Preventive and Reactive based TNL Congestion Control Impact on the HSDPA Performance

Thushara Weerawardane, Andreas Timm-Giel, Gennaro C. Malafronte, Durastante Gianluca, Stephan Hauth, Carmelita Görg

Abstract—High Speed Downlink Packet Access (HSDPA) is an extension of the Universal Mobile Telecommunications System (UMTS) technology of 3GPP Rel-99, with the objective to increase the data rate and reduce the latency in the downlink. The main focus of the work presented is to analyse the effect of congestion at the lub interface on the HSDPA performance. The data flows should be adequately controlled in order to avoid congestion in the transport network. The 3GPP (3rd Generation Partnership Project) Rel. 5 specifications highlight two congestion detection mechanisms which are based on the frame sequence number (FSN) and the Delay Reference Time (DRT) fields of HSDPA data frame. In addition to these, a third congestion detection mechanism based of Checksum of HSDPA data frame is considered. This paper discusses a congestion control scheme deploying all three congestion detection methods. It is shown, that a congestion control algorithm can effectively work using these congestion detection triggers and can control the offered load to the transport network. The simulation results presented in this paper confirm that the congestion in the transport network can be avoided, and hence the performance of HSDPA network can be significantly improved in all aspects.

Keywords—HSDPA, Congestion Detection, Congestion Control

I. INTRODUCTION

HSDPA is an upgrade to the 3GPP Rel-99 UMTS technology meeting the ever increasing user requirements. It is an asymmetric solution that enables higher downlink throughput to the end user. Several enhancements to WCDMA R’99 were considered in 3GPP release 5 [1]. Moving the scheduler from the RNC (Radio Network Controller) to the NodeB and reducing the Transmission Time Interval (TTI) from 10ms to 2ms allows adapting better to the varying radio channel. The Adaptive Modulation and Coding (AMC) can adapt the modulation rank and coding rate to the current radio channel condition of each user. With a constant transmit power, the effective data rate can so be significantly enhanced for users with good radio channel conditions. The other prominent improvement of the physical layer is the use of Hybrid ARQ (HARQ), where the soft bits from retransmission requested by the UE are combined with soft information from the original transmission prior to the decoding [2] leading to an improved BER performance.

In addition to the physical layer improvements, currently the research is focused on enhancements and performance measures on lub interface [3], [4], [5] and [6]. From the operator point of view, the lub is usually a scarce resource. The HSDPA transport channel, HS-DSCH [1], [2] is controlled by the MAC-hs scheduler which is a new entity located in the Node B. Packet switched traffic via HS-DSCH is scheduled in the Node B and then transmitted over the air interface. Moving the scheduler to the Node B enables more efficient implementation of the scheduling by allowing the scheduler to work with the most recent channel information. The required data for radio channels should arrive with low delay at the Node B side in order to minimize the data buffering in the MAC-hs. The Transport Network Layer (TNL) is optimized by deploying suitable flow control and congestion control techniques minimising the lub link bandwidth requirements, providing better QoS to the end user and also reducing the operator costs [4], [5], [6] and [8]. Compared to Rel-99, HSDPA uses two buffering points at the Node B and at the RNC. Since the transport capacity is limited and the demand of the air interface varies rapidly according to the channel condition, the data flow over the lub interface should be adapted accordingly. The credit based flow control scheme quantifies the required air interface capacity and informs the RNC as a resource allocation for each user flow separately. However, when the demanded user traffic over the air interface is greater than the available lub capacity, the congestion occurs at the UTRAN transport. Due to lub congestion, the packet loss probability increases significantly, causing more retransmissions at higher layers such as RLC and TCP. Hence the offered load is further staggered leading the system into congestion collapse and wasting network resources. To avoid such circumstances, a proper congestion control scheme is required in addition to the credit based flow control mechanism.

A credit based flow control mechanism has been developed in [4] to optimise the lub utilisation while providing required QoS to the end user in the HSDPA network. It smoothes down the HSDPA traffic and reduces the burstiness over the lub interface. The implemented credit allocation algorithm is based on the periodically updated provided bit rate (PBR) of the user-specific priority queues in the Node B.

As mentioned above when the lub link is overloaded, congestion occurs and cannot be protected by the flow control mechanism and a proper TNL based congestion control mechanism is required. 3GPP has introduced two basic mechanisms working at the FP layer (i.e. transport media independent) to perform lub congestion detection (see TS 25.435): frame loss detection by means of FSN (Frame Sequence Number) supervision and delay build up detection by means of DRT (Delay Relative Time) supervision. In addition to these, a third congestion detection mechanism based of Checksum of HSDPA data frame is considered.

The remainder of this paper is structured as follows. The preventative and reactive congestion detection mechanisms are discussed in details in the second and Third chapters. Next, the congestion control procedures which are activated upon the congestion detection triggers are described in details. Then the
simulation model and simulation scenarios are presented in following chapter. Finally, the simulation results and analysis are presented along with the conclusion at the end of the paper.

II. REACTIVE CONGESTION DETECTION

The reactive Iub congestion detection mechanism detects the congestion upon detection of packet losses at the transport network. Since the TNL uses the ATM as the transport technology between Node B and RNC, small cell losses can occur when the system is overloaded. This either leads to corrupted frames or loss of complete frames, the later due to the bursty losses. When the header CRC check fails, the packet cannot be identified and is discarded completely. In many situations, the FP layer receives frames with a valid header but invalid payload. Such errors are identified as the payload CRC errors. The most common frame errors can be categorised and described as follows.

1. Missing a last segment or tail of the FP frame

   When the last segment or the tail of the FP PDU is lost, the receiver waits until it receives the last segment of the next frame to be reassembled. This results in creating a large frame with an invalid CRC during the reassembly process. Both original frames are lost.

2. Missing any segment except the tail of a frame

   The receiver can reassemble a FP frame having an invalid CRC due to a missing segment or due to an insertion of a foreign segment. Since the tail of the frame is preserved, the next frame can be detected again.

3. Loss of complete frames

   When a burst of cell losses occurs at the transport network, one FP frame or several FP frames can be lost. These kind of bursty losses are very common for HSDPA PS networks when congestion persists and are detected monitoring the FSN.

   Another important reason besides Iub congestion for missing or invalid frames is given physical layer errors; Such physical layer errors are filtered out by the Iub CD algorithm by means of a probabilistic discriminator.

   The error type 1 (missing a last segment or tail of the FP frame) and 2 (missing any segment except the tail of a frame) can be easily identified by the FP payload CRC check. The type 3 (Loss of complete frames) can be identified by using a frame sequence number (FSN).

III. PREVENTIVE CONGESTION DETECTION

The preventive congestion detection is in charge of detecting Iub packet losses before they occur. The preventive DRT based delay build-up algorithm monitors the FP PDU delay variation for the correctly received FP frames through the transport network for each MAC-d flow. The figure below shows the delay variation for several FP PDU transmissions with respect to the RNC and Node B reference counters.

According to the Figure 1, the delay build-up algorithm monitors the TNL delay variations for each arrival of an FP PDU at the Node B. The figure also shows the two inputs to the congestion control algorithm: one is provided by the FC module and the other is provided by the congestion detection module. The FC and the congestion detection algorithms work independently and provide information to the congestion control module promptly.

IV. CONGESTION CONTROL

The functional block diagram of the flow control, congestion detection and congestion control are shown in Fig. 2.

Both congestion detection and congestion control algorithms are implemented in a cascade with the credits based flow control mechanism which is used to estimate the current air interface capacity in the Node B for HSDPA [4]. The congestion detection is triggered upon each arrival of an FP PDU at the Node B. The figure also shows the two inputs to the congestion control algorithm: one is provided by the FC module and the other is provided by the congestion detection module. The FC and the congestion detection algorithms work independently and provide information to the congestion control module promptly.

The MAC-d flow can be either in congestion state or in flow control state. Therefore, the system can be represented by a flow state machine as shown in the following figure.
events are triggered by the detection module the MAC-d flow goes to Congestion Control State (CC) by sending the CA message to the RNC. At the congestion control state, the MAC-d flow is controlled by the AIMD (a,b) algorithm [9], [10]. The credits allocation is done on top of the flow control credits and the amount of credit reduction is done according to the severity of the congestion trigger which is received by the CC module. The AIMD (a,b) rate control mechanism is used to reduce the rate by b⋅ \( \frac{c}{a+b} \) where c is a constant that depends on the severity of congestion. The credits allocation is done on top of the flow control credits and the amount of credit reduction is done according to the severity of the congestion trigger which is received by the CC module.

The comprehensive HSDPA simulation model is developed under the OPNET simulation environment. The simulation model is specially designed for the TNL feature analysis, but allows in addition performance analysis is common type of HSDPA packet traffic consists of web and FTP traffic which uses the TCP (Transmission Control Protocol) as a reliable end-to-end transmission protocol. These scenarios are defined according to the types of traffic models they use: simulation scenario 1 is based on the FTP traffic model, whereas simulation scenario 2 is based on the ETSI traffic model. Each simulation scenario is simulated with two different configurations, each with 20 UEs in the cell. Configuration 1 is defined by using flow control and congestion control mechanisms together and Configuration 2 only uses the FC mechanism. The basic idea of these two configurations with two different scenarios is to analyse the effect of the congestion control for two different offered load situations.

For configuration 1, both FC credits and CC credits, and sends the lowest between them to the RNC in a CA message. The condition for the MAC-d flow to exit the CC state is that the value of credits calculated by AIMD (a,b) either comes back to the value of credits assigned to the MAC-d Flow when it entered the CC state or it is exceeded by the flow control credits.

V. SIMULATION MODEL AND SIMULATION SCENARIOS

The comprehensive HSDPA simulation model is developed under the OPNET simulation environment. The simulation model is specially designed for the TNL feature analysis, but allows in addition performance analysis is common type of HSDPA packet traffic consists of web and FTP traffic which uses the TCP (Transmission Control Protocol) as a reliable end-to-end transmission protocol. The web traffic model which is defined by the ETSI standards [1], [7] is selected to evaluate the performance of HSDPA traffic under moderate network load. The FTP traffic model is used to overload the network therefore, it provides the worst case traffic scenarios to analyse the impact of the above features on the end user performance as well as the overall network performance. Under this traffic configuration, all users in the cell download a large file and utilize the network resources up to the maximum available capacity. The parameters of the web traffic model defined by ETSI and the FTP traffic model are given in the Table I. Both ETSI and FTP traffic models are adapted as application layer traffic models which include the TCP protocol adaptation as well. The windows 2000 based TCP configuration that uses the TCP Reno version, is used for all simulations. Therefore, TCP congestion control functionality is included for these analyses.

Two simulation scenarios are defined according to the types of traffic models they use: simulation scenario 1 is based on the FTP traffic model, whereas simulation scenario 2 is based on the ETSI traffic model. Each simulation scenario is simulated with two different configurations, each with 20 UEs in the cell. Configuration 1 is defined by using flow control and congestion control mechanisms together and Configuration 2 only uses the FC mechanism. The basic idea of these two configurations with two different scenarios is to analyse the effect of the congestion control for two different offered load situations.

The effect of the credit based flow mechanism was tested and presented in [4]. The key results presented in this paper are used to validate the performance of the TNL congestion control algorithm on the performance of the HSDPA network.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Distribution and values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packet call interarrival Time</td>
<td>Geometric distribution ( \mu = 5 ) seconds</td>
</tr>
<tr>
<td>Packet call size</td>
<td>Pareto distribution ( a=1.1, k=4.5 ) Kbyte, ( m=2 ) Mbyte</td>
</tr>
<tr>
<td>FTP Traffic Model Parameters</td>
<td></td>
</tr>
<tr>
<td>File size</td>
<td>Constant Distribution ( \mu = 12 ) Mbyte</td>
</tr>
</tbody>
</table>

Table I: Traffic Models for HSDPA

ATM link throughput

Fig. 4 shows the ATM link utilisation probability distributions. The offered traffic to the system is higher without Congestion Control.

Since both scenarios use the same traffic models, the additionally offered load to the TNL network is assumed to be caused by higher layer retransmissions. To validate the
assumption and exclude that the CC mechanism throttles down the traffic unnecessarily, further results are analysed at the higher layer protocols in the following.

**AAL2 loss ratio**

The AAL2 loss ratio (LR) is shown in table II. The AAL2 buffer discards the packets having a waiting time longer than 50 milliseconds.

<table>
<thead>
<tr>
<th>Without_CC</th>
<th>With_CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.50</td>
<td>0.01</td>
</tr>
</tbody>
</table>

From the figure, it can be seen that there is a clear advantage of using a congestion control scheme which significantly reduces the loss ratio. The loss ratio for the scenarios without CC is about 29% and it is reduced to 0.01% with congestion control.

**FP end-to-end delay**

The FP PDU end-to-end delay probability distribution for the downlink is shown Fig. 5. The FP end-to-end delay is measured from the point of time when an FP PDU is sent from the FP layer in the RNC to the point of time when it is received by the FP layer in the Node B.

Fig. 5 shows the average FP end-to-end delay which is about 45 ms for the scenario without CC and about 10 ms for the scenario with CC. This indicates about 77% reduction of the FP average delay at the transport network layer.

**IP throughput**

The overall average IP throughput for all users is shown in Fig. 6. The IP throughput is measured between end user entities; it shows the overall performance of the UTRAN as well. The average IP throughput is about 1257 kbit/sec for CC based configuration, and about 654 kbit/sec without CC. These results confirm a significant gain of average IP throughput for the scenario which uses the congestion control algorithm compared to the scenario which does not use the CC algorithm.

Since the achieved end user performance is clearly improved, it can be concluded that the Iub CC algorithm provides actual benefits in terms of end user application throughput as well as better UTRAN transport resource utilization.

**B. Simulation scenario 2 ETSI traffic**

Simulation scenario 2 is tested using the moderate ETSI based traffic model. The ATM Iub bandwidth is configured to 1 Mbit/s and the simulation duration is 2000 seconds. The following figures show the results to evaluate the performance of the congestion control scheme for the bursty traffic model.

**ATM Link throughput**

As in the previous simulation, these two simulation scenarios show the ATM utilisation with and without the CC algorithm.

The offered load increases mainly due higher layer retransmissions

Fig. 7 shows the ATM link throughput probability distribution for both simulation scenarios. The scenario with the congestion control scheme shows less offered traffic to the TNL network compared to the scenario without the congestion control scheme. However the analysis of the AAL2 frame loss and IP throughput shows that the application end user is improved, as it is reported below.
AAL2 Loss Ratio

The AAL2 loss ratio is shown in table III for the ETSI based simulations.

<table>
<thead>
<tr>
<th></th>
<th>Without_CC</th>
<th>With_CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOSS RATIO (%)</td>
<td>25.10</td>
<td>0.07</td>
</tr>
</tbody>
</table>

As in the FTP based simulations, the figure shows a clear reduction of losses at the transport network for the scenario using the CC mechanism.

FP end-to-end delay

Fig. 8 shows the probability distribution of the end-to-end FP delay for both scenarios.

The average FP delay which is about 35 ms for the scenario without CC and about 8 ms for the scenario with CC. This indicates about 77% reduction of the FP average delay at the transport network layer. As in the FTP based simulations, this indicate that the CC control based scenario provides much better TNL delay performance compared to the scenario without CC control.

IP throughput

The overall average IP throughput for all users is shown in Fig. 9. The IP throughput is measured between end user entities. As in the FTP based simulations, the achieved throughput for the CC based simulation is significantly higher compared to the scenario without CC control. This shows that end user throughput can be enhanced by avoiding congestion at the transport network.

VI. CONCLUSION

This paper shows the theoretical and modelling aspects of a preventive and reactive congestion control algorithm. The performance of a combined congestion control along with a flow control mechanism is presented, simulated and validated using different HSDPA simulation environments [4].

In this simulation analysis, the simulation scenarios are configured with the same set of network and air interface parameters in order to validate the effect of the congestion control feature clearly. The simulation results show that for both traffic models (ETSI and FTP traffic) the usage of preventive and reactive congestion control algorithms allows to effectively control TNL congestion and minimize the network resource waste due to higher layers retransmissions.

REFERENCES