BANNARI AMMAN INSTITUTE OF TECHNOLOGY
An Autonomous Institution Affiliated to Anna University - Chennai • Approved by AICTE • Accredited by NAAC with "A" Grade
SATHYAMANGALAM - 638401 ERODE DISTRICT TAMILNADU INDIA
Ph : 04295-226000/221289 Fax : 04295-226666 E-mail : stayahead@bitsathy.ac.in Web : www.bitsathy.ac.in

22IS303- COMPUTER ORGANIZATION AND ARCHITECTURE

UNIT – V
ACCELERATORS- FPGA, ASIC

FPGA – INTRODUCTION

- An FPGA (Field-Programmable Gate Array) is an integrated circuit that can be programmed or reconfigured after its initial manufacturing. It consists of an array of programmable logic blocks and interconnections that can be customized to perform various digital tasks

BASIC PRNCIPLE OF FPGA OPERATION

- **Logic-block Architecture**: A Configurable Logic Block (CLB) serves as a fundamental component within an FPGA, providing the essential logic and storage capabilities necessary for a given application design. This core component can range from fine-grained transistors to coarse-grained processors, with trade-offs between resource efficiency and performance
- The spectrum of basic logic blocks includes options like NAND gates, multiplexers, lookup tables (LUTs), and PAL-style wide input gates. LUT based CLBs, employed by commercial vendors like Xilinx and Altera, strike a balance between granularity extremes, offering flexibility without excessive resource wastage.
- A basic logic element (BLE) typically consists of a LUT and a Flip-Flop, with a LUT-4, for instance, utilizing 16 SRAM bits to implement 4-input Boolean functions. CLBs can form clusters of interconnected BLEs, facilitating local routing networks for inter-BLE communication. While CLBs constitute the foundation, modern FPGAs often incorporate specialized blocks like memory, multipliers, and DSP blocks, optimized for specific tasks, but potentially leading to resource inefficiencies when unused
- **Routing Architecture**: In FPGA architecture, the computing capabilities are provided by programmable logic blocks interconnected through a configurable routing network. This network, comprising wires and programmable switches, is essential for implementing user-defined circuits. To cater to the diverse routing needs of different digital circuits,

FPGA routing interconnects must be highly adaptable. While routing requirements vary, many designs exhibit a preference for short wires due to their locality, but they also require some longer connections.

- Designing FPGA routing interconnects must strike a balance between flexibility and efficiency. The organization of routing resources in relation to logic block placement, referred to as global routing architecture, greatly impacts overall efficiency. Additionally, the specific configuration of switch blocks at a detailed level, known as detailed routing architecture, plays a crucial role.

ARCHITECTURE OF FPGA

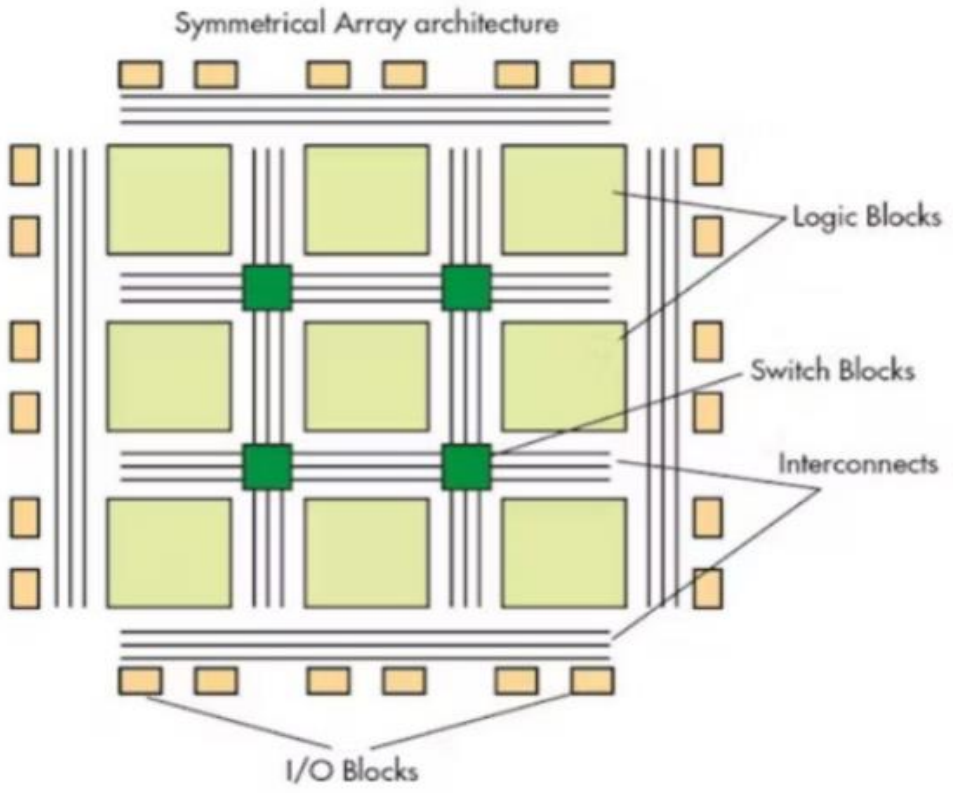

Fig 1.1 Architecture of FPGA

- FPGAs excel in time-critical applications like software-defined radio, medical devices, and mil-aero systems, making FPGAs more versatile for prototyping than ASICs. FPGAs also shine in parallel processing, ideal for tasks involving extensive data processing, such as embedded vision and measurement systems
- FPGAs are reconfigurable chips, allowing their internal logic to be modified via software updates, bug fixes, or support for new protocols
- But FPGAs are costlier than microcontrollers or ASICs, especially in high-volume production scenarios, as the per-unit cost tends to be higher. FPGAs typically consume more power than ASICs and microcontrollers, which may be a concern in power-sensitive applications

ASIC – INTRODUCTION

- An application-specific integrated circuit is an integrated circuit (IC) that's custom-designed for a particular task or application.

BASIC PRNCIPLE OF ASIC OPERATION

- The working principle of ASIC chips is similar to regular integrated circuits, as they are both composed of multiple logic gates (such as AND gates, OR gates, NOT gates, etc.) and memory units (such as registers, TPS2051BDBVR memory, etc.).
- ASIC chips can be designed and manufactured according to specific application requirements, which allows them to have higher performance and lower power consumption. The workflow of an ASIC chip is roughly as follows: Input Signals: The ASIC chip receives input signals such as data, clock, control signals, etc; Logical Operations: The ASIC chip performs logical operations such as AND, OR, NOT, etc;
- **Data Storage**: The ASIC chip stores the calculation results in registers, memory units, and so on;
- **Output Signals**: The ASIC chip outputs the computed results, such as data, status signals, etc.
- The design of an ASIC is carried out step by step, following a sequence known as the ASIC design flow
- The following flowchart outlines the steps of the design process:

  1. **Design Input**: In this step, the design's microarchitecture is implemented using hardware description languages like VHDL, Verilog, and System Verilog.

  2. **Logic Synthesis**: During this stage, a netlist of the logic units to be used, along with interconnects and all other required components, is prepared using HDL.

  3. **System Partitioning**: In this step, the larger-sized chip is divided into ASIC-sized portions.

  4. **Pre-layout Simulation**: Simulation testing is performed at this stage to check for design errors.

  5. **Layout Planning**: The netlist blocks are arranged on the chip in this step.

  6. **Placement**: The positions of the units within the blocks are determined.

  7. **Routing**: Connections are established between blocks and units in this step.

  8. **Extraction:** Electrical characteristics like resistance and capacitance values of the interconnects are determined. 9. Post-layout Simulation: This simulation is carried out to ensure that the system operates correctly under interconnect loads before manufacturing models are submitted.

DESIGN STEP OF ASIC:

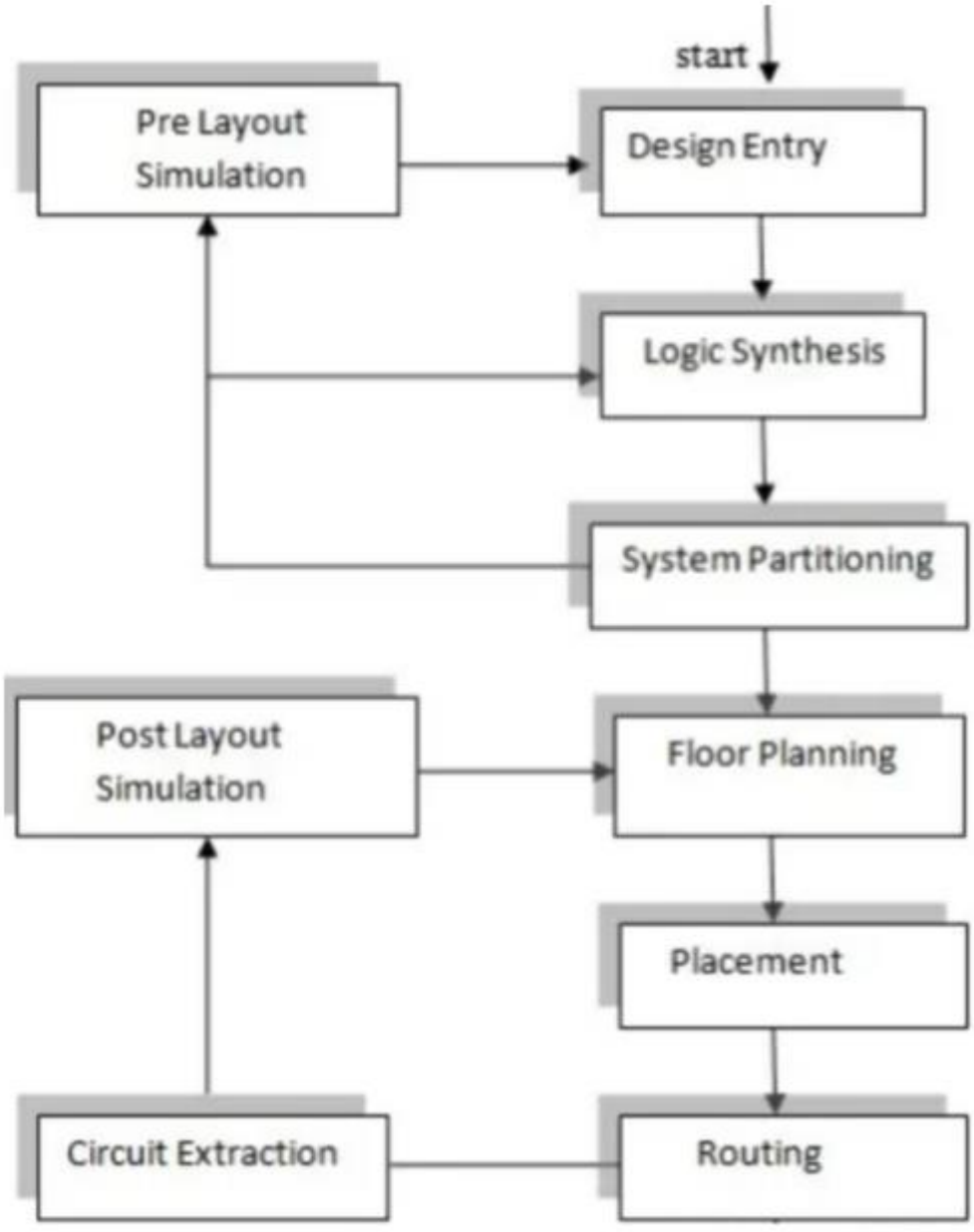

Fig 1.2 ASIC Design Step

Comparing GPU, FPGA and ASIC:

- GPU, FPGA, and ASIC stand out as the primary AI acceleration technologies, each distinguished by its distinct operational principles and a set of strengths and weaknesses that render them suitable for diverse applications. GPU CUDA programming has revolutionized medical image analysis, empowering tasks such as image denoising, image registration, segmentation, and visualization
- On the FPGA front, neural network implementation underscores FPGA's prowess in optimizing computational blocks for hardware-based neural networks. Additionally, FPGA finds practical application in aerospace, addressing radiation induced faults through fault identification, mitigation, and correction strategies, ensuring uninterrupted system operation.

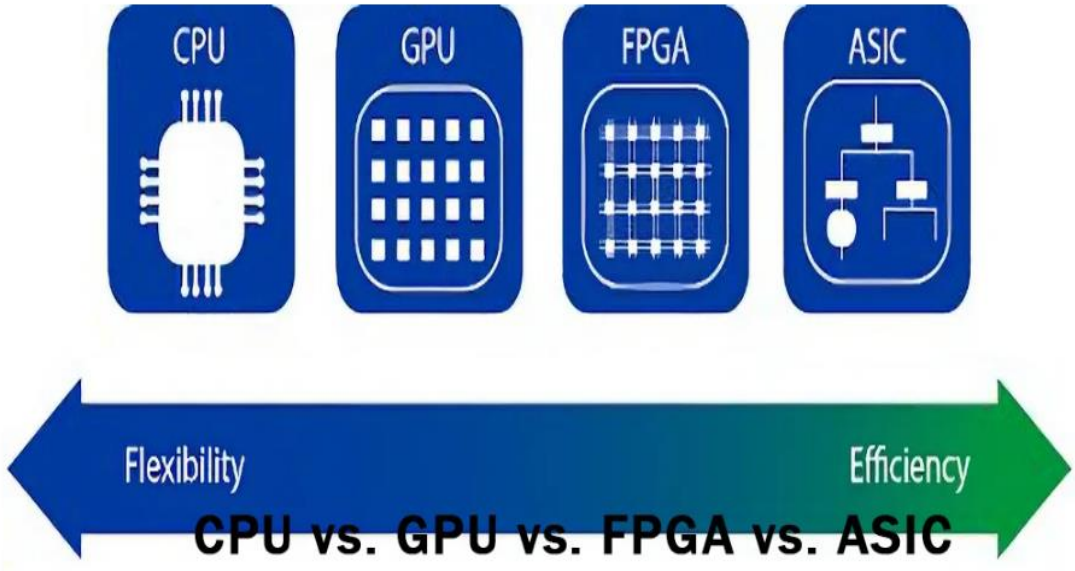

Fig 1.3 CPU Vs GPU Vs FPGA Vs ASIC

- ASICs have showcased their prowess in Bitcoin mining, boasting remarkable energy efficiency and high-performance capabilities. This ASIC represents a significant leap forward, offering a cost-effective solution for miners aiming to reduce energy consumption while achieving high mining throughput.

Accelerator	GPU	FPGA	ASIC
Overview	Initially intended for graphics and videos, these chips are now employed in a diverse array of computationally demanding tasks.	An integrated circuit that can be programmed or reconfigured after its initial manufacturing.	An integrated circuit (IC) that's custom-designed for a particular task or application.
Strengths	Parallel computing; Flexible and programmable; Widespread support	Low latency; Adaptable; Parallel computing; reconfigurable	Lower unit cost; optimum performance and power consumption
Weaknesses	Bottleneck; High power consumption; Generate substantial heat; May not offer lowest latency for individual task	Expensive in large production; Consume more power than ASIC;	Take long time to the market; Expensive design tools; Some design issues

Fig 1.4 Comparing GPU, FPGA and ASIC