

# Enhancing Packet Data Access in WCDMA

Janne Peisa<sup>a</sup>, Stefan Parkvall<sup>b</sup>, Erik Dahlman<sup>b</sup>, Pål Frenger<sup>b</sup>, Per Beming<sup>b</sup>

<sup>a</sup>*Ericsson Research, FIN-02420 Jorvas, Finland*

<sup>b</sup>*Ericsson Research, 164 80 Stockholm, Sweden*

**Key words:** WCDMA, Packet data, HSDPA

**Abstract:** In this paper, the first step of evolving WCDMA beyond IMT-2000 is outlined and elaborated upon. It is shown how techniques such as fast link adaptation, fast hybrid ARQ and fast scheduling, among others, can be applied to WCDMA, leading to increased throughput for best-effort services, lower delays, and downlink peak rates in the order of 10 Mbit/s. The end-to-end protocol architecture is discussed, and using protocol simulations the benefit of having a hybrid type II ARQ scheme in Node B is shown.

## 1. INTRODUCTION

Wideband CDMA (WCDMA) as specified by the 3<sup>rd</sup> generation Partnership Project (3GPP) is rapidly emerging as the leading global 3G (IMT-2000) standard. Current WCDMA specification fully satisfies the IMT-2000 requirements, including support of data rates up to 2 Mbit/s in indoor/small-cell-outdoor environments and up to 384 kbit/s with wide-area coverage, as well as support for both high-rate packet data and high-rate circuit-switched data. However, to satisfy also the future increasing demands on packet-data services, studies are already ongoing on the evolution of the WCDMA air interface beyond the IMT-2000 requirements. The first step of this evolution is the introduction of high-speed downlink packet access (HSDPA), currently studied within 3GPP in the HSDPA study item. The main goal of HSDPA is to allow WCDMA to support downlink peak data rates up to

approximately 10 Mbit/s for packet-data services, i.e., far beyond the IMT-2000 requirement of 2 Mbit/s. Furthermore, HSDPA should also enhance the WCDMA packet-data capabilities in terms of lower delay and improved overall capacity.

In an earlier paper [1], the system performance of the HSDPA has been studied. In this paper we present the first protocol results of the HSDPA.

A multitude of new technologies are currently considered for this first step of the WCDMA evolution, for example higher order modulation and fast link adaptation, fast hybrid ARQ, fast scheduling, fast cell selection and multiple-input-multiple-output (MIMO) antenna solutions.

Higher-order modulation, e.g., 16-QAM and 64-QAM, provides higher spectral efficiency in terms of bit/s/Hz compared to QPSK and can thus be used to provide peak data rates in the order of 10 Mbit/s within the current 5 MHz WCDMA bandwidth. Unfortunately, higher-order modulation schemes are significantly less robust to noise, interference and other channel impairments. Hence, higher-order modulation should be combined with fast link adaptation where the coding and modulation scheme employed is rapidly adapted to the instantaneous channel conditions. By using fast link adaptation, users experiencing favorable channel conditions, e.g., close to the cell site, can use higher-order modulation and high-rate coding, e.g., 16-QAM and rate 3/4, thus offering higher peak rates to those users. Similarly, users with less favorable conditions, e.g., users close to the cell border or users experiencing a fading dip, need to use robust QPSK modulation and low-rate coding.

In packet-data applications, the receiver typically detects and requests retransmission of erroneously received data units. Combining the soft information from both the original transmission and any subsequent retransmissions prior to a decoding attempt, generally known as hybrid ARQ type-11, will reduce the number of retransmissions required. This will increase the throughput and improve the delay observed by higher layer packets. Hence, a fast hybrid ARQ scheme tightly coupled to the link adaptation mechanism is beneficial. The link adaptation serves the task of selecting a good initial estimate of the amount of redundancy needed in order to avoid an excessive number of retransmissions. The hybrid ARQ mechanism serves the purpose of fine-tuning the effective code-rate (implicit link adaptation) and counteracts any prediction errors resulting from the link adaptation scheme.

Fast scheduling is the mechanism determining which user to transmit to in a certain time instant. It is a key element in the design of a packet-data system as it to a large extent determines the overall behavior of the system. Maximum system throughput is obtained by assigning all resources to the user with the currently best radio environment, while a practical scheduler

should include some degree of fairness. By selecting different scheduling algorithms, the operators can tailor the behavior of the system to suit their needs.

## 2. EVOLVING WCDMA WITH HSDPA

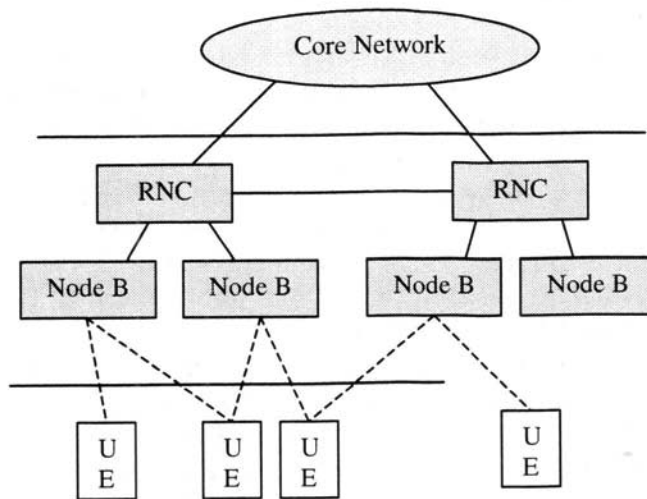


Figure 1. The architecture of UTRAN

The current UTRAN (Universal Terrestrial Radio Access Network) architecture is illustrated in Figure 1. A number of RNCs (radio network controllers) are connected to the core network. Each RNC controls one or several Node Bs (base stations), which in turn communicates with the UEs (user equipment). The basic principles of HSDPA as described in the previous section relies on rapid adaptation of transmission parameters to the time-varying channel conditions and the corresponding functions, such as fast link adaptation and fast scheduling, should therefore be placed close to the air interface, preferably in the Node B. Thus for HSDPA, certain functionality is proposed to be added the Node B as part of a new Node B entity MAC-HSDPA. The current RNC entities such as RLC and MAC should still remain for several reasons, e.g.:

- Cipherring and in-order delivery of the data are examples of functions provided by the RLC of the RNC

- In soft handover scenarios between two different Node Bs, the RNC can guarantee no loss of data if the hybrid ARQ mechanism in the Node B fails, for example due to a limited number of retries.

Hence, with this background, the extended features of the Node B should not be seen as a replacement of the RNC, but rather as a complement, providing, from the RNC point of view, a highly reliable channel supporting high data rates.

## 2.1 General Channel Structure

The core of the high-speed packet data evolution is the addition of a new channel, HS-DSCH (High Speed Downlink Shared CHannel), focusing on best-effort traffic and complementing already existing channels. Similarly to the current DSCH, the HS-DSCH corresponds to a common channelization code resource shared among several users, primarily in the time-domain. To simplify the design of the UE and to reduce the amount of downlink signaling, a fixed spreading factor is preferred as no major advantages is seen with a varying spreading factor. As for the current DSCH, soft handover is not used for the HS-DSCH.

Allocation of the HS-DSCH is done on an HS-DSCH TTI (transmission time interval) basis. The use of a short TTI, shorter than the current minimum WCDMA TTI of 10 ms (one frame length), for the HS-DSCH has several advantages. First, it reduces delays, which is especially important if one or several retransmissions are required before the data is correctly received. Secondly, it allows for a finer granularity in the scheduling process, facilitating better tracking of the channel variations. It should also be noted that there are no interleaver gains associated with longer TTIs as the channel variations are exploited by the link adaptation, hybrid ARQ and scheduler instead of the decoder. Hence, a short TTI is preferred and setting the HS-DSCH TTI equal to the slot duration (0.67 ms) is one possibility. Within each TTI, a shared code resource, such as the HS-DSCH code resource, is preferably allocated to a single user at a time, i.e. the primary way of sharing the HS-DSCH code resource should be in the time domain. Additional support of HS-DSCH code multiplex implies that data to two or more users could be transmitted in parallel within the same TTI, using distinct part of codes from the set of all channelization codes allocated for HS-DSCH. Code multiplexing increases the amount of signaling in each TTI and should only be supported if the increased granularity in the range of supported payload sizes is found beneficial.

In addition to the transmission of user data, there is a need for control signaling to the UE. First, the UE (UEs in case of code multiplex) scheduled for the HS-DSCH in the upcoming HS-DSCH TTI must be notified. In

addition, lower layer control information, e.g., the transport format, including modulation and coding scheme, to be used and hybrid ARQ related information needs to be communicated to the corresponding UE(s). The control information is only relevant for the UE(s) about to receive the HS-DSCH and can therefore be placed on one or several shared control channels, shared in the time domain similarly to the HS-DSCH. Each UE being involved in the scheduling process has an associated dedicated physical channel, DPCH, set up, carrying a flag indicating when to read the HS-DSCH. In addition to this flag, the associated DPCH is used to carry power control commands for the uplink and, if desired, other services, e.g., circuit-switched voice.

## **2.2 Adaptive Coding and Modulation**

As discussed in the introduction, higher order modulation in conjunction with link adaptation provides a tool for maximizing the instantaneous utilization of the fading radio channel. By transmitting the HS-DSCH at a (in principle) constant power, i.e., no fast downlink power control, the modulation and coding scheme can be selected to maximize the downlink data rate given the instantaneous channel conditions. Note that the use of link adaptation instead of fast power control does not mean that the HS-DSCH power cannot vary, for example due to the variations in the power used by other channels. One interesting aspect with link adaptation is the benefit in terms of supportable data rate for the end user from more advanced receiver designs, contrary to traditional systems, where the performance gain of an advanced receiver design is a benefit for the network operator, in terms of improved capacity (lower power needed), and does not benefit the user in terms of better service quality.

A number of transport blocks, originating from the RNC, are concatenated, followed by the addition of a CRC for error detection, Turbo encoding, modulation and multicode spreading. The use of a single CRC for all transport blocks in the TTI reduces the overhead compared to using a CRC per transport block. Furthermore, in most cases, when a TTI is received in error, most of the transport blocks are erroneous and must be retransmitted. Hence, there seems to be no major gains associated with being able to identify erroneous transport blocks separately.

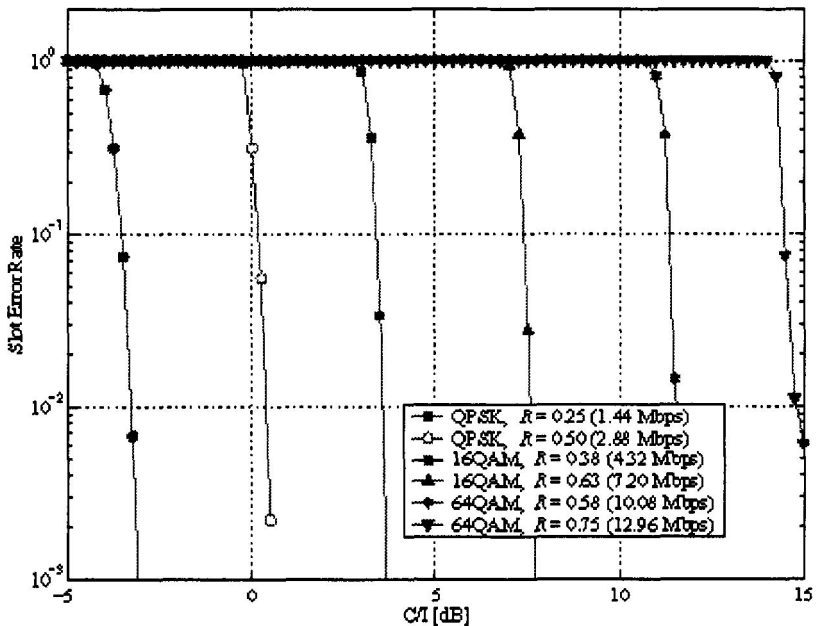


Figure 2. AWGN link performance

In Figure 2, the performance for some different examples of modulation and coding schemes that might be used for HSDPA are illustrated, covering a wide range of different carrier-to-interference ratios. Scenarios with carrier-to-interference ratios below the lowest MCS can be efficiently supported by relying on the use of soft combining in the hybrid ARQ mechanism. Support of the highest order modulation schemes, such as 64-QAM, is of less interest in a traditional cellular scenario due to the high requirements on both channel quality and radio electronics. However, it can be useful in other scenarios, e.g., fixed wireless applications or indoor networks, and should therefore be part of HSDPA as a network/UE option.

Selecting the instantaneous modulation and coding scheme can be done in several ways. A scheme where the UE controls the link adaptation is possible, but would prevent the possibility taking traffic aspects such as the amount of data to transmit into account. It is therefore preferred to let the Node B control the link adaptation. Explicit measurement reports from the UE based on the signal-to-interference ratio of the common pilot channel (CPICH) is one possibility to base the selection of modulation and coding

scheme upon. However, it is not clear how such an approach would work in case of adaptive antennas, where the Node B directs a narrow beam to each user. As the CPICH is broadcast over the entire cell, measurements on the CPICH does not reflect the propagation conditions in the narrow beam. This problem can be resolved by, for example, letting the instantaneous transmit power of the power-controlled associated DPCH control the link adaptation. A high transmit power indicates a less favourable radio environment and vice versa.

## **2.3 Hybrid ARQ**

Hybrid ARQ, where the soft information from retransmissions requested by the UE, is combined with soft information from the original transmission prior to decoding greatly improves the performance and adds robustness to link adaptation errors.

Retransmission requests can be made either per transport block or per HS-DSCH TTI (or combinations thereof), but as previously stated, the advantages by performing retransmissions individually on transport blocks are limited since, in most cases, either most of the transport blocks in a slot are erroneous or all of them are correctly decoded. Hence, retransmission requests per HS-DSCH TTI simplify the required uplink signaling. If a TTI is correctly decoded, an acknowledgement is sent to the Node B, using the uplink dedicated physical control channel. On the other hand, if the TTI is erroneously received, a retransmission is requested within a few milliseconds after receiving the erroneous TTI. Once the retransmission occurs, the UE can perform soft combining of the previously received versions of the data with the retransmission before decoding, thereby greatly increasing the probability of successful decoding. Retransmissions are requested until the TTI is either correctly decoded or a predetermined number of attempts have occurred. Furthermore, as the hybrid ARQ mechanism resides in the Node B, retransmissions can be rapidly requested, allowing for the small delays required at the high data rates envisioned.

The link layer ARQ mechanism cannot provide fully reliable end-to-end data transfer. If a reliable end-to-end data transfer is needed, it is envisioned that a further retransmission protocol is implemented at transport layer. One example of such a protocol is Transport Control Protocol (TCP) encountered at the Internet protocol suite. Thus the the link layer retransmission scheme should be viewed as an optimisation, rather than a completely reliable data transfer mechanism. Furthermore, in certain situations, e.g. in hand over, it may not be possible to make the retransmissions from the NodeB. In those

cases it may be beneficial to resort to link layer retransmissions from RNC. These two link layer retransmission schemes are depicted in Figure 3.

The utilization of soft information from both the original transmission and the retransmission attempts, if any, can be performed in different ways, e.g., Chase combining or incremental redundancy (R). Chase combining, where an identical encoded packet is transmitted in response to a retransmission request, is the simpler of the two and the receiver only has to form a weighted sum of the multiple copies (maximum ratio combining) before decoding. Incremental redundancy, on the other hand, implies that additional parity bits are transmitted in response to a retransmission request. The additional parity bits can be used in conjunction with the previously transmitted ones, effectively forming a lower rate code. The performance of incremental redundancy can be superior to Chase combining in several cases, e.g., high-rate coding and modulation schemes, especially if the link adaptation is less accurate. Incremental redundancy requires careful attention when used in conjunction with Turbo coding, or any other systematic code, as the systematic bits are more important than the parity bits for those codes.

For the hybrid ARQ scheme to work properly, the erroneous retransmission entity must be identified in a robust fashion. One possibility could be to assign a number to each transmitted entity, transmit this number to the UE separately from the user data, and include the number in any retransmission requests. This allows the implementation of a conventional selective-repeat protocol at the cost of signaling the identity. Another possibility is to use a stop-and-wait protocol, which only requires a single bit acknowledge in the reverse signaling. To be able to fully utilize the channel for a single UE,  $N$  staggered stop-and-wait protocols can be operated in parallel. The slot number can be used to identify which of the  $N$  channels that is currently receiving information. By imposing a strict requirement on the timing of the acknowledgement relative to the data transmission, there is no need for explicitly signaling which of the  $N$  channels that is acknowledged as this can be deducted from the slot number.



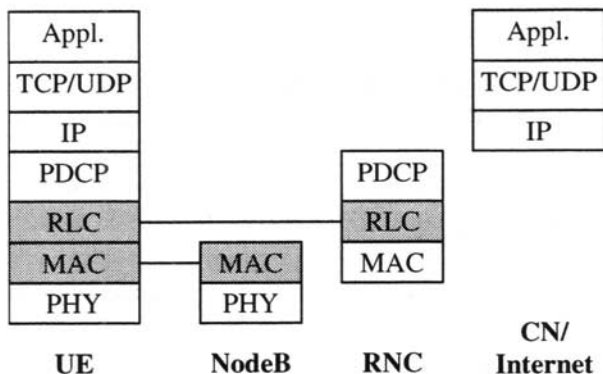


Figure 3. Protocol stack illustrating the two link layer retransmission protocols

## 2.4 Fast Cell Selection

For services such as speech that poses strict requirements on continuously receiving data, soft handover offers the advantage of a diversity gain and consequently less transmitted power at the cost of allocating resources in several cells. The HS-DSCH, on the other hand, is used for best effort services, where the diversity gain to a large extent is exploited by the scheduler. Furthermore, soft handover is cumbersome to implement as the scheduling is decentralized. Hence, similarly to the current DSCH, soft handover is not used for the HS-DSCH. Instead, fast cell selection, where the UE continuously selects the servicing cell with the instantaneously best radio conditions can be applied. To a limited degree, fast cell selection is similar to the site selection diversity transmission (SSDT) mechanism available in WCDMA. The main difference is the centralized SSDT-scheduling in the RNC, while fast cell selection does not involve the RNC.

Fast intra-Node B cell selection is fairly straightforward, while inter-Node B cell selection requires a mechanism to transfer the hybrid ARQ state between the two Node Bs involved. Transferring the hybrid ARQ state can be performed either via the RNC, or over the air, controlled by the UE. Through-the-RNC transfers introduced additional delays, while the over-the-air approach can introduce robustness problems as not necessarily one uplink to each Node B of sufficient quality is available. The power control mechanism of the uplink channels in WCDMA only ensures that the signal is received by at least one of the Node Bs, not necessarily both. One viable approach could therefore be not to include fast cell selection or only fast intra-Node B cell selection, provided the performance gains outweighs the increase in signaling required. Higher layer signaling could be used to switch servicing Node B on a slower basis.

## 2.5 Fast Scheduling

The scheduler is a key element in the design as it controls the allocation of the channel among the users and to a large extent determines the overall behavior of the system. As mentioned in the introduction, different schedulers can be used for different scenarios as there is no need to standardize the scheduling algorithm used, but instead different vendors can choose different criteria. Information upon which the scheduler can base its decisions on includes, but is not limited to, the predicted channel quality, as used by the link adaptation, the current load of the cell and different traffic priority classes. The scheduler also operates closely to the fast cell selection mechanism, which can be seen as a spatial scheduling complementing the temporal scheduling. This means that some of the advantages obtained by the fast cell selection scheme can be achieved by tweaking the scheduling algorithms instead.

## 3. PROTOCOL SIMULATION RESULTS

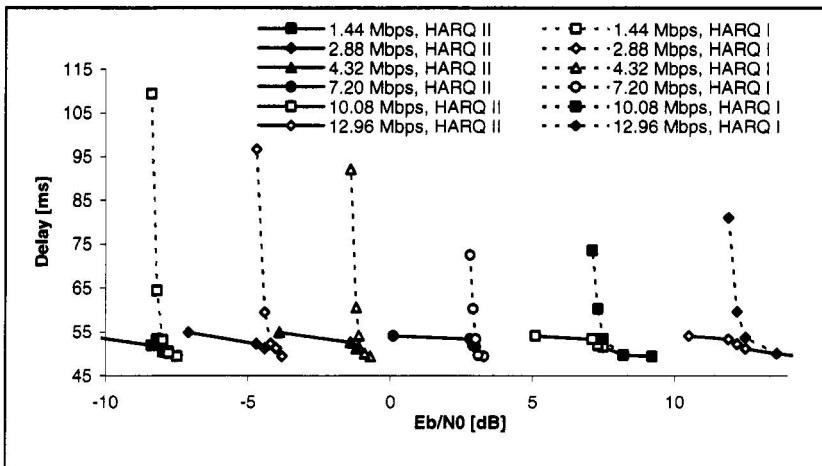


Figure 4. Higher layer packet delay as a function of  $E_b/N_0$

To illustrate the packet data performance in an HSDPA-evolved WCDMA system, a simplified system simulation was set up. The packet data performance was evaluated using the delay of a single higher layer packet as a measure of the performance. This delay is one of the major factors influencing the end-to-end performance of the system.

The link results in Figure 2 were used to model the link performance. Using AWGN link results is reasonable as the channel changes slowly compared to the short HS-DSCH TTI of one slot. Four staggered stop-and-wait ARQ protocol queues were used.

The delay of a single higher layer packet of 1460 bytes (e.g. an IP packet) is shown in Figure 4 for different modulation and coding schemes. The delay consists of a fixed network delay (e.g. the Internet and UMTS Core network delays) and radio access network delay. The fixed network delay was set to 37 ms and the radio access delay was 21 ms. Thus the smallest possible delay is 48 ms, if there are no errors at all on the radio link and if the IP packet can be transmitted on a single TTI.

The IP packet delay for HSDPA scheme is shown in Figure 4. For comparison also a scheme without soft combining (i.e. a HSDPA with hybrid ARQ Type-I) is also shown. The largest benefit of using hybrid ARQ type-II is obtained at high (several tens of percents) slot error rates. As the slot error rate approaches 100%, the delay of the hybrid ARQ type I diverges. It is well known that hybrid ARQ type II can operate efficiently at much higher error rates, which can also be seen clearly from the Figure 4.

Even a small error on the C/I value (caused e.g. by an error on the link adaptation) can cause a dramatic increase in the retransmission delay if hybrid ARQ type I is used. The hybrid ARQ Type II is much more robust against such errors.

## 4. CONCLUSIONS

The first step of the evolution of WCDMA beyond IMT-2000 has been described. By utilizing techniques such as fast link adaptation, fast hybrid ARQ and fast scheduling, the possibility to offer peak rates in the order of 10 Mbit/s to users in favorable conditions was illustrated. Protocol simulations further illustrated the packet data performance of the concept and demonstrated the benefit of having Hybrid type-II ARQ scheme. Clearly, WCDMA is a very flexible standard, which in the future can be evolved beyond third generation requirements to satisfy the future demands on packet-data services.

Fast cell selection, where the UE rapidly selects the cell with the most favorable radio conditions, is one possibility to increase the utilization of the shared radio resources, and can be used as an alternative to soft handover. Fast cell selection can be seen as a complement to the scheduler, providing a spatial dimension in the scheduling in addition to the temporal one.

MIMO, multiple-input multiple-output antenna systems, is a technique for increasing the data rates in rich scattering environments through the

exploitation multiple independent channels, implemented by means of multiple transmit antennas *and* multiple receive antennas.

With the basic principles above, there is a possibility for unequal service provision, offering higher data rates to users in favorable conditions. It can be noted that some of the above technologies have been applied to GSM for the evolution towards EDGE. Thus, to some extent, HSDPA can be seen as an EDGE-like evolution of WCDMA towards 3.5G and beyond, similar to EDGE being an evolution of GSM towards 2.5G and 3G. In the following sections, it will be shown how the key technologies discussed above can be added to WCDMA in order to facilitate a smooth evolution towards 3.5G.

## REFERENCES

- [1] Stefan Parkvall, Erik Dahlman, Pål Frenger, Per Beming, Magnus Persson, “The Evolution of WCDMA Towards Higher Speed Downlink Packet Data Access”, to be presented at VTC Spring 2001