Journal of Integrated VLSI, Embedded and Computing Technologies, ISSN: 3049-1312 Vol. 2, No. 3, 2025 (pp. 12-22)
RESEARCH ARTICLE ECEJOURNALS.IN

Challenges and Solutions in Low-Power ASIC Design for Edge Computing Applications

Ahmed Khan*

*Department of Computer and Network Engineering, College of Information Technology, United Arab Emirates University, Al Ain P.O. Box 15551, United Arab Emirates

Keywords:
ASIC Design; Edge Computing; Low-Power Design; Power Efficiency; Resource Optimization; System-on-Chip (SoC)

Corresponding Author Email:
ahmedkh.an@uaeu.ac.ae

DOI: 10.31838/JIVCT/02.03.02

Received : 06.02.2025
Revised : 10.03.2025
Accepted : 19.04.2025

Abstract

On the tail side, edge computing has become the technological innovation of importance due to the increasing demand for real time data processing in which the proliferation of Internet of Things (IoT) devices is one of the main drives. As sophistication of these applications increases, never ever has there been a greater requirement for energy efficient hardware solutions. Thus, among the powerful technologies that came into existence as a result of the novel challenges that edge computing environment brings, ASICs have also emerged. In this article, we explore challenges faced in coming up with low power ASIC for edge computing applications followed with an overview of innovative solutions to tackle these problems. By processing data closer to the source in the edge, edge computing provides lower latency, lower bandwith usage and a much better privacy and reliability. This distributed approach gives new constraints, foremost of which are power consumption and thermal management. ASICs, inherently matchable to target applications, are one of the ways to further optimize performance and reduce energy usage. saying that in this landscape, we examine some of the important consideration in the low power ASIC design and how this is affecting the way edge computing hardware is changing.

How to cite this article: Khan A (2025). Challenges and Solutions in Low-Power ASIC Design for Edge Computing Applications. Journal of Integrated VLSI, Embedded and Computing Technologies, Vol. 2, No. 3, 2025, 12-22

The Edge Computing Paradigm and its Hardware Requirements

Systems that perform analysis and processing of data in a distributed manner are now called edge computing. Edge computing is an option to the centralized cloud structure that is meant to address some of the drawbacks of the centralized cloud including latency, bandwidth limitations and privacy issues by moving the computation close to a source of data. This decentralized approach is required for real time processing applications like Autonomous vehicles, industrial automation and smart city infrastructure. The hardware requirements for edge computing device are much different from traditional server in data center. In the end, most node just work with very strict power budgets of battery or energy harvested. However, they must first perform at high speeds for a set of given tasks under energy constraints. These devices also need to be hardy in a variety of environments and used with minimal maintenance and low maintenance requirements.[1-3]

The meeting of the requirements mentioned above depends on ASIC. Unlike general purpose processors, ASICs have been specialized for specific workload and typically provide much better performance per watt. Thanks to The ASICs for edge computing application, designers can achieve the best tradeoff between computational capability and power consumption, which provides a perfect one. ASIC design technique evolution limits the ability of edge computing hardware to evolve. As edge applications becomes more complex, embedding machine learning, computer vision and sensor fusion, special hardware accelerators are needed. Thanks to the presence of these accelerators which can be additionally

integrated to standard processing units, SoCs proven to be very attractive systems for edge deployment, can be readily included.[4-5]

Challenges in Approach for Low Power ASIC Design in Edge Computing

One of the main challenges in ASIC design for edge computing applications lies in that the system demands to be highly performant whilst being eupheuristically power hungry. The challenges mentioned above are simply aggravated further by the diversity and sometimes variability of edge computing workloads. Therefore, the main problem is to balance the performance and power efficiency. This is due to heavy computational tasks that occur for such edge devices, particularly video processing in real time or machine learning inference, on strict power constraints. That means you cannot simply optimize optimally architecture level to transistor level.[6-8]

The thermal management of edge ASICs is another essential problem. Active cooling is impractical in the environments where many edge devices work and the passive methods of heat dissipation are used instead. This additional limitation adds an even heavier constraint to put with the designers in reducing power consumption as well as distributing heat evenly over the chip. Variability of edge computing workloads is another substantial challenge. For edge devices, unlike in a data center where there can often be predictable usage patterns, the workloads of the edge devices are often wildly varied, depending on environmental conditions or user behavior. These demands impose the requirement on ASICs that energy efficiency is ensured in operation while they can adapt to changing demands.[9-11]

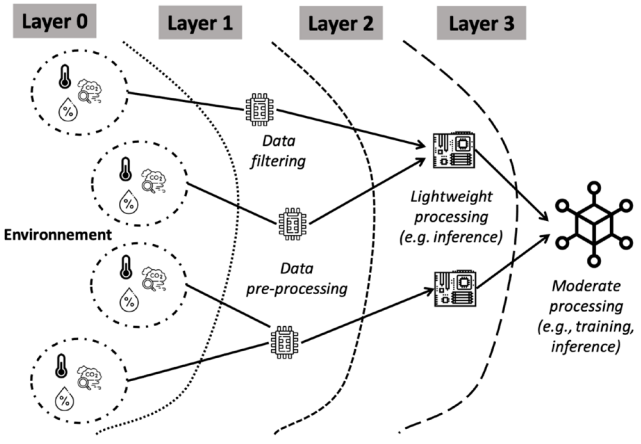

Fig. 1: Challenges in Approach for Low Power ASIC Design

These environments are also security and reliability critical. The security features in ASICs in need to protect sensitive data and unauthorized access. Yet they have to be resistant to environmental imperatives: temperature gradients, electromagnetic interference and physical tampering. To face these challenges we will need to tackle all aspects of ASIC design such as innovative architectures, advanced process technologies as well as sophisticated power management techniques. We then discuss in the rest of this paper some of the essential strategies and solutions for generating very efficient ASICs for edge computing application which are capable of overcoming these issues.[12]

Low-Power ASIC Design for Edge Computing Applications: Challenges and Solutions

On the tail side, edge computing has become the technological innovation of importance due to the increasing demand for real time data processing in which the proliferation of Internet of Things (IoT) devices is one of the main drives. As sophistication of these applications increases, never ever has there been a greater requirement for energy efficient hardware solutions. Thus, among the powerful technologies that came into existence as a result of the novel challenges that edge computing environment brings, ASICs have also emerged. In this article, we explore challenges faced in coming up with low power ASIC for edge computing applications followed with an overview of innovative solutions to tackle these problems.[13-14]

By processing data closer to the source in the edge, edge computing provides lower latency, lower bandwith usage and a much better privacy and reliability. This distributed approach gives new constraints, foremost of which are power consumption and thermal management. ASICs, inherently matchable to target applications, are one of the ways to further optimize performance and reduce energy usage. saying that in this landscape, we examine some of the important consideration in the low power ASIC design and how this is affecting the way edge computing hardware is changing.[15-16]

The Edge Computing Paradigm and its Hardware Requirements

Systems that perform analysis and processing of data in a distributed manner are now called edge computing.

Edge computing is an option to the centralized cloud structure that is meant to address some of the drawbacks of the centralized cloud including latency, bandwidth limitations and privacy issues by moving the computation close to a source of data. This decentralized approach is required for real time processing applications like Autonomous vehicles, industrial automation and smart city infrastructure.[17]

The hardware requirements for edge computing device are much different from traditional server in data center. In the end, most node just work with very strict power budgets of battery or energy harvested. However, they must first perform at high speeds for a set of given tasks under energy constraints. These devices also need to be hardy in a variety of environments and used with minimal maintenance and low maintenance requirements.[18-19]

The meeting of the requirements mentioned above depends on ASIC. Unlike general purpose processors, ASICs have been specialized for specific workload and typically provide much better performance per watt. Thanks to The ASICs for edge computing application, designers can achieve the best tradeoff between computational capability and power consumption, which provides a perfect one. ASIC design technique evolution limits the ability of edge computing hardware to evolve. As edge applications becomes more complex, embedding machine learning, computer vision and sensor fusion, special hardware accelerators are needed. Thanks to the presence of these accelerators which can be additionally integrated to standard processing units, SoCs proven to be very attractive systems for edge deployment, can be readily included.[20-22]

Challenges in Approach for Low Power ASIC Design in Edge Computing

One of the main challenges in ASIC design for edge computing applications lies in that the system demands to be highly performant whilst being eupheuristically power hungry. The challenges mentioned above are simply aggravated further by the diversity and sometimes variability of edge computing workloads. Therefore, the main problem is to balance the performance and power efficiency. This is due to heavy computational tasks that occur for such edge devices, particularly video processing in real time or machine learning inference, on strict power constraints. That means you cannot simply optimize optimally architecture level to transistor level [23] (Fig. 2).

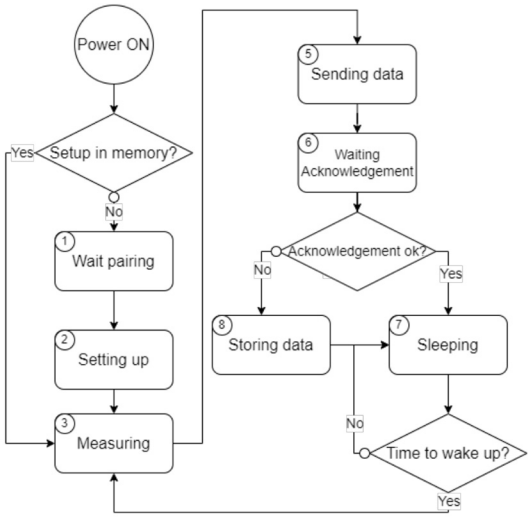

Fig. 2: Low Power ASIC Design in Edge Computing

The thermal management of edge ASICs is another essential problem. Active cooling is impractical in the environments where many edge devices work and the passive methods of heat dissipation are used instead. This additional limitation adds an even heavier constraint to put with the designers in reducing power consumption as well as distributing heat evenly over the chip.

Variability of edge computing workloads is another substantial challenge. For edge devices, unlike in a data center where there can often be predictable usage patterns, the workloads of the edge devices are often wildly varied, depending on environmental conditions or user behavior. These demands impose the requirement on ASICs that energy efficiency is ensured in operation while they can adapt to changing demands.[24]

These environments are also security and reliability critical. The security features in ASICs in need to protect sensitive data and unauthorized access. Yet they have to be resistant to environmental imperatives: temperature gradients, electromagnetic interference and physical tampering. To face these challenges we will need to tackle all aspects of ASIC design such as innovative architectures, advanced process technologies as well as sophisticated power management techniques. We then discuss in the rest of this paper some of the essential strategies and solutions for generating very efficient ASICs for edge computing application which are capable of overcoming these issues.[25]

The other strategy is integrating hardware accelerators for the popular functions. Dedicated

implementation of machine learning or signal processing engines in the form of neural network engines or DSP blocks can give orders of magnitude better energy efficiency in applications of edge computing that need to run these. As a result, asynchronous design techniques turn out to be an opportunity to achieve a substantial amount of power savings. By removing the global clock, asynchronous designs can reduce dynamic power consumption and thus improve the overall energy efficiency. Although challenging to implement, asynchronous architectures are becoming popular for low power ASIC design of edge applications.[26]

Power Optimization by Circuit-Level Techniques

However, circuit related techniques play a crucial role in the area of power consumption in ASICs for edge computing as well. Here the techniques are to maximise overall energy efficiency on the chip by individually optimising the components and circuits. Dynamic voltage and frequency scaling (DVFS) is one of the most widely used circuit level technique. Moreover, the ASIC can also tune its voltage and frequency on demand, or functionally of the contents of the workload running, within DVFS, increasing the power down during low concentration. However, the overhead for DVFS and the sensitivity of power supply systems have caused instability in the system at different points in voltage and frequency.[27]

It also needs to reduce dynamic power consumption, whereby one of the techniques used is clock gating. It is the concept of clock gating when clock signals are disabled to circuit blocks that are in inactive state, so as to save energy by inhibiting unnecessary switching activity. Clock gating can be implemented from coarse granularity (say at functions unit level) to fine granularity (say flip flop level). The low power ASIC design is important and, as the scale factors shrink, leakage current reduction becomes one of the important factors contributing to the overall power consumption of the static power. Reducing leakage current in standby modes might be achieved with methods such as multi-talk CMOS (MTCMOS) as well as adaptive body biasing and not affect the performance during active operation.[28]

On chip interconnects can obtain great power reduction benefit from low swing signaling. For this case, designers can decrease the dynamic power consumed in due to signal transitions by reducing the voltage swing signals passed from one region of the chip to another. However, signal integrity and noise margins to the method are very important. Another key area of low power ASIC design is standard cell, custom circuit design. Power at the gate level is addressed by means of transistor sizing, stack forcing and path balancing. Furthermore, such power savings result from using custom designed, energy efficient memory cells and register files in data intensive edge computing. To minimize resistive losses over the chip, we implement power aware place and route. For ASICs, the power overhead can be significantly influenced by various power grid design including activity based power grid design and intelligent clock tree synthesis .[29]

An investigation of Power Management and Control Strategies

In edge computing applications for ASICs, the efficiency of energy depends on an effective power management. To accomplish that we need sophisticated control strategies that make use of power reduced design techniques, but also make control power dynamically to take into account workload as well as the environment. Power management uses multiple power modes. As such, ASICs can define different operational states with different levels of functionality and power consumption to support adaptation to workload changing requirements. In these cases, the typical power modes encompass the full performance active power mode to an ultra low power sleep state with a number of these in between for other trade offs between performance and power consumption [30] (Fig. 3).

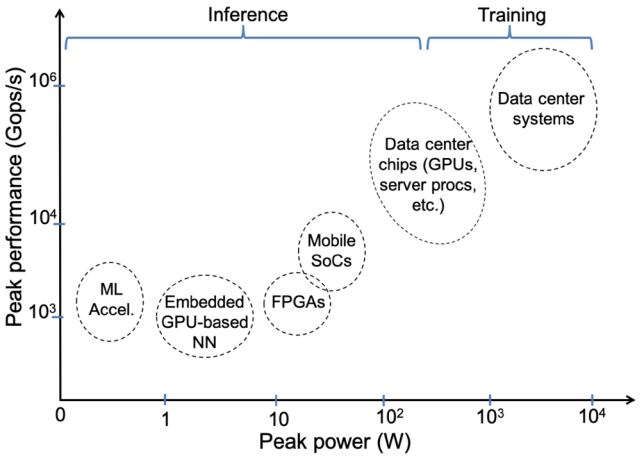

Fig. 3: Power Management and Control Strategies

Task scheduling and workload prediction both have an important feature for Power management.

By doing this, the ASIC can actively cut energy usage by the means of task scheduling and predicting future workload pattern. Some amount will include race to sleep techniques where we run to exhaustion some tasks to quickly put the system in a low power state, or with just in time computing where the processing is deferred as far as possible to avoid wasting power. Closely coupled are edge computing ASICs, which include thermal management as part of power management. To prevent overheating, the ASIC is able to adjust performance and power consumption in the appropriate granularity making use of on chip temperature sensors and thermal management units that are provided by the ASIC. Thermal constraints can be particularly critical and damaging in passively cooled edge devices in which they can be a critical limiting factor in long term reliability.

The other important aspect in good power management is in power aware software optimization. ASICs give software makers the offer of providing APIs (and hardware hooks) to do power optimization at the application level. This may include the ability to selectively drive or not hardware accelerators according to application needs or power aware migration of tasks among heterogeneous cores. It is expected that future edge computing ASICs that have energy harvesting capabilities and adaptive power management will achieve even more energy efficiency. By introducing the circuits to extract energy from the ambient energy sources like light, vibration or RF signals, ASICs can offer longer operation time and the reduction of the dependence on battery power. This harvested energy can then be adaptively managed attuned to the prospected, as well as the current, energy availability and power such harvested energy using adaptive power management algorithms so that energy use is adaptedively optimized. The robust power integrity and noise management techniques also enable reliable operation across different power modes. Also, it brings in the details of careful design of power delivery network, use of decoupling capacitors, and use of specialized power gating to attenuate noise and offer higher, clean power to the most critical circuits (Table 1).[31]

Fully depleted silicon on insulator (FD-SOI) technology as an appealing technology option towards low power edge computing ASICs has been developed. The FD-SOI performs well at low voltages is a good match for applications that require energy saving. In FD SOI processes, it is possible to back bias, and thus, fine grained power management of dynamic power to enable more dynamic power management, and thus, more fine grained control of performance and power consumption. Compelling advantages in ultra low power edge devices will continue to be provided by mature process nodes (28nm and above). Most of the time, these nodes will have smaller leakage current and less variation problems than more advanced nodes. In addition, due to the well established design ecosystems of mature nodes and lower mask costs, the more competitive edge computing solutions for some edge computing applications could be realized.

Non volatile memory technologies embedded MRAM or ReRAM are becoming more central to the design of the low power ASIC design for use in the edge computer. High speed and low power are the advantage offered by these technologies, compared with traditional volatile memories that have lower standby power, faster wake up times and better data retention. By integrating these non volatile memories onto the ASIC, we are able to take advantage of the fact that they are non volatile, and thereby save

Table 1: ASIC Design Methodologies and Tools

Challenge	Description	Impact
Power Consumption	Reducing power consumption while maintaining performance is a significant challenge.	Increases operational costs and limits battery life in portable devices.
Heat Dissipation	Effective thermal management is required due to compact design constraints.	Can lead to overheating, damaging components and reducing system longevity.
Signal Integrity	Maintaining high-quality signals in a noisy, low-power environment is difficult.	Decreases system reliability and data transmission efficiency.
Design Complexity	Designing for low power involves complex trade-offs and design iterations.	Prolongs development time and increases the risk of design flaws.
Manufacturing Cost	High-performance ASICs often come with high manufacturing costs due to custom fabrication.	Limits widespread adoption due to high upfront costs.

enormous system level power. As future low power ASICs are expected to rely on current and up coming technologies, such as carbon nanotube FETs (CNFETs) and 2D material based transistors, the two dimensions are synergistic. Due to the very low voltage operation and low leakage current offered by these technologies, ultra low power and ultra low leakage edge computing devices would be made possible in a new generation. Manufacturability and integration with current CMOS processes are major hurdles, but they still exist. Finally, advanced packaging technology such as 2.5 D and 3 D integration makes power optimization possible for edge computing ASICs. These technologies allow one to put together dissimilar dies and components and thus interpose the best process technologies to deal with the dissimilar functions in a single package. It can improve system performance as a whole and decrease the power consumption.[32]

Low Power ASIC Design Methodologies and Tools

However, designing low power ASICs for edge computing applications is very difficult and thus requires advanced design methods and tools. All of these methods and techniques are very crucial in the trade off between power, performance, area in order to satisfy the tight specifications of edge computing environment. Power Aware design flows are needed for creating energy efficient ASICs. These power analysis and optimization flows are greatly incorporated in the design stages from RTL development to implementation in physical stage. Provide capabilities to estimate early power mainly for early architecture and implementation decisions in the design cycle.

High level synthesis (HLS) tools are now popular for low power ASIC design for edge computing. HLS allows designers to think at the level of higher functionality abstraction (typically C/C++ or SystemC) and to synthesize automatically its RTL implementation. This approach enables designers to easily evaluate different architectural options which, in turn, itself might lead to implementation on a chip with lower power consumption. UPF and CPF are power intent specification languages that help in managing the complexity of multi voltage designs. It lets the designers define power domain, power modes and power management strategies independent of the functional RTL to offer a more structured approach to low power design.

The power performance of an ASIC is to be validated under a real workload and the advanced simulation and emulation are needed to accomplish that task. Power aware simulation tools, based on dynamic and static power analysis techniques, help the designers in finding power hotspots and reduces the overall energy efficiency in the system. Hardware emulation platforms offer us the ability to estimate the more accurate power estimates as well as develop and validate the software on the realistic hardware model. Machine learning assisted design optimization for low power ASIC design is a developing trend. ML algorithms can also predict the power consumption, placement and routing can optimize if it is done using the goals and constraints, and finally, ML algorithms can even suggest architectural improvements for the design goal and constraints. They can cut down significantly the design time and further reduce overall power consumption dramatically.

Adapted formal verification techniques are gaining increasing importance for low power design. They help in verifications of how much correct power management is used, how much the power domains isolate each other and how far the design power intent is different from the actual power intent. In any circumstances, if employed early in the design process, formal methods can catch power related bugs at a point where the botches in consequence are far less than if they are stressed out in silicon. Finally, methods to characterize the power are necessary to develop complete power modeling and characterization. Methodologies for power characterization of complex IP blocks and subsystems as well as a user's manual for creating detailed power models for standard cells, memories and custom blocks are included.[33]

Verification and Validation Challenges in Low Power ASIC Design

Low power ASICs for edge computing applications are very hard to verify and validate under the combination of functionality, performance and power management. It is important that the ASIC performs correctly over all power modes and transitions given stringent power and performance requirements, and the ASIC must be verified completely in this case. Validation process including the verification power aware is a critical part of validation process. It was also found to be interesting that the ASIC worked fine in all power modes, as well as while going from one power mode to another. For example, specialized verification or formal methods are used to verify the power management logic so as to ensure that it does not corrupt the system behavior,

Table 2: Solutions for Low-Power ASIC Design

Solution	Description	Benefit
Dynamic Voltage and Frequency Scaling (DVFS)	Adjusting the voltage and frequency of components dynamically to reduce power usage.	Improves energy efficiency while maintaining required performance levels.
Thermal Management Techniques	Utilizing heat sinks, active cooling, and thermal pads to manage heat dissipation.	Prevents overheating, enhancing device longevity and reliability.
Signal Conditioning	Implementing filters and shields to minimize signal interference.	Enhances data transmission quality and reduces error rates.
Optimized Design Tools	Using advanced CAD tools to optimize low-power designs and ensure efficiency.	Speeds up development time and reduces risk of design errors.
Cost-Effective Manufacturing Processes	Adopting advanced manufacturing technologies to reduce costs and enhance scalability.	Reduces production costs, enabling more affordable solutions.

nor corrupt the power management logic itself (Table 2).

The ASIC dynamic power modeling faces another substantial challenge of the simulation and verification of the behavior under real workloads. Therefore, representative test scenarios should be developed in order to simulate the anticipated use in edge computing environment. Through advanced power analysis tools in conjunction hardware emulator platforms, dynamic power consumption and optimization opportunities are thoroughly analyzed. The completion of thermal validation for the low power ASICs for edge computing is verified. Thermal behavior of the chip for different workloads and environmental conditions is simulated to remain in temperature limits. This process involves use of advanced thermal modeling and analysis tools as well as on chip temperature sensors.

The power delivery network robustness and power integrity breakdown are to be certified at all operational modes. Additionally, it confirms that voltage drops and noise levels do not fall outside of allowed ranges in mode transition or under different workloads. Advanced power integrity analysis tools and methodologies address challenges. Finally it is essential to prove that power management strategies are effective in real word scenarios assurance that power is sufficient to fulfill ASICs power efficiency targets. This may include generating power models of the whole of the edge computing system consisting of sensors, actuators and communication interfaces so as to realistically estimate the power dissipation and battery life while actually using it.[34-35]

Security Consideration in Low Power ASIC for Edge Computing.

Given that edge computing application security is very important because these devices often process sensitive data, disseminate it, and may work in hostile environments, solomon's balancing the cycles time through simple cycle balancing in soccer games in the league, combining cycles through simple cycles in soccer games in the league, and time propagation through simple cycles in soccer games in the league is of paramount importance. At the level of the ASIC, designing low power ASICs requires that it will also provide security features that come at a cost in terms of power consumption. To take care of the safeguarding of data processing along with the communication in edge gadgets, robust cryptographic engines are utilized. Nonetheless, cryptographic operations are computationally and power intensive. Any low power, crytographic integrated ASIC design requires these energy efficient cryptographic accelerators. It may involve lightweight cryptographic algorithms or using dedicated hardware blocks with crypto blocks optimized for performing common crypto operations.

Secure boot and runtime checking integrity are necessary for providing secure boot and runtime integrity to the edge devices in case they boot up in an unexpected state and remain compromised. These features usually require hardware root of trust implementations and the root of trust implementations must be powered naturally in order to consume minimum power and still provide relatively good

security guarantees. The use of one time programmable (OTP) memory to store cryptographic keys and security boot code and other means of reducing power overhead of these security controls is shown .[36]

Another consideration in low power ASIC design to use for edge computing is its side channel vulnerability resistance. Power patterns of a device can be used to extract sensitive information sensitive such as using differential power analysis (DPA). Just like power efficiency countermeasures to such attacks can be carried over onto the ASIC in terms of power efficiency for instance via randomized clock generation or power balancing, but considering security requirements, the ASIC must meet the security requirements with overheads in terms of power. Confidentiality and integrity of sensitive data require secure key management and storage to be done in edge devices. Often, such secure key storage requires hardware memory technologies or dedicated secure elements, which results in power hungry ASIC. To implement these features, security and power efficiency are tradeoffs that need to be well thought out by designers.

The use of physical primitives places security features in the most secure part of the design, and by building on edge computing ASICs which must also have physical security such as anti tamper circuits and sensors because of physical attack they are not vulnerable to these attacks. Increased power consumption in the device and these features must be designed very aggressively for energy efficiency. Power impact of these measures can be diminished by techniques which adapt the sampling rates of tamper detection sensors, and which enable security features only under high threat levels. The third requires that secure firmware update mechanisms be devised for securing the edge devices at all stages of its life cycles. One of the important aspects to be designed is the secure patch update process, such as authentication, integrity verification and experimenting as small amount of energy as possible in the update process, however, preserving the integrity of patches and firmware without breaks.

In Low Power ASIC Design, Future Technologies and Emerging Trends

New semiconductor technology, computing paradigms and newly emerging edge applications are leading to a very fast evolution of area of ASIC design in low power domain in the context of the edge computers. Several emerging technologies and trends will shape it, and further shape it, increasing the levels of energy efficiency and performance. At present there is a huge push for low power ASIC design for edge AI applications that aim to be inspired by neuromorphic computing architectures very much like our brain. These architectures, which are inspired from structural and functional principles of biological neural networks, offer very energy efficient sensory data and machine learning processing. Neuromorphic circuits implemented in ASIC offer substantial power savings in comparison with previous von Neumann Architectures and such implementation can have high impact to applications such as pattern recognition and sensory processing.

The integration of the novel memory technologies (magnetoresistive random access memory (MRAM) and resistive random access memory (ReRAM)) are going to revolutionize the field of low power ASIC design. A rapid read or write plus near zero standby power characterizes these non volatile memory technologies, making them perfect for edge computing uses. The integration of these memories directly into the ASIC promises large reductions in overall system power consumption as well as new techniques for data processing and storage through novel architectural approaches. Approximate computing techniques are increasingly being looked into for use in low power ASIC designs aimed for the edge that tolerate imprecision. Relaxing the exact computation requirement in some part of the algorithm or circuit allows them to achieve significant power savings. This technique is most relevant to edge AI and signal processing applications where small errors in intermediate computations do not matter.

However, energy harvesting and wireless power transfer technologies become precious in the design of ultra low power ASICs supporting edge computing as they continue to evolve in the form of matriculation. For edge devices, much can be gained by harvesting energy from ambient sources, for instance, light, vibration or RF signals, which can help extend energy lifetimes or remove dependence on batteries. However, ASIC integration of efficient energy harvesting circuits and adaptive power management algorithms is needed for long term deployment of autonomous edge devices. New computing paradigms for future low power ASICs will include in memory computing, photonic computing, and the increasingly machine learning driven ASICs. The power of data movement in a computer can be quite significant, so

computations can be run directly in the memory arrays in a particular in-memory computing architecture to reduce power. These photonic computing advantages —fast and low power – carry a prohibitive price, in terms of integration with existing CPMOS technologies.

Hence, the advent of quantum computing and the opportunity to combine it with existing classical architecture ASIC hybrids will be found in new edge computing ASIC hybrid architectures. At this point in time, quantum computers of large scale have a basic level of development, but that is not the case for currently small scale quantum processors (or quantum inspired algorithms implemented on ASICs) that may have important advantages in some edge computing applications, such as optimization and cryptography. Hardware designers, software developers and system architects will cooperate in designing truly optimized solutions for edge computing as the field progresses. Traditionally, the lines of dimension between hardware and software are redone, and hardware/software co design is an important attribute to reach power efficiency with good functionality. Low power ASIC technology for edge computing is a hard and promising future. As will the possibilities continue to grow, the next generation of edge devices will get more and more powerful, more and more able to do more and more, but more and more sustainable and more energy efficient. Since they will be used for developing the technology that stands to bring the next big wave of innovation in IoT, AI etc, helping us inch a bit closer to a truly intelligent world where intelligent and connected devices everywhere will seamlessly come into play.

Conclusion: Edge Computing ASIC Design: The Future

Adding to that, it becomes more apparent that next generation devices will place a significant emphasis on low power ASIC design as they incorporate the features and efficiency of distributing computing to the edge. This document has identified all of the challenges that we have identified from architectural consideration to the circuit level optimization, power management strategies to the implementations of security – all these have demonstrated the complexity and cross disciplinary nature of designing ASICs for the edge. The promise of energy efficiency coupled with performance is fueled by ever improving semiconductor technology coupled with ever bolder design techniques, and ever adopting new computing paradigms. As the number of edge devices increases and more complex use cases arise, there will still be even more necessity for very optimized, highly reapplied hardware solutions. Still, adopting an holistic approach over the chip may be the way forward with low power ASIC design for the future of edge computing. In other word, we desire optimization for some particular workloads, integration with energy harvesting technologies and adaptation to unknown environmental conditions. This will allow us to create more and more intelligent and adaptive power management strategies as the sophistication/performance of the ASIC continues to push from the edge to the fabric into the AI/machine learning capabilities. The future designs will need a compromise of power efficiency and security and the security will always be a key concern in the future designs involving the more sophisticated hardware based security features that offer strong protection while minimizing power efficiency. New memory technologies and novel packaging solutions will make possible novel architectural techniques that will seek to increase performance and reduce power.

References:

1. Niknam, S., Dhillon, H. S., & Reed, J. H. (2020). Federated learning for wireless communications: Motivation, opportunities, and challenges. *IEEE Communications Magazine, 58*(6), 46–51. https://doi.org/10.1109/MCOM.001.1900649
2. Liu, Y., Yuan, X., Xiong, Z., Kang, J., Wang, X., & Niyato, D. (2020). Federated learning for 6G communications: Challenges, methods, and future directions. *China Communications, 17*(6), 105–118. https://doi.org/10.23919/JCC.2020.06.009
3. Rabaey, J. M. (1997, April). Reconfigurable computing: The solution to low-power programmable DSP. *Proceedings of the 1997 ICASSP Conference, Munich, Germany.*
4. Zhang, H., et al. (2000). A 1–V heterogeneous reconfigurable DSP IC for wireless baseband digital signal processing. *IEEE Journal of Solid-State Circuits, 35*(11). https://doi.org/10.1109/4.881202
5. Lodi, A., et al. (2003). A VLIW processor with reconfigurable instruction set embedded applications. *IEEE Journal of Solid-State Circuits, 38*(11). https://doi.org/10.1109/JSSC.2003.818287
6. Khan, L. U., Saad, W., Han, Z., Hossain, E., & Hong, C. S. (2021). Federated learning for the Internet of Things: Recent advances, taxonomy, and open challenges. *IEEE Communications Surveys & Tutorials, 23*(3), 1759–1799. https://doi.org/10.1109/COMST.2021.3075439
7. Bankman, D., Yang, L., Moons, B., Verhelst, M., & Murmann, B. (2018, February). An always-on 3.8J/86in 28nm

CMOS. *2018 IEEE International Solid-State Circuits Conference (ISSCC)*, 222–224. https://doi.org/10.1109/ISSCC.2018.8310267

8. Yin, S., Ouyang, P., Tang, S., Tu, F., Li, X., Zheng, S., Lu, T., Gu, J., Liu, L., & Wei, S. (2018). A high energy-efficient reconfigurable hybrid neural network processor for deep learning applications. *IEEE Journal of Solid-State Circuits, 53*(4), 968–982. https://doi.org/10.1109/JSSC.2018.2797258
9. Amravati, A., Nasir, S. B., Thangadurai, S., Yoon, I., & Raychowdhury, A. (2018, February). A 55nm time-domain mixed-signal neuromorphic accelerator with stochastic synapses and embedded reinforcement learning for autonomous micro-robots. *2018 IEEE International Solid-State Circuits Conference (ISSCC)*, 124–126. https://doi.org/10.1109/ISSCC.2018.8310198
10. Meenakshi, S., Ahamad, S., Vallabhuni, R. R., Malmarugan, R., Rathiya, R., Yadlapalli, K., Rayi, V. K., Gargelwar, A. P., Pithadiya, B. H., Maurya, S., Verma, R., & Rauthan, M. S. (2023). Blockchain for IoT security and privacy. *The Patent Office Journal, 01*, India. Patent Application No. 202241077458 A.
11. Arikumar, K., Prathiba, S. B., Alazab, M., Gadekallu, T. R., Pandya, S., Khan, J. M., & Moorthy, R. S. (2022). FL-PMI: Federated learning-based person movement identification through wearable devices in smart healthcare systems. *Sensors, 22*(4), 1377. https://doi.org/10.3390/s22041377
12. Cordeiro, M., Markert, C., Araújo, S. S., Campos, N. G., Gondim, R. S., da Silva, T. L. C., & da Rocha, A. R. (2022). Towards smart farming: Fog-enabled intelligent irrigation system using deep neural networks. *Future Generation Computer Systems, 129*, 115–124. https://doi.org/10.1016/j.future.2021.10.003
13. Xiao, Y., Zhang, X., Li, Y., Shi, G., Krunz, M., Nguyen, D. N., & Hoang, D. T. (2023). Time-sensitive learning for heterogeneous federated edge intelligence. *IEEE Transactions on Mobile Computing, 23*, 1382–1400. https://doi.org/10.1109/TMC.2023.3251589
14. Gonzalez-Huitron, V., León-Borges, J. A., Rodriguez-Mata, A., Amabilis-Sosa, L. E., Ramírez-Pereda, B., & Rodriguez, H. (2021). Disease detection in tomato leaves via CNN with lightweight architectures implemented in Raspberry Pi 4. *Computers and Electronics in Agriculture, 181*, 105951. https://doi.org/10.1016/j.compag.2020.105951
15. Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al. (2017, July 21–26). PhotoRealistic single image super-resolution using a generative adversarial network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA*. https://doi.org/10.1109/CVPR.2017.19
16. Kvatinsky, S., et al. (2014). MAGIC—Memristor-aided logic. *IEEE Transactions on Circuits and Systems II: Express Briefs, 61*(11), 895–899. https://doi.org/10.1109/TCSII.2014.2357292
17. Giordano, M., et al. (2021). Chimera: A 0.92 TOPS, 2.2 TOPS/W edge AI accelerator with 2 MB on-chip foundry resistive RAM for efficient training and inference. *2021 Symposium on VLSI Circuits, 1–2*. https://doi.org/10.23919/VLSICircuits52068.2021.9492345
18. Podobas, A., Sano, K., & Matsuoka, S. (2020). A survey on coarse-grained reconfigurable architectures from a performance perspective. *IEEE Access, 8*, 146719–146743. https://doi.org/10.1109/ACCESS.2020.3014358
19. Karunaratne, M., Mohite, A. K., Mitra, T., & Peh, L.-S. (2017, June 18–22). HyCUBE: A CGRA with reconfigurable single-cycle multi-hop interconnect. *Proceedings of the 54th ACM/EDAC/IEEE Design Automation Conference (DAC), Austin, TX, USA*. https://doi.org/10.1145/3061639.3062326
20. Lopes, J. D., & de Sousa, J. T. (2016, June 28–30). Versat, a minimal coarse-grain reconfigurable array. *Proceedings of the International Conference on Vector and Parallel Processing, Porto, Portugal*. https://doi.org/10.1007/978-3-319-43659-3_11
21. Hazra, J., et al. (2021). Optimization of switching metrics for CMOS integrated HfO_2-based RRAM devices on 300 mm wafer platform. *2021 IEEE International Memory Workshop (IMW)*, 1–4. https://doi.org/10.1109/IMW51030.2021.9439721
22. Anandaram, H., Kannan, V., Radhakrishnan, R., Vallabhuni, R. R., Vineetha, K. R., Deepa, A., Pramod, K., Divya, P., Sumi, M., Ashish, L., & Logeshwaran, J. (2023). AI/ML for network management and orchestration at the edge of future networks. *The Patent Office Journal, 01*, India. Patent Application No. 202241077460 A.
23. Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., & Kalenichenko, D. (2018, June 18–22). Quantization and training of neural networks for efficient integer-arithmetic-only inference. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA*. https://doi.org/10.1109/CVPR.2018.00285
24. Liu, S.-H., Lin, C.-B., Chen, Y., Chen, W., Huang, T.-S., & Hsu, C.-Y. (2019). An EMG patch for the real-time monitoring of muscle-fatigue conditions during exercise. *Sensors, 19*(14), 3108. https://doi.org/10.3390/s19143108
25. Murmann, B. (2020). Mixed-signal computing for deep neural network inference. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 29*(1), 3–13. https://doi.org/10.1109/TVLSI.2020.3023409
26. Rajpurkar, P., Hannun, A. Y., Haghpanahi, M., Bourn, C., & Ng, A. Y. (2017). Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint arXiv:1707.01836*. https://doi.org/10.48550/arXiv.1707.01836

27. Yousefpour, A., Fung, C., Nguyen, T., Kadiyala, K., Jalali, F., Niakanlahiji, A., Kong, J., & Jue, J. (2019). All one needs to know about fog computing and related edge computing paradigms: A complete survey. *Journal of Systems Architecture, 98*, 289–330. https://doi.org/10.1016/j.sysarc.2019.02.009
28. Miettinen, A. P., & Nurminen, J. K. (2010). Energy efficiency of mobile clients in cloud computing. *Proceedings of HotCloud 2010,* 4.
29. Stone, J. E., Gohara, D., & Shi, G. (2010). OpenCL: A parallel programming standard for heterogeneous computing systems. *Computing in Science & Engineering, 12*(3), 66–73. https://doi.org/10.1109/MCSE.2010.69
30. Fang, W., Zhang, Y., Yu, B., Liu, S., & December, D. (2017, December). FPGA-based ORB feature extraction for real-time visual SLAM. *Proceedings of the International Conference on Field-Programmable Technology (ICFPT),* 275–278. https://doi.org/10.1109/ICFPT.2017.8280143
31. Vanithamani, S., Chellapandi, N., Suganya, S., Vallabhuni, R. R., Satheesh, S., Amudha, K., Siddique, M., Rohini, A., Balu, S., Palanivel, K., Kannan, V., & Logeshwaran, J. (2022). Banana leaf disease detection using CNN – OpenCV-deep learning approach. *The Patent Office Journal, 52,* India. Patent Application No. 202241073393 A.
32. Yu, S., Liu, Y., & Tan, S. (2021). Approximate divider design based on counting-based stochastic computing division. *Proceedings of the 3rd Workshop on Machine Learning for CAD (MLCAD'21),* 1–6. IEEE.
33. Zanandrea, V., Borges, D., da Rosa, V. S., & Meinhardt, C. (2021). Exploring approximate computing and near-threshold operation to design energy-efficient multipliers. *Proceedings of the 34th Symposium on Integrated Circuits and Systems Design (SBCCI'21),* 1–6. IEEE.
34. Uguen, Y., Forget, L., & de Dinechin, F. (2019, September). Evaluating the hardware cost of the posit number system. *29th International Conference on Field-Programmable Logic and Applications (FPL), Barcelona, Spain.* Retrieved from https://hal.inria.fr/hal-02130912
35. Andrysco, M., Kohlbrenner, D., Mowery, K., Jhala, R., Lerner, S., & Shacham, H. (2015, May). On subnormal floating point and abnormal timing. *2015 IEEE Symposium on Security and Privacy (SP),* 623–639. IEEE. https://doi.ieeecomputersociety.org/10.1109/SP.2015.44
36. Dooley, I., Kale, L. V., & Goodwin, N. (2006). Quantifying the interference caused by subnormal floating-point values. *Technical Report*.
37. Tang, U., Krezger, H., & LonnerbyRakob. (2024). Design and validation of 6G antenna for mobile communication. *National Journal of Antennas and Propagation, 6*(1), 6–12.
38. Antoniewicz, B., & Dreyfus, S. (2024). Techniques on controlling bandwidth and energy consumption for 5G and 6G wireless communication systems. *International Journal of Communication and Computer Technologies, 12*(2), 11-20. https://doi.org/10.31838/IJCCTS/12.02.02
39. Sathish Kumar, T. M. (2023). Wearable sensors for flexible health monitoring and IoT. *National Journal of RF Engineering and Wireless Communication, 1*(1), 10-22. https://doi.org/10.31838/RFMW/01.01.02
40. Geetha, K. (2024). Advanced fault tolerance mechanisms in embedded systems for automotive safety. *Journal of Integrated VLSI, Embedded and Computing Technologies, 1*(1), 6-10. https://doi.org/10.31838/JIVCT/01.01.02
41. Sadulla, S. (2024). Techniques and applications for adaptive resource management in reconfigurable computing. *SCCTS Transactions on Reconfigurable Computing, 1*(1), 6-10. https://doi.org/10.31838/RCC/01.01.02
42. Borhan, M. N. (2025). Exploring smart technologies towards applications across industries. *Innovative Reviews in Engineering and Science, 2*(2), 9-16. https://doi.org/10.31838/INES/02.02.02
43. Velliangiri, A. (2024). Security challenges and solutions in IoT-based wireless sensor networks. *Journal of Wireless Sensor Networks and IoT, 1*(1), 8-14. https://doi.org/10.31838/WSNIOT/01.01.02
44. Uvarajan, K. P. (2024). Advanced modulation schemes for enhancing data throughput in 5G RF communication networks. *SCCTS Journal of Embedded Systems Design and Applications, 1*(1), 7-12. https://doi.org/10.31838/ESA/01.01.02
45. Uvarajan, K. P. (2024). Integration of artificial intelligence in electronics: Enhancing smart devices and systems. *Progress in Electronics and Communication Engineering, 1*(1), 7–12. https://doi.org/10.31838/PECE/01.01.02