Freescale Semiconductor & Effnet AB
White Paper

Document Number: HDRCMPWP
Rev. 0, 02/2008

Compressed Time-to-Market for Compressed Headers—With Effnet ROHC™ Software on Freescale Devices

by Colin Cureton
Networking Multimedia Group Applications
Freescale Semiconductor, Inc.
East Kilbride, Scotland

and Aniruddha Kulkarni
Effnet AB
Stockholm, Sweden

Contents

1 Introduction

The proliferation of the demand for always-on, high-speed, wireless data connections by the end-consumer has driven the development of several new wireless air interface standards. These standards which include WiMAX, 3G LTE, and UMB mean that competition in the market place is becoming even fiercer. As a result, cost of deployment, operational cost and maximum throughput on the air interface are becoming ever more important for operators as they choose their next-generation network.

Several developments have taken place on these standards when compared to the traditional first-, second-, and third-generation networks. These networks are based upon circuit-switched technology with a backhaul connection to the core network based on E1/T1, ATM, and some Ethernet-carrying circuit-switched payload. The first significant shift from this trend was with High-speed Downlink Packet Access (HSDPA), where a packet-switched environment is put on a circuit-switched baseline wireless air interface.

The move to HSDPA was a fine intermediate step but the increasing market requirements have led to the new wireless standards being designed from the ground up to provide optimum performance for packet-switched traffic. This in turn puts a number of challenges on the wireless infrastructure community trying to get the most performance out of the air interface.

Several changes can be made to layer 1 and layer 2 to introduce such techniques as Orthogonal Frequency Division Multiplexing (OFDM), Automatic Repeat re-Quest (ARQ) and implementation of a Quality of Service (QoS) aware scheduler. This can help to increase throughput rates and solves some of the challenges, but it still does not solve the problem of ensuring that there is reduced overhead or "wasted" bandwidth on the air interface.

This overhead is caused by several factors such as the signaling information between the scheduler and user equipment where the bursty nature of the payload and associated signaling overhead is much higher for packet-switched than for circuit-switched networks. Of course, there is also the overhead for IP stack related headers that are carried over the wireless link to the user equipment.

The ability to minimize the amount of this non-critical data consuming bandwidth is a key challenge to the operators and therefore equipment suppliers. The solution to this problem is to implement an efficient, advanced header-compression algorithm such as Robust Header Compression (ROHC), which is the topic of this article.

2 Overview of Robust Header Compression

The Robust Header Compression (ROHC) standard (RFC 3095) was developed in 2001 by the IETF. This compression scheme, which includes a compression algorithm and an information exchange protocol, can compress IP/UDP/RTP headers to, on average, just over one byte, and is robust even in the presence of severe channel degradation. It can also compress IP/UDP and IP/ESP packet flows.

The earlier header compression protocols and algorithms were developed for wired links and TCP-based applications, although recently, support for even UDP/RTP-based application has been added. ROHC on the other hand has been developed from the ground up for efficiency and robustness considering wireless link conditions.

ROHC reduces header overhead, which leads to reduced packet loss. The smaller packets are less susceptible to BER than bigger packets. The smaller packets can fit into exact available frame size, which avoids segmentation on the link layer. The reduced possibility of packet loss means the system does not often need to use ARQ mechanisms, and that it can even be used experimentally with lower levels of FEC. All these benefits lead to significant savings in system usage and spectrum usage. The end user experiences improved interactive response times and enjoys an improved experience of the applications they are using. The operator benefits by increased capacity of the system and satisfied users.

Header compression is used in many types of networks, including cellular networks, mobile broadband networks, satellite communications networks, low bandwidth wired networks, and public safety wireless networks such as Terrestrial Trunked Radio (TETRA).

The 3rd Generation Partnership Project (3GPP) standards group for the GSM/UMTS (W-CDMA) type of cellular network details IPHC and ROHC standards as part of Release 4 and onwards. Since Release 5, the IPHC and ROHC standards are part of HSPA and IMS specifications. Since Release 6, they are used together with MBMS services. From Release 8, they will be part of LTE specification. The 3GPP2 standards group for the CDMA type of cellular network also specifies various QoS techniques, including

ROHC. The CDMA2000 EV-DO Rev A specification, for example, uses header compression to help enable operators to move circuit-switched (CS) voice network infrastructure over to VoIP. Other standard bodies such as WiMAX Forum or the IEEE have also adopted ROHC for efficient use of radio resource. Figure 1 shows where header compression is implemented in LTE and WiMAX network.

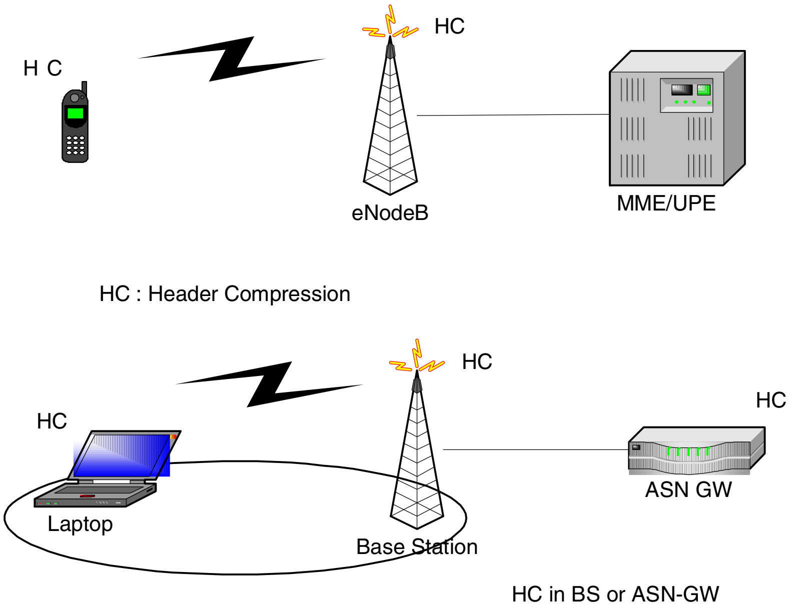

Figure 1. Header Compression in LTE and WiMAX Equipment

3 Process and Benefits of ROHC

Header compression is the process of reducing protocol header overhead in order to improve link efficiency while maintaining the same end-to-end transparency, as shown in Figure 2.

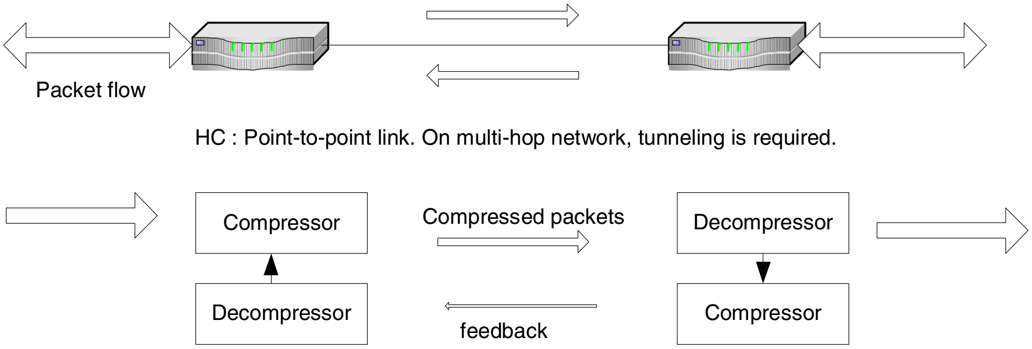

Figure 2. Process and Benefits of ROHC

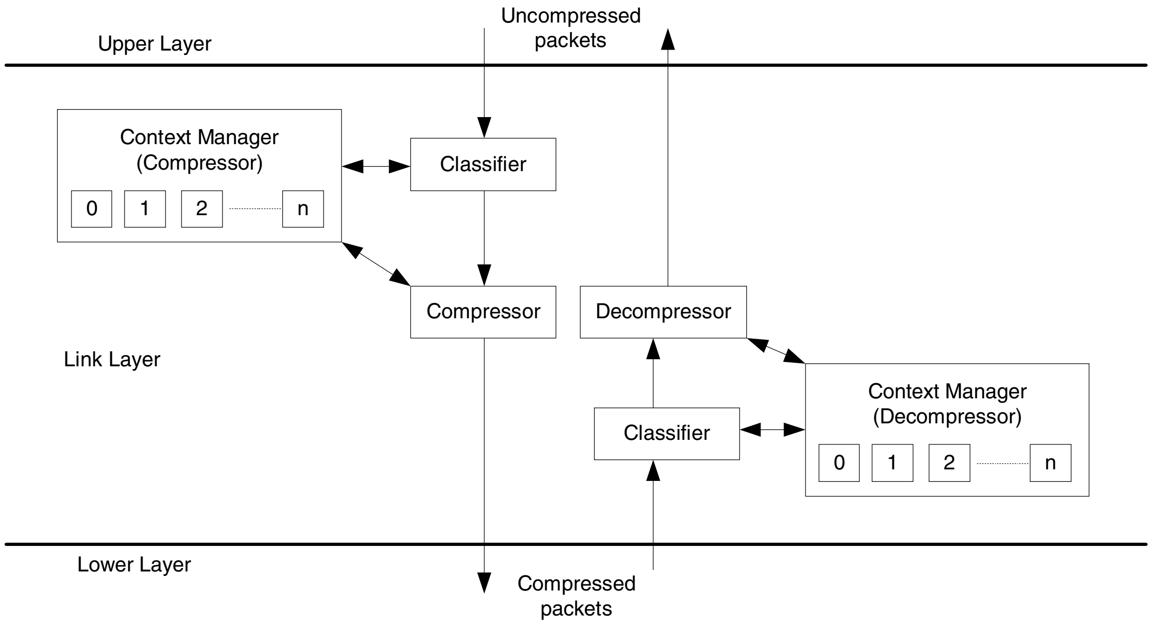

Figure 3. Header Compression Functional Blocks

Figure 3 shows the major functional blocks involved in header compression. Flow context is a set of values of specific packet header fields and the patterns of changes in those fields over time. This context is formed on the compressor and the decompressor side for each packet flow. The first few packets of a newly identified flow are sent uncompressed so a context can be determined on both sides. The number of these first few packets varies depending on the bit error rate (BER) and round trip time (RTT) of the specific link. After the context is established on both sides, the compressor compresses the packets as much as possible until the link contexts change or link error rates start to increase.

The W-LSB algorithm together with the feedback mechanism makes ROHC very robust against packet loss; the CRC in packets makes it robust against bit errors on the link and also against errors introduced due to long round trip times. It also results in very high compression ratios, which increase effective link bandwidth and decrease packet-processing requirements in power-sensitive downstream devices such as cell phones.

Figure 4 diagrams the flow of packets over a hypothetical link. It shows that once the context is established, ROHC sends packets with maximum compression. If the field pattern changes (as shown in the diagram by a silence period, that is, when neither party is talking), or if error conditions such as bit errors or packet loss occur, the context should be resynchronized. So more information is sent to the decompressor using slightly larger, but still compressed, packets. Once the context is re-synced the link reverts to fully compressed packets. The mechanisms such as W-LSB encoding, intelligent pattern detection, and feedback make ROHC robust.

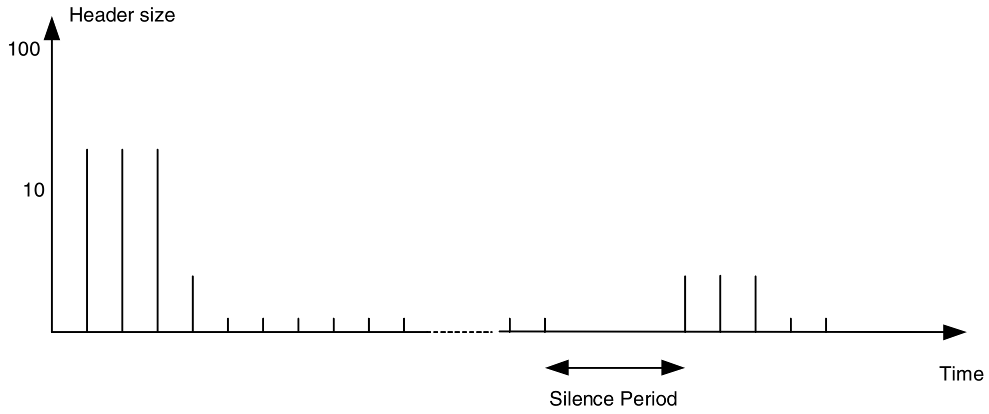

Figure 4. Packet Size Variation During VoIP Flows

ROHC can compress packets down to just 1 byte in most cases. When a compressed packet must carry the UDP checksum, the size increases to 3 bytes as shown in Table 1.

Table 1. Header Compression Grains

Protocol Headers	Total Header Size (Bytes)	Minimum Compressed Header Size (Bytes)
IP4/UDP	28	1
IP4/UDP/RTP	40	1
IP6/UDP	48	3
IP6/UDP/RTP	60	3

Table 2 shows the advantages of using ROHC in wireless networks. The data in Table 2 is derived from a test scenario based on a VoIP packet flow using a header chain of IPv6 (40 bytes) and UDP / RTP (total 20 bytes) carrying one frame of an AMR-NB codec at 12.2 kbps (31 bytes). Such packets were sent every 20ms over a simulated wireless link that used an uncorrelated BER model with a BER level of 10^{-3}. The simulation was run for 120 seconds, during which time 6000 packets were processed. As the table shows, ROHC reduced the total number of packets lost from 52% to 24%. It also decreased the required call bandwidth by 63%, reduced header sizes by 94% and reduced the number of packets lost because of the header field errors from 38% to 3%.

Table 2. Advantages of ROHC

	No ROHC	ROHC
Packets lost	3125/6000	1448/6000
Call bandwidth (kbps)	35.5	12.9
Average header size	60	3.1
Packets lost due to error in header	2309/6000	188/6000

ROHC's benefits are most apparent in codecs that can handle bit errors in their payload, as the data in Table 2 shows. But even without using such a CODEC, ROHC can reduce packet loss by about 50%. On certain wireless links, in fact, the compressed packets can be sent in just one link frame, thus using the radio resource highly efficiently. These characteristics make ROHC the VoIP enabler on wireless networks.

As we have seen, header compression greatly reduces the amount of data that must be sent so the effective link bandwidth is increased. This is extremely important because, as RFC 3095 states “Bandwidth is the most costly resource in cellular links. Processing power is very cheap in comparison. Implementation or computational simplicity of a header compression scheme is therefore of less importance than its compression ratio and robustness.”[1]

ROHC uses various encoding schemes including W-LSB and other specific field pattern-based encoding algorithms. CRC is used in most or all of the packets conveying context information to protect against residual errors and to maximize context validity. Packet loss and reordering, the most common error events on wireless networks, are handled by the compression/decompression method outlined previously.

ROHC must save a larger amount of context information than the length of any single non-compressed header. Depending on how much information is saved and how much information is processed for each packet, the balance between memory utilization and CPU processing load can be achieved. A large context can help to reduce processing per packet but can be costly for memory access (read and write) operations. The processor’s L2 cache also helps to limit the total number of processor cycles required since accessing cached data is more efficient than accessing data from system memory. But the trade-off for reduced processor cycles is that a given system will require larger amounts of memory and higher memory access bandwidth to support the larger amounts of context information.

In addition to saving context, a complete header compression solution also includes packet classification and context management modules. Flow classification can be based on hash-table algorithms and hash-keys are used to identify context information. Flow classification and context management are processor intensive and require suitable memory bandwidth so that the identified context can be accessed as quickly as possible.

Therefore a balance must be achieved between minimizing processing cycles while keeping memory sub-system requirements within bounds. This can be achieved by using sophisticated methods to control context memory size. The goal is to keep memory size and bandwidth within acceptable limits while still ensuring that the processor cycles required for processing and memory access are also within the limits required to attain a specific processor load factor at a high number of flows.

ROHC workloads are very different from typical desktop or server workloads. Unlike graphics or video-processing, for example, ROHC does not use any floating-point operations. Most of the processing involves branch (not loop) processing using logical and integer operations. Maintaining state after each compression and decompression operation is essential, so memory access (read, write, or both) happens frequently – but in a highly random fashion, with few if any sequential memory accesses. The most suitable hardware platform is therefore one in which ample processor performance is combined with a large L2 cache, where memory bandwidth is high, and where branch prediction is highly optimized.

The previous discussion has been focused primarily on network infrastructure equipment with high-end hardware architectures. As previously noted, ROHC is also used in mobile devices, which obviously have

[1] C. Bormann, et al., “Robust Header Compression (ROHC): Framework and Four Profiles: RTP, UDP, EXP and Uncompressed,” RFC3095, July 2001.

far lower processor and battery power as well as far more limited memory size and bandwidth. Mobile devices actually handle a very small number of contexts, since they are dealing only with those on their specific air-interface link, rather than those in effect for the thousands or tens of thousands of flows a Base Station Controller must handle. A good ROHC implementation for user equipment is highly optimized and configured to limit the number of processor cycles it requires, and to use the limited memory resources extremely efficiently.

4 Effnet Solution for Header Compression: Effnet ROHC™

Effnet has been active in IETF for development of header compression standards. It has contributed towards the development of RFC2507 (IPHC), RFC3095 (ROHC) and many others. Its expertise and experience is reflected in the development of products for header compression including header compression simulation environment (Effnet HCSim) and contribution to IETF standardization activities.

Effnet ROHC™ has been developed considering the needs of wireless networks including cellular, satellite and military applications. It includes support for standards within each segment such as 3GPP, 3GPP2, MANET, and so on. As it is used in wide variety of networks and devices, it has been written from ground up to be highly portable and efficient. A mobile terminal is limited by its battery power and processing capabilities whereas system side nodes need to support a very large number of users or connections.

Effnet provides the complete implementation, including all of the features mentioned in the header compression standards. The feedback from customers and partners is reflected via new features, enhanced algorithms and improved efficiency on various platforms.

5 System Partitioning

Typical base station implementations are fully software driven, with only a small portion of the processing steps implemented in hardware (coprocessors) – typically computational intensive blocks such as for example encryption and channel decoding. This makes partitioning of the processing over the different devices available in the market place flexible and subject to the specific vendor requirements. Two typical examples are shown (for an LTE stack) in Figure 5.

The top half of this figure shows a split between a network interface card and a number of channel cards. A channel card in this figure handles a fixed number (typically 1 or 3) sectors worth of OSI Layer 2 and Layer 1 processing and connects to remote RF heads over an OBSAI or CPRI interface. One or multiple channel cards interface to a network interface card (Ethernet backplanes becoming prevalent) that terminates the backhaul connection to the network. As we mentioned before, typically, this backhaul connection is based on Ethernet, but packet protocols over ATM/TDM are also used.

Different partitioning options of the software over the different cards are possible using such a system split, with in this example the hard real-time software (driven by the 1 msec TTI imposed by the LTE requirements) running on the channel card and the soft real-time software running on the network interface card. This soft real-time portion includes ROHC, GTP (de-)tunneling including classification, and air interface encryption, together forming the PDCP sublayer as defined by 3GPP. Safety of the connection between the base station and the network is ensured by running IPsec towards the network. This solution ensures that the packets transported over the backplane between channel card and network interface card are encrypted, leaving no packet snooping possibilities for hackers that physically break into the base station.

A second partitioning option is shown in the bottom half of Figure 5. Here, the network interface card is simplified to a simple Ethernet switch, or removed altogether (we are assuming here that any ATM/TDM network connectivity is handled by an optional external hub that translates to Ethernet). The network traffic is terminated immediately to the channel card, collapsing the architecture of the base station and thus potentially lowering the overall device count. This type of architecture suits well to micro- and picocells.

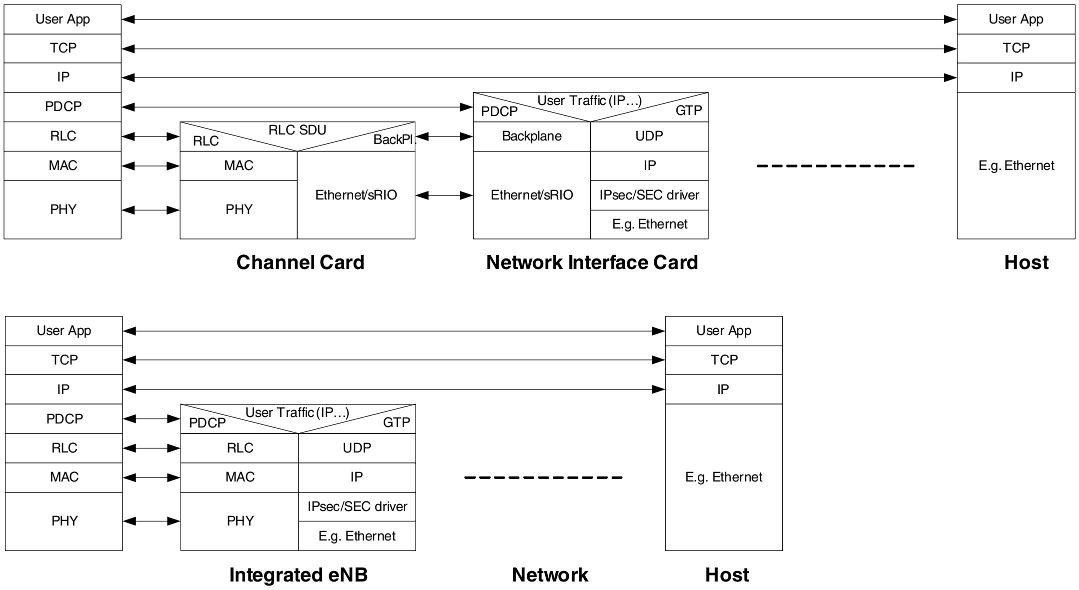

Figure 5. Some System Partitioning Options—Integrated and Distributed Design

6 Freescale Solutions Targeted at Wireless Infrastructure

Freescale has developed an extensive range of digital solutions for the Wireless Infrastructure market. From the StarCore®–based DSP family to the PowerQUICC™ families of communications processors built on Power Architecture™ technology, Freescale has developed a legacy in the infrastructure space which other silicon vendors cannot match.

This portfolio of devices is designed to optimize the performance of the digital processing of wireless infrastructure solutions. Freescale's current multicore DSP the MSC8144 is the third generation of multicore DSP Freescale has produced for Wireless Infrastructure and targets the PHY layer processing for wireless standards such as W-CDMA, CDMA, WiMAX and LTE.

The PowerQUICC families offer a range of scalable Power Architecture based devices optimized for Layers 2, 3 and above and with a range of price, power and performance points to offer an architecture that can be scaled from low end consumer devices such as Femto base stations up to multi-core devices suitable for high end core network applications.

The PowerQUICC devices are designed for optimal system partitioning providing the ideal feature set without wasted silicon and therefore wasted cost. They have several features which optimize the performance of wireless base stations and are currently the only range of processing solutions to incorporate these techniques with low latency interconnect technologies such as Serial RapidIO®, PCI-Express and SGMII. These SERDES technologies are implemented in such a way that there is minimal cost penalty for using one interface but not the other two helping to ensure the solution is as cost effective as possible.

Using the Network Interface or Transport block of a modern base station as an example, the strength of the PowerQUICC devices becomes apparent. From a high level, the requirements for a device aimed at the 3GPP network interface are as follows:

- Enough data path performance to handle PDCP processing for 3 sectors of LTE traffic
- Enough data path performance to handle the E1/T1 requirements for at least 3 sectors of W-CDMA
- The ability to integrate the W-CDMA/LTE control processing into the same device
- Power consumption of ideally 7W or below
- Serial RapidIO, SGMII or PCI-Express connectivity for the backplane

Obviously, this is a very high level summary of what is really required and gives little consideration to the multiple different options that are available for implementing the network interface solution, but straightaway some of the challenges facing network interface designers become apparent. For example, how to design in a solution that provides the required multi-protocol support without wasting silicon by having separate IP blocks for each protocol or standard. What is required here is a programmable approach.

In the case of the PowerQUICC family, to support true multi-protocol functionality, then the QUICC Engine provides an ideal solution. The QUICC Engine can be programmed to support multiple protocols for either termination or interworking. It also provides an excellent means for offload or acceleration of key tasks from the core. So a network interface architect can provide E1/T1 support for W-CDMA and then when LTE becomes the focus can re-program the QUICC Engine to handle some of the PDCP processing, for example the IPSec processing in conjunction with the integrated cryptographic block.

This in itself is not necessarily something revolutionary and in fact can be achieved by several silicon vendors with their Network Processor solutions. What differentiates the PowerQUICC solutions are other key features. Most important of which is the integration of this multi-protocol network functionality with an e500 Power Architecture Core. This provides significantly more processing performance than other such devices within an overall device power window matching other lower performance solutions. The e500 core allows the integration of control plane processing or for the increased processing required by the PDCP layer of 3G LTE.

In addition there is the aforementioned integrated Security Engine. As the new wireless standards evolve security is becoming more and more important. Within LTE the network interface may have to handle encryption over the IP backhaul to the core network as well as the encryption to the UE. This means that providing an integrated security engine capable of handling the throughput of the IPSec on the backhaul and the encryption out to the UE can be a distinct performance advantage.

As not only the security requirements but also the throughput levels increase one of the simplest and most effective things that can be done is the implementation of an efficient memory architecture. This includes optimized cache memories and integration of the memory controller inside the PowerQUICC device. This reduces the latency between the core and the external memory saving critical clock cycles in situations where a large number of subscriber stations need to be processed by the base station.

ROHC provides an ideal example here, whereby if the active flows can all be maintained in cache then the system will obviously achieve greater performance. So looking at a 3 sector LTE system and assuming 1000 active ROHC flows for 1000 active users. Then we can select a device targeted for exactly this level of performance. In this case, Figure 6 shows the performance impact of supporting up to and then more than 1000 active users on the chosen PowerQUICC3. As can be observed anything over this number of active users would require access to external memory. Given that cache memory can be expensive in terms of silicon real estate and therefore cost of the device there are two things to consider here. Firstly, there has to be a balance between the correct size of cache for the exact system requirements and the resulting silicon size and associated cost. The approach of throwing huge caches and high performance cores employed by some silicon vendors is not always the most appropriate. The second thing to consider is what happens when there is a requirement to go to external memory. By integrating the memory controller the performance hit of having to access the external memory is reduced. Of course this impact can also be reduced further by the use of smart pre-fetching algorithms, which is exactly what would be done for systems where the number of users looked to be increased beyond the caching capability.

As a result Freescale has developed a range of different cache options within the PowerQUICC family of devices, allowing OEM’s to cater their solution towards the targeted number of active users in the system. All of these devices however have an integrated memory controller.

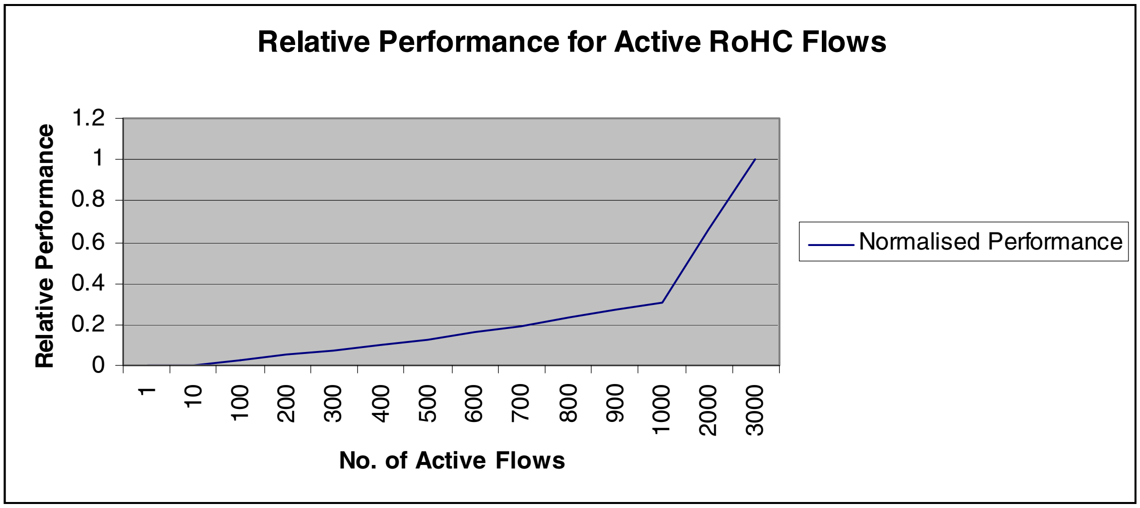

Figure 6. Relative Performance for Active ROHC Flows

A secondary feature in the caching capabilities of the PowerQUICC devices is the cache stashing of arriving packets from the Network Interface. Cache stashing lowers both the memory latency as well as the memory interface throughput requirement, keeping more bandwidth available for other applications running in parallel. If the network interface card also processes IPSec and air interface security, a high throughput offload co-processor is needed for performing these cryptographic operations without heavy core dependency. If this offload block is integrated on the same device it further reduces the latency required to process the data, as well as reducing system cost.

Freescale and Effnet have cooperated to ensure that the Effnet software runs efficiently on the Freescale PowerQUICC range of devices. The e500 Power Architecture core is an ideal embedded platform providing a very high level of efficiency, low power consumption and low pricing. Effnet have created a highly optimized Robust Header Compression solution, this combined with the PowerQUICC architecture makes for a compelling solution for next generation network interface cards. The Effnet and Freescale development teams have worked to integrate the ROHC software with the existing Freescale LTE software solution and fully understand the optimum partitioning of the system for power consumption, data throughput and price. This can support and in many cases accelerate development time of wireless base station solutions.

As discussed previously the ROHC workloads are ideally suited to an architecture developed for wireless infrastructure and with the right choice of silicon there is often not a need for high performance and high power devices aimed at desktop or server applications. For solutions where cost, power consumption and flexibility is of primary importance the requirement is for a highly optimized, highly integrated, programmable and low cost silicon solution, where silicon real estate is kept at a minimum, power consumption is kept as low as possible and it is enabled by a highly efficient software solution. This makes the combination of Freescale and Effnet ideal for network interface solutions for current and future wireless standards.

7 Additional Information

Details on the performance data, system partitioning, the best devices for the profile and throughput rates required as well as reference software for the PDCP, MAC, RLC and PHY later of 3G LTE are available on request under NDA from your local Freescale.

8 References

1. C. Bormann, et al., “Robust Header Compression (ROHC): Framework and four profiles: RTP, UDP, EXP and uncompressed,” RFC3095, July 2001.

9 Revision History

Table 3 provides a revision history for this application note. Note that this revision history table reflects the changes to this application note template, but can also be used for the application note revision history.

Table 3. Document Revision History

Rev. Number	Date	Substantive Change(s)
0	02/11/2008	Initial release.

Freescale Contact Information:

Home Page:
www.freescale.com

Web Support:
http://www.freescale.com/support

USA/Europe or Locations Not Listed:
Freescale Semiconductor, Inc.
Technical Information Center, EL516
2100 East Elliot Road
Tempe, Arizona 85284
+1-800-521-6274 or
+1-480-768-2130
www.freescale.com/support

Europe, Middle East, and Africa:
Freescale Halbleiter Deutschland GmbH
Technical Information Center
Schatzbogen 7
81829 Muenchen, Germany
+44 1296 380 456 (English)
+46 8 52200080 (English)
+49 89 92103 559 (German)
+33 1 69 35 48 48 (French)
www.freescale.com/support

Japan:
Freescale Semiconductor Japan Ltd.
Headquarters
ARCO Tower 15F
1-8-1, Shimo-Meguro, Meguro-ku
Tokyo 153-0064
Japan
0120 191014 or
+81 3 5437 9125
support.japan@freescale.com

Asia/Pacific:
Freescale Semiconductor Hong Kong Ltd.
Technical Information Center
2 Dai King Street
Tai Po Industrial Estate
Tai Po, N.T., Hong Kong
+800 2666 8080
support.asia@freescale.com

For Literature Requests Only:
Freescale Semiconductor
Literature Distribution Center
P.O. Box 5405
Denver, Colorado 80217
+1-800 441-2447 or
+1-303-675-2140
Fax: +1-303-675-2150
LDCForFreescaleSemiconductor
@hibbertgroup.com

Document Number: HDRCMPWP
Rev. 0
02/2008

Effnet Contact Information:

Effnet AB, Head quarter:
Gustavslundsvägen 151C
Box 15040
SE-167 15 BROMMA
Sweden
Phone: +46 (0)8 564 605 50
Fax: +46 (0)8 564 605 60
E-mail: info@effnet.com

Website: www.effnet.com

Information in this document is provided solely to enable system and software implementers to use Freescale Semiconductor products. There are no express or implied copyright licenses granted hereunder to design or fabricate any integrated circuits or integrated circuits based on the information in this document.

Freescale Semiconductor reserves the right to make changes without further notice to any products herein. Freescale Semiconductor makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Freescale Semiconductor assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Freescale Semiconductor data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Freescale Semiconductor does not convey any license under its patent rights nor the rights of others. Freescale Semiconductor products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Freescale Semiconductor product could create a situation where personal injury or death may occur. Should Buyer purchase or use Freescale Semiconductor products for any such unintended or unauthorized application, Buyer shall indemnify and hold Freescale Semiconductor and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Freescale Semiconductor was negligent regarding the design or manufacture of the part.

Performance tests are done using a specific Effnet ROHC™ release and test application and reflect performance of Effnet ROHC™ on a specific Freescale product using specific testcases and test environment. Any difference in system hardware or software design or configuration may affect actual performance. For more information on performance tests and Effnet products, visit Effnet website (www.effnet.com).