Compilers, Parallel, Grid, and Cloud Computing

Dr. Clayton S. Ferner
Professor
Department of Computer Science
University of North Carolina Wilmington

 © copyright 2011, Clayton S. Ferner, UNC Wilmington

Outline

- Compilers
- Parallel Computing
- Automatic Parallelizing Compilers
- Grid Computing
- Cloud Computing

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Compilers

- Computers understand Machine Code (instructions in binary)

```
0000000 457f 464c 0102 0001 0000 0000 0000 0000
0000010 0002 003e 0001 0000 0440 0040 0000 0000
0000020 0040 0000 0000 0000 0da8 0000 0000 0000
0000030 0000 0000 0040 0038 0008 0040 001d 001a
0000040 0006 0000 0005 0000 0040 0000 0000 0000
0000050 0040 0040 0000 0000 0040 0040 0000 0000
```

- Humans can understand Machine Code, but is it extremely difficult to work with

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Compilers

- Humans understand natural languages (English, French, German, etc.)
- Computers do not understand natural languages
- Natural languages are too ambiguous to be useful to program a computer:
  - “Add a pinch of salt.”
  - “I'll meet you at 4ish.”
  - “How about a cup of coffee?”
  - “What's the matter?”

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Compilers as translators

- A high-level language is one that humans can easily work with but is unambiguous

```
int i;
for (i = 0; i < 10; i++)
   array[i] = 2*i;
if (a < b)
   a = b;
return a;
```

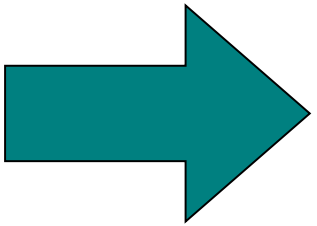
- Compilers are essentially translators

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Compilers as translators

- A compiler is a software program that translate a high-level programming language to assembly language

```
if (a < b)
   a = b;
```



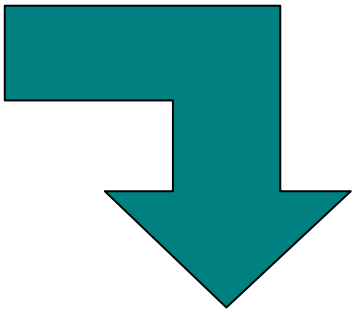
```
movl    %esi, -8(%rbp)
movl    -4(%rbp), %eax
cmpl    -8(%rbp), %eax
jge     .L2
.L2:
movl    -8(%rbp), %eax
movl    %eax, -12(%rbp)
```


© copyright 2011, Clayton S. Ferner, UNC Wilmington

Compilers as translators

- Assembler is software that translates a program in assembly language to machine code

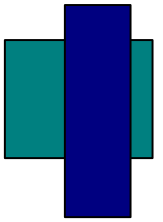
```
movl    %esi, -8(%rbp)
movl    -4(%rbp), %eax
cmpl    -8(%rbp), %eax
jge     .L2
.L2:
movl    -8(%rbp), %eax
movl    %eax, -12(%rbp)
```



```
0000000 457f 464c 0102 0001 0000 0000 0000 0000
0000010 0002 003e 0001 0000 0440 0040 0000 0000
0000020 0040 0000 0000 0000 0da8 0000 0000 0000
0000030 0000 0000 0040 0038 0008 0040 001d 001a
0000040 0006 0000 0005 0000 0040 0000 0000 0000
```


© copyright 2011, Clayton S. Ferner, UNC Wilmington

Courses related to compiler technology



- CSC 360 Formal Languages and Computability
- CSC 434 Programming Languages
- CSC 457 Compiler Construction

© copyright 2011, Clayton S. Ferner, UNC Wilmington

What is parallel computing?

- Parallel computing is the use of multiple processors to solve a single problem or to work on a single program
- The use of multiple processors to work in separate independent programs is not parallel computer

© copyright 2011, Clayton S. Ferner, UNC Wilmington

What is parallel computing?

- Human beings are very good at parallel work:
  - Cutting grass
  - Building a house
- The more processors there are to work on a problem the faster it gets done
- Linear Speedup – if you have a program that takes *T* seconds to complete on one processor, with *N* processors dividing up the work, it could take as little as *T*/*N* seconds

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Parallel computing is a challenge

- Hindrances to real speedup - synchronization and communication
- The processors must be programmed correctly to perform their share of the work and to synchronize and communication properly
- Writing a correct program to run on multiple processors is much more challenging that writing one to run on a single processor

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Parallel computing is important

- We've pushed the limits of processor speeds
  - Speed of light
  - Heat
- In order to increase the speed of computers they need to use multiple processors

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Parallel computing is important

- Many machines nowadays have dual- and quad-core processors
- Intel will soon be making machines with hundreds of processors
- What do we do with all those processors?

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Courses Related to Parallel Computing

- CSC 337 Parallel Computing
- CSC 437/537 Parallel Computing II
- CSC 446/546 Grid Computing

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Automatic Parallelizing Compilers

- Wouldn't it be great if you could just write a program to run on one processor, then have a compiler figure out how to make it work correctly for multiple processors?
- Automatic Parallelizing Compilers attempt to do just that

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Paraguin Compiler

- The Paraguin Compiler is a compiler I built using the SUIF Compiler from Stanford Univ.
  - (http://people.uncw.edu/cferner/Paraguin/)
- Automatic parallelizing compiler
- Produces MPI code for execution on Distributed Memory systems

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Grid Computing

- Grid Computing - Using geographically distributed and interconnected computers together for computing and for resource sharing.
- Usually, grid computing involves teams working together on a common goal, sharing computing resources and possibly experimental equipment.
- The geographically distributed grid computing team and their resources is called a virtual organization.

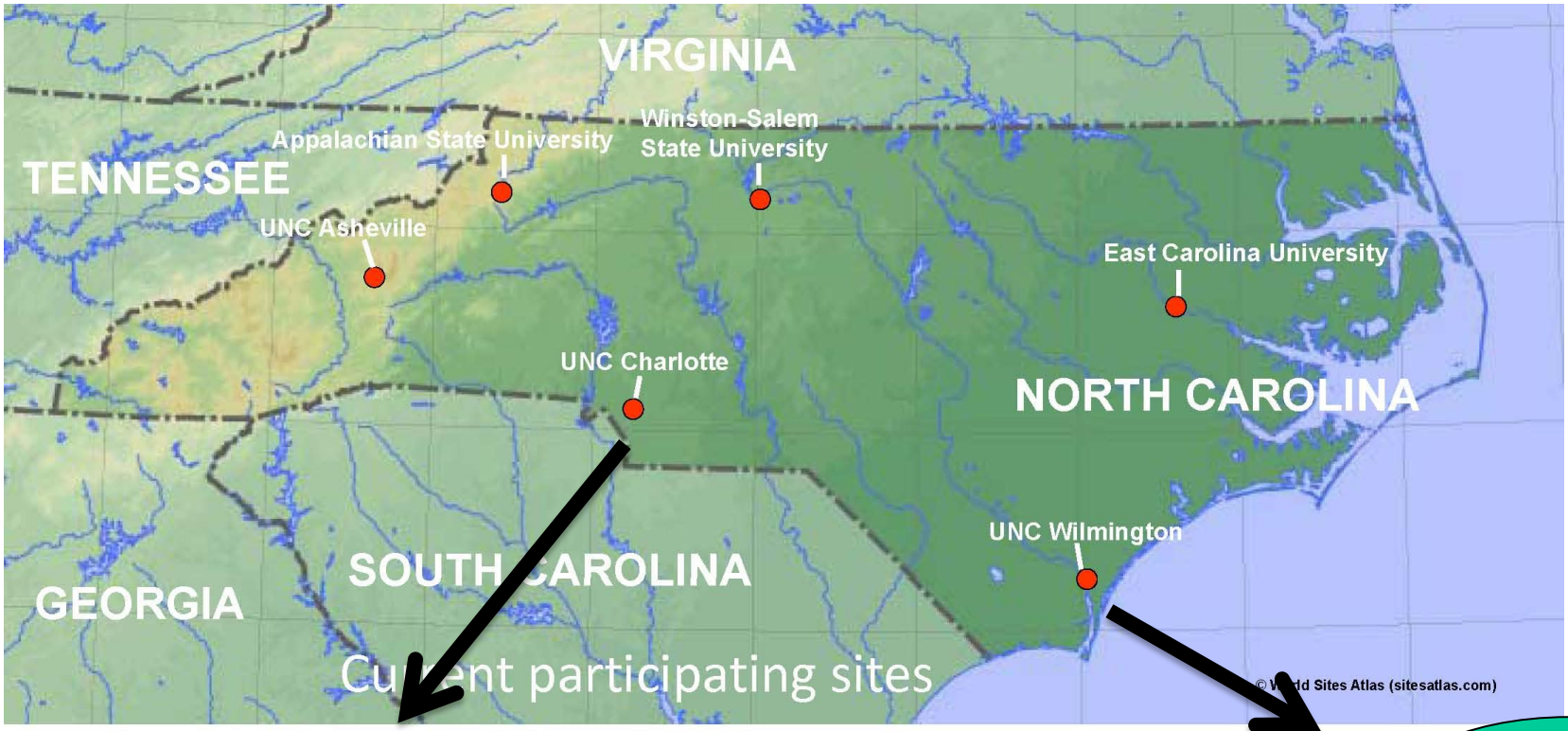
© copyright 2011, Clayton S. Ferner, UNC Wilmington

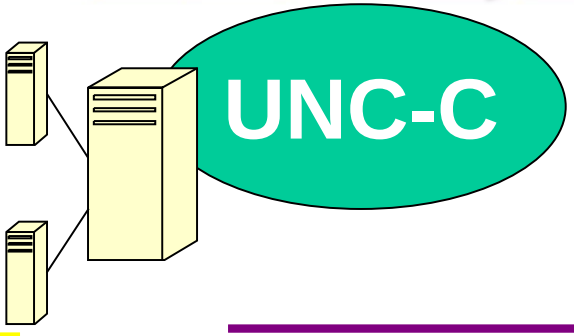
Courses Related to Grid and Cloud Computing

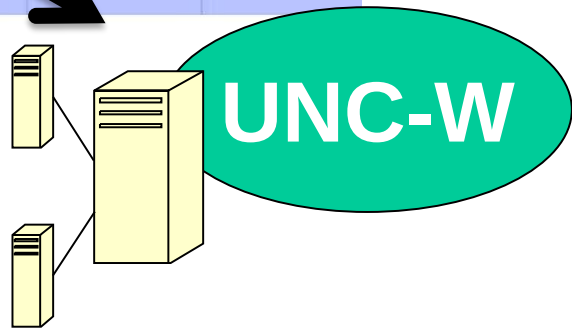
- CSC 446/546 Grid Computing
  - Appalachian State University
  - East Carolina University
  - University of North Carolina, Ashville
  - University of North Carolina, Charlotte
  - University of North Carolina, Wilmington
  - Winston-Salem State University

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Grid Computing Class this Semester



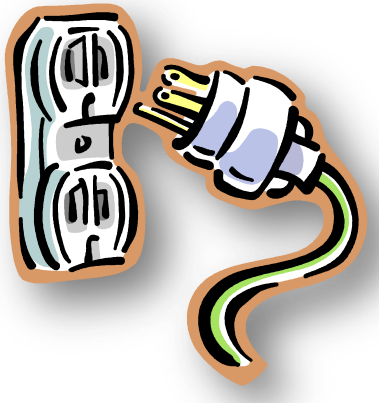




© copyright 2011, Clayton S. Ferner, UNC Wilmington

Cloud Computing

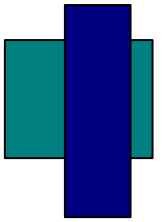
- Cloud computing provides computation, software, data access, and storage services that do not require end-user knowledge of the physical location and configuration of the system that delivers the services.*
- The basic idea is the “rent” computing resources instead of owning/maintaining
- Both models of computing have been compared to using a power grid



* Wikipedia. http://en.wikpedia.org/wii/Cloud_computing

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Cloud Computing versus Grid Computing



- Grid Computing is about resource sharing - resources are owned by the participants
- Grid Computing creates the abstraction of a “virtual organization”
- Cloud computing resources are owned by some organization (Google, Amazon, Microsoft, etc.)
- Cloud computing uses virtualization

© copyright 2011, Clayton S. Ferner, UNC Wilmington


My Experience
Using Amazon's
EC2 Cloud

Need

- My children's school has an annual auction to raise money
- Auction software is not cheap, especially for non-profit organizations
- I developed a set up PHP pages and a corresponding mysql database to server their needs

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Need (continued)

- Server is one of my laptops at home, open through firewall via port 80.
- During the actual auction an take the laptop and router to the auction and setup a Intranet.
- This spring, United Way of Roanoke paid me to use my software for their auction.
- I was nervous about the performance of my laptop with Internet connectivity through RoadRunner.
- What if my aging laptop fails?

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Need (continued)

- What if my aging laptop fails? – It actually happened
- I had a backup of everything on my new laptop and switched servers.
- But that stopped working briefly when I couldn't boot the machine until I got through the fsck to check and repair the disk.
- How was I going to provide decent worry-free response during the auction in Roanoke, VA?

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Solution

- Solution – to rent a server on Amazon's EC2 cloud
- I setup up a Fedora Linux based machine
- I installed
  - Apache (html client)
  - Mysql database
- Copied over:
  - Database backup
  - PHP pages

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Solution

- My new laptop had identical PHP pages and database to serve as a backup
- I setup an automatic backup every minute during the auction
- Backups were automatically copied to my laptop (which served as a mirror) within a minute
- All changes made during the auction were updated to my server within 2 minutes

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Solution

- I had two machines with identical interface and identical (within two minutes) data
- Each machine was in a different location on different networks:
  - My laptop at home in Wilmington via RoadRunner (174.106.1.201)
  - Cloud machine someplace in Virginia (50.16.64.61)
- Any problems with one machine and I could have the auction volunteers switch servers by simply switch IP addresses.

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Cost

- While I had the machine up to learn, install software, experiment, the cost was approximately $2
- I had the server up for several days before, during, and after the auction (~ week).
- Total cost ~$12

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Conclusion

- Amazon's cloud gave me the resources to run this auction with:
  - Reliability
  - Low cost
  - Decent bandwidth
  - Full access (root) to the virtual machine
  - Peace of mind

© copyright 2011, Clayton S. Ferner, UNC Wilmington

Discussion

- Questions?

http://people.uncw.edu/cferner/papers/CSC10011F.pdf

© copyright 2011, Clayton S. Ferner, UNC Wilmington