### **Beowulf Clusters**



**Robert Whitten Jr** 





MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

### What is beowulf?

- Mythical Old-English hero who defeats Grendel, the green dragon
- Movie starring a CG-animated Angelina Jolie
- Cluster architecture





### **Beowulf cluster history**

- 1993 Becker & Sterling started Beowulf project
- 1994 Becker & Sterling constructed a 16 node cluster
- 1996 DOE and NASA demonstrate beowulf cluster that surpasses 1GFlop/s
- 1997 Caltech demonstrates beowulf cluster that surpasses 10GFlop/s
- Present beowulf clusters are used at many universities, industries, and government sites





### What is a beowulf cluster?

- Collection of stand-alone computer networked together
- Network equipment to connect computers
- Software
  - Operating system software
  - Application software
  - Communication software (MPI, OpenMP)



### Hardware

- Any old computer will do
  - Laptops
  - Desktops
  - Playstation 3s ??
- Network
  - Switches / hubs
  - Cable
  - Ethernet adapters (most computer already have this, including the PS3)







**Network Topologies** 



Reference: <u>www.teach-ict.com/</u>, uva.ulb.ac.be/





### **Star Topology**



- Ethernet
- IEEE 802.3
- CSMA / CD
- Switch vs. Hub
- Broadcast vs. Cut-Through
- Cable: CAT V, UTP, RJ45





### **Model for Networking**

OSI Model – Developed by ISO



I International

S Standards

O Organization





### **OSI Model**

Layer	Name	Mnemonic
7	Application	<mark>≜</mark> II
6	<b>P</b> resentation	People
5	<u>S</u> ession	<u>S</u> eem
4	<u>T</u> ransport	<u>T</u> o
3	<u>N</u> etwork	Need
2	Data-Link	<b>D</b> ata
1	Physical	Processing





#### **Data Flow**

### The Seven Layers of OSI



Reference: http://catalyst.washington.edu/help/computing\_fundamentals/networking/img/osi\_model.jpg





### **The Model - Examples**

Layer	Description		
7. Application	HTTP, FTP, SMTP, POP3, HTTPS	6	
6. Presentation	Data Representation : ASCII		
5. Session	Starts and Stops the Session (Log	gon / Logoff)	
4. Transport	TCP – reliable / Handshake Transmission Control Protocol	UDP – unreliable (streaming data – mp3) User Datagram Protocol	
3. Network	IP – Internet Protocol, Router, Be	st Path Selection	
2. Data-Link	MAC (Media Access Control) Sub layer – Ethernet Switch	LLC (Logical Link Control) Sub layer	
1. Physical	Wire, Cable, The bits going across the wire, The NIC		











### **TCP/IP: Addressing Scheme**



IP = Internet Protocol Protocol is What? Set of Rules

- **1.Dotted Decimal Notation**
- 2.No Decimal Number can be >255
- 3.Must Have Subnet Mask
- 4.Broken Down into Classes based on High order of bits in the first octet.





### **Dotted Decimal Notation**

Example of IP address:

192.16.32.5 255.255.255.0 IP SM Dotted . Decimal . Notation . Rule



Remember – 4 Octets , a Decimal number that represents the value of an 8 bit Binary number





### No Number can be 255

Sample IP Address:

**192**.16.32.5 **255**.255.255.0

#### IP The first Octets = 192 & 255

SM What is 192 in Binary? What is 255 in Binary?

27	2 <sup>6</sup>	<b>2</b> <sup>5</sup>	24	<b>2</b> <sup>3</sup>	<b>2</b> <sup>2</sup>	<b>2</b> <sup>1</sup>	2 <sup>0</sup>	100
128 +	64 +	0 +	0 +	0 +	0 +	0 +	0 =	192
1	1	0	0	0	0	0	0	
27	<b>2</b> <sup>6</sup>	<b>2</b> <sup>5</sup>	<b>2</b> <sup>4</sup>	<b>2</b> <sup>3</sup>	<b>2</b> <sup>2</sup>	<b>2</b> <sup>1</sup>	<b>2</b> <sup>0</sup>	055
128 +	64 +	32 +	16 +	8 +	4 +	2 +	1 =	255
1	1	1	1	1	1	1	1	



### **IP Address Classes**

Class	1 <sup>st</sup> Octet	High Order Of Bits	Number Bits Used for Network	Default Subnet Mask
Class A	1-126	00000001 01111110	8	255.0.0.0
Class B	128 - 191	<b>10</b> 000000 <b>10</b> 111111	16	255.255.0.0
Class C	192 -223	<b>110</b> 00000 <b>110</b> 11111	24	255.255.255.0
Class D	224 - 239	<b>1110</b> 0000 <b>1110</b> 1111		Reserved







## SAMPLE IP Address: <u>Sample IP Address:</u> <u>192.16.32.5</u> 255.255.0 N H By Performing a Binary AND Function <u>IP Address AND Subnet Mask</u>

11000000.00010000.00100000.00000101 <u>11111111111111111111111111100000000</u> 11000000.00010000.00100000.00000000 <u>192. 16. 32. 0</u>





### **Sample IP Address:**

#### 

192.16.32.0 = Network ID 192.16.32.255 = Broadcast ID 192.16.32. 1 – 254 = Host Addresses



### **MAC Mini Network**















### **MAC Mini Network**









### **Steps to Configuring the MAC Mini's**



### 1) Create Accounts

- a. Name Machine
- b. User Name
- c. Password
- 3) <u>Configure IP Address</u>

- 2) <u>System Preferences</u>a. File Sharingb. Remote Sharing
  - c. Energy Saver off
  - d. Firewall off





### **Configure System – Set hostname**

#### **Click on Sharing**

#### 000 0 0 Sharing System Preferences ◄ ► Show All Q a < > Show All Personal Computer Name: arc1 File New One Computers on your local network can access your computer at: Q (3) Edit... 10 arc1.local ..... Dock Appearance Desktop & Exposé & International Security Spotlight Screen Saver Spaces On Service O Printer Sharing: Off $\checkmark$ DVD or CD Sharing Hardware Printer Sharing allows other people to use printers connected to this Screen Sharing computer. \* $\checkmark$ File Sharing 1 W **Printer Sharing** CDs & DVDs Bluetooth Displays Energy Keyboard & Print & Fax Sound Web Sharing Saver Mouse $\checkmark$ Remote Login $\checkmark$ Remote Management Internet & Network ✓ Remote Apple Events Xgrid Sharing Ø ۲ Internet Sharing MobileMe Network QuickTime Sharing Bluetooth Sharing System Open Print & Fax Preferences († 1 0 Click the lock to prevent further changes. ? Accounts Date & Time Parental Software Startup Disk Time Machine Universal Speech Update Controls Access

Name: ARC1, ARC2, etc





### **Configure System – Add user**

#### **Click on Accounts**

000	Accounts	6	
<ul> <li>▶ Show All</li> <li>My Account</li> <li>mh1 Admin</li> <li>Other Accounts</li> <li>Admin Standard</li> <li>Bobby Admin</li> <li>Cuest Account Sharing only</li> </ul>	Automatic login: Disabled Display login window as: O List of u Name ar Show the Restart, Sleep, and Shut Show Input menu in login window Show password hints Use VoiceOver at login window Enable fast user switching View as: Name	sers nd password Down buttons	
Login Options	langes.		(?

### Click on lock to unlock Click on +

000	Accounts		
Show All		Q	
My Account         mh1         Admin         Standard         Image: Solar Account Sharing only         Image: Sharing only	Automatic login:       Disabled         Display login window as: <ul> <li>List of t</li> <li>Name a</li> <li>Show the Restart, Sleep, and Shut</li> <li>Show Input menu in login window</li> <li>Show password hints</li> <li>Use VoiceOver at login window</li> <li>Enable fast user switching</li> <li>View as:</li> <li>Name</li> <li>\$</li> </ul> <li>Mame</li>	sers nd password Down buttons	
+ -			
Click the lock to prevent	further changes.		(?)





### **Configure IP Address**

#### **Click on Network**

000	Network
Show All	Q
Loca	tion: Automatic
Bluetooth     Not Connected	Status: Not Connected
Ethernet     Not Connected	
● FireWire Not Connected	Configuration: Default
e AirPort	Telephone Number:
	Account Name:
	Password:
	Connect
	(Set Up Bluetooth Device)
+ - *-	Show modem status in menu bar
Click the lock to prevent f	urther changes. Assist me Revert Apply

#### **Click on Ethernet**

000	Network
Show All	Q
L	cation: Automatic 🗘
Bluetooth Not Connected     Solution     Ethernet Not Connected	Status: <b>Cable Unplugged</b> The cable for Ethernet is not plugged in.
● FireWire 🍲	Configure: Using DHCP \$
e AirPort	IP Address: Subnet Mask: Router: DNS Server: Search Domains:
+ - *-	Advanced ?
Ilick the lock to prever	t further changes. Assist me Revert Apply





### **Configure IP Address**

Machine	IP Address	Subnet Mask
ARC1	172.30.220.1	255.255.255.0
ARC2	172.30.220.2	255.255.255.0
ARC3	172.30.220.3	255.255.255.0
ARC4	172.30.220.4	255.255.255.0
ARC5	172.30.220.5	255.255.255.0
ARC6	172.30.220.6	255.255.255.0

000	Network	
Show All		(Q )
Lo	ation: Automatic	\$
Bluetooth Not Connected     Section     Ethernet Not Connected     Section	Status:	Cable Unplugged The cable for Ethernet is not plugged in.
FireWire      Not Connected     On	Configure: IP Address: Subnet Mask: Router: DNS Server: Search Domains:	Manually     \$       172.30.220.1     255.255.255.0
+ - *-		Advanced) (?)
Click the lock to prevent	further changes.	Assist me Revert Apply





### Server vs. Nodes

In the next steps, we will first setup the Server (ARC1). Then to setup the nodes all we need to do is a secure copy to the nodes.







### **Configure a node**



- 1. Setup Host Files
- 2. Setup SSH
- 3. Install & Configure GCC
- 4. Install & Configure MPICH
- 5. Configure MPI to work with SSH
- 6. Test SSH, GCC and MPI





### **Setup Host Files**

#### cd to root (cd /)

#### Type: sudo pico /etc/hosts

○ ○ ○ Terminal - bash - 53×14	○ ○ ○ Terminal — bash — 53×14
Last login: Mon Jul 6 17:11:37 from arc1 arc1:~ mh1\$ cd / arc1:/ mh1\$	Last login: Mon Jul 6 17:11:37 from arc1 arc1:~ mh1\$ cd / arc1:/ mh1\$ sudo pico /etc/hosts





### **Setup Host Files**

Add the following at the bottom of the file:

172.30.220.1 ARC1 172.30.220.2 ARC2 172.30.220.3 ARC3 172.30.220.4 ARC4 172.30.220.5 ARC5 172.30.220.6 ARC6

$\odot \odot \odot$	Terminal — nano — 53×22
GNU nano 2.0.	1 File: /etc/hosts
##	
# Host Database	
#	
# localhost is	used to configure the loopback interf\$
# when the syste	em is booting. Do not change this en\$
##	11
127.0.0.1	
255.255.255.255	broadcastnost
::1	localnost
fe80::1%L00	Localhost
172.30.220.1	arcl
172.30.220.2	arc2
172.30.220.3	arc3
172.30.220.4	arc4
172.30.220.5	arc5
172.30.220.6	arc6
rg Get HAU Write	erk Read of Prev ok Cut Inc Cur Pos
X Exit Just	1 W Where V Next V Uncut fo Spell
	- <sup>12</sup>

What is SSH?

Why do we Need it?

#### The Seven Layers of OSI







We will be modifying 3 files:

- 1. /etc/ssh\_config
- 2. /etc/sshd\_config
- 3. /etc/hosts.equiv







#### Type: sudo pico /etc/ ssh\_config

host \*

HostbasedAuthenticat ion yes

# EnableSSHKeysign yes

GNU nano 2.0.1

Terminal — nano — 71×32 File: /etc/ssh\_config

# This is the ssh client system-wide configuration file. See # ssh\_config(5) for more information. This file provides defaults for # users, and the values can be changed in per-user configuration files # or on the command line.

# Configuration data is parsed as follows:

- 1. command line options
- 2. user-specific file
- 3. system-wide file

# Any configuration value is only changed the first time it is set.

# Thus, host-specific definitions should be at the beginning of the # configuration file, and defaults at the end.

# Site-wide defaults for some commonly used options. For a comprehens\$
# list of available options, their meanings and defaults, please see t\$
# ssh\_config(5) man page.

- #
  - Modified by ARC Curriculum Team June 30, 2009
- # The folling three lines are REQUIRED for SSH
- host \*

HostbasedAuthentication yes

EnableSSHKeysign yes

#Options for SSH are below. To activate any option, simply
#remove the # mark from the beginning of the line.

- # ForwardAgent no
- # ForwardX11 no

∧G Get Help∧O WriteOut∧R Read Fil∧Y Prev Pag∧K Cut Text∧C Cur Pos ∧X Exit ∧J Justify ∧W Where Is∧V Next Pag∧U UnCut Te∧T To Spell





Type: sudo pico /etc/sshd\_config

#### HostbasedAuthentication yes

```
Terminal — nano — 73×19
  GNU nano 2.0.1
                            File: /etc/sshd_config
# For this to work you will also need host keys in /etc/ssh_known_hosts
#RhostsRSAAuthentication no
# similar for protocol version 2
# The HostbasedAuthentication line below must be activated
# (the # is removed for SSH to work with MPI)
HostbasedAuthentication yes
# Change to yes if you don't trust ~/.ssh/known_hosts for
# RhostsRSAAuthentication and HostbasedAuthentication
#IgnoreUserKnownHosts no
# Don't read the user's ~/.rhosts and ~/.shosts files
#IgnoreRhosts yes
# To disable tunneled clear text passwords, change to no here! Also,
∧G Get Help ∧O WriteOut ∧R Read File∧Y Prev Page∧K Cut Text ∧C Cur Pos
∧X Exit     ∧J Justify  ∧W Where Is ∧V Next Page∧U UnCut Tex∧T To Spell
```





Type: sudo pico /etc/hosts.equiv

#### Add the following hostnames:



### **Install and Configure MPICH**

What is MPICH?

MPICH-V1





Reference: http://en.wikipedia.org/wiki/MPICH





### **Download MPICH**

#### http://www.mcs.anl.gov/research/projects/mpich2/downloads/ index.php?s=downloads

MPICH2	<u>)</u>		
home abo	ut downloads documentation	publications support release	information
OWNLOADS			
>Downloads			
SVN			
License			
Previous Versions			
Nightly Snapshots	home > downloads > downloads		
	Current Stable Version: 1.1		
	Platform	Download	Size
	UNIX and Windows (source)	[mpich2-1.1.tar.gz]	17 MB
	Windows IA32 (binary)	[mpich2-1.1-win-ia32.msi]	10 MB
	Windows EM64T/AMD64 (binary)	[mpich2-1.1-win-x86-64.msi]	13 MB







### **Install MPICH**

#### Type:

cd /arc\_share/mpich1.2.7p1

./configure - -with-device=ch\_p4 - prefix=/usr/local/mpich-1.2.5/ch\_p4 -rsh=ssh





#### ○ ○ ○ Terminal - bash - 80×24

```
ARC1:mpich=1.2.7p1 arc1$ ./configure ==with=device=ch_p4 =prefix=/usr/local/mpic
h-1.2.5/ch_p4 -rsh=ssh
Configuring with args --with-device=ch_p4 -prefix=/usr/local/mpich-1.2.5/ch_p4 -
rsh=ssh
Configuring MPICH Version 1.2.7 (release) of : 2005/11/04 11:54:51
Use the environment variable RSHCOMMAND instead of configure argument -rsh=ssh
checking whether filesystem respects case in file names... no
Warning: You may encounter problems with the commands mpice and MPICC
checking for current directory name... /arc_share/mpich-1.2.7p1
checking for architecture... freebsd_ppc
checking for install
checking for ar... found /usr/bin/ar ()
checking for ranlib
checking for gnumake... no
checking whether make supports include... yes
checking for OSF V3 make... no
checking for virtual path format... VPATH
checking for xIC
checking for q++
checking whether g++ returns correct error code... yes
Compiling C++ interface with q++
checking whether selected C++ compiler can compile iostream.h... yes
```

Include C++ bindings for MPI from http://www.osl.iu.edu/research/mpi2c++





### **Building the Cluster**

Modify the file that contains our hosts:

cd /arc\_share/mpich-1.2.7p1/util/machines/sudo pico machines.freebsd\_ppcChange the list to ARC1 through ARC6 and remove the .local

Ctrl o to write

Ctrl x to exit





### **Machines File**







### http://www.nccs.gov





Oak Ridge National Laboratory

U. S. Department Of Energy 42