

Next-Generation Wireless Infrastructure

Requirements and Implementation

2014

Introduction

The use of wireless technology has exploded over the last decade, driven by the extraordinary growth in the use of smartphones around the world. Increasing network capabilities have made possible greater data functionality on these phones – and that functionality has created a hunger by users to consume even more data.

Current wireless network capacity is insufficient for meeting these demands, and the challenges are such that simply attempting to scale the current architecture will not work. New architectural approaches are needed, and they will require new SoCs; demanding the highest possible performance IP cores. What follows is a discussion of the issues motivating the need for change, what those changes should look like, and the kinds of cores and IP that can lead to success for providers of network equipment.

Market Dynamics

Wireless data traffic is being driven by two major contributors: smartphones and tablets. Market forecasts predict rapid smartphone growth, and although smartphones currently represent only 18 percent of total global handsets, they generate 92 percent of total global data traffic. The 7 billion smartphones that will have been sold by 2017 will have a substantial effect on the overall amount of data traffic.

The tablet market, meanwhile, increased 2.5-fold in 2012. Each tablet generates 2-3 times more data traffic than the average smartphone.

This growth has driven many of the changes forced on the networking and infrastructure markets. Estimates show traffic rising over the cellular network by 7x between now and 2017 (Figure 1).

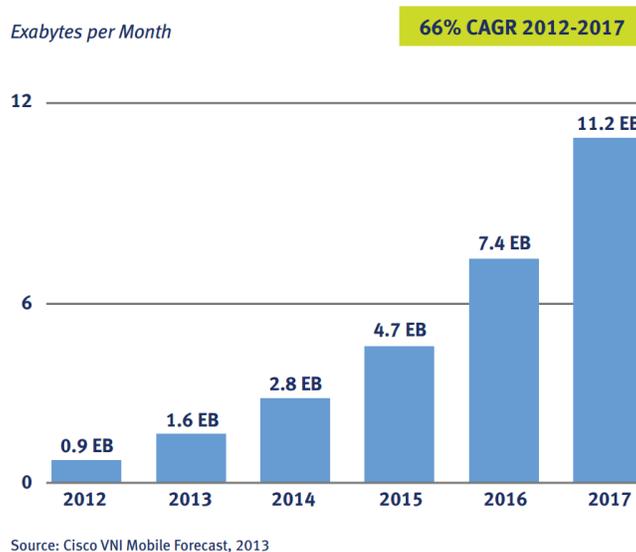


Figure 1. Global Mobile Data Traffic, 2012 to 2017 (Source: [Cisco VNI](#))

Social media and video traffic are the main drivers of data volume growth. This will be further compounded by a growing desire by subscribers to be “always on,” with contextual data being exchanged even when the subscriber is not actively using the device. In addition, wire-line subscribers are migrating towards more balanced usage of wired and wireless services.

A [report from Cisco](#) highlights that mobile video traffic will exceed 50 percent for the first time in 2012, and it forecasts that, by 2017, two thirds of the world's mobile data traffic will be video traffic. This will be driven by the availability of higher bandwidth and better-quality connections on the cellular network.

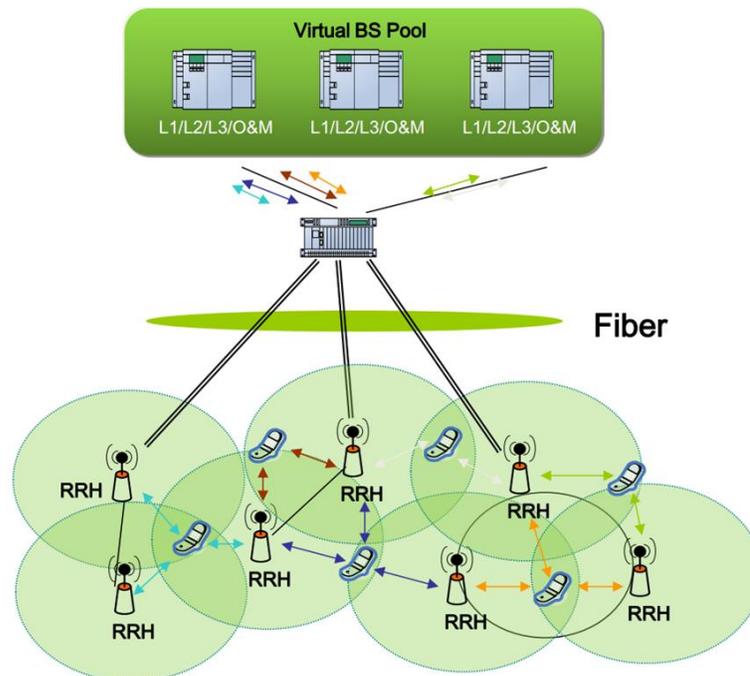
Since subscriber use of video will be largely responsible for the majority of this increase in data traffic, mobile network operators must figure out how to distribute this content over their cellular networks. And, in doing so, they must provide the “quality of experience” that their subscribers will demand.

Today's wireless network topology has base stations at the edge of the network providing the air interface connectivity and management, access and aggregation networks providing the pipe into the network, and servers provisioning the content from the opposite side of the core network. This architecture makes simply scaling the network infeasible, both from an operating and a capital standpoint.

The newer LTE-Advanced (LTE-A) and HSPA+ wireless data standards offer more spectrum with wider channel bandwidth, better spectral efficiency, and WiFi offloading. Future networks leveraging these technologies will move in one of two directions: towards a mix of heterogeneous small cells (the so-called HetNet) or towards a Cloud RAN (C-RAN) architecture, depending on which will offer the best coverage and capacity for the subscriber base.

This choice allows service providers to implement different clusters of baseband capability, utilizing either local (HetNet) or remote (C-RAN) antennas both as a medium for servicing the air interface as best as possible and for delivering content as close as possible to the edge of the network. This allows the network operators to provision the best schemes for policy, quality of service (QoS), classification, packet delivery schemes, and management of the air interface across the range of radio access technologies possible and, most importantly, to try and improve the cell edge performance.

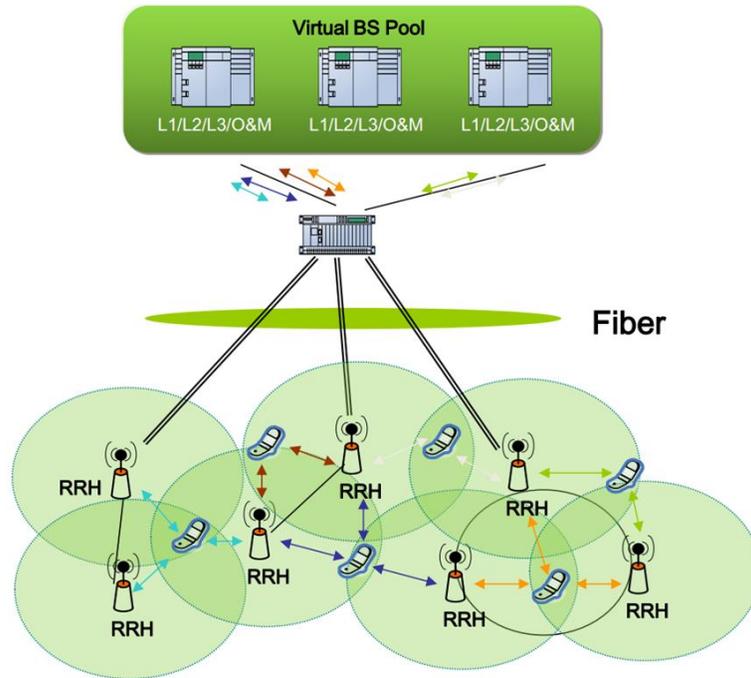
The C-RAN architecture is driven by geographies where the cellular operators are also the national operators, with access to significant dark fiber in the infrastructure combined with large, dense urban populations. Such regions include China, South Korea, Japan and other parts of the Far East. China Mobile, SK Telecom, KDDI, and NTT DOCOMO are all examples of operators who seem to be aggressively adopting this approach.



C-RAN, illustrated in

Source: China Mobile

Figure 2, allows cellular operators to collapse multiple macro BTS platforms into one centralized BTS “hotel.” This combines baseband capability (L1 air interface, air interface security, air interface scheduling, and control of the baseband and transport interfaces) with content delivery server platforms in one location. The air interface is then provisioned through distributed antennas, called remote radio heads (RRH), that may be located even tens of kilometers from the C-RAN BTS.



Source: China Mobile

Figure 2. C-RAN Architecture

This is the opposite model of HetNet approaches where macro, micro, metro and pico base stations utilize local antenna resources with local baseband processing, and then each individual BTS utilizes microwave or fiber links to backhaul traffic through the access and aggregation network to the core. The C-RAN deployment model centralizes all of this capability into one box.

The architectural change can be seen by comparing Figure 3, which illustrates the current network, with

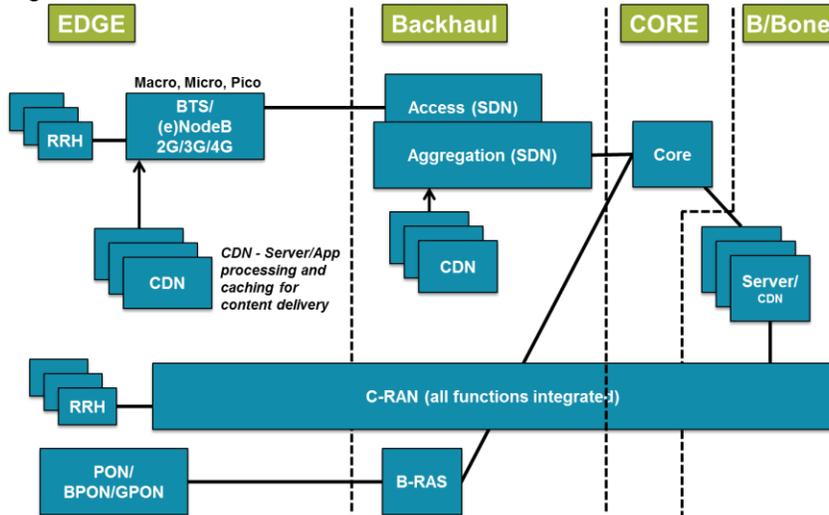


Figure 4, which shows how the network is likely to evolve. The addition of C-RAN is clearly evident, along with greater HetNet access complexity.

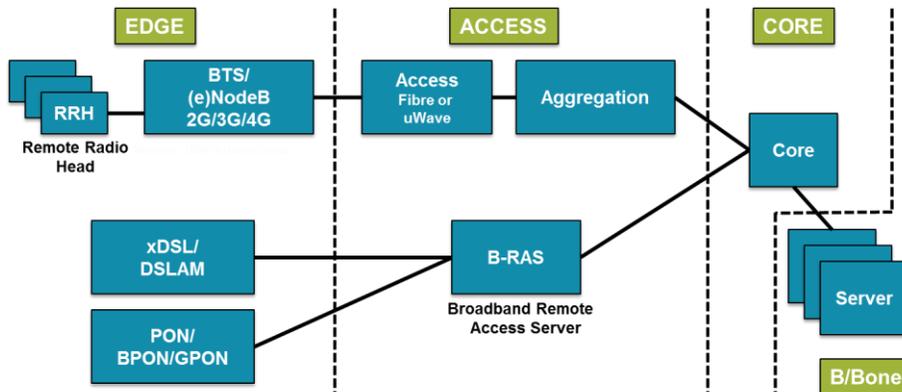


Figure 3. Infrastructure Segments

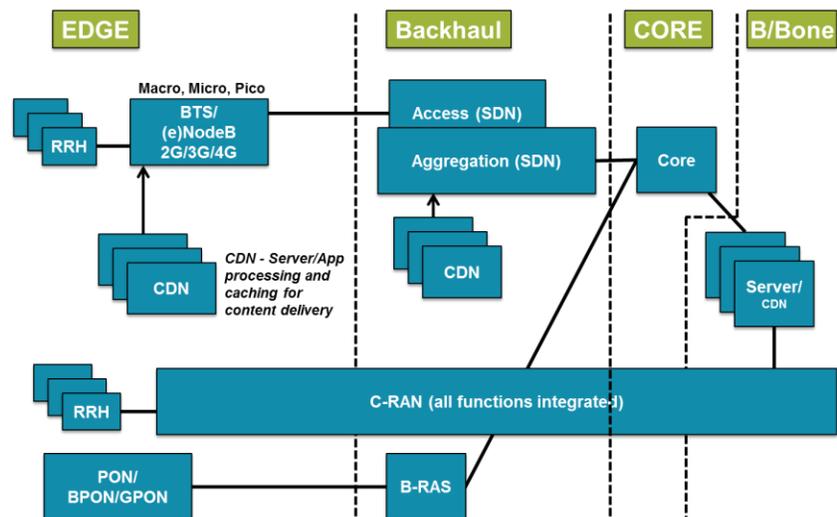


Figure 4. Infrastructure Migration: 2014 Onward

Network equipment hardware will have to be flexible enough to support this wide mix of compute functionality required in the network from the core to the edge. Where required, it should offer a balanced mix of data processing and control processing, along with the ability to accommodate additional functionality such as Layer 1 air interface processing, which has until today required discrete DSPs and FPGAs for execution.

The complexity of the wireless network has been compounded by various releases in the 3GPP specification. LTE-A was first introduced in 3GPP Release-10 in 2011 and offers several key enhancements to LTE capabilities, including increased multiple-In/multiple-out (MIMO) antenna capabilities, coordinated multi-point techniques (CoMP), and sophisticated wireless signal interference management. The 3GPP Release-11, which has just been ratified, adds improved carrier aggregation, MIMO, relay nodes, and new frequency bands to LTE-A.

The combination of these new techniques and the forecast traffic increases by the year 2017 will dramatically increase the complexity and volume of control and signalling traffic. This impacts the complexity of the equipment handling the traffic, and ultimately the complexity of the processors within the equipment.

Base Station Processing Requirements

The anticipated network changes are influenced by the need to provide higher data rates at the edge of the network. Content delivery closer to the subscriber will be required; this will necessitate caching and local storage of data local to the base station in some cases. It may also include provisioning for better subscriber quality of experience according to priority and demand.

For example, some subscribers may be allocated higher priority – perhaps bronze, silver or gold status. Or perhaps certain traffic types will be prioritized, such as video and voice, due to their real-time

requirements, as compared to web browsing. In some cases, it may be necessary to support storage of data in the cloud network, perhaps as part of the access and aggregation cloud. All of these changes will have a profound effect on the types of traffic to be supported in each of the different equipment types, and as a result, the component technology used in the equipment is likely to be fundamentally different from that used today.

With C-RAN for example, the traffic density in a single box is significantly higher than that passing through a macro BTS platform. Many of the control elements found in the core network today may thus be accommodated adjacent to the baseband circuitry, and the functions may be partitioned differently. Virtualization technologies will likely be employed to support allocation of resources on general-purpose processor blades while maintaining close integration with data plane components. Activities such as network function virtualization (NFV), sponsored by the European Telecommunications Standards Institute (ETSI) and supported by many of the main network operators, will be utilized in these types of systems.

Enterprise infrastructure applications blend different levels of three distinct function: control plane processing, packet or backhaul processing, and scheduling over the several clusters of cores available. Base station equipment will have a very different mix of these functions from core equipment. Designers have to determine the best mix of processor capability to handle the processing required, selecting from the available technology choices to deliver their designs in the time required to match market opportunity.

Control Plane:

Around 10% of traffic requires the control plane; 90% is processed in the data plane. Control-plane functions require the maximum amount of processing per packet, typically involving tens of thousands of instructions per packet, usually allocated in a 'run to completion' mode. Out-of-order and multi-stage pipelines can be utilized very effectively.

Performance relies on ready access to high-bandwidth memory. If access stalls, then the control packet is typically not flushed from the pipeline; the processor will stall until the data is available, because control processing must complete before the decision can be made as to what to do next in the string of control packets.

High-performance cores having virtualization capability can meet the needs of the control plane, the content delivery network (CDN), and other functions that require high single-thread performance. Applications running in the control plane include NFV, CDN for cloud and edge networks, and potentially new remote access technology requiring significantly more performance (such as LTE-A for example).

Data Plane:

The edge of the network sees data rates in the range of hundreds of Mbps or perhaps a Gbps; the access/cloud segment experiences data from one to the tens of Gbps; and the core processes from twenty to hundreds of Gbps. Unlike the control plane, the challenge here is in handling bursts of backhauled traffic, processing the headers, and placing the data into buffers without dropping any packets.

This involves a completely different blend of processing. Many data-plane designs use dedicated DSPs for this capability, and they interface these data-plane processors to the SoC system through the ARM® AMBA® interconnect. The DSP offers a dedicated and optimized instruction set for data plane processing and offloads the CPU from power-hungry and computation-intensive functions.

Instead of the tens of thousands of instructions per packet used for control processing, packet processing may use only hundreds of instructions per packet. Access to cache (instruction, data, L2, and L3) and external memory are also different for data packet processing.

A key distinction can be made between data and control processing. ARM uses the terms 'stateless' and 'stateful' to distinguish between them. Stateless processing uses a sea of small cores to handle streams of packets coming into the system-on-chip (SoC). Each core executes in a 'run to completion mode' to classify headers and dump packets into memory. Each packet is handled independently; the core knows nothing about any prior packet. The number of cores and the size of the interconnect scale simply according to the speed of the interface. Stateful processing, by contrast is used in a higher level of decision making, where the history of packets matters. This is where flows and sessions can be managed, typically in the control plane.

Scheduling:

The other challenge in wireless systems is orthogonal to the first two. For user access scheduling, where users have to be scheduled according to air interface bandwidth available, latency is key. In the case of LTE for example, where there may be hundreds of subscribers on the air interface to be scheduled into their own timeslot; all of this has to be calculated over potentially multiple cores in accordance with the LTE standard: within 1 ms. This involves a lot of priority calculation, scheduling of receive and transmit tasks, and signaling to and from the DSPs, processors and memory. So, the ability to use multiple cores and then to switch contexts between them is essential.

Technology Requirements:

As the data consumption on smart connected devices spirals up, the burden on system designers to deliver more performance using the same power and equipment footprint is creating new design paradigms and decision criteria. The availability of higher-performance, multicore-capable processors, coherent interconnects, and optimized performance-enhancing physical and logical IP, have catapulted performance/watt and performance scalability to the top of any selection criteria for SoC component technology. In addition, in an R&D budget-challenged world, an industry-standard instruction set architecture (ISA) with a well-supported software and tools ecosystem enables SoC design managers to deliver their products to market faster and to conserve R&D dollars for developing value-added and differentiated application-specific features.

New platforms for meeting the challenges outlined above will need a mix of heterogeneous CPU, DSP and function-specific accelerator cores. More and more capabilities will be integrated onto single SoCs, and these will typically process multiple traffic types, including data channel payload, control plane traffic, front-end processing, and user scheduling.

With this trend towards integration and higher performance SoCs, a mix of processing elements supports bursty high-speed traffic payloads and latency-sensitive traffic via smart signal processing partitioning between the CPU and DSP.

Figure 5 shows a multicore SoC architecture typical of what might be designed today. At the center are the RISC CPU cores that include Layer 1 and 2 caches. These cores are connected through a switch to a shared Layer 3 cache, external DDR3 memory and the I/O subsystem. Accelerators for security, compression/decompression, and packet acceleration are also connected through the switch.

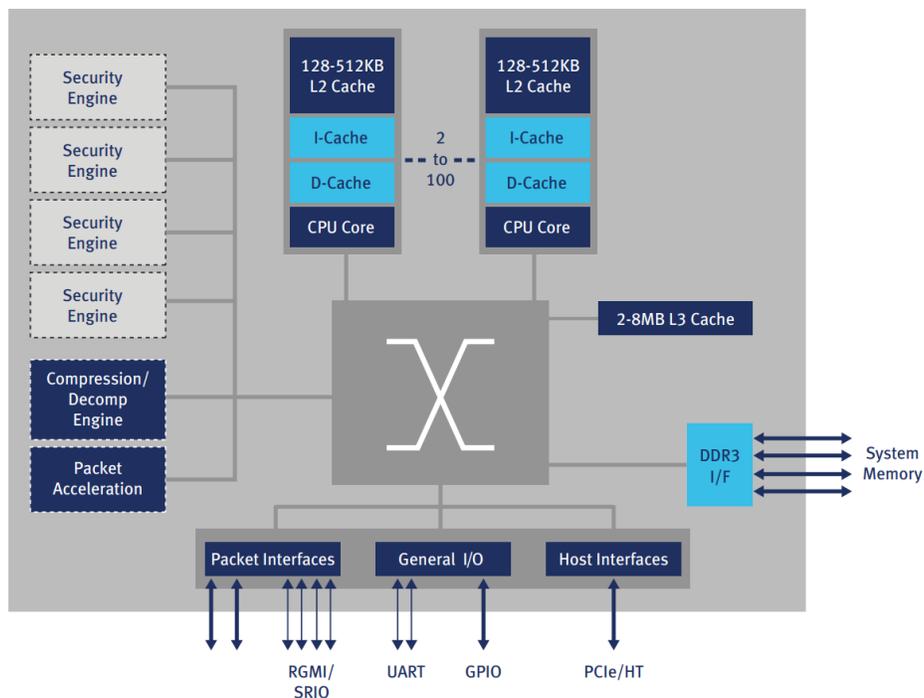


Figure 5. Typical Multicore Processor

In this architecture, a crossbar switch is used to provide a low-latency connection between the individual components. Typically, other functions (for example, DSPs for Layer 1 processing) are added as separate components through the same crossbar.

The challenges OEMs face in designing their network equipment are widespread. They include defining the right SoC and defining equipment architectures with the ability to meet performance, cost and power targets. OEMs need to scale their designs to:

- Support different BTS sizes;
- Provide efficient HW-SW partitioning;
- Provide acceleration for some of the compute-intensive functions including backhaul packet processing, security, backhaul authentication, and air interface;
- Support data sharing and the memory hierarchy, including L2 and L3 caches;
- Support dynamic resource allocation in real time, based on system loading;
- Support multi-mode for multiple radio access technology (RAT) standards;
- and ensure ease of programmability in a complex multicore environment.

Support for these features and many others is essential in today's networks.

Heterogeneous Multicore Architecture

One solution for meeting the above challenges is a heterogeneous multicore SoC with the right mixture of CPUs, DSPs and other hardware accelerators that meet the exact performance and power requirements of the target applications. Designing these kinds of multicore SoCs poses a challenge to the system architect due to the complexity of these SoCs. They require the right high-performance and low-power CPUs and DSPs, as well as the means to connect them together. It is also often important to maintain data coherency between all of the components in the system, which makes the problem even harder.

System and SoC developers address the complexity of these designs by making wide use of licensed IP, particularly for standard RISC cores and interconnect IP. This can provide best-of-breed components for the differing needs of the many components in the heterogeneous structure. However, each of those components can be further scaled by replacing an individual processor with a cluster of cores structured for symmetric processing.

The use of such clusters allows fully-flexible assignment of any task to any of the cores in the cluster while varying the number of utilized cores according to the system load in order to minimize power consumption. Clusters of CPUs and DSPs must be accommodated by the interconnect switch, which includes maintaining cache coherency between cores and providing a low-latency path between the CPU cores, DSP cores, caches, external memory and networking I/O.

The typical internal switch in existing multicore processors is designed to provide a connection between CPU cores, memory and I/O. It is not designed to support cache coherency across multiple cores or handle the wide range of bandwidth and QoS connections between the mix of general-purpose CPU cores, DSP cores, and packet processing engines, each with direct access to Layer 3 caches, external memory, and network I/O. Thus, the internal switch needs to be replaced by a coherent interconnect that will support these requirements.

Figure 6 illustrates an example system that can be built using ARM and CEVA IP. This shows a macro base station platform and utilizes an ARM big.LITTLE™ technology approach where the larger processor cores can be used for applications processing and control functions, while the smaller processor cores can be utilized for more real-time activities including L2 scheduling and packet processing.

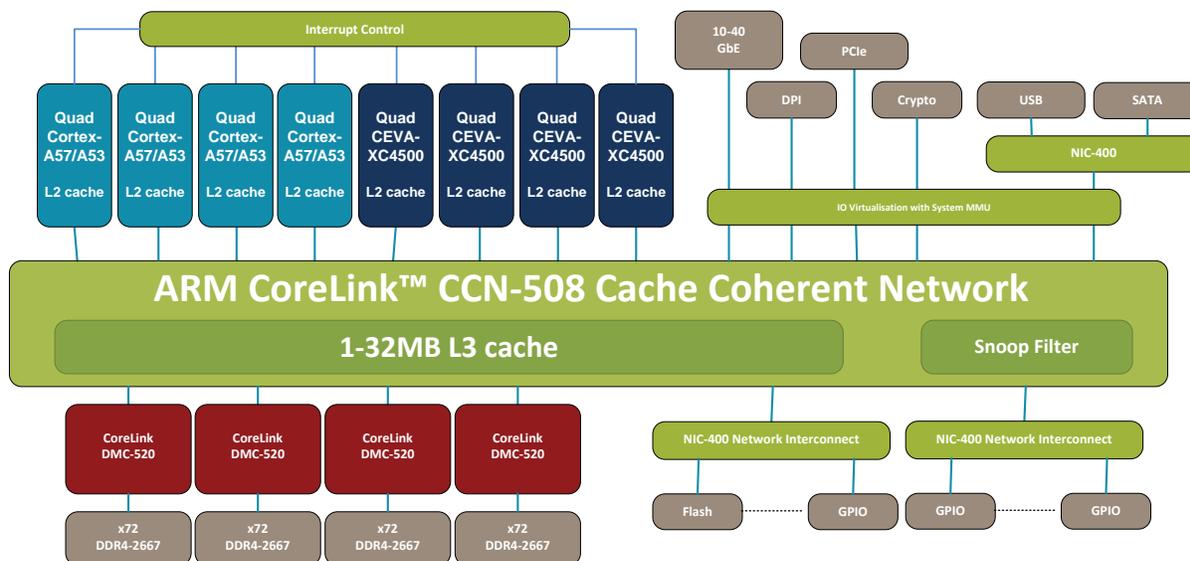


Figure 6. Typical Heterogeneous System Solution

More advanced multicore processors support cache coherency across multiple cores – not only CPUs, but also DSPs, GPUs, and accelerators. All can be closely coupled with the network I/O or separate processing subsystems.

The new interconnect in next-generation integrated multicore processors must:

- Provide a low-latency connection between cores and other elements in a transparent manner using a standard interface;
- Handle the real-time processing required for time-sensitive functions such as air interface scheduling;
- Enable using shared memories with multiple level cache to improve SoC performance and minimize memory area and power;
- Support full coherency across multiple cores and external memory interfaces; and
- Enable full utilization of system resources with minimal traffic management overhead.

The performance of the coherent interconnect should be deterministic and consistent across different implementations. This allows system architects to predict performance and scale solutions from as few as 2 to as many as 128 cores in a seamless manner. The coherent interconnect should also provide quality of service (QoS) features to manage latency while ensuring that the highest priority traffic has the lowest latency, without unduly slowing low-priority traffic.

IP for Next-Generation Wireless Infrastructure

Licensed IP can dramatically reduce the cost and risk in SoC integration and associated software developments. The majority of newer integrated multicore solutions use licensed cores; few SoCs will use a proprietary core. The following sections describe IP from ARM and CEVA that is particularly well-suited to the needs of these new wireless infrastructure enhancements. Their descriptions are followed by examples of equipment implementations.

ARM® Cortex®-A50 Processor Series for Wireless Infrastructure Applications

The ARM Cortex-A57 processor is ARM’s highest performing processor, designed to further extend the capabilities of future mobile and enterprise computing applications, including compute-intensive 64-bit applications such as high-end computers, tablets, and server products. Mobile applications can be addressed with 2 or 4 cores. Higher-performance systems like wireless infrastructure typically require 4, 8, 16 or 32 cores interconnected by the ARM® CoreLink™ CCN Cache Coherent Interconnect described below.

The ARM Cortex-A57 delivers significantly more performance than the ARM Cortex-A15, with as much as a 44% improvement in 32-bit code per-cycle performance, at a higher level of power efficiency. This is illustrated in Figure 7. The performance increase on web browsing and integer workloads are shown below. The performance for memory intensive workloads increases even more due to the enhanced micro-architecture of ARM Cortex-A57. Dedicated cryptography extensions improve the performance of cryptography algorithms by ten times over current-generation processors.

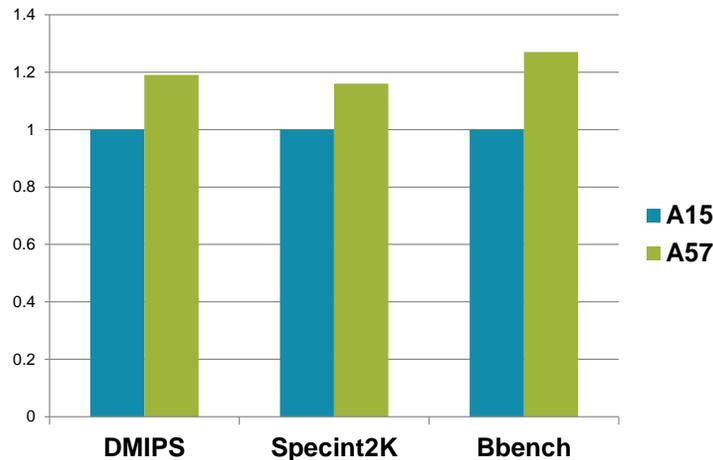


Figure 7. 64-bit ARM Cortex-A57 CPU performance as compared to the ARM Cortex-A15 CPU

The processor can be used exclusively for the CPU, as illustrated in Figure 8, or it can be paired with the ARM [Cortex-A53](#) processor into an ARM [big.LITTLE](#) implementation configuration that enables scalable performance and optimal energy-efficiency.

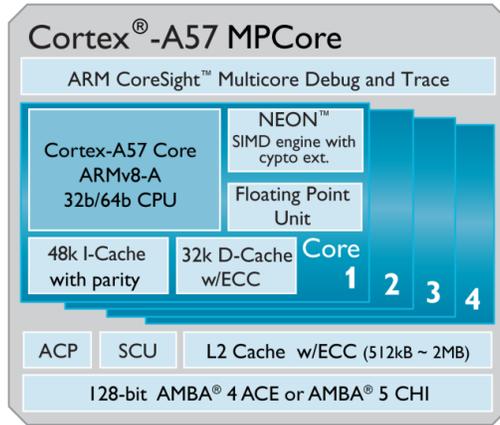


Figure 8. ARM Cortex A57 symmetric multicore implementation

The ARM Cortex-A53 processor is ARM's most efficient application processor, delivering today's mainstream infrastructure experience in a quarter of the power of other processors made using the same process node.

The ARM Cortex-A53 ARMv8 processor is capable of supporting 32-bit ARMv7 code and 64-bit code. It delivers more performance at higher power efficiency than the ARM Cortex-A9 processor, and it is capable of deployment as a standalone main applications processor that defines today's high-end mobile platforms.

The performance graph below shows ARM Cortex-A9 and ARM Cortex-A53 processors on a multicore 'rate-style' benchmark that tests integer and floating-point performance with a mix of large and medium data sets. Rate benchmarks like this evaluate the ability of a multi-processing system to handle memory traffic and coherence requirements. They stress the memory system of each CPU and the L2 cache. The rate portion of the benchmark duplicates the same code on the 2nd, 3rd, and 4th CPUs in each system, and measures the delivered aggregate performance.

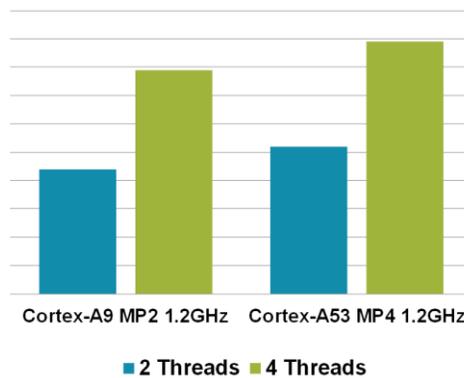


Figure 9. ARM Cortex-A53 processor performance as compared to the ARM Cortex-A9 processor

For infrastructure and networking equipment manufacturers, ARM's higher performance and ARM big.LITTLE processor configurations of cores and interconnect are ideal for innovative designs in applications such as cellular base stations, cloud equipment and core networking equipment. The ARM cores are augmented with dedicated engines such as the CEVA-XC advanced vector DSPs to offload ARM cores from DSP tasks and Layer 1 data processing.

ARM provides silicon partners with the building blocks to build solutions that can be scaled across multiple densities of platforms – from small-cell base stations to small-office/home-office (SOHO) switch routers, to micro BTS platforms and small to medium size offices, on up to macro BTS platforms and enterprise switches and routers. This satisfies the highest performance requirements of server, evolved packet core (EPC) nodes, and core network equipment.

This allows OEM customers to utilize common software platforms from ARM's broad ecosystem to run on ARMv7 or ARMv8-based processor products. ARM is working closely with many of our networking partners to broaden the software focus to include networking-specific requirements such as real time pre-emptive support for multicore processor designs, virtualization support, and specific APIs that can standardize processing requirements like inter-cluster communications and packet processing. These software ecosystems and processor cores are developed in such a way as to support multiple-processor and multiple-cluster designs based on the ARMv7 (32-bit) or ARMv8 (32/64 bit) cores.

ARM continues to pursue all of this activity with a strong focus on the processing/power efficiency of the building blocks. ARM's roadmap includes processor cores that meet a number of these different processing needs including A, R and M series devices.

CEVA-XC4500 Multi-Core DSP Processor for Wireless Infrastructure Applications:

The CEVA-XC family of digital signal processor (DSP) cores features a combination of very long instruction word (VLIW) and single instruction, multiple data (SIMD) engines that enhance typical DSP capabilities with advanced vector processing capabilities. The scalable CEVA-XC architecture, with four CEVA-XC processor generations to date, offers a selection of advanced communication processors to support software-defined modem design with minimal hardware acceleration.

The fourth generation of the widely-licensed CEVA-XC architecture, the CEVA-XC4500 is specifically optimized for next-generation wireless infrastructure applications. The CEVA-XC4500 delivers highly powerful fixed-point and floating-point vector capabilities; supplying the performance and flexibility demanded by these applications.

The CEVA-XC4500 processor incorporates an advanced memory subsystem including:

- L1 caches with full cache coherency based on ARM® AMBA® 4 ACE interconnect,
- Master and slave AXI4 ports, and
- Low-latency fast interconnects and advanced DMA-based data traffic mechanisms utilizing queues and buffers.

This advanced DSP subsystem enables advanced multicore system architectures incorporating CEVA's DSPs alongside ARM CPUs based on ARM's latest multicore interconnect IP.

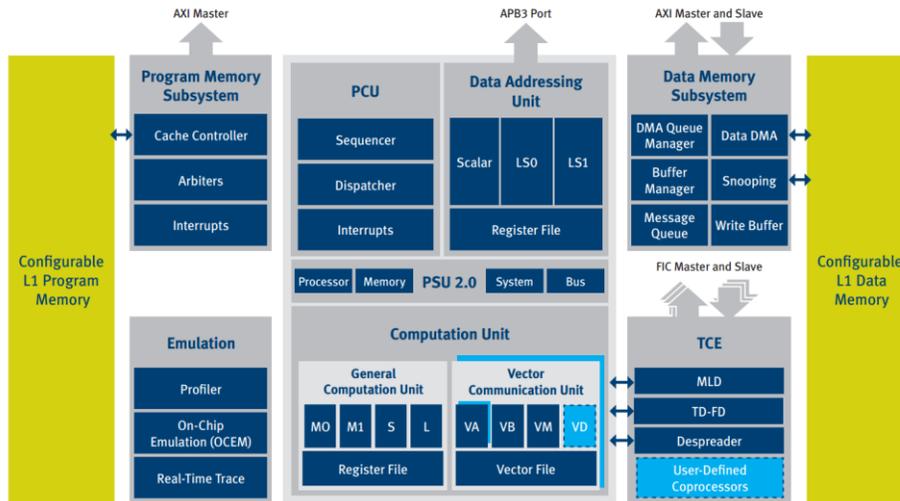


Figure 10. CEVA-XC4500 Block Diagram

With the ability to create a coherent cluster of CEVA-XC4500 cores and the addition of system-specific hardware accelerators, a powerful cluster of two or four DSPs can form a high-processing power unit that can support the L1 processing of an entire BTS remote radio head. Figure 11 shows an optional example of a DSP cluster that includes four CEVA-XC4500 cores. In addition to the DSPs, this cluster also includes several CEVA tightly-coupled extensions (TCE) that offload the DSP from common cycle-consuming tasks. Among the TCEs are the maximum likelihood detector (MLD), descrambler, FFT and DFT accelerators, Viterbi decoder, and hybrid automatic repeat request (HARQ) accelerator.

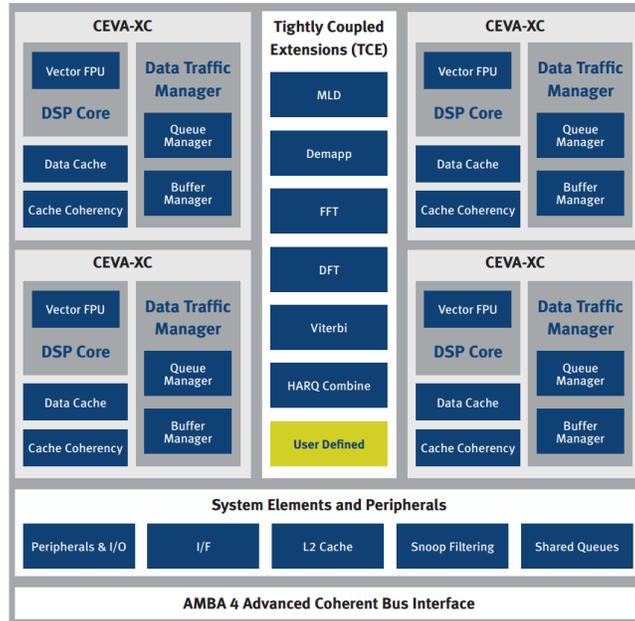


Figure 11. CEVA-XC4500 Cluster Example

Target applications for the CEVA-XC4500 include:

- Wireless cell baseband processing, scalable from pico and metro cells up to macro cells and C-RAN;
- Remote radio heads, targeting digital frontend processing and handling advanced DSP functions like digital pre-distortion (DPD), up/down sampling filters, up/ down conversion, and quadrature modulation correction;
- Wireless backhaul, including wideband spectrum point-to-point communication supporting up to 4096 QAM; and
- Enterprise WiFi and WiFi-cellular offload, addressing WiFi 802.11ac Access Points and Small Cells requiring 3x3 or 4x4 MIMO.

The relationship between the first three of these applications is illustrated in Figure 12.

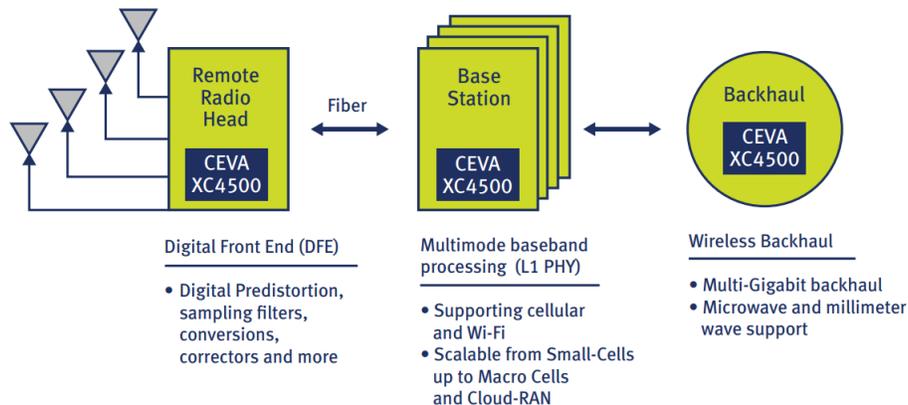


Figure 12. CEVA-XC4500 target applications

ARM CoreLink CCN Cache-Coherent Interconnect

The ARM CoreLink CCN-504 cache coherent interconnect supports ARM Cortex-A15, ARM Cortex-A53, ARM Cortex-A57, and next-generation ARMv8 processor cores. The support for these industry-standard cores and ARM AMBA interfaces makes this solution very flexible. Using these IP blocks, developers can design very high-performance multicore SoCs for a wide range of applications, including wireless infrastructure and servers.

ARM’s CCN technology not only supports interconnection of ARM cores, but also DSPs like the CEVA-XC4500, as well as hardware accelerators. This allows cache coherency to be maintained across the entire heterogeneous structure, simplifying the architecture and improving performance.

The ARM CoreLink CCN-504/508 and future versions of cache coherent interconnect also provide integrated QoS regulation and priority management with sophisticated QoS-based protocol-layer flow control. The QoS support provides prioritized arbitration from ingress to egress, and eliminates head-of-line (HOL) blocking where a stalled task can hold up all other queued tasks. Optional QoS regulation at all ingress points allows users to set QoS requirements per I/O and per accelerator. Prioritized arbitration is provided for CPU cores, coherent accelerators, I/O, Layer 3 cache and external memory access.

Future versions of CCN will extend these capabilities by increasing the number of cores that can be supported, the amount of integrated L3 cache, and the number of accelerators and DMC ports that can be managed.

Not only does this offer sufficient core performance, but the networking control and data plane applications are dependent on the ability to access memory very quickly with minimal delay. This allows implementation of the architecture shown in Figure 6, which has an L1 cache per core, shared L2 cache per cluster, and shared L3 cache per SoC – with coherency maintained at each level. These coherency support features are fundamental to providing excellent system performance. Figure 13 illustrates CCN’s role in managing mixed network traffic.

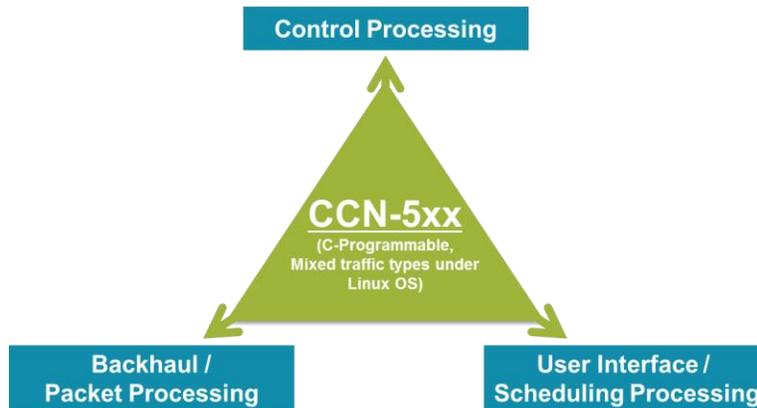


Figure 13. Mixed traffic use cases

Using ARM and CEVA IP to Implement Wireless Infrastructure

Given the CPU cores from ARM and the DSP cores from CEVA, and interconnected by ARM's CCN cache coherent interconnect, the diverse components of wireless infrastructure can be readily implemented with optimal power and performance. The following sections provide conceptual examples.

Conceptual BTS Architecture

An LTE-A macro-cell base station can be built with L1 PHY processing. This is done using clusters of CEVA-XC4500 processors, while the L2, L3 and upper layers are processed by clusters of ARM Cortex-A50 series processors. Each DSP or ARM cluster can be scheduled to process the traffic of the uplink and/or downlink data of the macro-cell sectors or component carriers when using carrier aggregation.

Each DSP cluster contains homogenous CEVA-XC4500 processors and each CPU cluster contains homogenous ARM Cortex-A15 cores. Each DSP cluster may also include hardware accelerators to offload the processor from low-flexibility processing tasks like FFT, error correction and cryptography, resulting in balanced hardware/software partitioning. The CPU and DSP processor clusters are fully synchronized through a cache coherent interconnect or network.

Small-Cell

Figure 14 shows a small-cell BTS using this combination of CEVA-XC4500 and ARM Cortex-A50 series cores. This system makes use of external hardware accelerators to offload the processor from well-known compute-intensive tasks, balancing the partitioning between software and hardware.

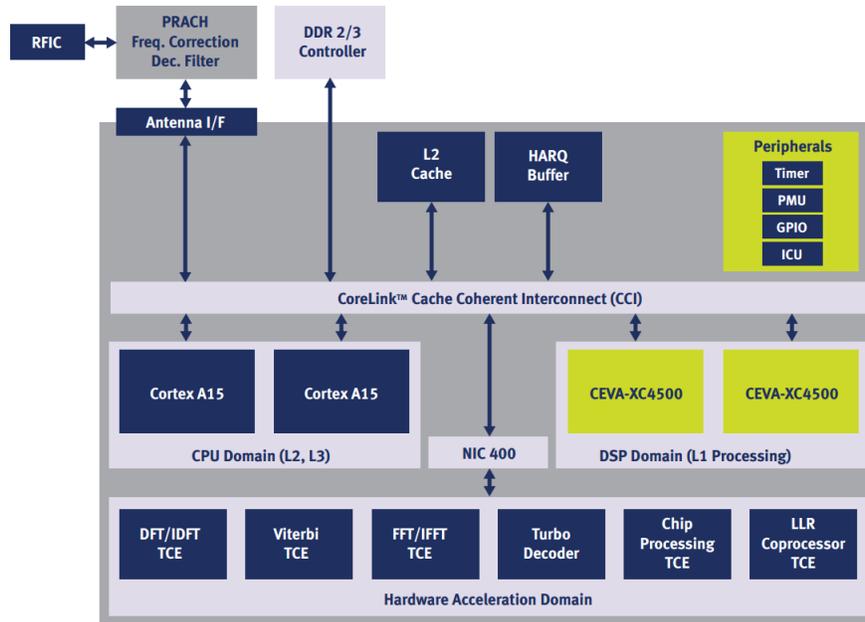


Figure 14. Small-Cell System Architecture Example

Cloud Radio Access Network

Figure 15 illustrates a C-RAN system implementation. C-RAN installations are built from centralized, software-based processing farms. These processing farms are built using heterogeneous computing systems, with ARM cores managing all of the transport and control functionality and CEVA-XC4500 cores handling all the baseband functions. The ARM and CEVA-XC cores are fully synchronized through a cache coherent network based on ARM CCN. The number of ARM and CEVA cores is scalable according to the performance requirement needed. An advanced memory hierarchy is used, where L1 caches are available to each individual core, L2 caches serve individual clusters, and an L3 cache is shared across the entire SoC. Coherency is maintained at each level using advanced ARM AMBA and CCN technologies.

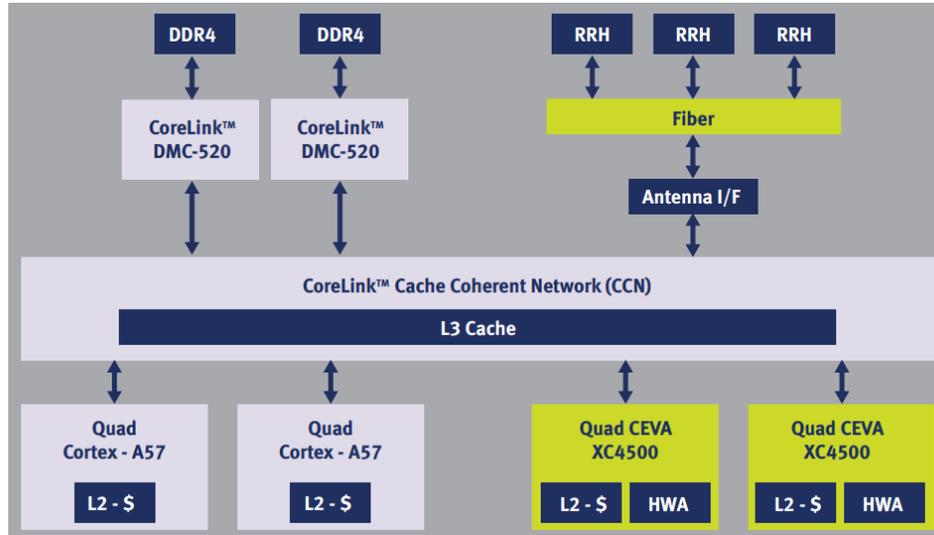


Figure 15. C-RAN System Architecture Example

Conclusion

The dramatic increase in the amount of traffic that the global wireless network must process is pushing the existing infrastructure beyond what it can manage. High-level architectural changes will include both HetNet and C-RAN approaches, with each being deployed in environments thought most suitable by service providers.

These added capabilities will demand extraordinary processing for the equipment to keep up with the traffic. The complexity arises not only from a simple compounding of the number of data packets, but also due to increasingly sophisticated control requirements driven by evolving standards. Advanced multicore architectures will be required to handle the load, and these will be heterogeneous at the highest level, with a mix of CPUs, DSPs, and other units dedicated to specific tasks. Each of those processing nodes, however, can be a symmetric, homogeneous cluster of cores, providing a level of scalability critical to adapting a given platform to a wide variety of specific installation requirements with minimal redesign.

Cores from ARM and CEVA play critical roles in such systems, and their power/performance characteristics will make them a preferred choice for new systems being built to handle tomorrow’s traffic. By mixing suitable numbers of ARM and CEVA cores, complemented by select hardware accelerators and interconnected by an ARM CCN interconnect, OEMs can design equipment that meets the needs of service providers, and service providers can build the networks that meet the needs of their subscribers.