Cross-Layer PEP-Spoofer Approach to improve TCP Performance in DVB-RCS Networks

Giovanni Giambene, Snezana Hadzic

CNIT - Research Unit of the University of Siena
Via Roma, 56, 53100 Siena, Italy

giambene@unisi.it
Summary

- Introduction to cross-layer design and *Broadband Satellite Multimedia* (BSM) standard
- Survey on *Performance Enhancing Proxies* (PEP)
- Cross-layering for PEP referring to the BSM model.
- Case-study for a DVB-RCS scenario (interactions PHY-MAC-TCP)
- Concluding remarks.
Introduction on cross-layer design and BSM standard
DVB-S2/-RCS

DVB-S2 is second-generation standard for forward link in broadband satellite networks

- Transmission is organized in blocks of 25 MHz of Ka or Ku bands
- Adaptive Coding and Modulation (AMC) for dynamic change of ModCod level depending on the channel conditions by means of $E_b/N_0$.
- ModCod thresholds are determined to have Frame Error Rate (FER below $10^{-7}$)

DVB-RCS specifications allow return path for satellite networks based on DVB-S-S2

- Return channel dynamically assign its time-frequency resources Multi-Frequency-Time Division Multiple Access, (MF-TDMA, air interface)
Cross-layering

The layered OSI approach for air interface design is based on the separate optimization of distinct protocols, thus reducing the complexity and allowing the interoperability among equipments of different manufacturers through the use of standardized interfaces.

Due to the dynamic nature of the radio channel, there exists tight interdependence between layers in satellite networks; hence, a strict modularity and layer independence may lead to a sub-optimal performance.

The interest is here on architectures where the reference layered architecture is enriched with interactions between protocols at non-adjacent layers. This is what is meant for cross-layering.
Characteristics of cross-layering

- Direct exchange of the control information among non-adjacent layers of OSI stack concept.

- Techniques to support exchange of signaling
  - *In-band signaling*, using enriched packet headers to notify internal state variables to other layers.
  - *Out-of-band signaling*, via the control plane using new primitives and suitable SAP.

- The coordination of signaling could be done by a protocol layer (i.e., MAC) or some external element.

- Signaling could be also classified as internal, (i.e., within a given network element) or external, (i.e., between two network elements)
**Possible cross-layer interactions (SatNEx II – ja2230)**

- **Downward signaling**: application and codec type, Quality of Service (QoS) requirements, priority, protocol type and internal protocol state;

- **Upward signaling**: propagation conditions, handover preparation measures, congestion notification, policing updates, and application scaling request.

Workshop on Satellite PEPs
Current Status and Future Directions, December 2, 2008, ESA, Noordwijk, Netherlands
Reference scenario based on BSM and DVB-S2/-RCS
BSM protocol stack

<table>
<thead>
<tr>
<th>External Layers</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UDP</td>
</tr>
<tr>
<td></td>
<td>TCP</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Satellite Independent</th>
<th>IPV4 / IPV6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Satellite Independent Adaptation Functions</td>
</tr>
<tr>
<td></td>
<td>SI-SAP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Satellite Dependent</th>
<th>Satellite Dependent Adaptation Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Satellite Link Control (SLC)</td>
</tr>
<tr>
<td></td>
<td>Satellite Medium Access Control (SMAC)</td>
</tr>
<tr>
<td></td>
<td>Satellite Physical (SPHY)</td>
</tr>
</tbody>
</table>
Reference scenario

The scenario is DVB-S2 (forward channel with ModCod adaptation) and DVB-RCS (return channel with DAMA) control by the NCC.

Our interest is on the return path:
- Satellite Terminals (STs) servers that send TCP flows towards the network.
- NCC dynamically assigns resources

PEP spoofing functionality is used at the NCC to regulate traffic injection of TCP flows, by freezing its cwnd value.
BSM layer 3 and layer 2 architecture (U- and C-plane)
L2 and L3 interactions through SI-SAP

Referring to an IP-based satellite network, it is commonly considered that at the IP level between 4 and 16 DiffServ queues are manageable to support different IP traffic classes.

While, these queues can be mapped into 2-4 MAC layer queues.

QoS support is achieved by active queue management at IP layer and by suitably mapping L3 queues to MAC ones (QIDs are used for mapping purposes; MAC layer queues are only used as FIFO transmission buffers).
Primitives for SI-SAP signaling between MAC and L3

The primitives that are used by the “L3 queue manager” for QIDs management are: SI-C-QUEUE_OPEN-***, SI-C-QUEUE_MODIFY-***, SI-C-QUEUE_CLOSE-***.

Moreover, SI-C-QUEUE_STATUS-*** is a cross-layer primitive for the interaction on the C-plane between the “L3 queue manager” and the layer 2 Satellite Terminal QID Resource Manager (STQRM).

STQRM functions are: mapping of QIDs, control of SD resource availability and allocation; STQRM also translates primitives arriving at the SI-SAP into lower layer primitives and vice versa, if necessary.

Each primitive can be of different types and for each type suitable parameters are set.

SI-C-QUEUE_STATUS-*** can be of different types (*** = -req, -cfm, and -res); this primitive is typically triggered by the “L3 queue manager” and can be used for MAC - L3 cross-layer exchange within a given terminal (use of req & cfm) for interactive C-plane signaling; however, it can be also employed for signalling with a remote terminal (*** = -res) for peer-to-peer C-plane signaling.
Exchange of signaling among non-adjacent layers and BSM

Existing BSM primitives permit interactions between layer 2 and layer 3

Innovative design with new primitives among non-adjacent layers is needed and some options already exists

- Concatenated primitives at different layers, but this approach does not have proper cross-layer meaning.
- Broadcast primitives in upward and downward direction on the “bus” shared and available to all the layers in the stack.

The following study focuses on adopting new combined approaches like PEP and cross-layer design in order to improve TCP performance in DVB-RCS satellite networks compliant with the BSM.
Survey on main PEP characteristics
PEP taxonomy

PEPs are network elements (or software protocols on them) used to improve performance of satellite links

Split connection
- TCP connection is terminated on PEP agent and another one is originated to the other end-system

Transparency
- End-users are unaware of PEP presence, thus requiring no modifications in end-points

Integrated/distributed implementation
- Integrated - if the implementation is only in one network element (our reference scenario is well suited for an integrated PEP implementation)
- Distributed – two or even more PEP elements are involved in the network

Symmetrical (both directions) or asymmetrical (one way) implementation
Cross-layer issues for PEP

**MAC-TCP interactions**
- Capacity at MAC layer (in the return path) should be distributed among competing TCP flows according to a DAMA scheme. This functionality could be supported by the PEP functionality in the NCC, in our scenario.

**PHY-TCP interactions**
- PEP could be used for controlling the sender window parameters (in the forward path) during short disconnection intervals (bad channel conditions) for mobile users receiving TCP data, like in the M-TCP approach. In this case the PEP can modify fields of ACK packets, so non-transparent approach is needed.

Both of these cross-layer approaches could be combined in our proposed DVB-RCS scenario.
Our case study for a PEP-Spoof approach and related cross-layer signaling in a BSM scenario
System architecture and traffic scenario (return path case)

In this architecture, we know that the satellite segment is the bottleneck due to the round trip propagation delay and the delay in the adopted DAMA mechanism to request capacity by STs.

We assume a TCP-aware RBDC-like scheme.

At MAC layer of the ST, TCP internal status data are provided (like the congestion window value) to determine the future need of the terminal.

The PEP is located at the NCC/gateway and is used to filter the layer 4 ACK flows.
PEP-Spoofer with cross-layer signaling

In ST, the TCP internal state information (i.e., cwnd, and TCP phase) should be propagated to MAC layer

- Usage of enriched headers or periodical primitive, which requires modification of the DVB-RCS standard for supporting such downward cross-layer action.
- DAMA capacity requests are sent on the basis of current queue occupancy and prospected TCP traffic injection for the next super-frame.

The NCC receives incoming DAMA requests

- If the resources the ST requested are not available, the NCC defines at MAC layer a corresponding limit \( \text{cwnd}^* \). If the ST buffer is congested, \( \text{cwnd}^* \) value is provided to the transport layer according to the cross-layer upward signaling.
- Notifies the allocation through the *Terminal Burst Time Plan (TBTP)*

NCC at transport layer operates as PEP in the forward direction

- In case of congestion indication from MAC layer, layer 4 ACK packets are spoofed by the PEP at the NCC.
- A \( \text{cwnd}^* \) value is set in a ACK* packet, while a flag in the ACK* notifies the receiving ST about the use of \( \text{cwnd}^* \).
Description of cross-layer interactions and PEP role

This cross-layer technique could be classified as explicit, decentralized and centralized, upward and downward, horizontal MAC-centric, in-band and out-of-band signaling.
Details of the DAMA controller at NCC with PEP-Spoofer

1. ST rate value sent in the capacity request together with cwnd

2. DAMA controller with PEP-Spoofer

3. ST buffer congestion indication

4. TBTP MAC layer signaling

5. ST MAC layer signaling (enhanced DAMA capacity request)

6. Can DAMA allocate this rate in the next superframe?

7. Yes, cwnd could be allocated for the next super-frame

8. Not. A maximum allocable capacity is identified that corresponds to some cwnd* that is lower than cwnd

9. Is the ST buffer close to congestion? and cwnd* has been set?

10. Yes, yes

11. Signaling towards transport layer triggered by MAC layer to notify the PEP about the presence of congestion and the need to stop the ST cwnd increase to cwnd*
Considerations for our PEP-Spoof

This technique does not use TCP splitting nor TCP spoofing at the NCC/gateway, because only one cwnd is present per flow and layer 4 ACKs are end-to-end.

However, the NCC/gateway needs to have PEP functionalities: on the downstream channel, the NCC/gateway operates a spoofing action analyzing the ACKs and setting the cwnd* field and option, if needed. This type of PEP use is non transparent since it entails a modification to the sender TCP version.
## Survey of characteristics for our PEP-Spoofer proposal

<table>
<thead>
<tr>
<th>PEP characteristics</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Involved layer</td>
<td>Our PEP proposal operates at transport layer</td>
</tr>
<tr>
<td>Distributed or Centralized</td>
<td>The implementation is centralized, co-located with the NCC gateway</td>
</tr>
<tr>
<td>Symmetric or Asymmetric</td>
<td>Asymmetric for the return path (our proposal if applied on the forward path would require modifications of transport layer in Internet servers and this is not the scalable solution )</td>
</tr>
<tr>
<td>Split connection</td>
<td>Not, only one cwnd value is present end-to-end</td>
</tr>
<tr>
<td>Transparent to Transport</td>
<td>Sender side modifications are needed</td>
</tr>
<tr>
<td>Transparent to Application</td>
<td>Yes</td>
</tr>
</tbody>
</table>
## Employed mechanisms in our PEP-Spoofer proposal

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP ACK spacing</td>
<td>Not necessary, depending on TCP version</td>
</tr>
<tr>
<td>Local TCP acknowledgements</td>
<td>We only modify ACK if there is congestion</td>
</tr>
<tr>
<td>Local TCP retransmissions</td>
<td>No, our TCP is end-to-end.</td>
</tr>
<tr>
<td>TCP ACK filtering and reconstruction</td>
<td>Yes TCP ACKs are filtered and modified if there is congestion</td>
</tr>
<tr>
<td>Handling link Disconnection</td>
<td>Yes, if a M-TCP-like approach is adopted</td>
</tr>
<tr>
<td>Caching, parse/prefetch or other priority based mechanism</td>
<td>No, in our case PEP is co-located with NCC, they can not be separated so any priority is reflected on the resource allocation performed by the NCC at MAC layer</td>
</tr>
</tbody>
</table>
Comments on signaling involved from the BSM standpoint

- MAC - TCP direct cross-layer relation at a given ST seems to be problematic.
  - No conventional mechanism is available to convey the TCP status (i.e., its state and the cwnd value) to the MAC layer.
  - If current primitives are used, a complex scheme of signalling should considered involving primitives from layer 4 to layer 3 that are coordinated with SI-SAP primitives from layer 3 to layer 2.
    - For instance, SI-C-QUEUE-MODIFY-req primitive could be used for this purpose for dynamically adjusting QID parameters and involving the STQRM to request an update of the capacity allocation.
    - In any case, this is not a so proper cross-layer implementation.
  - Otherwise, there would be the need of primitives that punch holes in the protocol stack in order to related directly non-adjacent layers.

- Similar considerations are valid for the cross-layer functionalities needed at the NCC/gateway that operates as a PEP, filtering ACKs.
Direct signaling among non adjacent layers

Many techniques could be used for signaling between non-adjacent layers. A possibility is represented by Cross-Layer Signaling Shortcuts (CLASS) method.

- Method of punching holes in the protocol stack and interchanging the limited set of cross-layer information between layers using ICMP.
- Light-weighted ICMP format for internal cross-layer exchange of information
- Standard ICMP format for external cross-layer exchange.

Note that ICMP could be used for setting the initial TCP window value, ACK filtering or even resetting the retransmission timers.
Comparison with the existing Integrated PEP proposal

- Integrated PEP usually involves splitting approach with enhanced TCP versions
  - Drawback is breaking end-to-end semantics and generated ACK are not usually correlated with real ACK packets.

- Our proposal does not rely on splitting; however, shared database, like in the split approach, should be considered for implementation.

- Typically *ad hoc* TCP versions are used (this is also in our proposal). Hence, the PEP is non-transparent.
Conclusions and future work

Possible cross-layer approaches referring to a DVB-S2/-RCS IP-based satellite network compliant with the BSM model have been presented.

The adoption of cross-layer design methods permits to exploit lower layer adaptation to optimize the performance at higher layers.

- We proposed a PEP-Spoofer technique at the NCC/gateway to improve the congestion control (TCP performance) in DVB-S2/-RCS GEO-based networks.

A further research work is needed to identify efficient BSM mechanisms for delivering cross-layer information among non-adjacent layers (e.g., transport and MAC layers) and to model the possible interactions between upward and downward cross-layer signalling.

Possible software implementation is needed for evaluation of cross-layer signaling and performance evaluation of our PEP-Spoofer approach.
5th International Workshop on Satellite and Space Communications 2009 (IWSSC 2009)

September 10-11, 2009
Siena-Tuscany, Italy

http://iwssc09.dii.unisi.it/

SatNEx II, 7th ja2230 integration virtual meeting, ISTI, Pisa, November 7, 2008
Thanks a lot

Any comments?

giambene@unisi.it
References


Spare slides
Cross-layer interactions under study: PHY, MAC, TCP

In the case of TCP NewReno and uncorrelated losses, we can use the square-root formula explaining cross-layer implications for TCP goodput performance:

\[
\Gamma = (1 - \text{PER}) \min \left[ \frac{I_{TCP}}{R(T)(B)} \times \sqrt{\frac{1.5}{\text{PER} \cdot B}} \right]
\]

Interactions between PHY (modulation and coding) and TCP

The selection of different modulation and coding combinations (impact on \(B, \text{PER}, \text{RTT}\)) at the physical layer (Adaptive Code and Modulation, ACM, air interface) could be related to TCP goodput: \(E_b/N_0\) thresholds should not be determined on the basis of PER requirements, but rather on the maximization of TCP goodput. Different TCP versions could require distinct optimizations of the trade-off between higher information bit-rates and lower coding protection that can be achieved by increasing the modulation and the coding level.

Interactions between MAC (resource allocation) and TCP

Resources should be assigned to the different TCP-based flows (impact on \(\text{RTT}\)) so that they have a fair sharing of air interface capacity. A better resource management is obtained by those TCP-aware scheduling schemes that manage flows taking into account both the information bit-rate variability due to PHY adaptation and the history in the service provided.
Signaling

Cross-layer signaling
Primitives of both *get* and *write* type are needed to exchange internal protocol state between adjacent or non-adjacent protocol layers.
This is valid for both the ST and the NCC/gateway/feeder.

One possibility is that there is **a protocol layer in charge of triggering primitives for cross-layer interactions**
For instance, MAC for a MAC-centric approach.
Explicit cross-layering and BSM protocol stack

The Broadband Satellite Multimedia (BSM) protocol stack defined by ETSI has been taken as a reference for our cross-layer investigations, thus considering the related SI-SAP interface between network layer (L3) and layer 2 (MAC).

We consider C-plane signaling of BSM protocol architecture.

ETSI TS 102 463 only defines primitives that are passed through SI-SAP between L3 and MAC adjacent layers on the control plane. These primitives could be used for cross-layer signaling between MAC and L3.

In the case of cross-layer interactions between non-adjacent layers involving SI and SD layers across SI-SAP, suitable new mechanisms should be envisaged in addition to SI-SAP primitives.
PEP taxonomy (cont’d)

} Spoofing

| Usually split connection approach used to isolate long-latency links
| The role of spoofer is to intercept, cache and acknowledge data from sender
| Our PEP spoofer is located close to the source and does not experience the high latency of satellite links.