Contents

Chapter 1

Micro-Architecture

Basic blocks of a Micro-Architecture:

Cache A high-speed unit to keep code and data.

- **ICache** holds instructions.
- **DCache** holde data.

IFU A unit to fetch instruction from cache.

IDU A unit to decode instruction after fetch.'

EU A unit to execute instruction:

- **ALU** for arithmetic and logic.
- **Branch Unit** for branching instruction.
- **L/S Unit** for loading and saving instruction

Register File A unit to save temporary results.

Program Counter A unit to locate next instruction.

Control Unit A unit to schedule all data movement.

1.1 Standard Benchmarks

1.1.1 Performance Metrics

Latency is the time between start and finish of a single task. The number of taks finished in a give unit of time **Throughput**. **Response time** is the total time to complete a task, also called. Response time consistes of

1. CPU time
   (a) User CPU time
   (b) System CPU time: time spent in OS doing tasks on behalf of a program.
2. I/O time

Then **System Performance** is the inverse time elapsed and **CPU performance** is the inverse user CPU time. From user perespective response time is more important and from system admin perespective throughput is more important.