INTRODUCTION TO PARALLEL COMPUTING

Dr. Emmanuel S. Pilli

Motivating Parallelism

- The role of parallelism in accelerating computing speeds has been recognized for several decades.
- Its role in providing multiplicity of datapaths and increased access to storage elements has been significant in commercial applications.
- The scalable performance and lower cost of parallel platforms is reflected in the wide variety of applications.

Motivating Parallelism

- Developing parallel hardware and software has traditionally been time and effort intensive.
- There are some unmistakable trends in hardware design, which indicate that uniprocessor (or implicitly parallel) architectures may not be able to sustain the rate of *realizable* performance increments in the future.
- This is the result of a number of fundamental physical and computational limitations.
- The emergence of standardized parallel programming environments, libraries, and hardware have significantly reduced time to (parallel) solution.

The Computational Power Argument

Moore's law states [1965]:

- *“The complexity for minimum component costs has increased at a rate of roughly a factor of two per year. Certainly over the short term this rate can be expected to continue, if not to increase.*
- *Over the longer term, the rate of increase is a bit more uncertain, although there is no reason to believe it will not remain nearly constant for at least 10 years.*
- *That means by 1975, the number of components per integrated circuit for minimum cost will be 65,000."*

The Computational Power Argument

- Moore attributed this doubling rate to exponential behavior of die sizes, finer minimum dimensions, and “circuit and device cleverness".
- In 1975, he revised this law as follows:
- “There is no room left to squeeze anything out by being clever. Going forward from here we have to depend on the two size factors - bigger dies and finer dimensions."
- He revised his rate of circuit complexity doubling to 18 months and projected from 1975 onwards at this reduced rate.

The Computational Power Argument

- If one is to buy into Moore's law, the question still remains - how does one translate transistors into useful OPS (operations per second)?
- The logical recourse is to rely on parallelism, both implicit and explicit.
- Most serial (or seemingly serial) processors rely extensively on implicit parallelism.
- We will now focus on explicit parallelism.

The Memory/Disk Speed Argument

- While clock rates of high-end processors have increased at roughly 40% per year over the past decade, DRAM access times have only improved at the rate of roughly 10% per year over this interval.
- This mismatch in speeds causes significant performance bottlenecks.
- Parallel platforms provide increased bandwidth to the memory system.

The Memory/Disk Speed Argument

- Parallel platforms also provide higher aggregate caches.
- Principles of locality of data reference and bulk access, which guide parallel algorithm design also apply to memory optimization.
- Some of the fastest growing applications of parallel computing utilize not their raw computational speed, rather their ability to pump data to memory and disk faster.

The Data Communication Argument

- As the network evolves, the vision of the Internet as one large computing platform has emerged.
- In many other applications (typically databases and data mining) the volume of data is such that they cannot be moved.
- Any analyses on this data must be performed over the network using parallel techniques.

Scope of Parallel Computing Applns

- Parallelism finds applications in very diverse application domains for different motivating reasons.
- These range from improved application performance to cost considerations.

Applications in Engg and Design

- Design of airfoils (optimizing lift, drag, stability), internal combustion engines (optimizing charge distribution, burn), high-speed circuits (layouts for delays and capacitive and inductive effects), and structures (optimizing structural integrity, design parameters, cost, etc.).
- Design and simulation of micro- and nano-scale systems.
- Process optimization, operations research.

Scientific Applications

- Functional and structural characterization of genes and proteins.
- Advances in computational physics and chemistry have explored new materials, understanding of chemical pathways, and more efficient processes.
- Applications in astrophysics have explored the evolution of galaxies, thermonuclear processes, and the analysis of large datasets from telescopes.
- Weather modeling, mineral prospecting, flood prediction, etc., are other important applications.
- Bioinformatics and astrophysics also present some of the most challenging problems with respect to analyzing extremely large datasets.

Commercial Applications

- Some of the largest parallel computers power the wall street!
- Data mining and analysis for optimizing business and marketing decisions.
- Large scale servers (mail and web servers) are often implemented using parallel platforms.
- Applications such as information retrieval and search are typically powered by large clusters.

Applications in Computer Systems

- Network intrusion detection, cryptography, multiparty computations are some of the core users of parallel computing techniques.
- Embedded systems increasingly rely on distributed control algorithms.
- A modern automobile consists of tens of processors communicating to perform complex tasks for optimizing handling and performance.
- Conventional structured peer-to-peer networks impose overlay networks and utilize algorithms directly from parallel computing.