UNIX, C, AND DATA STRUCTURES

Lecture 3:
More Unix
What is Data?
Beginning on Memory
Taylor...
Reading

• None for now
What should you be doing?

• Due today: 2A
• Assigned today: 1B
Office Hours: Weeks 2-10

- Monday: 1030-1130 (HC), 5-6 (ZH)
- Tuesday: 1130-1230 (PK), 5-6 (VL)
- Wednes: 2-3 (ZH), 5-6 (VL)
- Thursda: 1230-130 (HC), 5-6 (VL)
- Friday : 1230-130 (PK), 4-5 (VL)
Plan for today

• Quick review of hello world in C
• Permissions
• What is Data?
• Beginning on Memory in C
Plan for today

- Quick review of hello world in C
- Permissions
- What is Data?
- Beginning on Memory in C
Hello World Example

```c
#include <stdio.h>
int main()
{
    printf("Hello world\n");
}
```

Hanks-iMac:Downloads hank$ gcc hello.c
Hanks-iMac:Downloads hank$ ./a.out
Hello world
Notes on previous example

• Ignore for now: #include <stdio.h>
• ”main” is the first function called in C.
  – (Sort of. There are other things called before main to set up the program. But you won’t be mucking with those.)
• gcc is a C compiler. It is short for GNU Compiler Collection. GNU is an open source effort.
• The output is called “a.out”
• We have to add “./”, as it “./a.out”
  – “.” is the directory the shell is in. Therefore “./a.out” says run the program called a.out in the current directory.
More C: variable types

• When you declare variables, you have to declare their type.
  – We learned one type: int (integer)
  – There are more types (floating point, etc). We will discuss them later today.
More C: functions

• As you would expect, you can define your own functions and call them.

```
#include <stdio.h>
int foo1() { return 3; }
int foo2() { return 4; }
int main()
{
    printf("%d\n", foo1() * foo2());
}
```

```
Hanks-iMac:Downloads hank$ gcc function.c
Hanks-iMac:Downloads hank$ ./a.out
12
```
Iteration is done with constructs “for” and “while”

```
#include <stdio.h>

int main()
{
    int X = 0;
    while (X < 83)
        X += 5;
    printf("X is %d\n", X);
}
```

```
Hanks-iMac:Downloads hank$ gcc while.c
Hanks-iMac:Downloads hank$ ./a.out
X is 85
```
For loop in practice
(with increment operator)

Hanks-iMac:Downloads hank$ cat sum.c
#include <stdio.h>
int main()
{
    int sum = 0;
    int i;
    for (i = 0 ; i < 10 ; i++)
    {
        sum = sum+i;
    }
    printf("Sum is %d\n", sum);
}

Hanks-iMac:Downloads hank$ gcc sum.c
Hanks-iMac:Downloads hank$ ./a.out
Sum is 45
Homework #2A: find primes between 5 and 100

Hanks-iMac:Downloads hank$ gcc prime_finder.c
Hanks-iMac:Downloads hank$ ./a.out
5 is a prime number
7 is a prime number
11 is a prime number
13 is a prime number
17 is a prime number
19 is a prime number
23 is a prime number
29 is a prime number
31 is a prime number
37 is a prime number
41 is a prime number
43 is a prime number
47 is a prime number
53 is a prime number
59 is a prime number
61 is a prime number
67 is a prime number
71 is a prime number
73 is a prime number
79 is a prime number
83 is a prime number
89 is a prime number
97 is a prime number
CIS 212: Project #2A
Assigned: September 27th, 2018
Due: October 2nd, 2018
(which means submitted by 6am on October 3rd, 2018)
Worth 3% of your grade

Assignment:
1) Write a C program that finds the prime numbers between 5 and 100.
2) Your program should have the exact same output as mine.
3) You can confirm this:
   a. Download my “correct_output” file
   b. Run your program as “./a.out > my_output”
   c. Run the diff program to difference the two:
      i. diff /path/to/correct_output /path/to/my_output
      ii. And put the proper paths in the place of /path/to
   d. If diff gives no output, they are identical.

This project will be graded by:
1) Running the diff program against your output (as per above)
2) Inspection of your code

If the diff program shows any difference, you will get less than half credit.

What should you upload?: Just a single file, which is your C source code.
Plan for today

• Quick review of hello world in C
• Permissions
• What is Data?
• Beginning on Memory in C
Files

• Unix maintains a file system
  – File system controls how data is stored and retrieved

• Primary abstractions:
  – Directories
  – Files

• Files are contained within directories
Directories are hierarchical

• Directories can be placed within other directories

• “/” -- The root directory
  – Note “/”, where Windows uses “\”

• “/dir1/dir2/file1”
  – What does this mean?

File file1 is contained in directory dir2, which is contained in directory dir1, which is in the root directory
Permissions: System Calls

• System calls: a request from a program to the OS to do something on its behalf
  – ... including accessing files and directories

• System calls:
  – Typically exposed through functions in C library
  – Unix utilities (cd, ls, touch) are programs that call these functions

Permissions in Unix are enforced via system calls.
Permissions: Unix Groups

• Groups are a mechanism for saying that a subset of Unix users are related
  – In 2014, we had a “330_S14” unix group on ix
  – Members:
    • Me
    • 2 GTFs

The commands for creating a group tend to vary, and are often done by a system administrator

CIS uses “groupctl”
Permissions

• Permissions are properties associated with files and directories
  – System calls have built-in checks to permissions
    • Only succeed if proper permissions are in place

• Three classes of permissions:
  – User: access for whoever owns the file
    • You can prevent yourself from accessing a file!
      – (But you can always change it back)
  – Group: allow a Unix group to access a file
  – Other: allow anyone on the system to access a file
Three types of permissions

• Read
• Write
• Execute (see next slide)
Executable files

• An executable file: a file that you can invoke from the command line
  – Scripts
  – Binary programs

• The concept of whether a file is executable is linked with file permissions
There are 9 file permission attributes

• Can user read?
• Can user write?
• Can user execute?
• Can group read?
• Can group write?
• Can group execute?
• Can other read?
• Can other write?
• Can other execute?

User = “owner”
Other = “not owner, not group”

A bunch of bits ... we could represent this with binary
Translating R/W/E permissions to binary

<table>
<thead>
<tr>
<th>#</th>
<th>Permission</th>
<th>rwx</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>full</td>
<td>111</td>
</tr>
<tr>
<td>6</td>
<td>read and write</td>
<td>110</td>
</tr>
<tr>
<td>5</td>
<td>read and execute</td>
<td>101</td>
</tr>
<tr>
<td>4</td>
<td>read only</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>write and execute</td>
<td>011</td>
</tr>
<tr>
<td>2</td>
<td>write only</td>
<td>010</td>
</tr>
<tr>
<td>1</td>
<td>execute only</td>
<td>001</td>
</tr>
<tr>
<td>0</td>
<td>none</td>
<td>000</td>
</tr>
</tbody>
</table>

Which of these modes make sense? Which don’t?

We can have separate values (0-7) for user, group, and other.
Unix command: chmod

• chmod: change file mode

• chmod 750 <filename>
  – User gets 7 (rwx)
  – Group gets 5 (rx)
  – Other gets 0 (no access)

Lots of options to chmod
(usage shown here is most common)
Manpage for chmod

• “man chmod”

CHMOD(1) BSD General Commands Manual CHMOD(1)

NAME
chmod -- change file modes or Access Control Lists

SYNOPSIS
chmod [-fv] [-R [-H | -L | -P]] mode file ...
chmod [-fv] [-R [-H | -L | -P]] [-a | +a | =a] ACE file ...
chmod [-fhv] [-R [-H | -L | -P]] [-E] file ...
chmod [-fhv] [-R [-H | -L | -P]] [-C] file ...
chmod [-fhv] [-R [-H | -L | -P]] [-N] file ...

DESCRIPTION
The chmod utility modifies the file mode bits of the listed files as specified by the mode operand. It may also be used to modify the Access Control Lists (ACLs) associated with the listed files.

The generic options are as follows:

-f Do not display a diagnostic message if chmod could not modify the mode for file.
Unix commands for groups

• **chgrp**: changes the group for a file or directory
  – `chgrp <group> <filename>`

• **groups**: lists groups you are in
```bash
C02LN00GFD58:~ hank$ mkdir CIS330
C02LN00GFD58:~ hank$ cd CIS330
C02LN00GFD58:CIS330 hank$ touch a
C02LN00GFD58:CIS330 hank$ ls -l
```
```
total 0
-rw-r--r-- 1 hank staff 0 Apr 3 08:14 a
```

How to interpret this?
Permissions and Directories

• You can only enter a directory if you have “execute” permissions to the directory.

• Quiz: a directory has permissions “400”. What can you do with this directory?

   Answer: it depends on what permissions a system call requires.
Directories with read, but no execute

Last login: Thu Apr  3 08:14:33 on ttys007
C02LN00GFD58:~ hank$ mkdir CIS330
C02LN00GFD58:~ hank$ touch CIS330/a
C02LN00GFD58:~ hank$ chmod 400 CIS330
C02LN00GFD58:~ hank$ ls CIS330
a
C02LN00GFD58:~ hank$ cd CIS330
-bash: cd: CIS330: Permission denied
C02LN00GFD58:~ hank$ cat CIS330/a
cat: CIS330/a: Permission denied
Unix scripts

• Scripts
  – Use an editor (vi/emacs/other) to create a file that contains a bunch of Unix commands
  – Give the file execute permissions
  – Run it like you would any program!!
Unix scripts

• Arguments
  – Assume you have a script named “myscript”
  – If you invoke it as “myscript foo bar”
  – Then
    • $# == 2
    • $1 == foo
    • $2 == bar
Project 1B

• Summary: write a script that will create a specific directory structure, with files in the directories, and specific permissions.
Project 1B

Assignment: Create a shell script that will create a directory structure and files within that directory structure, all with the specified file permissions. The script should be named “proj1b.sh”. (A consistent name will help with grading.)

Note: you are only allowed to use the following commands: mkdir, touch, cd, chmod, mv, cp, rm, rmdir. (You do not need to use all of these commands to successfully complete the assignment.)
Project 1B

The directory structure should be:

```
<table>
<thead>
<tr>
<th>Root dir</th>
</tr>
</thead>
</table>
| Dir1  
  Permissions: 770 |
|   File1  
  Permissions: 400 |
|   Dir3  
  Permissions: 000 |
|   Dir4  
  Permissions: 750 |
|   File4  
  Permissions: 666 |
| Dir2  
  Permissions: 775 |
|   File2  
  Permissions: 640 |
```

Key:
- **Files:** Name of file Permissions
- **Directories:** Name of directory Permissions

The script should take an argument, and the argument should be the location to create the directory. So, if the script is run as `proj1b/tmp`, then it would create directories `tmp/Dir1`, `tmp/Dir2`, `tmp/Dir1/File1`, `tmp/Dir2/File2`, etc.
Plan for today

• Quick review of hello world in C
• Permissions
• What is Data?
• Beginning on Memory in C
To a Computer, Data Means Lots of Bits. So What is a Bit?

• Short for “binary digit”
• 2 states:
  – on (1)
  – off (0)
• Like a light switch
• Computers can turn their bits (light switches) on and off
• Computers implement this using electricity in a “capacitor”:
  – has electrical charge: on / 1
  – no electrical charge: off / 0
What is a Byte?

• A group of 8 bits

• We could write the values of the 8 bits in a row.
• Example: 01000001
  – This is a binary number
Binary Numbers

• Normal system (called decimal):
  – Values for any digit are: 0,1,2,3,4,5,6,7,8,9
  – What we use every day

• Binary:
  – Values for any digit are: 0,1
  – What computers use
Conventions Surrounding Binary Numbers

- 01000001 is 65 in decimal
- By convention, 65 represents ‘A’
  
  — Called “ASCII”

<table>
<thead>
<tr>
<th>Dec</th>
<th>Hx</th>
<th>Oct</th>
<th>Char</th>
<th>Dec</th>
<th>Hx</th>
<th>Oct</th>
<th>Char</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NULL</td>
<td>32</td>
<td>20</td>
<td>0</td>
<td>Space</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>SOH</td>
<td>33</td>
<td>21</td>
<td>01</td>
<td>!</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>STX</td>
<td>34</td>
<td>22</td>
<td>02</td>
<td>&quot;</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0</td>
<td>ETX</td>
<td>35</td>
<td>23</td>
<td>03</td>
<td>#</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0</td>
<td>EOT</td>
<td>36</td>
<td>24</td>
<td>04</td>
<td>$</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0</td>
<td>ENQ</td>
<td>37</td>
<td>25</td>
<td>05</td>
<td>%</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0</td>
<td>ACK</td>
<td>38</td>
<td>26</td>
<td>06</td>
<td>&amp;</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>0</td>
<td>BEL</td>
<td>39</td>
<td>27</td>
<td>07</td>
<td>'</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>0</td>
<td>BS</td>
<td>40</td>
<td>28</td>
<td>08</td>
<td>(</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>0</td>
<td>HT</td>
<td>41</td>
<td>29</td>
<td>09</td>
<td>)</td>
</tr>
<tr>
<td>10</td>
<td>A</td>
<td>0</td>
<td>LF</td>
<td>42</td>
<td>2A</td>
<td>05</td>
<td>:</td>
</tr>
<tr>
<td>11</td>
<td>B</td>
<td>0</td>
<td>VT</td>
<td>43</td>
<td>2B</td>
<td>05</td>
<td>;</td>
</tr>
<tr>
<td>12</td>
<td>C</td>
<td>0</td>
<td>FF</td>
<td>44</td>
<td>2C</td>
<td>05</td>
<td>&lt;</td>
</tr>
<tr>
<td>13</td>
<td>D</td>
<td>0</td>
<td>CR</td>
<td>45</td>
<td>2D</td>
<td>05</td>
<td>=</td>
</tr>
<tr>
<td>14</td>
<td>E</td>
<td>0</td>
<td>SO</td>
<td>46</td>
<td>2E</td>
<td>05</td>
<td>&gt;</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>0</td>
<td>SI</td>
<td>47</td>
<td>2F</td>
<td>05</td>
<td>?</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>DLE</td>
<td>48</td>
<td>30</td>
<td>06</td>
<td>@</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>0</td>
<td>DC1</td>
<td>49</td>
<td>31</td>
<td>06</td>
<td>A</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>0</td>
<td>DC2</td>
<td>50</td>
<td>32</td>
<td>06</td>
<td>B</td>
</tr>
<tr>
<td>19</td>
<td>3</td>
<td>0</td>
<td>DC3</td>
<td>51</td>
<td>33</td>
<td>06</td>
<td>C</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>0</td>
<td>DC4</td>
<td>52</td>
<td>34</td>
<td>06</td>
<td>D</td>
</tr>
<tr>
<td>21</td>
<td>5</td>
<td>0</td>
<td>DC5</td>
<td>53</td>
<td>35</td>
<td>06</td>
<td>E</td>
</tr>
<tr>
<td>22</td>
<td>6</td>
<td>0</td>
<td>DC6</td>
<td>54</td>
<td>36</td>
<td>06</td>
<td>F</td>
</tr>
<tr>
<td>23</td>
<td>7</td>
<td>0</td>
<td>DC7</td>
<td>55</td>
<td>37</td>
<td>06</td>
<td>G</td>
</tr>
<tr>
<td>24</td>
<td>8</td>
<td>0</td>
<td>DC8</td>
<td>56</td>
<td>38</td>
<td>07</td>
<td>H</td>
</tr>
<tr>
<td>25</td>
<td>9</td>
<td>0</td>
<td>DC9</td>
<td>57</td>
<td>39</td>
<td>07</td>
<td>I</td>
</tr>
<tr>
<td>26</td>
<td>A</td>
<td>0</td>
<td>DCA</td>
<td>58</td>
<td>3A</td>
<td>07</td>
<td>J</td>
</tr>
<tr>
<td>27</td>
<td>B</td>
<td>0</td>
<td>DCB</td>
<td>59</td>
<td>3B</td>
<td>07</td>
<td>K</td>
</tr>
<tr>
<td>28</td>
<td>C</td>
<td>0</td>
<td>DCC</td>
<td>60</td>
<td>3C</td>
<td>07</td>
<td>L</td>
</tr>
<tr>
<td>29</td>
<td>D</td>
<td>0</td>
<td>DCD</td>
<td>61</td>
<td>3D</td>
<td>07</td>
<td>M</td>
</tr>
<tr>
<td>30</td>
<td>E</td>
<td>0</td>
<td>DCE</td>
<td>62</td>
<td>3E</td>
<td>07</td>
<td>N</td>
</tr>
<tr>
<td>31</td>
<td>F</td>
<td>0</td>
<td>DCF</td>
<td>63</td>
<td>3F</td>
<td>07</td>
<td>O</td>
</tr>
</tbody>
</table>

Also conventions for supporting decimal numbers (3.456)...

Source: www.LookupTables.com
• There are 10 types of people in the world – those that know binary and those that don’t
How do Bytes relate to computers?

• Computers have “memory”
  – Places to store data and then retrieve it later
• The memory is made up of many, many Bytes
Von Neumann Architecture

Input Device → Central Processing Unit → Arithmetic/Logic Unit → Memory Unit → Output Device
There is so much data that we need lots of prefixes...

• 1 KiloByte (KB) = 1024 Bytes
• 1 MegaByte (MB) = 1024 KB = ~1million Bytes
• 1 GigaByte (GB) = 1024 MB = ~1billion Bytes
• 1 TeraByte (TB) = 1024 GB = ~1trillion Bytes
• 1 PetaBytes (PB) = 1024 TB = ~1quadrillion Bytes
• 1 ExaByte = 1024 PB = ~1 quintillion Bytes
  = 1 billion billion Bytes
Analogy

• 1 story hotel
• 6 rooms, side by side
  – Room 1, Room 2, Room 3, Room 4, Room 5, Room 6
• Light left on == 1
• Light left off == 0
What is the Data?

1 0 0 0 1 0
What is the Data?
Mental Model For Memory

• 1 story hotel
• First room is 10000000
• Next room is 10000001
• And so on...
Reserving Memory

• You go to the hotel front desk and ask for a room
• When you are done, you check out and tell them you have left the room
Plan for today

• Quick review of hello world in C
• Permissions
• What is Data?
• Beginning on Memory in C
Why C?

• You can control the memory
• That helps get good performance

• If you don’t control the memory (like in other programming languages), you are likely to get poor performance

• ... so let’s talk about memory
Automatic vs Dynamic Memory

• In C, there are two types of variables
  – Actually three, but ignoring “static” for now
• Automatic: taken care of for you automatically
• Dynamic: you manage it

• So, of course, everyone wants automatic, right?
Minimum Reservation

• When reserving memory, the minimum request is 1 byte
  – You can access individual bits, but the hotel front desk will only let you reserve 1 byte at a time
  – So, minimum reservation is a block of 8 rooms
Types in C

• char: ASCII character, 1 byte
  – Values: -128 to 127

• unsigned char: 1 byte
  – Values: 0 to 255

Both char and unsigned char treat this as 127

Unsigned char: 128
Char: -128 (“two’s complement”)
Hotel Rooms

• You must reserve rooms in blocks of 8 – 8, 16, 24, 32, ...

• The data is just data

• But there are conventions on how to interpret it
Known types in C

- char – 1 byte
- unsigned char – 1 byte
- int – 4 bytes
- float – 4 bytes
- double – 8 bytes
- more...
unsigned int

- 4 bytes -> 32 bits
- Values: 0 to $2^{32}-1$
int

- 4 bytes -> 32 bits
- Values: \(-2^{31}\) to \(2^{31}-1\)
• Essentially like scientific notation

• -534563E-6

• Store 4 things:
  – Sign of number (1 bit)
  – “Mantissa” (534563)
  – Sign of exponent (1 bit)
  – Value of exponent (6)

• Inexact!!
Requesting automatic memory

• This is done “automatically”...

```c
#include <stdio.h>

int main()
{
    int A;
    A = 4;
    printf("Value of A is %d\n", A);
}
```
Arrays

```c
#include <stdio.h>

int main() {
    int A[21];
    int i;
    for (i = 0; i < 21; i++)
        A[i] = 3;
    printf("Value of A[10] is %d\n", A[10]);
}
```

This is just 84 (4*21) bytes. We interpret these 84 bytes as 21 integers, following known conventions.
Important Memory Concepts in C (1/9): Automatic Vs Dynamic

• You can allocate variables that only live for the invocation of your function
  – Called automatic/stack variables (will talk more about this later)

• You can allocated variables that live for the whole program (or until you delete them)
  – Called dynamic/heap variables (will talk more about this later as well)
Important Memory Concepts in C (2/9): Pointers

• Pointer: points to memory location
  – Denoted with ‘*’
  – Example: “int *p”
    • pointer to an integer
  – You need pointers to get to heap memory

• Address of: gets the address of memory
  – Operator: ‘&’
  – Example:
    ```c
    int x;
    int *y = &x;
    ```
Important Memory Concepts in C (3/9): Memory allocation

• Special built-in function to allocate memory from heap: `malloc`
  – Interacts with Operating System
  – Argument for malloc is how many bytes you want

• Also built-in function to deallocate memory: `free`
free/malloc example

Enables compiler to see functions that aren’t in this file. More on this next week.

```c
#include <stdlib.h>
int main()
{
    /* allocates memory */
    int *ptr = malloc(2*sizeof(int));

    /* deallocates memory */
    free(ptr);
}
```

sizeof is a built in function in C. It returns the number of bytes for a type (4 bytes for int).

don’t have to say how many bytes to free ... the OS knows
Bonus Slides
How a Computer is **Not** Like Monopoly

- **Actions during cycle/turn:**
  - Monopoly: roll, buy, build, trade
  - Computer: other

- **Who’s turn?**
  - Monopoly: passes between players
  - Computer: always the computer’s turn

- **Time spent per cycle/turn:**
  - Monopoly: variable
  - Computer: fixed

- **Duration of turn:**
  - Monopoly: sometimes over a minute
  - Computer: much faster!!!
“First” computer: ENIAC

- Year: 1946
- Location: Pennsylvania
- Purpose: military
- Cost: $487K
  - ($6.9M today)
- Technology:
  - very different than today
  - ... but still the same
Vacuum Tubes

• Vacuum tubes:
  – Glass tubes with no gas
  – Used to control electron flow in early computers

• Occasionally, a bug would get stuck in the tube and cause the program to malfunction

• We no longer have vacuum tubes, but the term bug has remained with us...
An ENIAC Computation

• Used for military calculations:
  – A-bomb design
  – Missile delivery

• ENIAC could do ~5000 calculations in one minute

• In one case:
  – ENIAC did a calculation in 30 seconds
  – Human being took 20 hours
  – 2400x increase in speed

source: wikipedia
Hertz (Hz) = unit of measurement for how fast you do something

- 1 Hertz = do something once per second
- KHz = 1024 Hz
- MHz = 1024 KHz
- GHz = 1024 MHz

- The ENIAC machine ran at 5000Hz, or about 5KHz.
  - Vocab term: “clock speed” → the number of cycles per second
    - the clock speed of the ENIAC was 5KHz
Today’s Desktop Computers Are Fast!

☐ Most computers run at ~1-3 GHz

☐ i.e., operates billions of instructions each second

☐ This is about one million times faster than the ENIAC

☐ … and the ENIAC was 2400X faster than humans

☐ (at least at tasks computers are good at)
What does a million-fold increase mean?

Distance: a 2” map of Oregon is 1:1,000,000 scale

Time: 1 second to 277 hours is 1:1,000,000 scale

Time: 1 minute to 694 days is 1:1,000,000 scale

Time: 1 hour to 114 years is 1:1,000,000 scale

Time: 1 day to 2738 years is 1:1,000,000 scale
1 million-fold increase!
How does this happen?

• Moore’s Law (old timer’s version)
  – Clock speed doubles every 18 months

• Moore’s Law (newer version but still for old timers)
  – Clock speed doubles every 24 months
Moore’s Law

• Moore’s Law (actual version)
  – Number of transistors doubles every 24 months
  – And clock speed is a reflection of number of transistors

• So what is a transistor?
  – Semiconductor device for amplifying or switching electronic signals/power
  – Fundamental building block of modern electronics
  – Replacement for vacuum tube
Moore's Law – The number of transistors on integrated circuit chips (1971-2016)

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.

The data visualization is available at OurWorldInData.org. There you find more visualizations and research on this topic. Licensed under CC-BY-SA by the author Max Roser.
But actually...

Source: maximumpc.com
The reason is power

• Desktop computer takes ~200W
  – There are multiple components that consume the power:
    • CPU
    • Monitor
    • Disk
    • Memory

• 200W * 1 year → ~$70
Relationship Between Power and Clock Speed

• Clock goes twice as fast → Power goes up by factor of 8
  – (Increase of X in clock speed → Increase of $X^3$ in power)

• Clock speeds haven’t changed in 12 years

• What if they had doubled every 2 years?
  • Then 64X faster
    – → 262144X more power (for the CPU)
    – → power bill now $18M
New vocab term: "core"

Lightweight version of a CPU

What Changed?
• We are getting double the transistors every two years
• … but clock speed is the same
• … so what is changing?

<table>
<thead>
<tr>
<th></th>
<th>INTEGRATED</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORES</td>
<td>GHZ</td>
</tr>
<tr>
<td>7290</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>7250</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>7230</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>7210</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Available beginning in September
2 Plus 15W for integrated fabric
3 Pricing shown is for parts without integrated fabric. Add additional $278 for integrated fabric versions of these parts. Integrated fabric parts available in October.

Choose Your Optimization Point

Intel Xeon Phi® inside

M - an optional SKU is available with support for up to 1.5TB memory per CPU socket
F - an optional SKU is available with integrated 100Gbps Intel Omni-Path fabric
How To Use Multiple Cores?

• Answer: parallel programming
  – Write computer programs that use all the cores
  – Ideally the coordination between the cores is minimal
Parallel Programming Concepts

• Usual goal:
  – if program takes $N$ seconds to run with one core
  – then take $N/2$ seconds to run with two cores
  – and $N/M$ seconds to run with $M$ cores

Let’s consider an example outside of computers
Example: paint a house

• One person: 6 days (1 day = 10 hours)
• Two people: 3 days
• Three people: 2 days
• Six people: 1 day

• Sixty people: 1 hour?
• Six hundred people: 6 minutes?
Example: paint a house, plan #2:

- One person: paint one house in 6 days
- Two people: paint two houses in 6 days
- Three people: paint three houses in 6 days
- One thousand people: paint 1000 houses in 6 days?

Parallel programming is hard, and smart people spend their whole careers figuring out how to make parallel programs be efficient.
GPUs: Graphical Processing Units (graphics)

• Historical:
  – Introduced to accelerate graphics (gaming!)
  – Boom with desktop PCs
  – Mid-2000’s: people start hacking interface to program a GPU to make it do things besides graphics
  – Late 2000’s: GPU makers jump on board and start encouraging folks to program GPUs directly
• GPGPU: General-purpose GPU programming
  – Mid 2010’s: GPUs used for *lots* of computing problems.
  – Machine learning workhorse!

Emergent Tech

Bitcoin heist with a twist: This time it's servers that were stolen

Icelandic cops cuff 11 on suspicion of data centre robberies

By Simon Sharwood, APAC Editor 5 Mar 2018 at 04:57

Icelandic police have cuffed 11 people in connection with four raids on data centres that targeted cryptocurrency mining equipment.

The raids started in December 2017 when three data centres were cracked in December. Another raid took place in January. 600 servers went missing in the heists.

Icelandic police kept the raids secret while they pursued their investigations. Those efforts culminated in 11 arrests and an appearance before the Reykjaness District Court last Friday. Two of the 11 were detained and the matter held over for another day.

The 600 servers are all still missing, the Associated Press reports. Which is no surprise: x86 kit is pretty generic. The real prize inside a bitcoin-mining rig is either GPUs, RAM or nicely fast solid-state disks. Those components are all tiny compared to servers and could probably have been posted out of Iceland piecemeal without much hassle.

Iceland has become something of a hub for demanding workloads like cryptocurrency mining because cheap energy and low ambient temperatures make it a low-cost destination to run data centres and the kit inside them. The nation also has a low crime rate.
Why Are GPUs So Good?

Market summary > NVIDIA Corporation
NASDAQ: NVDA - Mar 5, 7:59 PM EST

235.65 USD ↓0.89 (0.38%)
After-hours: 236.50 ↑0.36%

1 day | 5 day | 1 month | 3 month | 1 year | 5 year | max

12.82 Mar 8, 2013

• NVIDIA: company that makes GPUs
• NVIDIA Volta: latest type of NVIDIA GPU
• Volta facts:
  – 5120 cores
  – 1200MHz clock speed
  – Can do 2000X more operations than my laptop
• Suggested MSRP: $2,999.00

This level of increase in computation is not just a quantitative change, it is a qualitative one too.