Concurrent and Distributed Devices (CDD101)

Dr. Joseph Kehoe

Performance Theory

- **Latency**
  - Total time it takes to compute a single result
  - Measured in units of time
- **Throughput**
  - The rate at which a series of results can be computed
  - Measured in units of work per unit time
- **Power Consumption**
  - The amount of power required to performa computation
  - Measured in Work per power unit

Speedup

- **Compares latency in solving an indentical problem on one hardware unit versus P hardware units**
  - Sp=T1/Tp
- **Absolute Speedup**
  - Sequential Algorithm Time used for T1
- **Relative Speedup**
  - Parallel algorithm used for T1

Efficiency

- **Speedup divided bu number of workers**
  - Sp/P
  - T1/P*Tp
- **Ideal efficiency is 1**
  - 100% efficiency

Superlinear Speedup

- **Parallel Algorithm makes better use of cache**
- **Parallel Algorithm is better**
- **Parallel Algorithm with multiple threads uses cache better than single parallel algorithm execution**

Amdahl’s Law

- **Work is of two types**
  - Serial Work
    - Cannot be parallelised
  - Parallel Work
    - Can be done in parallel
- **$T1 = Work_{serial} + Work_{parallel}$**
- **$Tp >= Work_{serial} + Work_{parallel}/P$**

Amdahl’s Law

- **If f is the fraction of the total work that is serial then (1-f) is the fraction that is parallel**
  - $W_{serial}=fT1$
  - $W_{parallel}=(1-f)T1$
- **Sp<= 1/ [f+(1-f)/p]**
- **As P -> infinity**
  - $Seedup_{inf}<= 1/f$