CLOUD COMPUTING APPLICATIONS

Summary

Roy Campbell & Reza Farivar

MapReduce Advantages/Disadvantages

- Now it's easy to program for many CPUs
  - Communication management effectively gone
    - I/O scheduling done for us
  - Fault tolerance, monitoring
    - Machine failures, suddenly slow machines, etc., are handled
  - Can be much easier to design and program
  - Can cascade several MapReduce tasks
- But … it further restricts solvable problems
  - Might be hard to express problem in MapReduce
  - Data parallelism is key
    - Need to be able to break up a problem by data chunks

MapReduce Conclusions

- MapReduce has proven to be a useful abstraction

- Greatly simplifies large-scale computations

- Functional programming paradigm can be applied to large-scale applications

- Fun to use: focus on problem, let the middleware deal with messy details