Energy efficiency in next-generation mobile networks

Harish Viswanathan, Stefan Wesemann, Jinfeng Du and Harri Holma

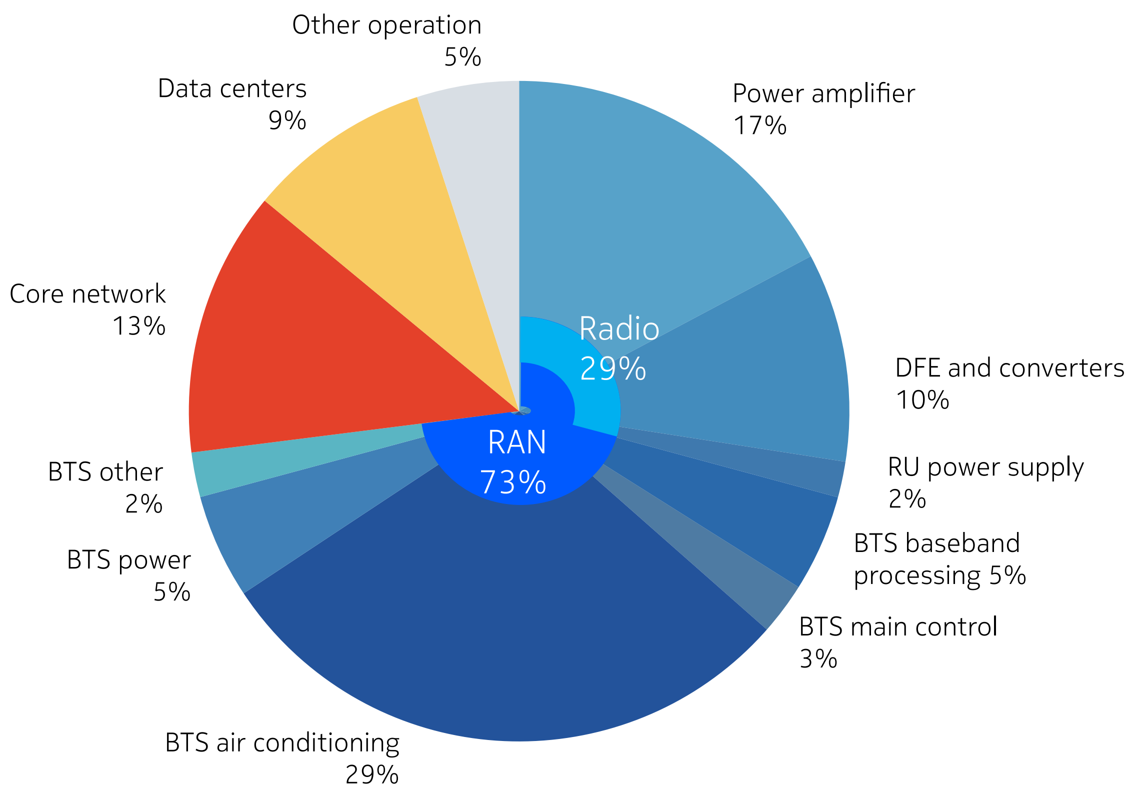
With the substantial negative implications of climate change looming ever closer, enabling sustainability is of paramount importance for any future technology. Minimizing energy use is one dimension of achieving sustainability, thus every aspect of 6G design, implementation and operation will be scrutinized to maximize energy efficiency. An analysis of the massive MIMO 5G radio energy consumption at different loads reveals under what specific conditions 6G should outperform 5G, setting qualitative energy efficiency design goals for 6G. Following this, we propose some design principles or essential features for the 6G air-interface and network architecture to minimize energy consumption.

Contents

Introduction

Worldwide, numerous companies are racing to achieve carbon neutrality and have established carbon footprint reduction goals. Communications service providers (CSPs) are no exception. In order to achieve the goal, the major focus will be on those aspects of operations that consume the largest amount of energy. Figure 1 shows the energy consumption for a typical CSP, based on a survey in 2021 of over 30+ networks with a mixture of multiple generation and multiple radio access technologies across different domains of operations [1]. It is clear that the majority of the energy consumption for a CSP is in network operations, and among the different parts of the network the RAN infrastructure consumes the most amount of energy.

Figure 1. Distribution of energy consumption for CSP operations (based on data from [1])



The importance of being energy efficient to minimize the carbon footprint cannot be underestimated. 6G will obviously have to achieve superior performance over the current generation on traditional metrics such as peak data rate, capacity, spectral efficiency, latency, and reliability that will be more demanding in terms of the energy the infrastructure requires. Nevertheless, energy efficiency will become a new metric that is as important as these other performance metrics. It is well known that it is not possible to simultaneously achieve peak performance across multiple metrics, such as latency and spectral efficiency or peak rate, and the system is typically designed to flexibly tradeoff between the different metrics depending on the use case. Likewise, we expect energy to start playing an important role, and the future network should be designed to optimize energy use against other performance metrics.

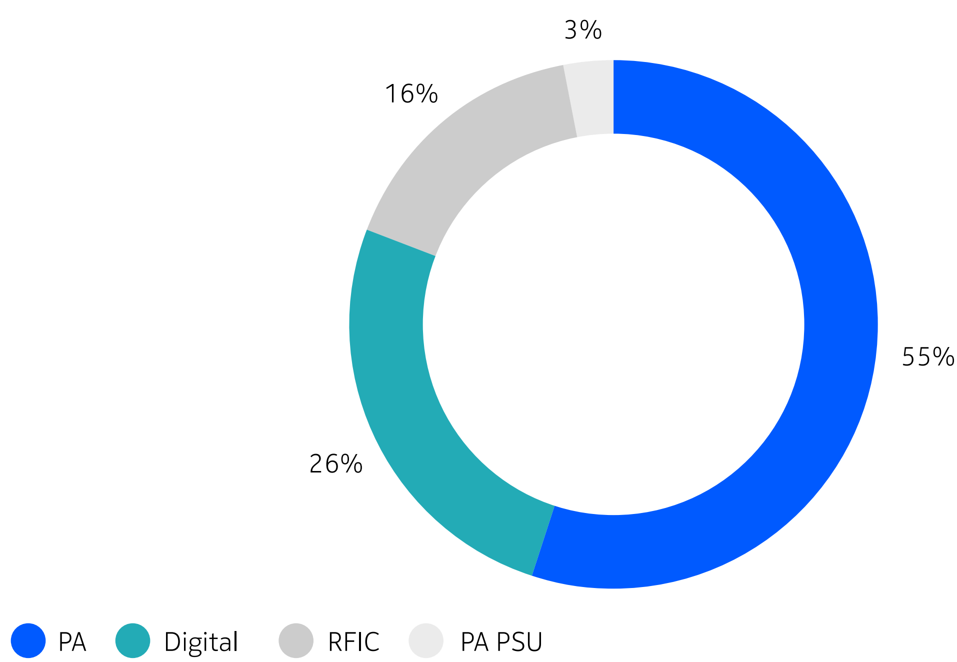
While much work has already been done to minimize the power consumption of devices, more attention and focus is required to minimize the power consumption of networks. The focus of this paper is to propose key features for the air-interface and network architecture of 6G that will enable increased energy efficiency and allow a smooth tradeoff between network power consumption and capacity. To this end, we begin with a brief look at the power consumption in massive MIMO 5G to understand the potential areas where 6G could be designed to be better. In the following sections, we will describe our air-interface design elements, architecture design principles and an emerging disruptive computing framework that should be exploited to ensure dramatic improvements in energy consumption at the component level.

Energy efficiency goals for 6G

Energy consumption of the network for any given deployment critically depends on the load in the system. The load varies substantially over the duration of the 24-hour period, with peak loading only occurring for a few hours during the day. Since network dimensioning is made to handle the busy hours, it is estimated [2] that, on average, up to 70% of radio resources are in idle, and 80% of sites only carry about 20% of traffic, permitting plenty of opportunities to improve energy efficiency.

Focusing on the massive MIMO radio unit that consumes the largest fraction of the energy of the network (29% in Figure 1), Figure 2 further breaks down the proportion of the power consumed by the different components at 100% load. The majority of the power consumption is spread across the power amplifier lineup (PA), the digital front end (RFIC), which includes the converters, the interpolation, digital pre-distortion and crest-factor reduction, and the digital processing integrated circuit (Digital) that performs L1-low signal processing functions including Fast Fourier transform (FFT) processing, beamforming, channel estimation and the first stage of MIMO detection. The actual proportions depend on the specific component technologies used such as ASIC or FPGA, the type of PA, and the choice of functions implemented in different platforms. Nonetheless, the figure highlights where most of the power is consumed, hence the processes to focus on for improving energy efficiency. The design of 6G should first and foremost consider solutions for maximizing the efficiency of the PA, whose typical efficiency is around 40% in today's 5G products, and then consider reductions in signal processing complexity to reduce the digital power.

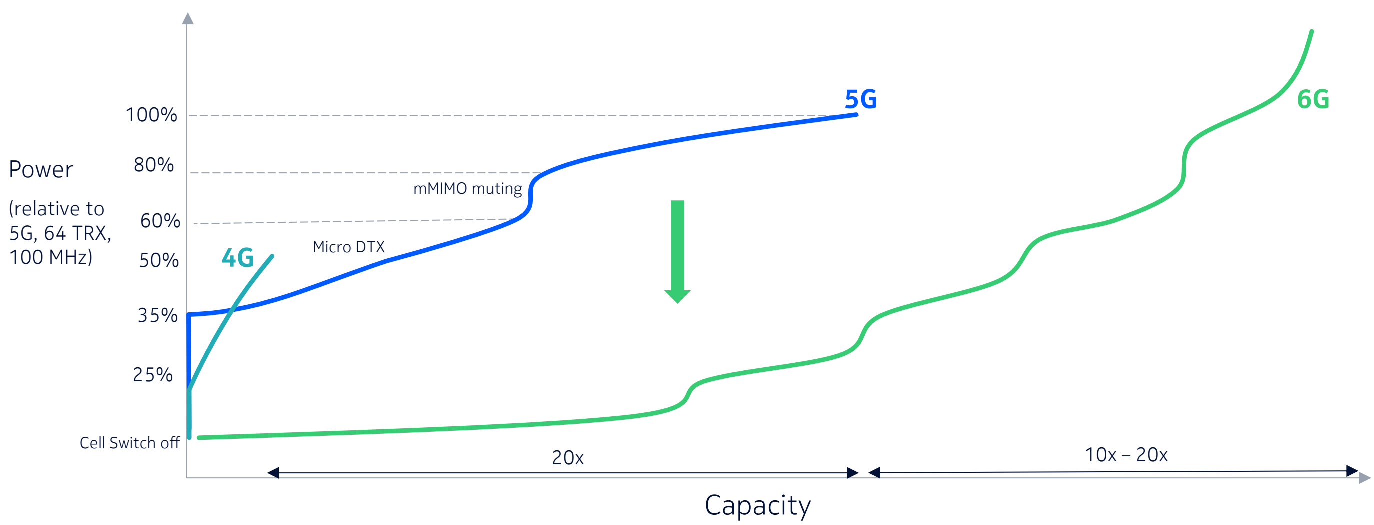
Figure 2. Massive MIMO (mMIMO) radio unit power consumption split at 100% load



The power consumed by the radio unit (RU) depends on the load of the system. At lower loads fewer physical resource blocks are used leading to lower output power for the same power spectral density. The lower output power translates to lower power consumed. However, the relationship is not linear since the amplifier efficiency is highest at its peak output power but reduces as the output power is reduced. For a typical state-of-art amplifier, the consumed power drops only as the square root of the output power. Furthermore, the RFIC and Digital powers do not diminish automatically with reduced load unless the digital circuits are pushed into sleep mode. This calls for novel operational and architectural innovations that can further improve energy efficiency.

Figure 3 shows, for illustration, the relative power consumption of a single carrier for a 5G mMIMO RU and that of a single carrier 4G LTE RU as a function of the cell capacity. Also illustrated in the figure is what we believe should be the target for 6G. The peak capacity of a 5G NR 64 TRX mMIMO 100 MHz cell is about 20 times that of the LTE 20 MHz FDD cell and is 90% more energy efficient in terms of watts per bits/sec [3]. In other words, 5G consumes 10 times less power per bits/sec, which translates to the LTE radio consuming 50% of 5G power at its peak capacity. We expect the same trend to continue in 6G, so that the peak capacity of a 6G 512 TRX extreme massive MIMO 400 MHz cell will be 10 to 20 times that of 5G NR [4], but peak capacity energy consumption will only be twice that of 5G power.

Figure 3. 6G power consumption goal versus 4G and 5G



For both 4G and 5G, power consumption decreases with reducing load because of lower output power by the PA. In addition, mMIMO muting — in which a portion of the array can be turned off at lower loads — can be implemented to reduce energy consumption leading to a step reduction at about half the load, illustrated in Figure 3. Another feature that is commonly implemented to reduce energy consumption is micro-discontinuous transmission (micro-DTX), where the analog components are pushed into micro-sleep mode when there is no data to transmit. For the lower loads, where the opportunity for sleep is higher, this appears as a steeper slope in the power consumption versus load curve. In the current generation of SoCs, while it is feasible to rapidly turn the PAs off and on during micro-DTX, the RFIC and Digital SoCs continue to consume the same amount of power. While future SoCs may be more advanced to allow rapid reduction in power consumption during short inactive periods through, for example, clock gating, the energy efficiency of current 5G systems is quite low as the load approaches zero because the substantial amount of power consumed by the digital frontend and baseband processing isn't required.

To bring power consumption close to zero, the cell essentially needs to be switched off. However, the time taken to become operational from this state is significant, hence it cannot be used for a temporary break in traffic. Different approaches, such as on-demand cell-activation, have been discussed in 5G Advanced to handle such “empty cell” cases more efficiently. An important observation here is that 5G consumes more power to provide the same capacity compared to LTE when the demand is very low, likely due to insufficient scalability and flexibility to maintain high energy efficiency at low loads. Clearly this indicates the need for technological and architectural innovations to further improve energy efficiency at any load.

The above considerations motivate the 6G target illustrated in Figure 3. First and foremost, when the offered load for the 6G cell is less than the peak capacity of a 4G or 5G cell, the 6G system should consume substantially less power than a 4G or 5G cell. In particular, the power consumption should gracefully approach the cell switch-off power with decreasing load. The goal should be to approach an exponential drop in energy consumption as the load decreases from the high loads and near linear scaling of energy consumption at low loads through a series of features that are activated at different loads. When activated, these features should lead to a step reduction in power consumption or a more continuous reduction with reducing load, as illustrated in the figure. Examples of such features are mMIMO muting, advanced micro-DTX with SoC power gating, waveform adaptation, advanced crest-factor reduction (CFR) utilizing increased guard bands and signaling reduction.

Note that achieving a significantly higher peak capacity than 5G to a factor of ~10x, may require higher power consumption. We allow for that in our target to support high-capacity use cases, however, in this regime the energy efficiency of 6G in bits/joule should be dramatically better than that of 5G. Considering that during most hours the loading is under 50%, the overall daily average energy consumption of 6G should be substantially less than that of 5G. For example, we should target that the ETSI 24-hour average power consumption of 6G RAN should be less than half of 5G.

Air-interface design principles

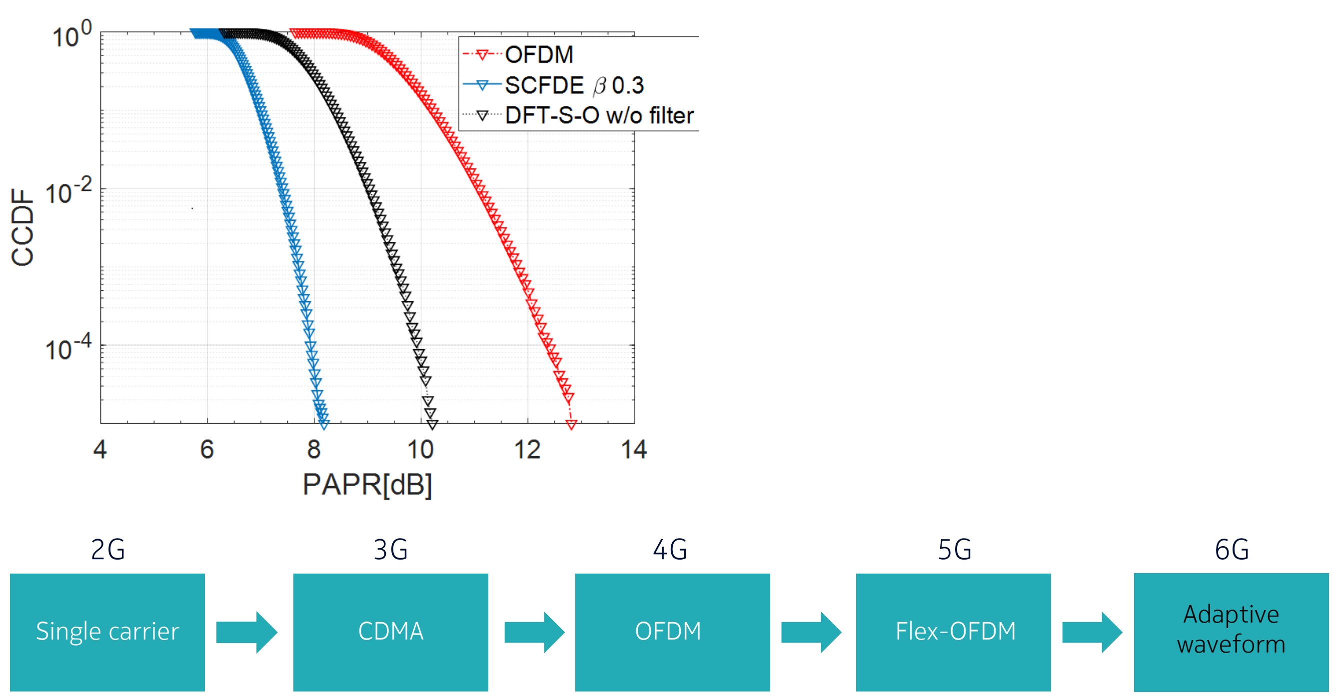
Adaptive waveforms

One of the hallmarks of each new generation is the choice of the waveform. 2G relied on single-carrier continuous phase waveform (GMSK) and time division multiple access (TDMA). As the system became wider in bandwidth, direct sequence spread spectrum (DSSS) was chosen as the waveform for 3G for handling multipath and code division multiple access (CDMA) augmented time division. In 4G, the orthogonal frequency division multiplexing (OFDM) waveform was used along with enhancements to compute capabilities to enable better spectral efficiency over wider bandwidth and facilitate adoption of MIMO. In 5G, OFDM was turned into "flex-OFDM" by enabling multiple sub-carrier spacings to accommodate different frequency bands, bandwidths and use cases, and to allow for the coexistence of different sub-carrier spacing if needed. A drawback of the OFDM waveform is its high peak to average power ratio (PAPR), which requires larger backoff in PA operation to achieve better PA linearity at the cost of lower PA efficiency.

With 6G, the flexibility should be further enhanced to allow multiple waveforms from which an appropriate one can be chosen depending on the load to, for example, facilitate energy savings or co-existence. Figure 4 shows the PAPR of different waveforms without clipping or digital pre-distortion (DPD) operations. It illustrates that single-carrier waveforms have substantially lower PAPR compared to OFDM. A lower PAPR will translate into better PA efficiency by operating closer to the peak output power of the PA. This can be translated into lower power consumption by adjusting the drain voltage and/or bias current of the PA.

At lower loads, there will be no need to multiplex multiple users across the frequency domain. Thus, it will be beneficial to switch from OFDM to a single carrier (SC) waveform such as SC-FDE or SC-FDMA to make energy consumption comparable to that of single carrier systems with narrow bandwidth. While OFDM with very good CFR and DPD technique may be able to achieve the same level of PAPR as the other waveforms, it is possible only with a significant amount of computing.

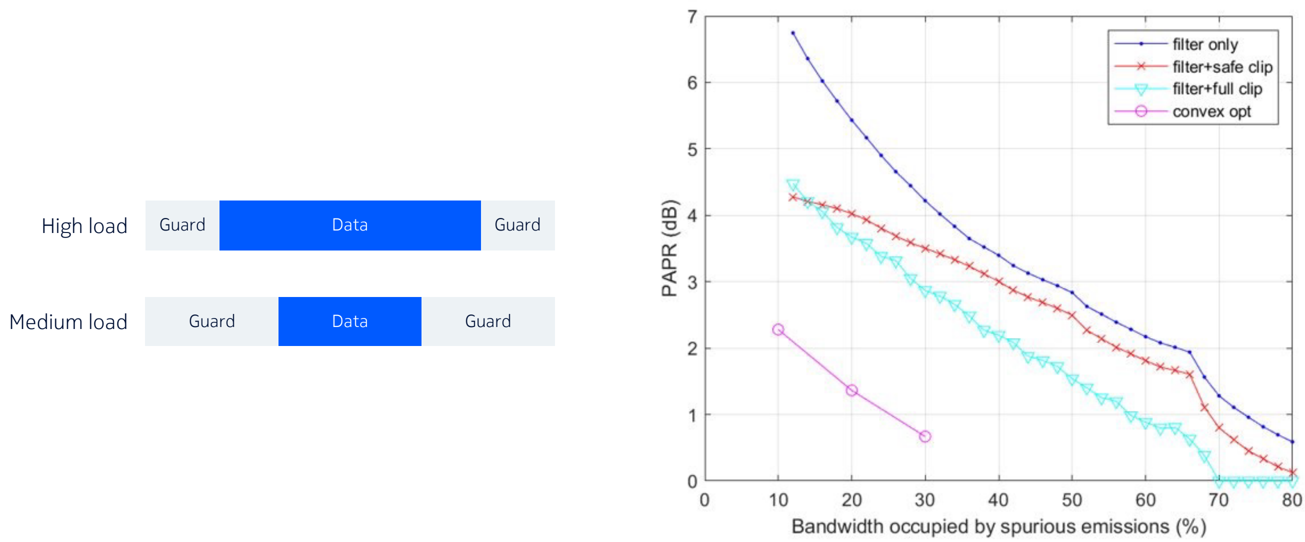
Figure 4. Waveform choices for 6G



Adaptive in-band guard sub-carriers

With OFDM signaling, it is possible to tradeoff energy efficiency and spectral efficiency by allocating some of the in-band tones as guard sub-carriers. As the load reduces, the system should be operated by reducing data transmission to an increasingly smaller subset of sub-carriers. It should also leverage the unused sub-carriers for transmission of signals dependent on the data signal thus reducing the overall peak amplitude of the time domain OFDM signal after the inverse FFT. The power consumption reduction from the reduced PAPR should more than compensate for the additional power required to transmit the compensating signals in the guard sub-carriers. The concept is illustrated in Figure 5.

Figure 5. Frequency-amplitude tradeoff



One approach is to iteratively calculate the signals transmitted in the guard sub-carriers [5]. The time domain signal peaks are identified and clipped, and the resulting error, computed. The frequency domain representation of the error signal at the guard sub-carrier locations is then computed, and the guard sub-carriers are set to compensate the error signal. This process is repeated multiple times. With 25% guard sub-carriers, PAPR reduction of about 3.5 dB is reported. A more fundamental approach to exploit additional guard-subcarriers is through the channel coding process. The codewords chosen for the code can be applied over the extended set of sub-carriers to result in low PAPR and the decoder can also exploit the guard sub-carriers as additional redundancy.

Database-driven spectrum mask adaptation

An alternative approach to increasing the efficiency of power amplifiers is to operate them in the non-linear regime with limited or no backoff. However, there are two major issues when transmitting signals through a non-linear PA. First, the signals will be distorted resulting in poor signal error vector magnitude (EVM), and the receiver may not be able to demodulate the signal. Second, the non-linearity will result in higher out of band emissions.

The first of these issues is beginning to be addressed using receiver processing based on machine learning (ML). It has been shown that the ML receiver can cope with non-linearity through special processing of the signal in the time domain followed by FFT operation and then followed by ML-based demodulation techniques [6].

The second problem is fundamentally introduced by the non-linearity and can be addressed to a certain extent through DPD techniques. However, that requires additional processing power at the transmitter and also limits the extent to which the PA can operate in the non-linear regime. One approach to address this second problem would be to relax the requirements on the out-of-band emissions whenever possible. For example, the adjacent band may not be utilized near the location of the cellular transmitter. In such scenarios, the power efficiency of the cellular transmitter can be vastly increased by allowing out-of-band emissions since no harm is caused. However, when the adjacent band receivers are in proximity, stricter emission masks should be applied, and the cellular transmitter should activate additional processing or backoff to ensure transmitted signals are not distorted.

A spectrum data base could be used, e.g., a simplified solution based on some of the principles used for the Citizens Band Radio Access (CBRS) bands for Priority Access and General Authorized Access to identify devices that are in use in different locations [7]. Transmitters would query the database to discover spectrum occupancy in the adjacent band in the surrounding area. In many cases, the bands adjacent to the cellular bands involve devices that are relatively static or known to follow deterministic trajectories, hence it should be possible to compute their presence in a given area and apply the spectrum mask accordingly.

Dynamically scalable arrays

6G can achieve a 10- to 20-fold increase in cell capacity [4] compared to the 5G system by leveraging larger bandwidths and larger antenna arrays in new spectrum bands such as 7–15 GHz. For example, going from 3.5 to 7 GHz will allow antenna arrays four times as large for the same array dimensions. Similarly, with four times as many transceiver chains, a large spatial multiplexing order can be achieved leading to higher spectral efficiency. There is, however, an important caveat: increasing the number of transceiver chains comes with a substantial increase in power consumption in the RFIC and digital sub-systems of the radio unit. This might be acceptable at heavy loads when 6G is delivering its peak capacity, however, at lower loads, the power consumption can and should be reduced considerably. mMIMO muting, which is already deployed in 5G, can reduce the number of transceivers by turning off subsets of transceiver chains as a function of the load.

One important aspect to consider when turning off transceivers is the resulting impact on coverage; the beamforming gain may be reduced when some portion of the antenna array is turned off. Typically, the coverage is uplink limited, hence it is possible to turn off antenna elements on the downlink but maintain the receiver chain to receive uplink signals. Beyond that point, muting will require coverage enhancement solutions such as repeated transmissions. Downlink and uplink coverage should be analyzed for various antenna configurations to determine what additional coverage solutions should be introduced for some of the configurations as needed. If control channel coverage is the limiting factor, then solutions, such as relying on control channels on a different carrier with carrier aggregation, can be considered.

NOKIA BELL LABS

Architecture design principles

Centralization

One of the key architecture decisions for implementation is where to locate the network functions. While the radio unit (RU) needs to be at the cell site, other network functions in the RAN could potentially be placed further back in a more central site such as the edge cloud or a RAN hub site. The Core functions could similarly be placed at the edge cloud, or a regional cloud or even a national cloud in some countries.

Several factors such as the availability of fronthaul, the availability of suitable edge cloud sites, operational savings, network agility to introduce new services and energy consumption will determine the optimal architecture choice and location of the functions. Here we focus only on the energy consumption perspective, assuming all other factors can be managed. Our main thesis is that network functions should be centralized to the maximum extent possible. Of course, the extent of centralization depends on the specific network function, but there are several that can potentially be centralized:

- RU at the cell site
- Base band processing at the edge data center or RAN hub within the required distance from the cell site to meet the target latency for optimal air-interface performance
- Higher layer RAN functions within the metro area
- Real-time core control and user plane at the regional level
- Subscriber database and other OAM functions.

One of the disadvantages of centralization from the energy consumption point of view is the energy required to transport the information to the central site for processing. On the other hand, central processing can lower energy consumption with statistical pooling gains and lower climate control energy consumption using centralized baseband liquid cooling solutions [8]. Energy consumption for processing and cooling of compute is expected to outweigh energy consumption for optical network transport [12].

Energy savings from pooling can be achieved in a centralized implementation by operating the least number of DUs/CUs to serve all the cells connected to the hub at a given time. For example, DUs/CUs that are provisioned to support peak throughput during busy hours could be turned off at night. To benefit from pooling, it is essential that the fronthaul link from the RUs can be flexibly routed to the different DU/CU instances. When traffic load is light, a single DU should be able to handle more cells than at times when traffic load is high, and hence the fronthaul links from a larger number of cells should terminate at that same DU when traffic is light.

When the traffic is heavy, on the other hand, more DUs are required and the fronthaul links from those cells will have to terminate at more DU instances. The migration to eCPRI in 5G already facilitates fast switching of the fronthaul link, and the specifications should be further solidified for 6G. When considering the period during which the DU serving a particular cell is migrated from one instance to another, it will be beneficial for any ongoing sessions to remain in the original source DU instance while new user sessions are homed in the target DU. Hence the next generation should support “mesh fronthaul” or other topologies where an RU is connected to multiple DUs, and a DU is connected to multiple RUs. This will also increase the resiliency of the network.

RAN acceleration for cloud implementation

While it is clear that centralization with the same custom hardware, as in distributed base stations, is beneficial to reduce energy consumption, implementation of the RAN DU and CU functions as microservices in the cloud has several other benefits. The cloud implementation enables use of COTS hardware and separates software from hardware. As a result, scalability is enhanced, and network automation can leverage cloud infrastructure deployment tools. Deployment in the cloud may thus reduce network operational expenses. A cloud-native network also enables rapid introduction of new services, for example, introducing a new network slice for a particular business or traffic type. Networks become multi-vendor resulting in less dependence on a single entity with potential for cost reduction. For these reasons, 5G is already experiencing a strong drive towards microservices-based cloud implementation [9].

However, implementation of signal processing intensive physical layer functions on a CPU or GPU is not energy efficient. This will especially be the case for a new generation system, since these are typically targeting very high data rates using large bandwidths and higher order mMIMO, thus they leverage special purpose computing to the greatest extent possible. RAN acceleration has already been considered for 5G cloud deployment using in-line acceleration with ASICs that are designed for the specific processing-intensive receive functions of the new generation. This approach provides the best performance for lowest power consumption, however, given the specialized nature of these devices, it also limits their applicability to general purpose computing. It is imperative, therefore, that there is a mechanism to completely power down the acceleration cards when not in use. As discussed, under light traffic conditions not all deployed acceleration cards will be required to stay active as multiple RUs will be pooled and terminated at the same DU.

Densification

An architectural trend that favors more energy efficient networks is the use of small cells or multi-TRP (transmission reception point) base stations with remote radio heads. Smaller cells transmit substantially less power than traditional macro base stations. The coverage area of small cells is relatively small and hence the bulk of the last mile connection is over the wireless optical network with a short wireless hop at the end. Since the power required for transmission over the optical link is substantially less than that of the wireless link, the transmission power for a given bit rate is smaller with small cells than with macro.

Since the macro supports a wide area, however, a large number of small cells each consuming a smaller amount of power is required to provide the same coverage. Each small cell needs to be designed only for a small capacity as the service area is much smaller, thus smaller arrays with fewer layers are sufficient. Similar to RF power, the digital power could also be small. At low loads, the small cells can be switched off dynamically according to traffic patterns within the cell, since only a subset of the small cells will have to serve users — although this is only possible where there is umbrella macro coverage. While prior research suggests that small cell deployments are more energy efficient compared to macros, their net energy efficiency should be re-evaluated under different loading conditions in light of high-capacity massive MIMO macros, which were not considered in the past.

Densification leads to new architectural possibilities such as cell-free massive MIMO [10] in which a user is not associated with any particular cell but is served by multiple surrounding antennas. In a dense network, users are surrounded by antennas instead of users surrounding antennas, as is normally the case. The user does not need to be associated with any particular radio cell and the user's state is maintained at a central location, potentially in the edge cloud. We expect that macro-based architectures will continue to remain to provide coverage and capacity in most locations with the small cell layer serving the dense areas.

Network exposure

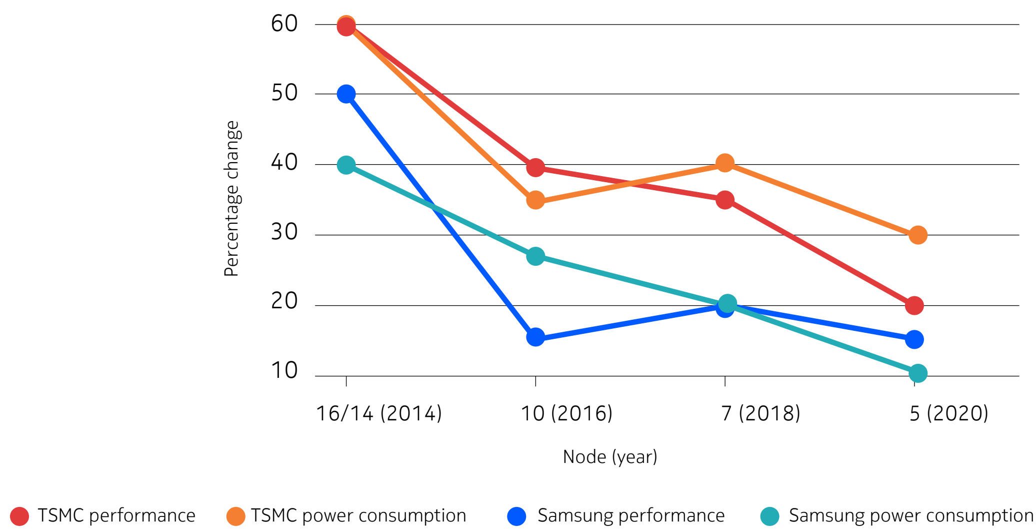
As we approach the 2030s, consumers will increasingly demand to understand the carbon footprint of what they consume, and networks will not be an exception. Already airlines publish the carbon impact of traveling on a flight. Rideshare companies such as Uber allow consumers the choice to pick the type of car for the ride based on their carbon footprint. It is more than likely that consumers will expect to know how their use of the network impacts their personal carbon footprint, for example, of a Metaverse gaming session that they may engage in with the next generation network, which may very well depend on the time of day when it happens.

Beyond consumers, network management decisions such as network configuration and optimization should use energy consumption as a metric. The network exposure framework that is already available in 4G and 5G networks should consider exposure of energy consumption-related metrics, which would allow computation of carbon footprint or allow better energy consumption management of the network.

Computing technology advances

Each new generation has relied on the evolution of semiconductor technology to reduce power consumption for the same amount of computing. However, published literature shows that incremental savings in power from going to the next generation of node technology has been declining steadily, as shown in Figure 6. At best marginal improvements can be achieved going down to 3 nm or 1 nm.

Figure 6. Evolution of power consumption of semiconductor node technology



Source: IEEE Spectrum

A major issue for compute-intensive multiply-accumulate operations such as what is needed for deep learning inference computations is the power consumption of data read-write operations from memory. In particular, the hyper parameter weights are stored in memory and must be fetched for the compute operation. This operation consumes a significant amount of power. Ongoing research on compute-in-memory (CIM) [11] aims to eliminate this power-hungry operation.

In analog CIM, a single flash transistor is used simultaneously for storing neural network weights and performing multiply-accumulate operations. Small electrical currents are steered through flash memory arrays that store neural network weights, and the matrix multiplication results are accumulated through a series of analog-to-digital converters. This provides very dense storage of neural network weights and high-performance AI processing in a single chip without the added cost of internal or external RAM and its associated components. Estimates from on-going research suggest that savings of up to two orders of magnitude are possible using higher compute intensity per watt measured in terms of Tera operations per watt (TOPS). Analog computing is somewhat inherently noisy. However, deep learning networks can be made robust against noise through appropriate offline training of the networks. Hence, analog CIM approaches can be leveraged better by implementing compute-intensive receiver processing though neural network models.

When load is low and portions of the SoC can be powered down depending on the amount of compute power needed, scalable compute in RU will allow a more flexible SoC deactivation, leading to further power savings in RU.

Summary

Energy efficiency will be an important metric for the next generation of networks, at least as important as the traditional performance metrics such as throughput, latency, reliability and scalability. While 5G improved over 4G substantially in terms of the bits/joule metric at peak load, it is highly inefficient during lightly loaded conditions. An important requirement for the 6G design should be to ensure that energy consumption scales down gracefully with traffic, approaching zero energy at zero traffic. With more than a 10- to 20-fold capacity increase expected at peak loads, 6G should leverage various technology advances to reduce overall average energy consumption by 50% and to improve the bits/joule metric by five to ten times at peak load, as compared to 5G.

We envision a broad range of potential solutions for minimizing energy consumption that will have a significant impact at different load levels and on different components of the end-to-end system. We summarize the different technologies across different load levels and components in Table 1, although the list is by no means exhaustive. We focus primarily on what we consider the most important ones to include for further, more detailed exploration. Note that some of the technologies can already be implemented in future product versions of 5G. The rough estimate of the benefits of the different technologies are also shown. In some cases, such as advanced computing and cloud RAN with pooling, the benefits have not been mapped directly to an energy or power metric. To understand the impact of these different technologies directly on the energy consumption in an end-to-end system, more detailed modeling is required to evaluate the energy consumption and efficiency of the different 6G design options. Further study is also needed to ensure that the 6G design enables fine control of the tradeoff between energy efficiency and QoS.

Table 1. Summary of key technologies

Load Level	RF	Digital	Network
All loads	PA efficiency (incl. waveform adaptation)	Compute-in-memory	Densification
	~10% higher PA efficiency	~10x higher TOPS/watt	Deployment dependent
High-to-mid	mMIMO muting	Scalable compute in RU	Central/Cloud RAN with pPooling
	~3x less RF power at 50% load than peak	~2x less at 50% load	30% pooling gain
Mid-to-low	Micro DTX	Micro DTX with SoC power management	Cell switch off
	10% of peak at 10% of load		Deployment dependent

To achieve dramatic reductions in energy consumption while achieving higher capacities when required, the industry should consider the following new approaches described in this paper in its next-generation standards:

- Adaptive waveforms to reduce PAPR while avoiding complex digital pre-distortion
- Frequency domain shaping or coding techniques to reduce PAPR by exploiting unused bandwidth and considering database-driven dynamic emission masks for exploiting more spectrum
- Facilitate mMIMO muting by ensuring coverage solutions for different array sizes
- Mesh fronthaul to ensure resilience and increase pooling gains for centralized functions
- Capability to power down RAN acceleration in the cloud during reduced RAN workloads
- Exposure of energy use statistics to consumer apps to promote energy awareness.

Abbreviations

3GPP	Third-generation partnership project
4G	Fourth-generation mobile network (3GPP)
5G	Fifth-generation mobile network (3GPP)
6G	Sixth-generation mobile network(3GPP)
AI	Artificial intelligence
ASIC	Application-specific integrated circuit
CBRS	Citizen broadband radio access
CDMA	Code-division multiple access
CIM	Compute-in-memory
CFR	Crest-factor reduction
COTS	Commercial off-the-shelf
CPU	Central processing unit
CSP	Communications service provider
CU	Centralized unit
DPD	Digital pre-distortion
DSSS	Direct sequence spread spectrum
DTX	Discontinuous transmission
DU	Distributed unit
eCPRI	Enhanced common public radio interface
ETSI	European Telecommunications Standards Institute
EVM	Error vector magnitude
FDD	Frequency division duplexing
FFT	Fast Fourier transform
FPGA	Field programmable gate array
GMSK	Gaussian minimum shift keying
GPU	Graphics processing unit
LTE	Long-term evolution mobile network (3GPP)
ML	Machine learning
MIMO	Multiple-input multiple-output
mMIMO	Massive multiple-input multiple-output
OFDM	Orthogonal frequency-division multiplexing
PA PSU	Power amplifier power supply unit
NR	5G new radio (3GPP)
OAM	Operations, administration and maintenance
PA	Power amplifier
PAPR	Peak to average power ratio
RAM	Random access memory
RAN	Radio access network
RFIC	Radio frequency integrated circuit
RU	Radio unit
SC	Single carrier
SC-FDE	Single-carrier frequency domain equalization
SC-FDMA	Single-carrier frequency division multiple access
SoC	System on a chip
TDMA	Time-division multiple access
TOPS	Tera operations per watt
TRP	Transmission reception point

References

[1] Green Future Networks: Network Energy Efficiency - NGMN, Dec. 2021. https://www.ngmn.org/publications/green-future-networks-network-energy-efficiency.html

[2] Further study on critical C-RAN technologies, NGMN, Mar. 2015. https://ngmn.org/wp-content/uploads/NGMN_RANEV_D2_Further_Study_on_Critical_C-RAN_Technologes_v1.0.pdf

[3] How 5G is bringing an energy efficiency revolution, Nokia White paper, 2020. https://onestore.nokia.com/asset/207360

[4] H. Holma, H. Viswanathan and P. Mogensen, "Extreme massive MIMO for macro cell capacity boost in 5G-Advanced and 6G," Nokia White Paper, Oct 2021.

[5] P. Boonsrimuang, P. Boonsrimuang, K. Mori, T. Paungma and H. Kobayashi, "PAPR Reduction Method for OFDM Signal by Using Dummy Sub-carriers," 2006 1st International Symposium on Wireless Pervasive Computing, 2006, pp. 1-5, doi: 10.1109/ISWPC.2006.1613571.

[6] D. Korpi, M. Honkala, J. Huttunen, F. Aoudia and J. Hoydis, "Waveform Learning for Reduced Out-of-Band Emissions Under a Nonlinear Power Amplifier," arXiv:2201.05524, Jan 2022.

[7] W. Gao and A. Sahoo, "Performance Impact of Coexistence Groups in a GAA-GAA Coexistence Scheme in the CBRS Band," IEEE Transactions on Cognitive Communications and Networking, vol. 7, pp. 184-196, Mar 2021.

[8] Nokia press release. "Nokia adds Liquid Cooling technology to latest AirScale Base Station portfolio outlining commitment to sustainability #MWC22." https://www.nokia.com/about-us/news/releases/2022/03/01/nokia-adds-liquid-cooling-technology-to-latest-airscale-base-station-portfolio-outlining-commitment-to-sustainability-mwc22/, Mar 2022.

[9] J. B. Moreira, H. Mamede, V. Pereira and B. Sousa, "Next generation of microservices for the 5G Service-Based Architecture," International Journal of Network Management, vol. 30, no. 6, Nov 2020.

[10] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson and T. L. Marzetta, "Cell-Free Massive MIMO Versus Small Cells," IEEE Transactions on Wireless Communications, vol. 16, pp. 1834-1850, Mar 2017.

[11] S. Yu, H. Jiang, S. Huang, X. Peng and A. Lu, "Compute-in-Memory Chips for Deep Learning: Recent Trends and Prospects," IEEE Circuits and Systems Magazine, vol. 21, pp. 31-56, Aug 2021.

[12] "The carbon benefits of cloud computing: A study on the Microsoft Cloud in partnership with WSP", Microsoft White Paper, 2018. https://www.microsoft.com/en-us/download/details.aspx?id=56950

About Nokia

At Nokia, we create technology that helps the world act together.

As a B2B technology innovation leader, we are pioneering the future where networks meet cloud to realize the full potential of digital in every industry.

Through networks that sense, think and act, we work with our customers and partners to create the digital services and applications of the future.

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: CID212810 (February)