

Approval: 14th Senate Meeting

Course Number	: CS507
Course Name	: Computer Architecture
Credits	: 3-0-2-4.
Pre-requisite	: CS201 – Computer Organization or Equivalent
Intended for	: BTech Computer Science Engineering (CSE) and Electrical Engineering (EE), MS, M. Tech. & PhD.
Distribution	: Elective for Third and Final year B. Tech (CSE/EE), MS, M. Tech. in VLSI/Signal Processing and Communication & PhD
Semester	: Even or Odd.

1. Preamble:

VLSI technology has been major driving force for phenomenal growth in computing power of processors, and for emergence of new applications. A few years ago, CPUs had nearly 1B transistors, while the recently announced Oracle SPARC M7 CPU will have more than 10B transistors on chip. Similarly, the GT200 GPU (2008) had 1.4B transistors, while the recent Geforce GTX TITAN X GPU has 8B transistors. Early applications were mainly compute intensive whereas newer applications are either data intensive or both data and compute intensive. These applications include weather forecasting, astronomical data analysis, multimedia and gaming applications etc. Since development VLSI technology has reached a level where no major breakthroughs are expected in near future, architects have been exploring alternative design and implementation strategies. The objective of this course is to provide detailed discussion on architectural mechanisms, which exploit parallelism available in programs at various granularities, and programming tools to work with such architectures.

2. Course Modules with Quantitative Lecture Hours:

1. **Introduction** – Defining Computer Architecture, Flynn’s Classification of Computers, Metrics for Performance Measurement **(4 Hours)**
2. **Memory Hierarchy** – Introduction, Advanced Optimizations of Cache Performance, Memory Technology and Optimizations, Virtual Memory and Virtual Machines, The Design of Memory Hierarchy, Introduction to Pin Instrumentation and Cachegrind, Case Study: Memory Hierarchies in Intel Core i7 and ARM Cortex-A8. **(8 Hours)**
3. **Instruction Level Parallelism** – Instruction-level Parallelism: Concepts and Challenges, Basic Compiler Techniques for Exposing ILP, Reducing Branch Costs with Advanced Branch Prediction, Dynamic Scheduling, Advanced Techniques for Instruction Delivery and Speculation, Limitations of ILP, Multithreading: Exploiting Thread-Level Parallelism

to Improve Uniprocessor Throughput, Modeling Branch Predictors using Pin Tool, Case Study: Dynamic Scheduling in Intel Core i7 and ARM Cortex-A8. **(10 Hours)**

4. **Thread Level Parallelism** – Introduction, Shared-Memory Multicore Systems, Performance Metrics for Shared-Memory Multicore Systems, Cache Coherence Protocols, Synchronization, Memory Consistency, Multithreaded Programming using OpenMP, Case Study: Intel Skylake and IBM Power8. **(10 Hours)**

5. **Data Level Parallelism** – Introduction, Vector Architecture, SIMD Instruction Set Extensions for Multimedia, Graphics Processing Units, GPU Memory Hierarchy, Detecting and Enhancing Loop-Level Parallelism, CUDA Programming, Case Study: Nvidia Maxwell. **(10 Hours)**

Computer Architecture Lab:

The lab experiments (assignments) will be designed to assist the theory covered in the class.

3. Text Books:

1. J.L. Hennessy and D.A. Patterson. Computer Architecture: A Quantitative Approach. 5 th Edition, Morgan Kauffmann Publishers, 2012.

4. References:

1. J.P. Shen and M.H. Lipasti. Modern Processor Design: Fundamentals of Superscalar Processors. McGraw-Hill Publishers, 2005.
2. D.B. Kirk and W.W. Hwu. Programming Massively Parallel Processors. 2 nd Edition, Morgan Kauffmann Publishers, 2012.
3. Pin – A Dynamic Binary Instrumentation Tool. <https://software.intel.com/en-us/articles/pin-a-dynamic-binary-instrumentation-tool>
4. Cachegrind: A Cache and Branch-Prediction Profiler. <http://valgrind.org/docs/manual/cg-manual.html>
5. OpenMP. www.openmp.org
6. CUDA. <https://developer.nvidia.com/cuda-zone>

5. Similarity Content Declaration with Existing Courses: Nil

6. Justification for new course proposal if cumulative similarity content is > 30%:N/A