t \) for the \( n \) service classes

\[
 v(t) = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad v(t), \quad b_i \geq 0, \quad \sum_{i=1}^{n} b_i \leq 1
\]

Several scheduling policies can be adopted, e.g.
- WFQ - Weighted Fair Queueing
- Priority Scheduling
- Weighted Round Robin
Scheduler

\( v(t) \): bandwidth assigned at time \( t \) to the SG from the NCC

\( v(t) \in \mathbb{R}^n \): transmission flows at time \( t \) for the \( n \) service classes

\[
v(t) = \begin{bmatrix}
    b_1 \\
    b_2 \\
    \vdots \\
    b_n
\end{bmatrix}
\]

\( v(t) \), \( b_i \geq 0 \), \( \sum_{i=1}^{n} b_i \leq 1 \)

Several scheduling policies can be adopted, e.g.

- **WFQ** - Weighted Fair Queueing
- **Priority Scheduling**
- Weighted Round Robin
Traffic Predictor

\( \hat{\mathbf{w}}(t + k|t) \): predictions, at time \( t \), of the input traffic flows at time \( t + k \), based on the available data \( \mathbf{w}^{t-1} \triangleq \{\mathbf{w}(k), k \leq t - 1\} \).

Several predictors can be used:

- **AR** predictor (*not self-similar*)
- **ARMA** predictor (*not self-similar*)
- **FARIMA** predictor (*self-similar*) based on a fractionally integrated ARMA model
- Non parametric statistical predictors
\( \hat{w}(t + k|t) \): predictions, at time \( t \), of the input traffic flows at time \( t + k \), based on the available data \( w^{t-1} \triangleq \{ w(k), k \leq t - 1 \} \).

Several predictors can be used:

- **AR predictor (not self-similar)**
- **ARMA predictor (not self-similar)**
- **FARIMA predictor (self-similar)** based on a fractionally integrated ARMA model
- Non parametric statistical predictors
The bandwidth request $u(t)$ for the time $t + \delta$, where $\delta$ is the Round Trip Delay, is based on the data already known at time $t$ and assuming that the request is always satisfied, i.e.

$$v(t) = u(t - \delta)$$

state: $x(t) = [y'(t), u(t-1), \ldots, u(t-\delta)]' \in \mathbb{R}^{n+\delta}$

Predictive model for time evolution of queue sizes:

$$\begin{cases}
  x(t + k + 1) = Ax(t + k) + Bu(t + k) + Ew(t + k) \\
  y(t + k) = Cx(t + k)
\end{cases}$$

Predictions $\hat{w}(t + k | t)$ are used in place of $w(t + k)$, $k \geq 0$. 
\( u(t) \) is chosen to trade-off small queue sizes (i.e. **low traffic delays**) vs. **bandwidth waste**

\[ J = \sum_{k=0}^{H-1} \left[ y'(t + \delta + k) Q y(t + \delta + k) + r u^2(t + k) \right] \]

\( Q = Q' \geq 0, \quad r > 0 \)

subject to the constraints:

\[
\begin{align*}
y_{\min} & \leq y(t + \delta + k) \leq y_{\max} \\
u_{\min} & \leq u(t + k) \leq u_{\max}
\end{align*}
\]

\( k = 0, 1, \ldots, H - 1 \)
Receding Horizon control strategy

Receding Horizon control is used to compensate the model uncertainties.

- Being $u(t), \ldots, u(t + H - 1)$ the optimal solution at time $t$
- we apply the request $u(t)$ discarding $u(t + 1), \ldots$
- at time $t + 1$ we repeat the process with the new data.
Other strategies

**Fixed allocation**

- **Med**: 110% of the average input bandwidth
- **Max**: 80% of the maximum input bandwidth

**Reactive controller** [F. Delli Priscoli and A. Pietrabissa, CDC 2002]

The bandwidth request is computed using a Smith Predictor:

\[
\begin{align*}
  u(t) &= w(t) + K \left[ y(t) - \sum_{k=t-\delta}^{t-1} \left( u(k) - \left( 1 - \frac{\delta^*}{\delta} \right) w(k) \right) \right]
\end{align*}
\]

where \( \delta^* \geq \delta \) is a parameter representing the desirable queueing delay while \( K \) is a gain which, for stability, must belong to the interval \((0, 1]\).
### Simulation and Control parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round Trip Delay ($RTD$)</td>
<td>500 [ms]</td>
</tr>
<tr>
<td>Number of input flows</td>
<td>3 (EF, AF, BE)</td>
</tr>
<tr>
<td>Average bandwidth 1 ($w_1$)</td>
<td>14.124 [kbit/frame]</td>
</tr>
<tr>
<td>Average bandwidth 2 ($w_2$)</td>
<td>6.693 [kbit/frame]</td>
</tr>
<tr>
<td>Average bandwidth 3 ($w_3$)</td>
<td>7.563 [kbit/frame]</td>
</tr>
<tr>
<td>$q_1$, $q_2$, $q_3$</td>
<td>0.1786, 0.1428, 1.0</td>
</tr>
<tr>
<td>$r$</td>
<td>1.0</td>
</tr>
<tr>
<td>Sampling period</td>
<td>125.0 [ms]</td>
</tr>
<tr>
<td>$\delta = \text{sampling period} / \text{RTD}$</td>
<td>4</td>
</tr>
<tr>
<td>Prediction horizon $H$</td>
<td>4</td>
</tr>
<tr>
<td>Order of AR model $h$</td>
<td>4</td>
</tr>
<tr>
<td>$y_{min}, y_{max}$</td>
<td>0, $\infty$ [bit]</td>
</tr>
<tr>
<td>$u_{min}, u_{max}$</td>
<td>0, 2 [Mbit/frame]</td>
</tr>
</tbody>
</table>
Comparable bandwidth waste between Med, RHC and SPC (not depicted)

Max policy exhibits huge wastes
Queue length

- RHC outperforms Med and SPC
- Q parameters allow fine-tuning of RHC performance
Queue length for different QoS classes

- DiffServ QoS is supported by differentiating queue length.
Queue length after a (simulated) congestion

1 Mbit initial values of queue lengths
- queues are served according to the QoS priority
- bandwidth is released only after the congestion recovery
Conclusions I

DBA for GEO satellite systems based on

- adaptive predictor for the input traffic flow
- formulation of DBA as an optimal control problem with cost trading off queue size vs. bandwidth waste
- receding-horizon strategy for bandwidth request generation

Interesting results in terms of

- handling multiple services with differentiated QoS requirements
- low bandwidth waste
- high QoS (low delays and packet loss probability)
Conclusions I

DBA for GEO satellite systems based on
- adaptive predictor for the input traffic flow
- formulation of DBA as an optimal control problem with cost trading off queue size vs. bandwidth waste
- receding-horizon strategy for bandwidth request generation

Interesting results in terms of
- handling multiple services with differentiated QoS requirements
- low bandwidth waste
- high QoS (low delays and packet loss probability)
Future developments

- adoption of more realistic “self-similar” models (e.g., FARIMA) for traffic prediction
- distributed DBA schemes taking into account the presence of multiple SGs competing for the bandwidth resource.