The Emerging Need for Fronthaul Compression

David Brubaker, Strategic Marketing Manager, Altera, Now Part of Intel
Wei Qian, FAE Manager, Altera, Now Part of Intel
Melissa Sussmann, Solutions Marketing Manager, Altera, Now Part of Intel
Yoshifumi Takafuji, Field Application Engineer (FAE), Altera, Now Part of Intel
Mohammad Akhter, Architecture Director, IDT
Trevor Hiatt, Product Manager, IDT

The Common Public Radio Interface (CPRI) standard has shown a significant increase in bandwidth requirements as carrier aggregation and multiple-input multiple-output (MIMO) features of LTE-A are deployed. With maximum line rates at 24 Gbps and an increase in bandwidth requirements, compression techniques become a necessary part of CPRI core moving into the 4.5G and 5G era. This paper reviews the increase in requirements since 2006 (3G era) to 2016 (4.5G era). Even a maximum of 24 Gbps line rate will not have enough capacity in the near future to service the maximum downlink speeds defined by 3GPP R12 and techniques like compression are necessary today. In the case of LTE-A CPRI, compression will enable the fronthaul network to support LTE-A with minimal impact on the optics and fiber network. This white paper then reviews two viable compression technology offerings from Altera and IDT.

Introduction

First and second generation (1G and 2G) wireless Base Transceiver Station (BTS) systems had all functions such as radio frequency (RF), baseband, and power supply in a cabinet. Usually these systems require building space equipped with uninterruptible power supplies and cooling and network backhaul access, which restrict locations and made deployment of a cell-site expensive because each site require full infrastructure. In this architecture, the RF power amplifiers were located with the BTS and RF was delivered to the antenna via low loss, expensive RF cables. With the transition to 3G, operators and infrastructure suppliers broke the BTS into a base-band unit (BBU) and remote radio head (RRH) to solve these issues and defined a standard interface[Note 1] between them: CPRI. The later architecture is also referred to as evolved Node B (eNB).

The CPRI standard, first released in 2006, is an Open-Standard-Interface which distributes BTS functions between a BBU and a RRH. As shown in Figure 1, the BBU and RRH are connected with an optical fiber, over which, the user data and control data are transmitted using the CPRI standard. The connection between the BBU and the RRH recently designated fronthaul to distinguish it from the connection between the BBU and core network or backhaul connections in the Centralized Radio Access Network (C-RAN) architecture.

☞ Note 1: the Open Base Station Architecture Initiative (OBSAI) was also defined for the BBU to RRH connection but has largely been replaced by CPRI and is a legacy standard today.

**Figure 1. Typical BBU and RRH Configuration with CPRI Interconnection**
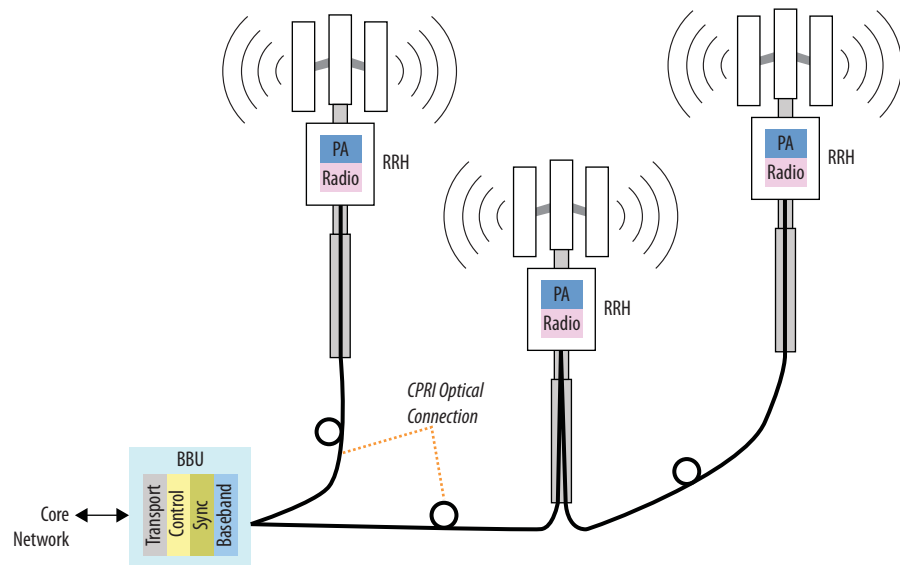
An advantage of the distributed BBU or RRH architecture is a full infrastructure such as sheltered building, cooling, and back-haul are only required at the BBU and the RRH can be placed in remote locations. The length of optical fiber can be several kilometers allowing optimum placement of the antenna. Compared with traditional BTS equipment, the RRH is light, flexible, low-cost and can be mounted at the top of the tower thus reducing RF loss between the PA and antenna. The BBU can be located near the network and serve multiple RRHs to further reduce network costs.

Fundamentally the CPRI connection provides a user plane (user data), control plane (control data), time and frequency information from the BBU to the RRH. The user data is provided in IQ format, which is used to directly modulate the RF signal. The control data is used to communicate between the BBU and RRH for configuration and monitoring. Time and frequency information are required to synchronize the RRH with the BBU and network enabling proper wireless operation.

CPRI for 3G and 4G Networks

The first generation CPRI standard (1.4) provided for ~2 Gbps data rate between the BBU and RRH, which was adequate for 3G systems. As 3.5G and 4G systems were introduced the CPRI BW requirement increased and today, a 2x2 MIMO system with one 20 MHz carrier requires 2.46 Gbps of CPRI bandwidth per sector (assuming 8B/10B encoding). A typical RRH will have three sectors and hence a 3-sector, 2x2 MIMO single 20 MHz RRU requires 7.38 Gbps data over the CPRI link. Figure 2 shows the timeline of CPRI releases with corresponding line rates.

LTE-Advanced (LTE-A), 3GPP, Release 12, approved in March of 2015, defines higher downlink speeds by implementing carrier aggregation and higher MIMO configurations. The LTE-A 3GPP R12 provides up to 1 Gbps downlink speeds; but as of 2016 networks implement few, if any, LTE-A features such as carrier aggregation and 8x8 MIMO. Today the average CPRI connection is between 5-6 Gbps and as operators deploy LTE-A features there will be a corresponding increase in CPRI data rate between the BBU and RRH.

**Figure 2. Timeline of CPRI Releases and Line Rates**
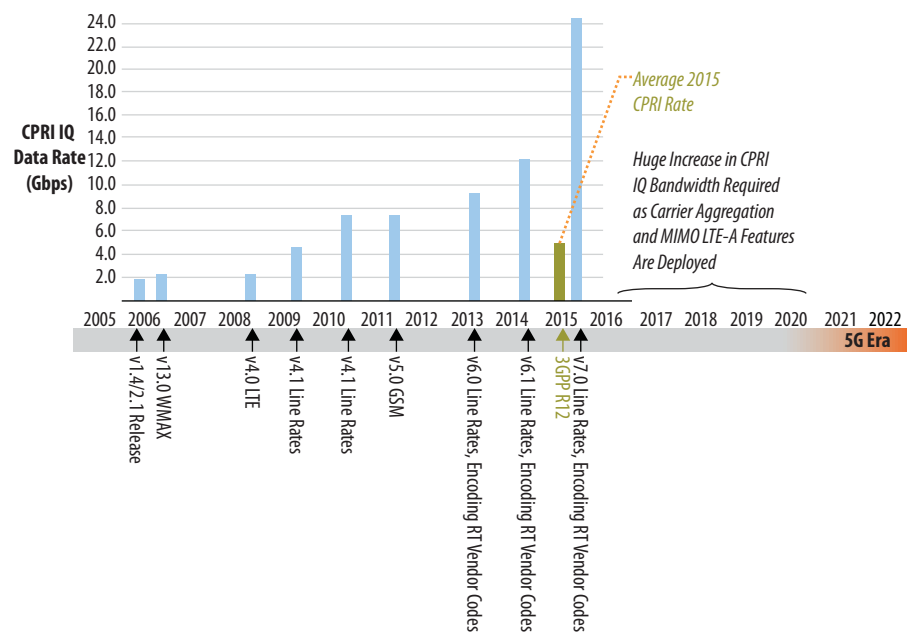
In anticipation of the required data rates for LTE-A the CPRI standards group released versions 6.0, 6.1, and 7.0 provide for more efficient encoding using 64B/66B and a maximum line rate of 24.33024 Gbps. Unfortunately even these line rates will not have enough capacity to service the maximum downlink speeds defined by 3GPP R12 and techniques like compression will be needed.

CPRI for LTE-A

We can determine the impact of LTE-A on CPRI line rates by selecting use cases that are consistent with 3GPP R12. The actual antenna and number of carrier is optional but a maximum of 8x8 MIMO and 5x20 MHz carrier aggregation is defined.

Figure 3 shows the required CPRI data rate based on typical use models without encoding. The appropriate encoding factor for 8B/10B or 64B/66B would increase the required data rate. As you can see from Figure 3, the maximum use case defined in 3GPP R12 requires a CPRI data rate of about 40 Gbps. This data rate is well beyond optics used in today's CPRI links, which on average support rates of 5-6 Gbps. Operators will avoid installing new fiber and will work to implement LTE-A with installed fiber and backhaul equipment.

Operators use the most cost effective optics for a given deployment scenario and higher-speed and single-mode optic modules are more expensive than multi-mode, lower-speed modules. Exceptions occur when a high-volume application optimizes costs for a particular line rate. Above 6 Gbps CPRI line rates similar to 10GE Ethernet have proven economical. Optic modules that support bandwidth greater than 10 Gbps are increasingly more expensive and as 25GE becomes widely deployed, a jump to 24 Gbps CPRI line rates may occur. Even with a jump to 24 Gbps CPRI, a single fiber will not support the maximum LTE-A configuration for a cell-site let alone typical ring or daisy chain architectures. A new approach is required; one method is CPRI compression.

Figure 3. CPRI Line Rates for Typical LTE-A- Use Cases

Raw CPRI Data Rate vs. MIMO and Carrier Aggregation (per Sector)
(without 8B/10B or 64B/66B Encoding; Scale Final Rates by Appropriate Encoding Factor)
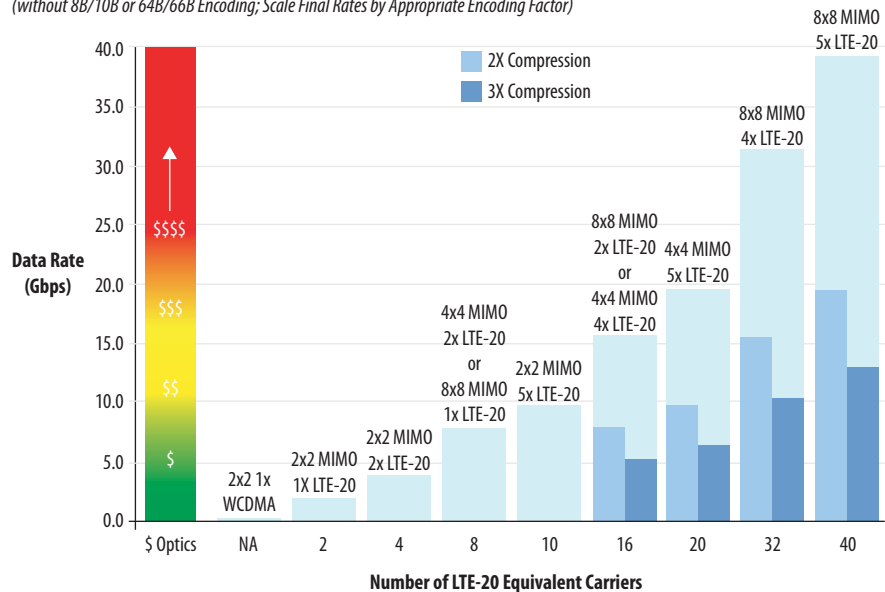
CPRI compression algorithms can be divided into lossy and lossless algorithms. A lossy CPRI compression algorithm will degrade signal quality by compressing the data in a manner that cannot be recovered but have the advantage of being simple to implement. Today, lossy compression algorithms of 2:1 can be implemented with minimal logic and cost. Compression ratios larger than 2:1 require more complex logic, which increases cost and power. At present, compression algorithms can reach 3:1 and still meet wireless standards. Lossless CPRI algorithms that will not degrade signal quality require even more exotic algorithms with corresponding increase in logic complexity.

Also shown in Figure 3 is the required CPRI data with 2:1 and 3:1 compression algorithms. The use case: 4x4 MIMO with 5x 20 MHz LTE carriers, a very realistic near term use case, requires a 20.37 Gbps CPRI data rate (64B/66B encoding) while a data rate of 10.16 Gbps and 6.77 Gbps would be required for 2:1 and 3:1 compression, respectively. Compression significantly reduces CPRI data rate requirements and provides system cost savings by using lower data rate optics.

The maximum use case in LTE-A is 8x8 MIMO and 5x 20 MHz LTE carriers requires a CPRI line rate of 40.5 Gbps (64/66B encoding) without compression. A 3:1 compression can reduce the required line rate to 13.5 Gbps which is achievable today at moderate cost.

CPRI compression techniques will be required to implement maximum LTE-A downlink speeds using cost-effective optics. Altera provides both an internal algorithm and supports partner solutions.

Altera Compression Algorithm

Altera offers, previously reported, lossy Mu-law compression algorithm with our IP-CPRI-V6 core. This algorithm uses the Mu_compand_val =8 to balance orthogonal frequency-division multiplexing (OFDM) error vector magnitude (EVM) and implementation size. A system block diagram showing the location of the optional compression IP is shown in Figure 4.

Figure 4. System Block Diagram Showing Optional Compression IP Blocks
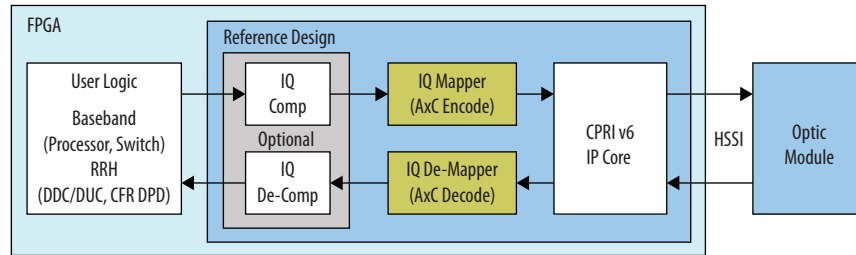
Altera compression algorithm permits varied input and output bit widths however our reference design assumes a 16 bit input with user selectable output bit widths of 12, 10, 8, 7, and 6. Figure 5 shows simulated and measured EVM along with compression ratio for various output bit widths. The measured EVM floor is 55 percent due to residual jitter and system impairments of the radio and measurement system. The EVM limit for 3GPP is eight percent so a 2:1 compression has minimal impact on overall system performance yet doubles the available CPRI capacity.

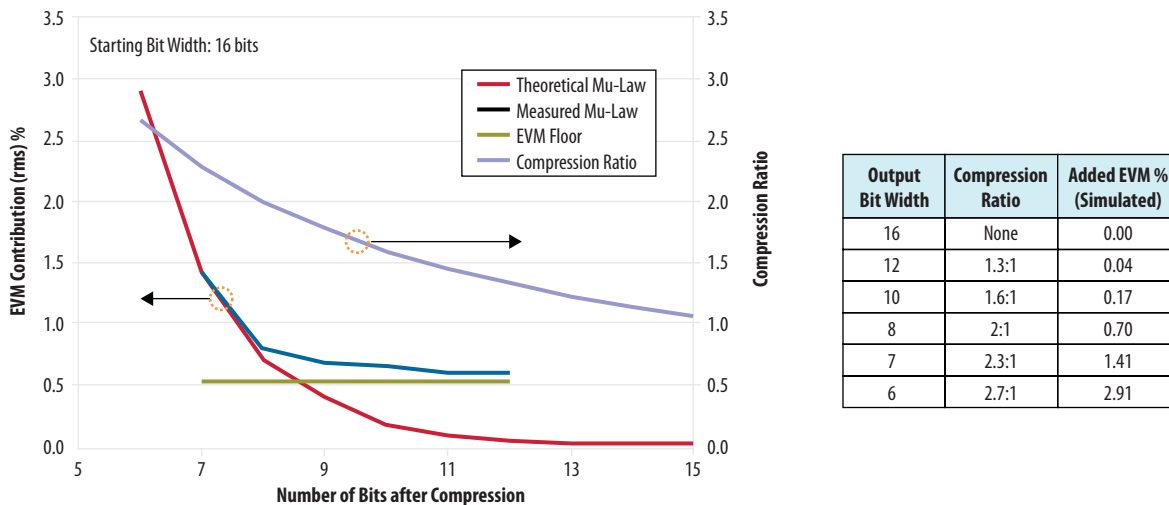**Figure 5. Theoretical Added EVM and Compression Ratio vs. Output Bit Width**
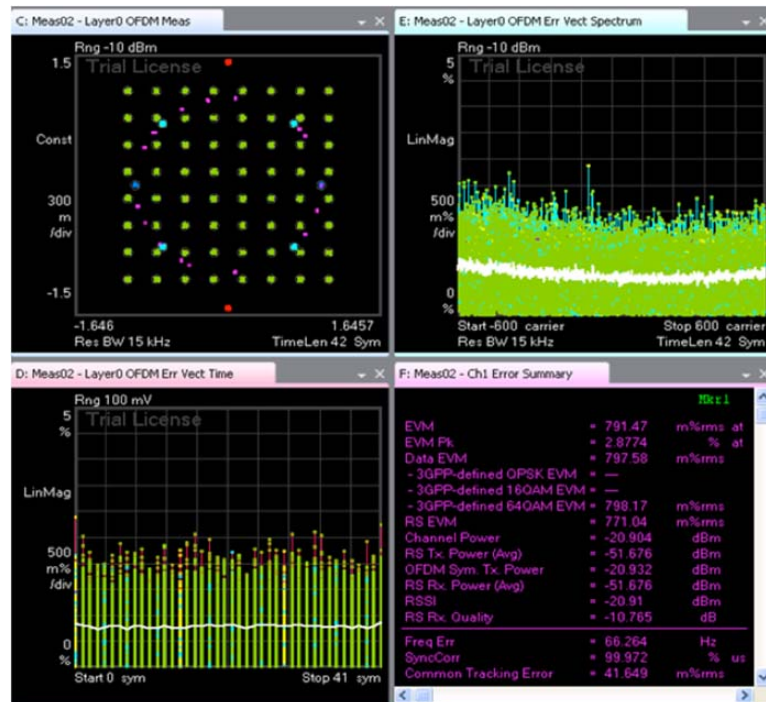

Output Bit Width	Compression Ratio	Added EVM % (Simulated)
16	None	0.00
12	1.3:1	0.04
10	1.6:1	0.17
8	2:1	0.70
7	2.3:1	1.41
6	2.7:1	2.91

Figure 6. Measured EVM for 2:1 Compression

A measured EVM of 0.8 percent at 2:1 compression (16 to 8 bits) is shown in Figure 6 which is reasonable as it includes residual EVM contributions from measurement equipment. For the Altera algorithm, compressor requires less than 15 adaptive logic module (ALM) for implementation and has negligible latency. Customers have a range of choices for Altera® FPGAs, including low-cost and mid-range products, such as Cyclone® V GX, Arria® V GX, and Arria 10 GX devices.

For compression ratios beyond 2:1 the EVM degradation of the Altera algorithm increases rapidly and will limit system performance. Algorithms that can provide adequate system performance at higher compression ratios are available from Altera partner IDT and can provide more CPRI capacity but generally require more digital hardware for implementation.

# IDT IQ Compression and Decompression IP Overview

IDT is a leader in wireless signal compression technology and low latency switching fabric technology, and has offered compression IP in production release worldwide since 2013. IDT's "I2Q Compression IP" leverages a unique algorithm optimized for wireless signals and has multiple protected patents worldwide.

The "I2Q Compression IP" offers high performance compressor and decompressor cores. This IP may be deployed in FPGAs or ASICs, and feature a small silicon footprint. This IP has been proven in Altera Cyclone and Stratix® FPGA families.

The IP features user-selectable compression ratios ranging from 1.5:1 to 3:1. With a wide range of compression ratios and a compression ratio granularity of 0.1 the IP can be used to support variable data rates and transceivers typically needed for wireless as well as optical fronthaul links. This IP is optimized to support GSM, WCDMA, and LTE/LTE-A signals at full CPRI data rates, keeping high signal quality at up to 3:1 compression rates. It is ideally suited for wireless fronthaul for base stations in

macrocell, C-RAN, small cell, wireless repeater, as well as distributed antenna system (DAS). In addition, the compression IP can be used to reduce the fabric link speed as well as the capacity requirement for sample storage in wireless base-station processing. The IP is agnostic to the link protocol and can be used for CPRI, Ethernet, or RapidIO® link protocols that are widely used in the evolving wireless access network.

Figure 7. Compression and Decompression Capabilities in Wireless Baseband and Remote Radio Head
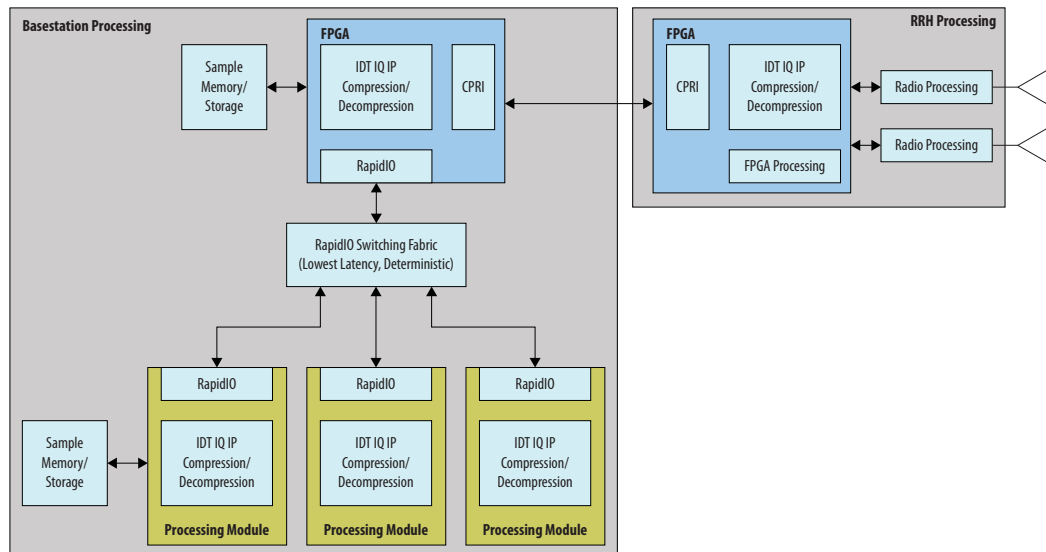
Features and Capability

The compressor and decompressor cover a wide range of packet sizes and IQ sample widths. It also handles critically sampled, oversampled signals as well as bursty signals.

IDT's IP allows performance optimization of EVM, latency, jitter, and complexity to best meet system requirements. Latency and latency jitter can be optimized based on the signaling protocol (e.g., GSM, WCDMA, and LTE) and the system deployment scenarios (e.g., C-RAN, small cell, macrocell). The IP can also handle multiple concurrent data streams.

The IP offers comprehensive error and event management capabilities for easy identification, notification of events and faults within the system.

This flexible IP is readily integrated in FPGA or ASIC with standard bus interfaces, and configurable buffer sizes for optimal utilization of available silicon resources.

The IDT IQ typical features and example performance results are summarized below:

- 1.5:1 to 3.0:1 compression ratios

- EVM performance 0.5 percent to 3.5 percent for typical 3G and 4G/LTE-A wireless signals for 2:1 to 3:1 compression

- Microsecond level latency and can be traded based on performance targeted

- High-performance core supporting uncompressed data rate up to 9.8304 Gbps with each instance of the IP

- Common clock rates of 61.44 MHz, 153.6 MHz for FPGA and 307.2 MHz for ASIC

Performance Results

Figure 8 demonstrates typical IDT IQ IP performance with various compression ratio and packet sizes.

Figure 8. Compression Ratio vs. Average EVM Percentage for IDT IQ IP
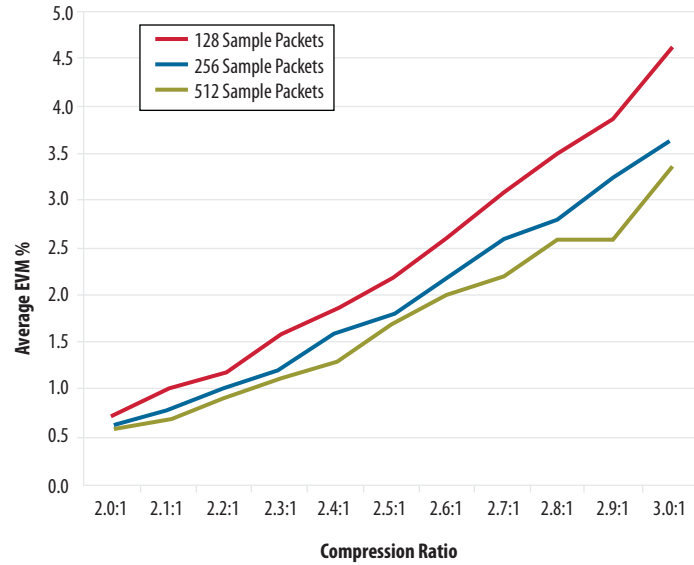
IDT IQ IP performance with 20 MHz LTE TM3.1 signal is shown in Table 1.

Table 1. IDT IQ IP Performance with 20 MHz LTE Carrier

| Parameters	Values
Downlink signal test mode	TM 3.1
Compression ratio	3:1 using IDT IQ compression
Average EVM	2.69% rms
Carrier bandwidth	20 MHz
Protocol	LTE 64 QAM

Figure 9 and Figure 10 shows the 64QAM LTE signal spectrum performance with IDT IQ IP

**Figure 9. LTE TM 3.1 Signal with 64 QAM without Compression**
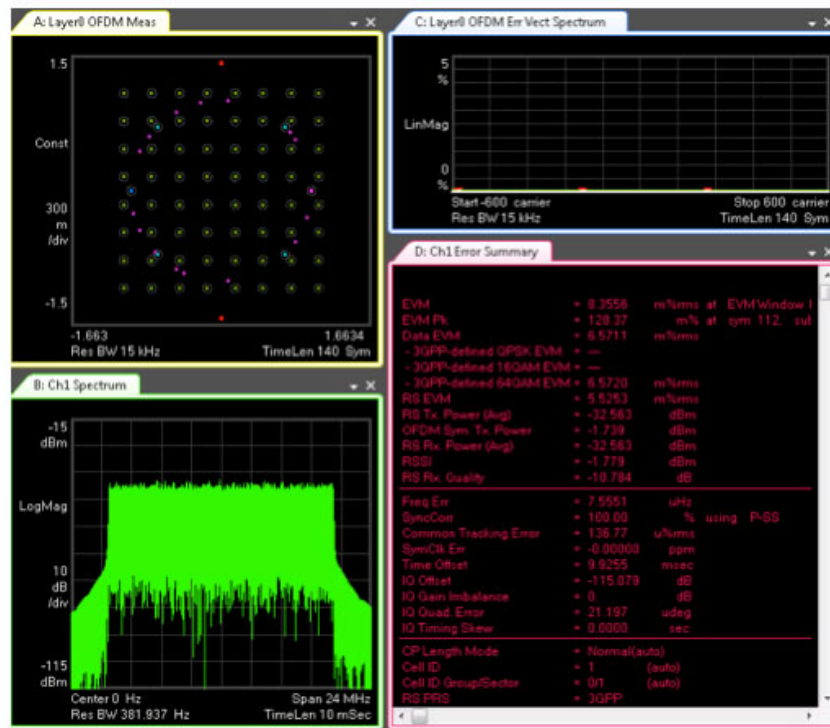
Figure 10. LTE TM 3.1 Signal with 64 QAM with 3:1 Compression for Average EVM of 2.69%
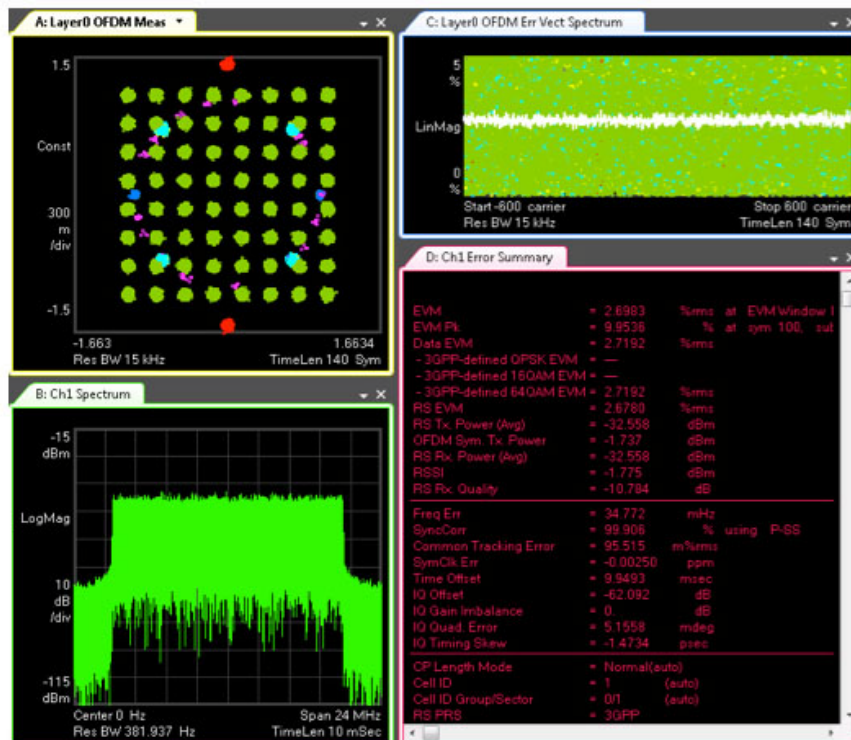
IDT 12Q IP Deliverables

IDT's I2Q compression IP deliverables include obfuscated compressor and decompressor register transfer level (RTL) and associated documentation, customer testbench (CTB) and associated user manual with example test signals and test results, and Synopsys® Design Constraints (SDC). IDT provides technical support of the IP throughout the development lifecycle and offers flexible licensing options.

Figure 11. IDT IQ IP and Testbench Deliverable as Part of the IP Package
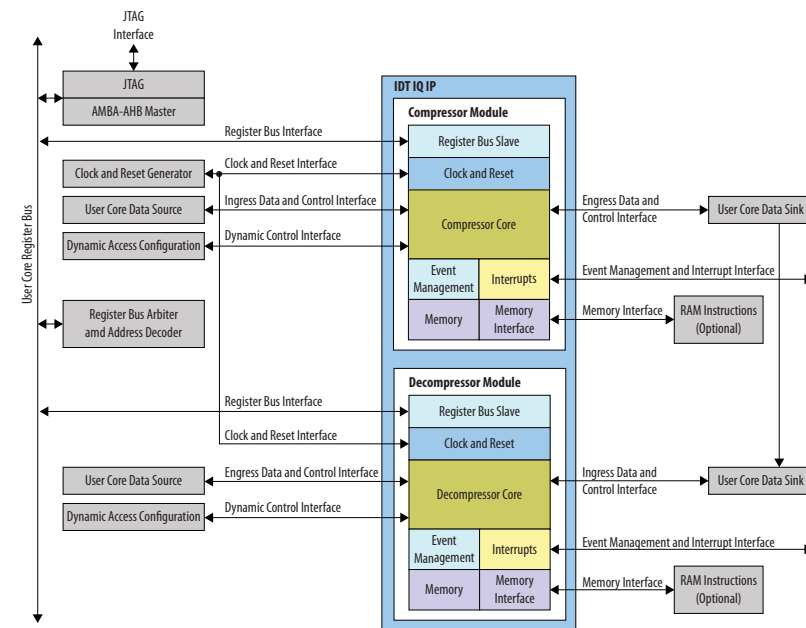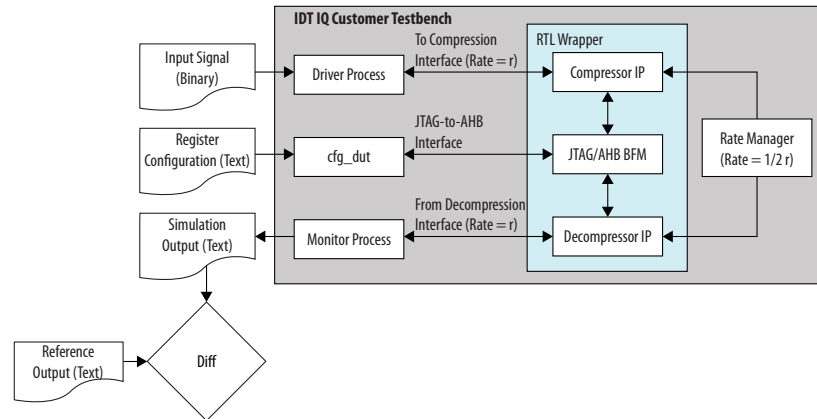
Impact of 5G on Fronthaul

As users are becoming comfortable with 4G speeds the wireless industry is beginning to develop 5G systems in anticipation of even higher download speeds and performance needs. 5G will provide up to ten times 4G user download speeds reaching 10 Gbps and much shorter latency. In order to accomplish these, data-rates techniques like massive MIMO will be used to improve spectral efficiency. In a massive MIMO architecture up to 128 individual antennas are used to improve SNR of each user while simultaneously providing interference cancellation. In addition to

massive MIMO, operators are planning for wider transmission bandwidth available higher frequency bands. The net result of massive MIMO and increased transmission bandwidth is an explosion in fronthaul data rates. Data rates of multiple hundreds of Gbps are easily projected for typical systems. Obviously the fronthaul network will not be able to support these data rates and we will see a major change in the fronthaul architecture for 5G systems.

This data rate limitation of current CPRI standard for 5G and beyond has been recognized by industry and there are several industry movements to address the problem. At the same time the basic 5G architecture is evolving in different directions with C-RAN, DAS, and small cell architectures; and a universal fronthaul solution may not be achieved.

Baseband Re-Partitioning

The most popular proposal to solve the fronthaul data-rate issue is baseband re-partitioning. In baseband re-partitioning, some of the processing currently in the BBU is moved to the RRH thus reducing fronthaul data rates and the residual processing in the BBU can be remote or moved to a centralized location in the C-RAN architecture. The fronthaul interface requires consistent, repeatable, and minimal latency along with high data rates and low error rates.

Re-partitioning moves some of the processing from the BBU to the RRH, and some back to the C-RAN. The optimum partition depends on many factors and is likely will vary depending on the overall architecture and application. For example a small cell head would likely have all the processing moved to the RRH with the data interface being an Ethernet connection while a full C-RAN may have more processing in the C-RAN location and less in the RRH.

Radio over Ethernet

Another proposal is to replace CPRI with Ethernet and transmit the IQ packets over Ethernet with each RRH looking more like a switch or router. This is called Radio over Ethernet (ROE) and is addressed by IEEE 1904.3. From a high-level view, this looks like a network simplification but there are several difficult problems to solve with ROE.

First, Ethernet is a best efforts communication protocol and is by nature bursty, rather than streaming like CPRI. This can be mitigated with buffers at the expense of latency. Second, traditional Ethernet does not have the time and frequency accuracy required by wireless systems. IEEE 1904.3 assumes the time-frequency requirements are solved by 802.1cm (Time Sensitive Networks for Fronthaul). Lastly, ROE using IQ does not inherently solve the bandwidth issue for 5G since the IQ infers transmission of user IQ data defined in the CPRI standard. IQ over Ethernet can strip off the 8B10B encoding and save 20 percent but line rates that use the 64B/66B encoding can only save three percent overhead. Hence, ROE will likely implement some form of baseband re-partitioning to move bandwidth expansion to the RRH.

Conclusion

Recent line rate increases in the CPRI standard are not sufficient to meet the coming demands of LTE-A or 5G. For LTE-A, CPRI compression will enable the fronthaul network to support LTE-A with minimal impact on the optics and fiber network. For 5G, data rates will increase by on order of magnitude to 100's of Gbps and a major overhaul of the fronthaul network will be required. For massive MIMO and C-RAN architectures, we expect that the processing that is done by the BBU will be moved to the RRH thus reducing the required fronthaul data rate.

Document Revision History

Table 2 shows the revision history for this document.

Table 2. Document Revision History

Date	Version	Changes
June 2016	1.0	Initial release.