ICDCIT 2025

AI-Based Power Optimization Techniques for VLSI Circuits

By

Bipasha Biswas, Khushi Singh

Kalinga Institute of Industrial Technology, BBS

2025-26

Problem Statement

Gate Array ICs, or FPGAs, are flexible chips that can be reprogrammed for different uses. However, they have some problems:

1. **Power Consuming**: They use more energy than other chips because they're flexible. This is a problem in places like data centers where energy is important.
2. **Hot Chips**: As they get more powerful, they also get hotter. This can damage them if not cooled properly.
3. **Expensive**: They're usually more expensive to make than other chips. This can make them less attractive for products that need to be cheap.

4. Security and Reliability Concerns:

FPGAs are at risk of being hacked because they can be changed. They must be very reliable in important devices like cars and medical equipment.

5. Size:

Making FPGAs smaller is difficult and expensive.

6. AI and ML:

AI and ML need special chips. Quantum computing may help, but it's not ready yet.

7. Environment:

Making chips is hazardous for the planet’s environment. We need to use less resources and recycle more.

8. Hacking:

Chips can be hacked at a very low level. This is a big problem with IoT devices.

Abstract

1. Solving *excessive power consumption* enhancing speed of chef processing while protecting hardware from external malicious attack for new generation computers and CPUs.
2. Integrating *Artificial Intelligence* and *Machine Learning* algorithms which *new generation ASIC and FPGA chips* to implement their parallel processing capabilities and flexibility for better supervision on abnormalities.
3. This enhances FPGA with *more optimized libraries tools and support* for complex deep learning frameworks.
4. This integration will be done using *High Level Synthesis (HLS)* in evening creation of AI and ML algorithms in high level programming language is like C/C++ or Python.
5. Integration of cyber security with FPGA using *Platform Firmware Resilience (PFR)* for detection and customization of system to protect against specific threats.
6. ML algorithm will be linked with *iOS app* or *Android app* along with *Web Portal* to provide *user* the *complete supervision* of their IC system condition.

Solution

1. **Hardware Design:**

Advanced Power Optimization:

Use dynamic voltage and frequency scaling (DVFS), clock gating, and power gating to adjust power usage dynamically based on workload.

Incorporate adaptive power management systems that leverage AI to predict workload patterns and adjust power usage accordingly.

Explore non-volatile memory (NVM) and 3D-stacked memory architectures to reduce power consumption in memory operations.

Use power-aware placement and routing algorithms to minimize signal propagation delays and reduce switching power.

Parallel and Distributed Processing:

Utilize distributed AI/ML workload management across multiple FPGA or ASIC units to optimize resource usage and reduce bottlenecks.
Optimize dataflow architecture for maximum throughput in AI/ML processing, with on-chip memory management to minimize external data transfers.
Develop custom hardware accelerators for key AI/ML operations (e.g., matrix multiplication, convolution layers) to offload intensive tasks from the CPU.

Hardware Security and Resilience:

Implement real-time hardware monitoring with AI-based anomaly detection to identify potential threats. Utilize side-channel attack mitigation techniques like noise injection or scrambling data to protect against vulnerabilities. Add self-healing mechanisms to detect and recover from potential attacks, ensuring system uptime.

2. **AI/ML Algorithm Integration:**

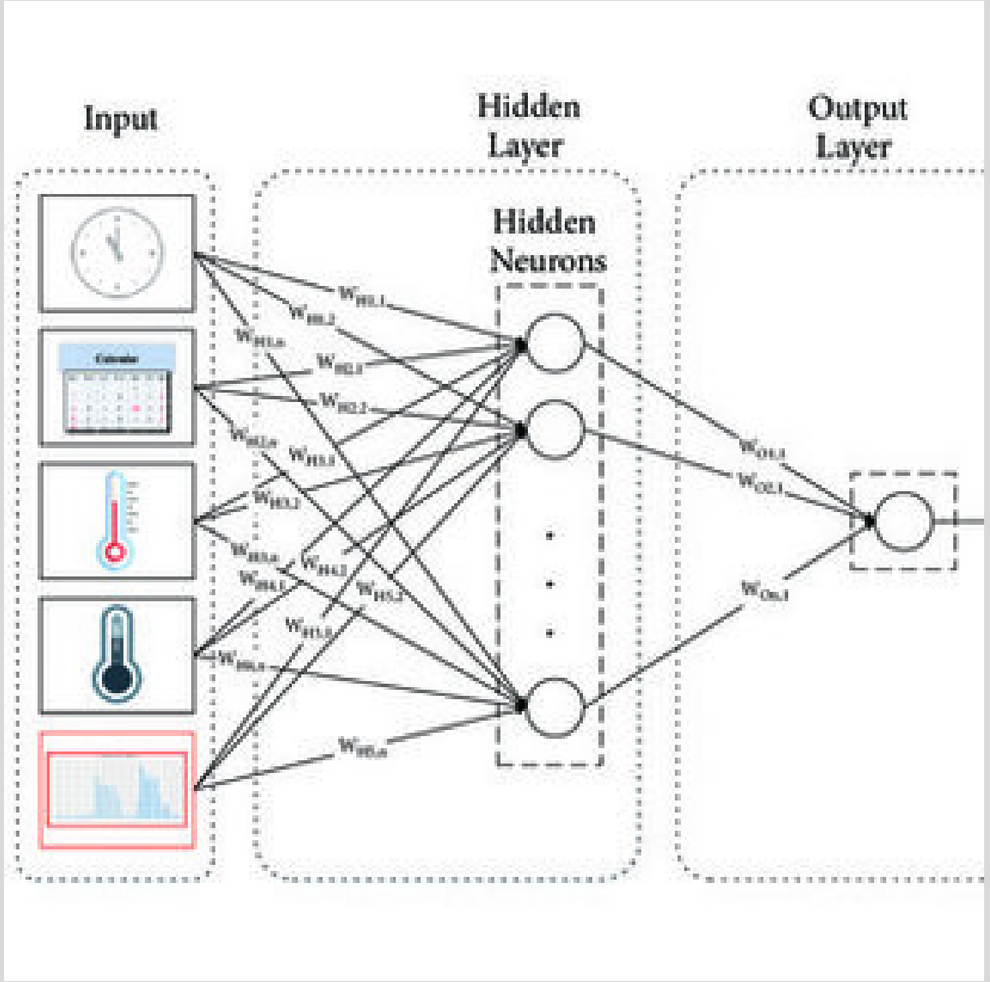
HLS-Based and AI-Driven Optimization:

Use AI-assisted HLS tools that automatically optimize AI/ML algorithms for power, performance, and area (PPA), offering suggestions based on past successful designs.

Implement auto-tuning algorithms to adjust precision dynamically during runtime, balancing performance with power consumption based on the specific workload. Optimize for edge AI scenarios, ensuring low-latency performance with minimal power overhead, crucial for mobile and IoT devices.

Algorithm Compression and Optimization:

Implement algorithm quantization and pruning techniques to reduce the computational complexity of AI models, saving power without sacrificing accuracy.
Explore the use of neuromorphic computing principles to model AI/ML tasks more efficiently on FPGA/ASIC architectures.
Incorporate hardware-friendly AI models like binary neural networks (BNNs) or low-bit precision neural networks that are optimized for FPGA/ASIC deployment.

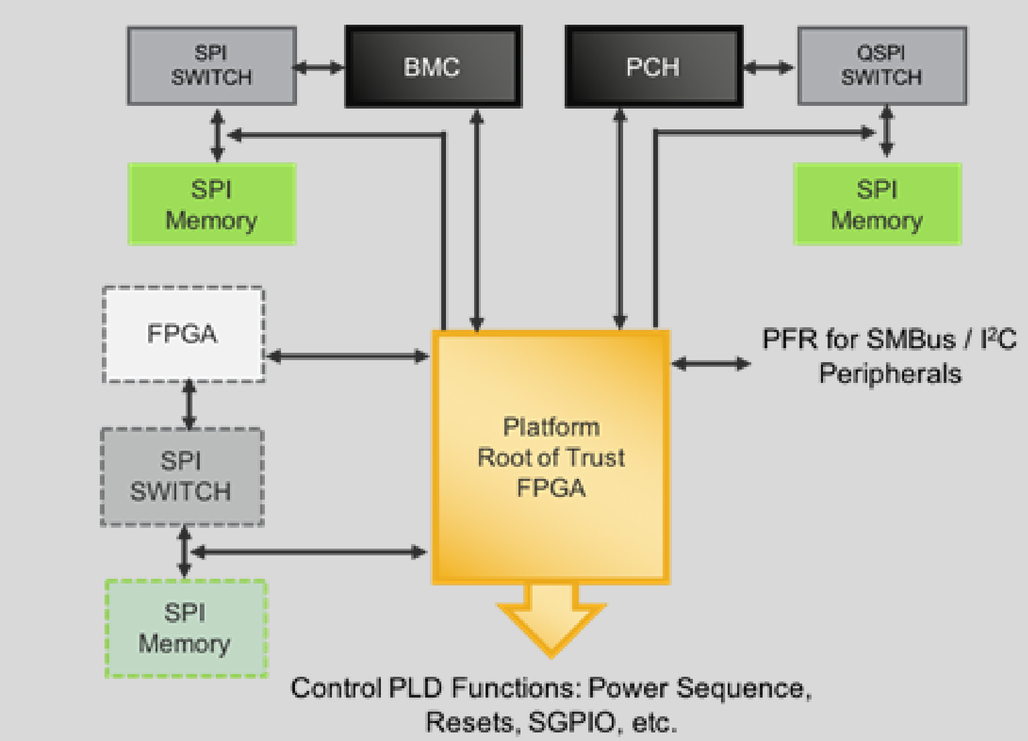


3. **Cybersecurity Integration:**
Advanced PFR-Based Protection:

Extend Platform Firmware Resilience (PFR) with AI-driven threat detection that identifies evolving attack patterns and adapts defenses in real-time.

Integrate Blockchain-based security mechanisms for secure firmware updates and device authentication to prevent unauthorized modifications.

Ensure compliance with industry security standards (e.g., FIPS 140-2, NIST SP 800-193) to meet the regulatory demands of aerospace, automotive, and healthcare applications.



DETECT → RECOVER → BOOT → PROTECT

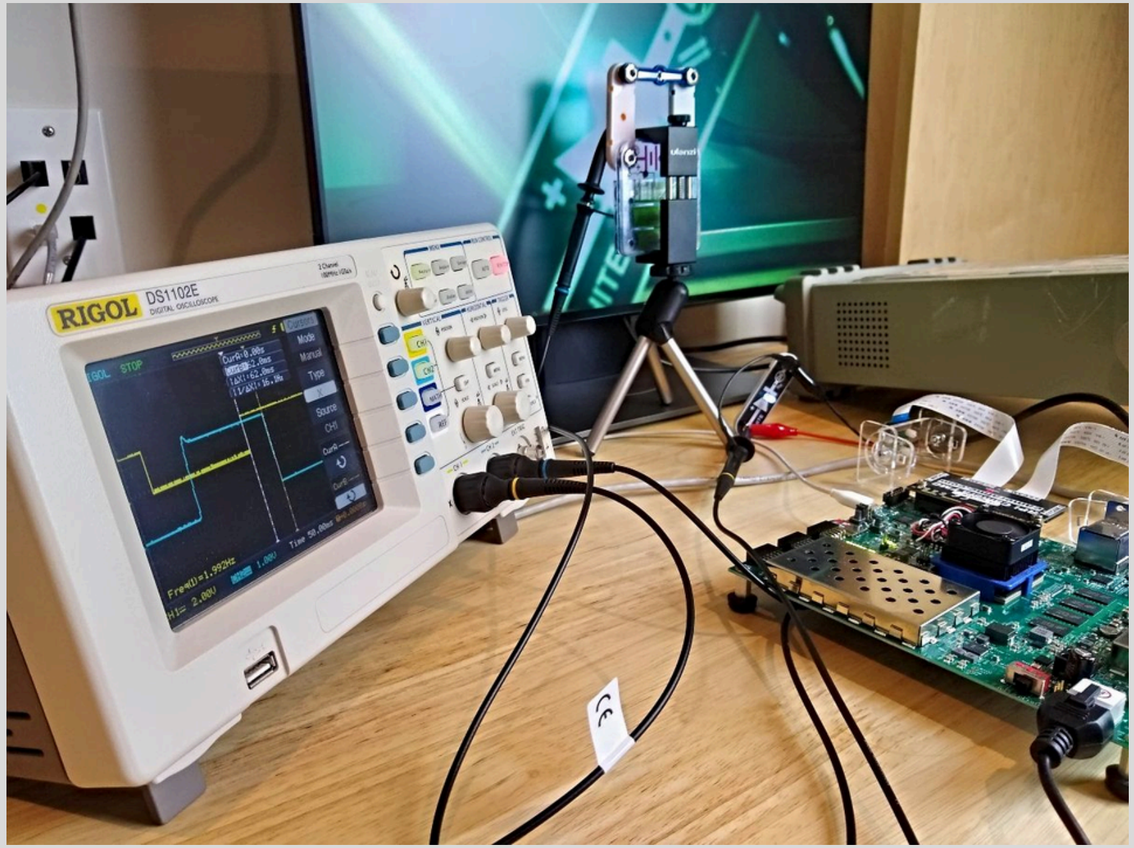
NIST SP 800 193 Based PFR Implementation for Servers

4. **Software and Cloud Integration:**
iOS and Android Apps with AI-Driven Monitoring:

Develop AI-powered mobile applications that allow users to monitor system performance, power consumption, and security risks in real time.

Provide automated anomaly detection and alerts via mobile apps, helping users identify potential issues early.

Enable AI-based recommendations for users to optimize performance or conserve power based on usage patterns.

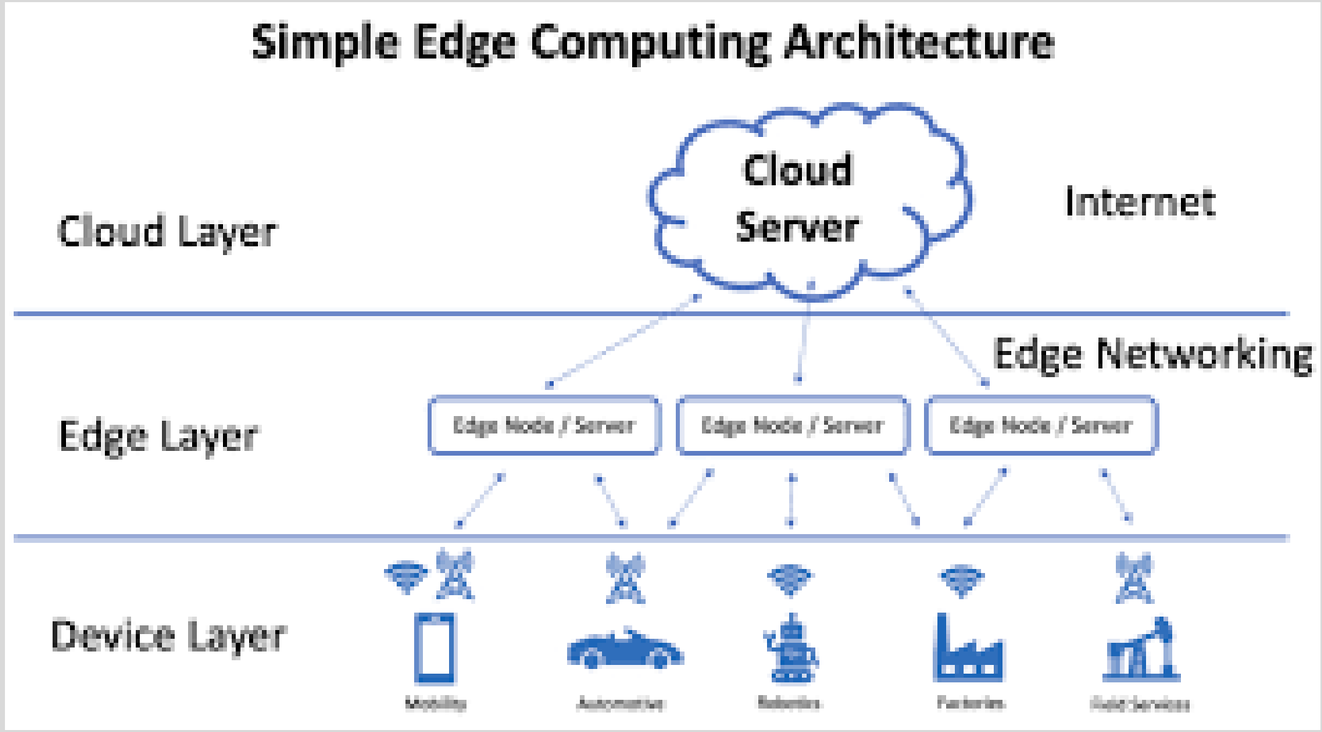


Additional Features:

Edge Computing Support: Optimize FPGA/ASIC designs for low-power, low-latency edge AI applications, allowing them to perform real-time AI inference without relying on cloud resources.

Energy Harvesting: Incorporate energy harvesting mechanisms (e.g., solar or thermal energy) for powering FPGA/ASIC chips in remote or low-power environments, reducing reliance on traditional power sources.

Green Computing Initiatives: Use recycled materials in chip production and design energy-efficient cooling systems to minimize the environmental footprint of the hardware.



Analysis of results, and compare with previous existed solution

The proposed solution for designing power-optimized AI/ML integrated FPGA and ASIC chips offers significant advancements in power efficiency, performance, security, and scalability. However, these advancements come at the cost of increased complexity, production costs, and potential overengineering for simpler applications.

Power optimization, AI/ML processing, security, scalability, and usability are all improved, but these improvements can also introduce additional complexity, costs, and potential overengineering. The solution is well-suited for high-performance, high-security, and scalable applications but may not always be practical for simpler use cases. Careful consideration of the specific requirements and trade-offs is essential when selecting the most appropriate design approach.

Limitations

- Complexity: The proposed solution adds considerable complexity, particularly for smaller or less demanding applications, potentially leading to longer design times, increased costs, and a higher likelihood of errors.
- Cost: The additional features and optimizations may raise production expenses, which could affect the solution's economic feasibility in certain markets or applications.
- Overengineering: For simpler use cases, the solution might be unnecessarily complex, resulting in extra costs and effort that aren’t needed.
- Environmental Impact: The emphasis on sustainability could drive up production costs, which may not be justifiable in markets with fewer regulations.

Future Work

- Custom Solutions: Develop more customized solutions tailored to specific use cases, emphasizing only the key features and optimizations necessary for each application.
- Cost Reduction: Investigate methods to lower production costs without compromising on performance and efficiency.
- Streamlined Design: Create simpler design methodologies or tools to help reduce complexity and speed up the development process.
- Hybrid Solutions: Explore hybrid approaches that combine the strengths of FPGA and ASIC technologies with specialized hardware accelerators for AI/ML tasks.
- New Technologies: Look into emerging technologies, such as neuromorphic or quantum computing, to enhance AI/ML capabilities and overcome current limitations.

By tackling these challenges and pursuing future research opportunities, even more sophisticated and efficient AI/ML-integrated FPGA and ASIC chips can be developed to meet the evolving demands of various applications.

Reference links

https://www.wipro.com/infrastructure/edge-computing-understanding-the-user-experience/

https://ieeexplore.ieee.org/document/777270

https://www.researchgate.net/publication/274078152_Optimization_of_Power_Consumption_in_VLSI_Circuit

https://vlsifirst.com/blog/using-artificial-intelligence-and-machine-learning-in-fpga-design

https://ieeexplore.ieee.org/document/777270

https://www.cyrene.eu/fpga-technology-and-supply-chain-cyber-security/