

FPGA vs ASIC: 5G changes the equation

Dan McNamara, [Mobile Experts](#)

For many years, there has been a tug-of-war between suppliers of FPGA and ASIC solutions. New features are introduced on FPGAs, and as they become well understood they were typically hardened onto ASICs for lower cost, lower power and high volume. This ongoing back-and forth generally works in the favor of the FPGA vendors in the early days of a new generation, and then turns in favor of ASIC vendors for longer term steady production.

In 2019, new products from Xilinx started to change these dynamics. Xilinx introduced an update to the Zynq family with an “RF System on Chip” (RFSoc) capability, which integrates ARM cores for some processor functions (reducing cost and power), enables programmable logic for beamforming and integrates the ADC and DAC functions into the RFSoc. The power savings and size shrinkage are very compelling, especially for Massive MIMO (mMIMO) radios.

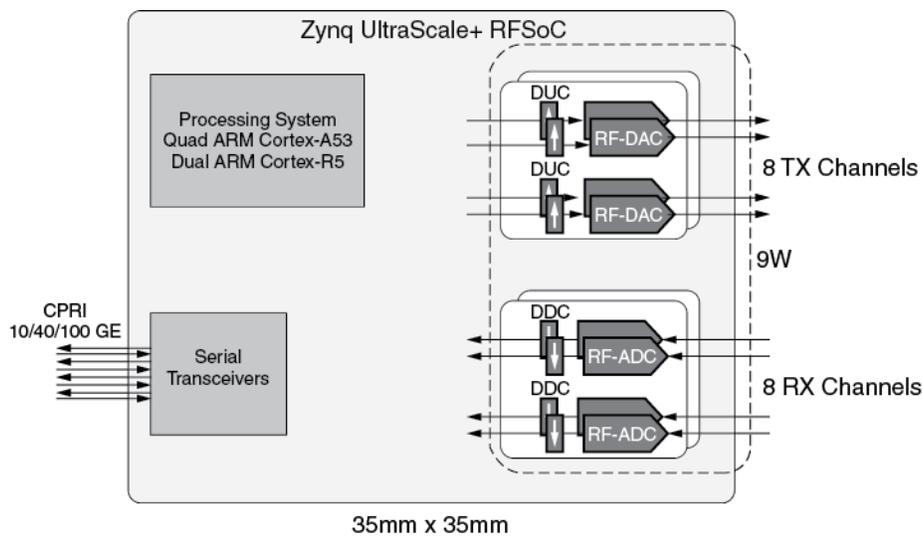


Figure 1: Integration of Processors and Data Converters

Source: Xilinx

If history was any indication, we might have expected ASICs to start showing up in 2020 that hardened these designs. That did happen in some cases. However, some larger forces combined to have a different outcome.

Architectural Trend: Integration of DFE and ADC/DAC

The integration of the data conversion functions offered an insight into huge thermal savings for the Remote Radio Unit (RRU) in 5G base stations. The savings comes out to roughly 1 W for each JESD204 lane or stated differently for each data stream coming from the antenna array. In the case of 64T massive MIMO, this can be more than 60W of power savings, out of a total heat dissipation budget of 250W. While an ASIC design can offer lower power consumption for the functions that are integrated, the overall solution savings are less clear.

This is where the math starts to get fuzzy. The metric that matters is total heat dissipation, not necessarily the benefit derived from one optimized section. With the constant pressure to manage heat in the RRU, the mass conversion to a traditional ASIC has been delayed. Even the holistic solutions offered by Tier 1 network equipment suppliers still have different partitions for the Digital Front End (DFE) and data conversion. There is not a single implementation approach yet.

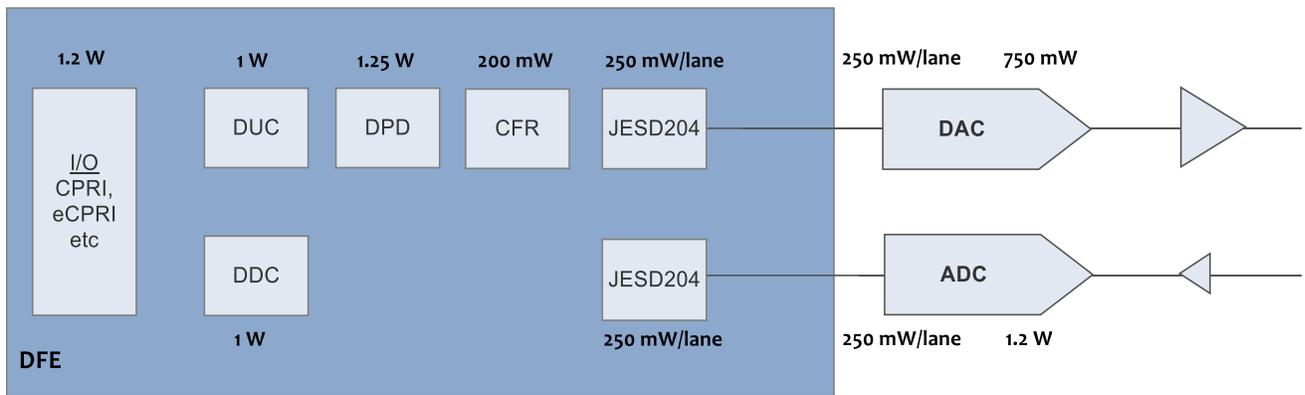


Figure 2: Power Consumption by Function in a Discrete 5G transceiver

Source: Mobile Experts

Note: Power consumption shown for a 100 MHz BW for a single transceiver chain.

JESD204B power is broken out separately from ADC/DAC power.

DUC: digital up-conversion; **DDC:** digital down-conversion; **DPD:** digital pre-distortion; **CFR:** crest factor reduction

In the Xilinx paradigm, a key problem to solve is the power consumption of the JESD204 interface, which is accomplished by the integration of the data converters with the digital front end, reducing the JESD204B processing dramatically. Xilinx integrates the digital front end (DUC, DDC, DPD) with the ADC and DAC to eliminate the overhead in power consumption.

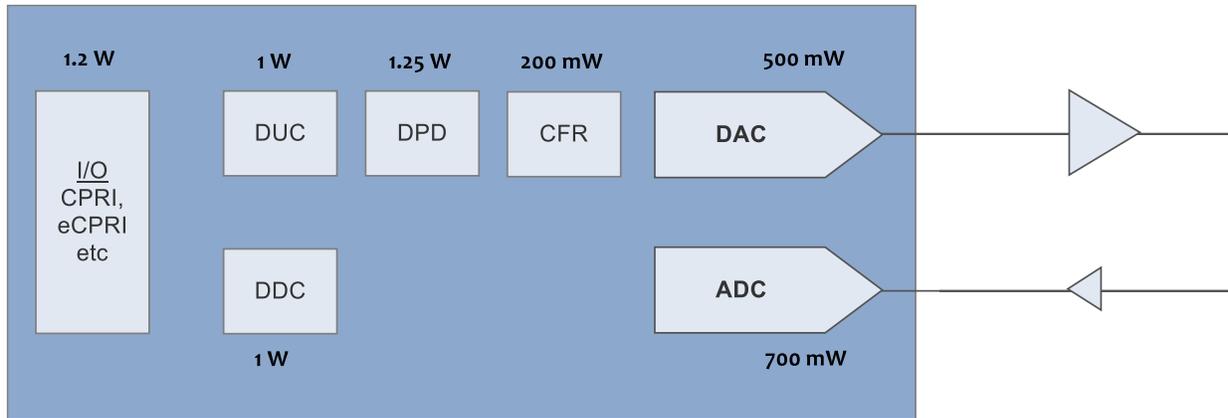


Figure 3: Power Consumption Impact of DFE/Data Converter Integration

Source: Mobile Experts

Note: Power consumption shown for a 100 MHz BW for a single transceiver chain.

Another issue with the integration of processor and data converters is related to the cadence of process node development. Processor logic on CMOS moves steadily down to the next process node (new ASICs will be moving to 7nm, 5nm or less), while data converters are generally developed on a process node that is behind the leading edge by two or three generations. (about 18 nm).

Considerations when Integrating

As stated earlier, the introduction of ASICs has historically followed once the performance requirements are locked down and the market ‘simply’ needs to scale up to large volumes. In the case of 5G there are still several variables that are not at a high level of maturity. The initial release of 5G under REL 15 (as Non-Stand Alone) came about 1 year sooner than was originally expected. The market adapted quickly, and we have concrete evidence that 5G is working. However, that does not mean solutions released for commercial operation in late 2019 had time to be optimized. There are several aspects to the 5G network that are still evolving:

- **Beamforming:** The implementation of beamforming is still undergoing improvements. Since the algorithms supporting this function exist in the RRU, it follows that a more flexible (i.e. FPGA) approach continues to have clear benefits.
- **Channel Bandwidths:** Bandwidths are expanding. Initial 5G systems commonly used 100MHz which corresponded to what individual Operators were able to access. However, the trend is to provide hardware supporting 200MHz or more to reduce product versions and streamline manufacturing (a benefit to both the Network OEM and the component supplier). Typically, 3.4-3.6GHz and 3.6-3.8GHz products can

meet most customer requirements. However, the upcoming C-Band auction in the US will allocate up to 280MHz for 5G use, further pushing the performance requirements in the RRU.

- **PA technology:** The technology selection of GaN for most of the mMIMO PAs also benefits from a flexible processor design. Channel bandwidths are quickly moving to 200MHz and beyond, up from the original configurations of 100MHz. In addition, power levels and Digital Predistortion need more optimization. The cycle time for PAs can be quicker than big silicon, so locking in features on the digital processors too soon will limit the capability to capitalize on where the performance of GaN will settle out.
- **O-RAN:** An emerging standard finding more traction is OpenRAN. While this architecture is getting a lot of attention, the specifics of implementation are still dynamic. While the definition of the Option 7.2x split is being finalized and there are yet other variations being considered by Operators that shift key functions in or out of the RRU such as Precoding, Layer Mapping, and Channel Estimation to name a few. The location of data computation makes all the difference in how the processors are defined and it will take a few more years to have all this settled out.

These issues will give a longer life to a more flexible design approach. With the long-term benefits of 5G still a little murky, flexibility will be the key to pivot as necessary for what is not yet known from a system requirements point of view.

FPGAs, ASICs, and RFSocS

In the previous section on technology trends we have already covered the major thrust toward RFSoc integration: It is all about reducing power consumption, cost, and size on the board.

The top three OEMs now use ASICs routinely in production of FDD and TDD LTE base stations. All have the volume to justify the ASIC, and the LTE application is stable enough to converge all customers onto a common ASIC platform. With the added complexities of 5G, the economics of ASICs have increased the volume threshold to more than 200,000 units per year to justify the expense of developing an ASIC for all of the top network OEMs.

For the 5G market, market scale is very lop-sided at the moment. Consider that for 2020, Mobile Experts predicts about 115M 5G transceivers to ship (each transceiver is a unique transmit path). Assuming there approximately 8 transmit chains per RFSoc (or ASIC), that puts the number of ASICs/FPGAs at 14M of which ~80% are consumed in China. For Huawei and ZTE, there are clear reasons to move to an ASIC solution based on supply options. However, for the suppliers focused mainly on regions outside China the answer might not be so clear. In 2020, outside China there are about 2.5-3M slots for ASICs/FPGAs and that are

split among many countries and operators. Most 5G deployments are on the scale of 100's to 10k's of base stations, so any single operator's deployment would not justify the use of an ASIC. Only the aggregated volume of a major OEM justifies investment in ASICs. With the market still learning what the performance of a 5G network looks like in the real world and with a transition to 5G SA (Stand Alone) starting in late 2020, the need for flexibility in system design is still real.

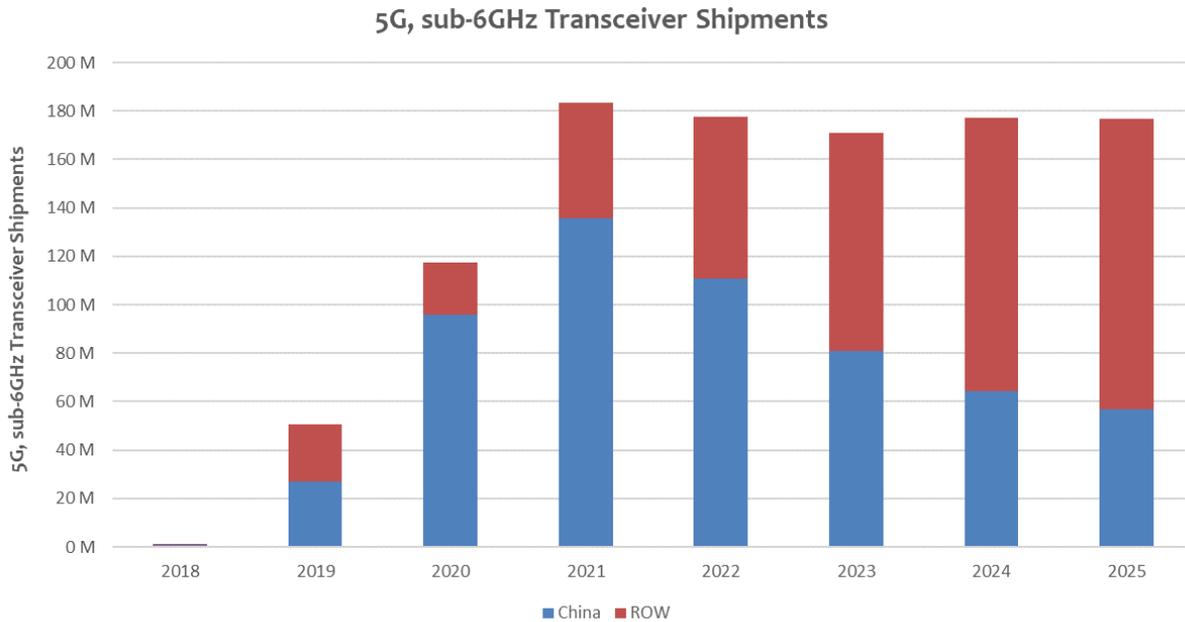


Figure 4: 5G, sub 6GHz transceiver shipments

Source: Mobile Experts

Additionally, there will be more flexibility needed for the RAN based on the new use cases that are being introduced with 5G. The initial roll-out of the 5G network from 2019 has been predominantly focused on the mobile users, the incorporation into smartphones. This is referred to as the eMBB (enhanced mobile broad-band) application. When talking about the overall 5G network, one of the big differences from 4G is the addition of two new applications: IoT and URLLC (ultra-reliable low latency connections). Longer term, these new applications are expected to drive a substantial amount of revenue in the B2B realm... but this has not happened yet. It will rely on the 5G infrastructure that is being deployed now and perhaps expand upon it. RRUs must accommodate the requirements needed to be deployed in buildings or as leased private networks that get connected to the main network via the C-RAN architecture. The data requirements for these many different use cases are not fully known yet but there is a good chance that it might be somewhat different than the approach (and design) for current eMBB systems. Flexibility with the management of data will allow smooth support and introduction of these new applications.

It is not hard to see the added complexity that comes with 5G networks. Every telecom generation has improved the user experience with better voice quality or more data and 5G is already doing this. However, the 5G network we see being deployed today is really the basic infrastructure that will be added to and expanded upon. This will look like Macro base stations, indoor RRUs, repeaters, and Small Cells. It is all 5G, but each can have a different requirement. The processors needed to support this diversity must be flexible based on each need as well as flexible to adapt to the still evolving use cases that we will not truly comprehend until the systems are deployed.

There is no doubt that 5G is bringing unprecedented amounts of data to be consumed, with greater complexity, to be accommodated by both the general consumer and business. The adoption of Massive MIMO (mMIMO) antennas dramatically improve performance and creates a huge amount of data. The processing of data is a key architectural consideration for network equipment suppliers. The typical transition from a more programmable solution to a hardened solution is not necessarily the optimal choice. The performance of these new systems is not yet fully understood in large, heavily loaded networks as 5G is still too new. Accounting for the new normal in network traffic will require continuous adjustments of many variables to arrive at a stable configuration. The ability to quickly adapt performance of key hardware components will be critical over the next several years.