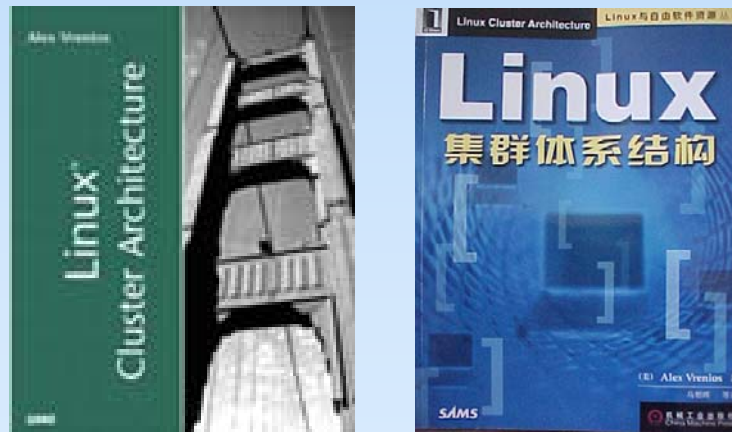


# Building a Linux Cluster

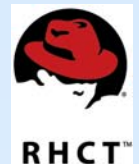
by

Dr. Alexander Vrenios

(Shameless Plug)



Slide # 1  
Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG



# Building a Linux Cluster

## Overview:

- What is a cluster computer?
- Why build a cluster computer?
- Assembling the PC computer hardware (on the cheap!)
- Configuring relevant Linux OS files for internetworking
- Custom software development can make PCs act as a team
- Three architectures for a application-layer cluster systems
- Hardware and Software Performance Considerations
- Recommended Material for Further Reading

Slide # 2

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## What is a cluster computer?

Networking Hardware  
Computer Hardware  
Operating System  
OS Configuration  
+ Custom Software  

---

Cluster Computer

A *cluster computer* is a collection of inter-networked computers that presents a ***single system image*** to its users. Its focus might be that of a data (or Web) server, a supercomputer, or a fault-tolerant High Availability server. The custom software gives the cluster its *personality*, architecturally speaking; these computers work together as a team toward a common goal.

Slide # 3

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Why build a cluster computer?

- Hobbyists:  
It's a new and interesting pathway to experience; and how many of your friends have a Linux cluster anyway?
- Professionals:  
Sophisticated systems are often developed in parallel, meaning the hardware won't be ready when you want to test your software. Having a test bed will get you past the hardware independent bugs, and put you in a position to polish your product when the platform is finally ready.
- Managers:  
This is all bleeding edge stuff; you'll want to prepare for the issues your people might face and the questions they might ask. Experience gives you the insight you will need.
- Academics:  
Analyze data from a live system, instead of questionable and potentially over-simplified software simulation output.

Slide # 4

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Assembling the Computer Hardware

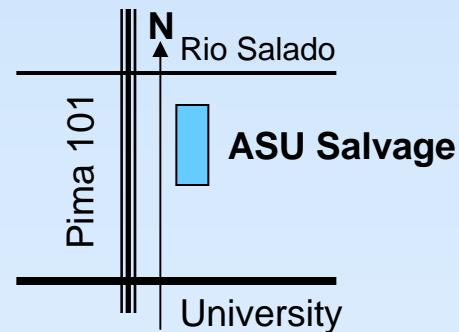
Slide # 5

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Finding (cheap) PC Computer Hardware:

- Small computer stores (Renaissance Computer, e.g.)
- Newspaper and club and organization newsletter ads
- Family, friends and neighbors (closets, garage sales)
- Large organizations? (hospitals, industry, schools etc.)
- Computer salvage outlets:

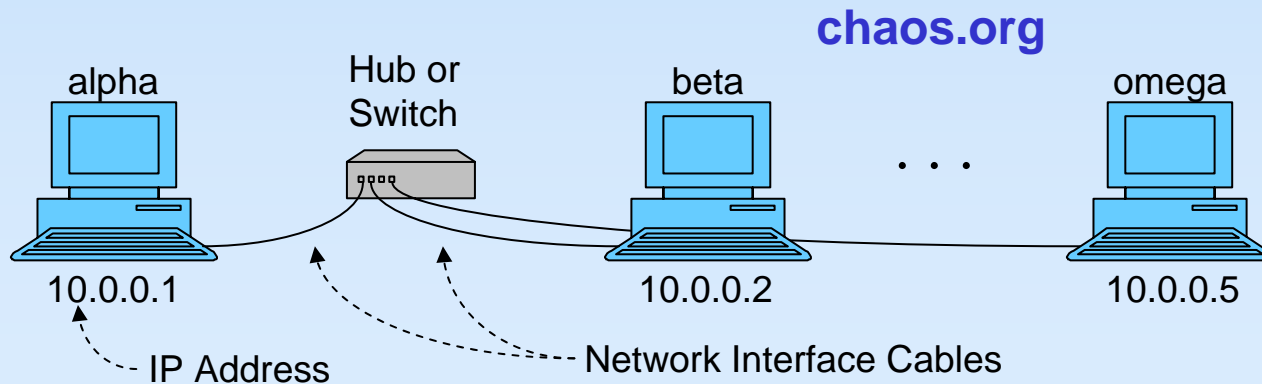
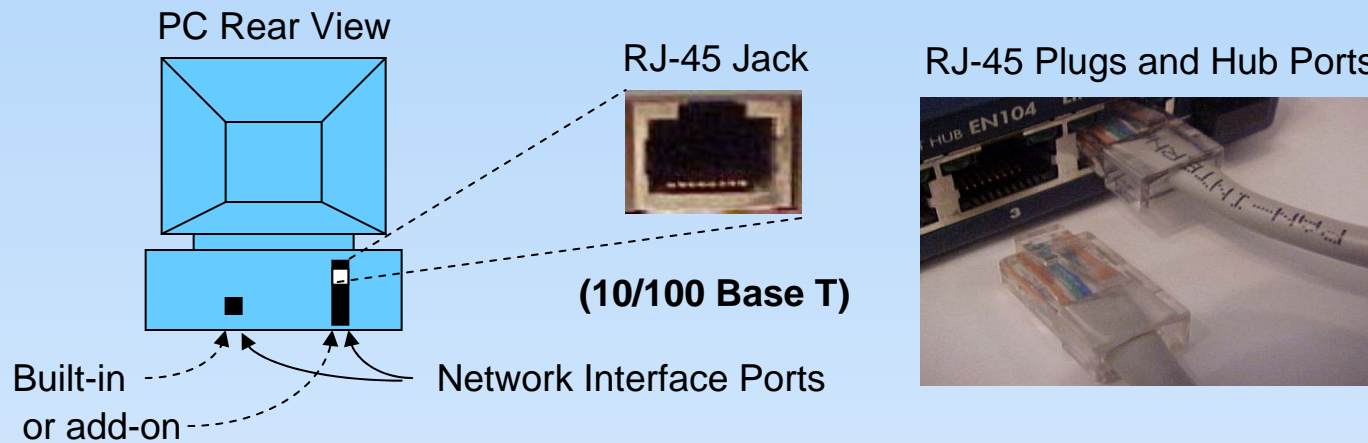


Slide # 6

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

Connecting PCs into a LAN is easy today:



Slide # 7

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## CHAOS: Cheap Array of Outmoded Systems:

- Finding and networking **N** personal computers:

4-way KV Switches  
(may be optional)

Network Activity  
Monitoring

Surge Protectors



Books!

Development System  
and NFS File Server

Slide # 8

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

**A major upgrade:**



**Slide # 9**

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Configuring the Operating System Files

Slide # 10

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Awareness of Other Computer Nodes:

- First, file `/etc/hosts` must exist on each of the network nodes:

127.0.0.1	localhost	localhost.localdomain
10.0.0.1	alpha	alpha.chaos.org
...		
10.0.0.5	omega	omega.chaos.org

Slide # 11

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Remote NFS Access to Local Files:

- File `/etc/exports` on 10.0.0.1, the NFS server named alpha:

Server → `/home` `(rw)`

- File `/etc/fstab` on each cluster node except the server named alpha:

Clients →

<code>/dev/hda1</code>	<code>swap</code>	<code>swap</code>	<code>defaults</code>	<code>0</code>	<code>0</code>
<code>/dev/hda2</code>	<code>/</code>	<code>ext3</code>	<code>defaults</code>	<code>1</code>	<code>1</code>
<code>alpha:/home</code>	<code>/home</code>	<code>nfs</code>	<code>defaults</code>	<code>0</code>	<code>0</code>
<code>/dev/fd0</code>	<code>/media/floppy</code>	<code>auto</code>	<code>noauto</code>	<code>0</code>	<code>0</code>
<code>none</code>	<code>/proc</code>	<code>proc</code>	<code>defaults</code>	<code>0</code>	<code>0</code>

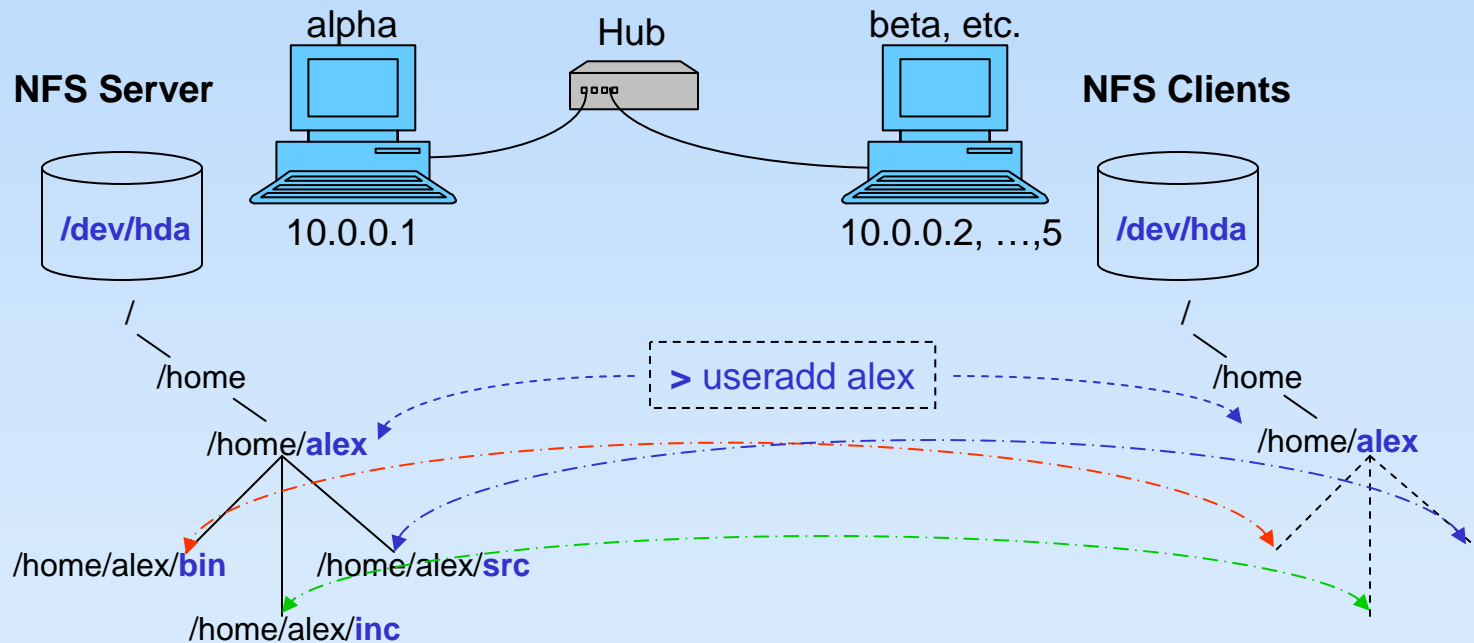
Slide # 12

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## The Network File System:

- NFS offers the *illusion of locality* via remote-mount points



Slide # 13

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

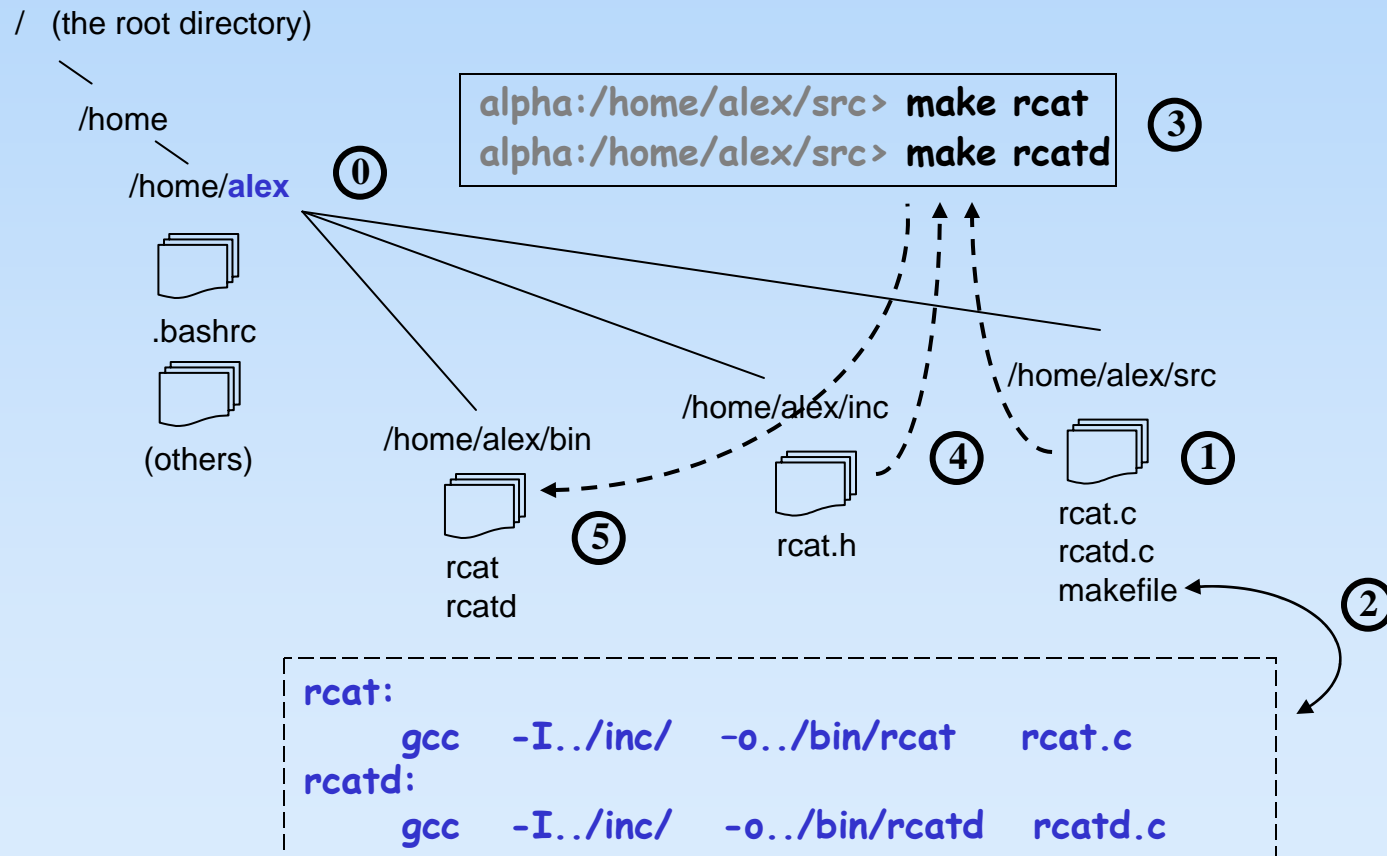
## Developing the Custom Software

Slide # 14

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Local User Files for Software Development:



Slide # 15

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Configuring Internetworking Services (xinetd):

- File /etc/xinetd.d/rcatd on each remote-server node:

```
service rcatd
{
    port                = 5000
    socket_type         = dgram
    protocol            = udp
    wait                = yes
    user                = alex
    server              = /home/alex/bin/rcatd
    only_from           = 10.0.0.0
    disable             = no
}
```

Refers to name of service

Means 10.0.0.\* (the network)

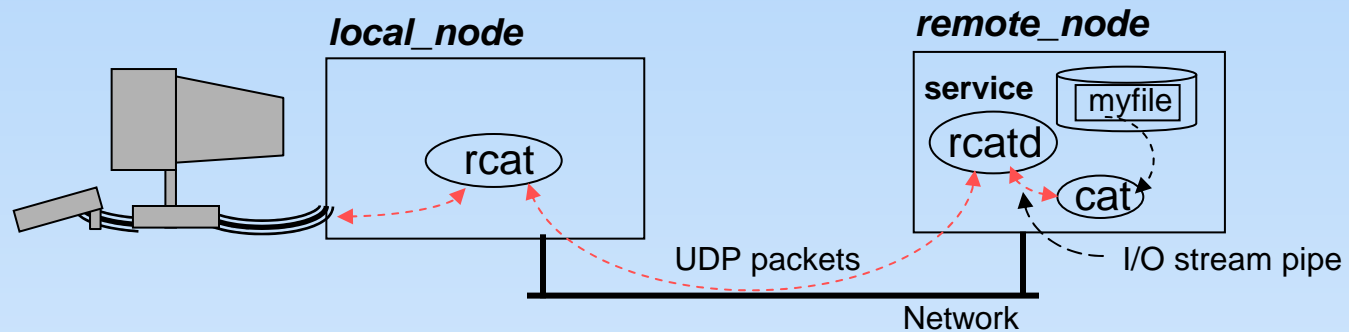
- [Reconfiguration] `omega:/root> /etc/rc.d/init.d/xinetd restart`
- [Red Hat] `omega:/root> service xinetd restart`

Slide # 16

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Example Internetworking Service (rcatd):



Screen output on *local\_node*:

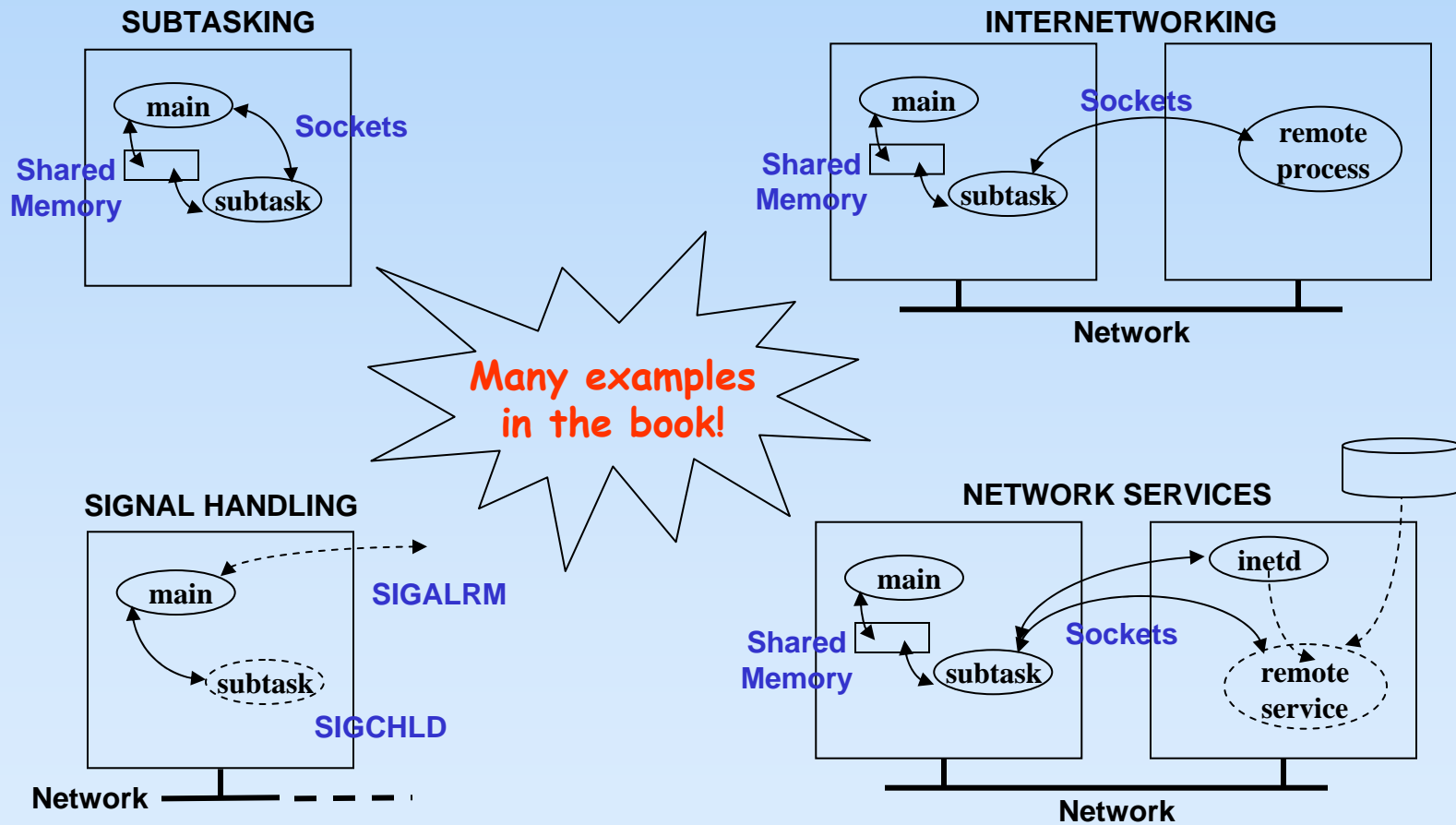
```
>rcat myfile remote_node
first line from myfile
second line from myfile
subsequent lines...
>
```

Slide # 17

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Distributed System C-Language Skills:



Slide # 18

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

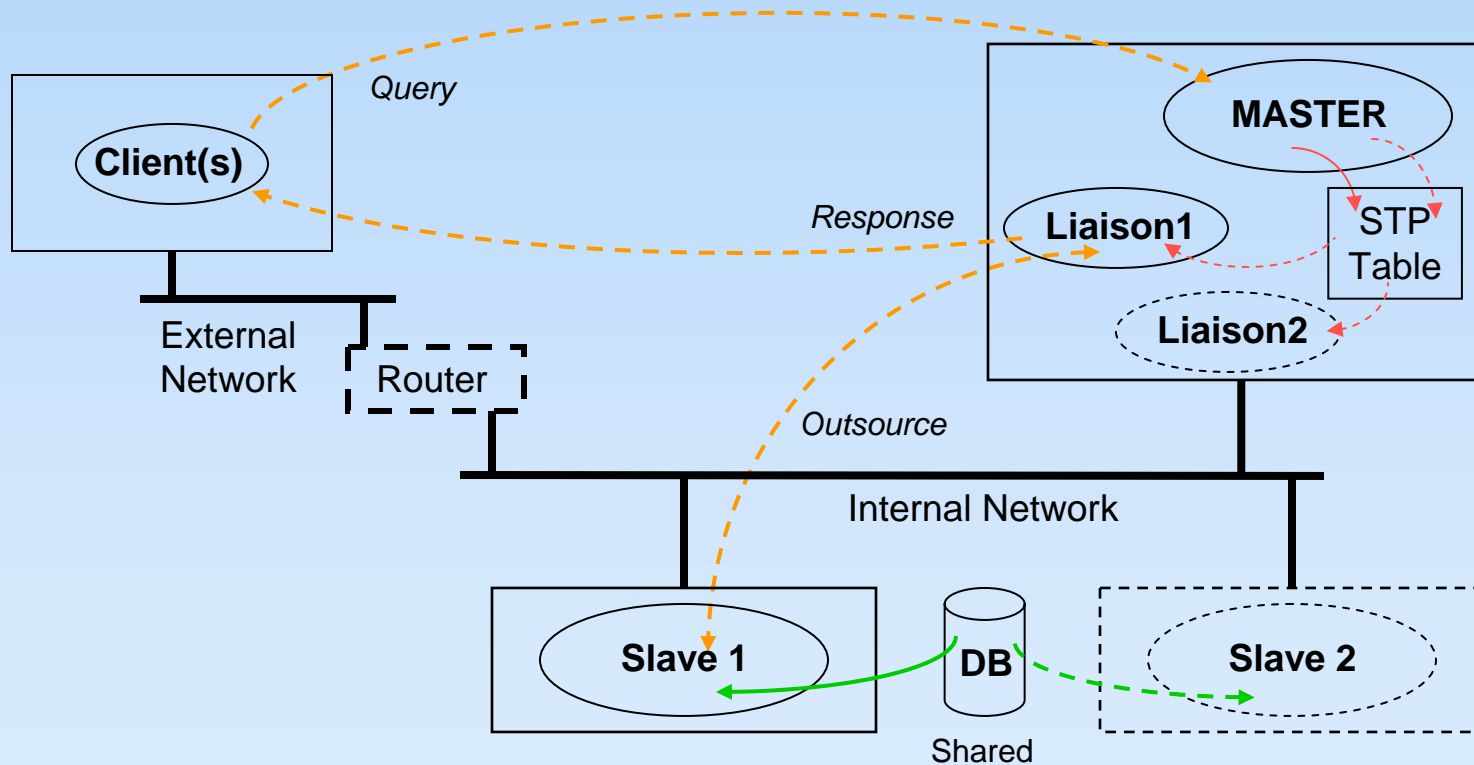
## An Example Master-Slave Cluster Server Architecture

Slide # 19

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Master-Slave Cluster Server - Architecture:



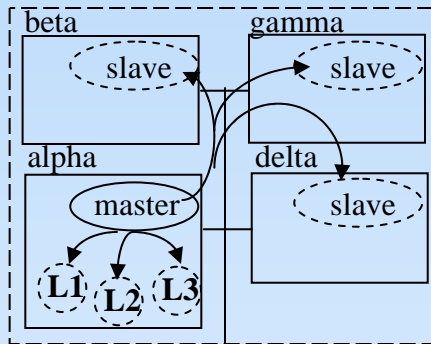
Slide # 20

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

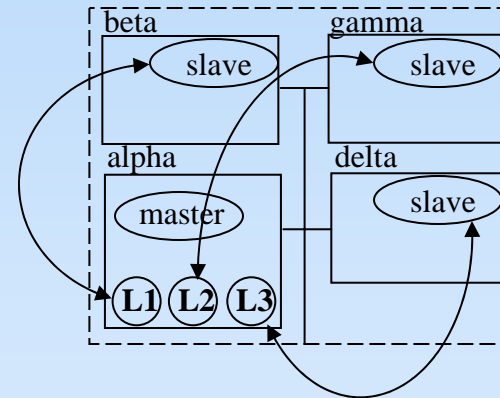
## Master-Slave Cluster Server - Initialization:

Broadcast starts *slave* tasks...



master starts local liaisons, one for each registering remote slave.

Local liaisons contact slaves...



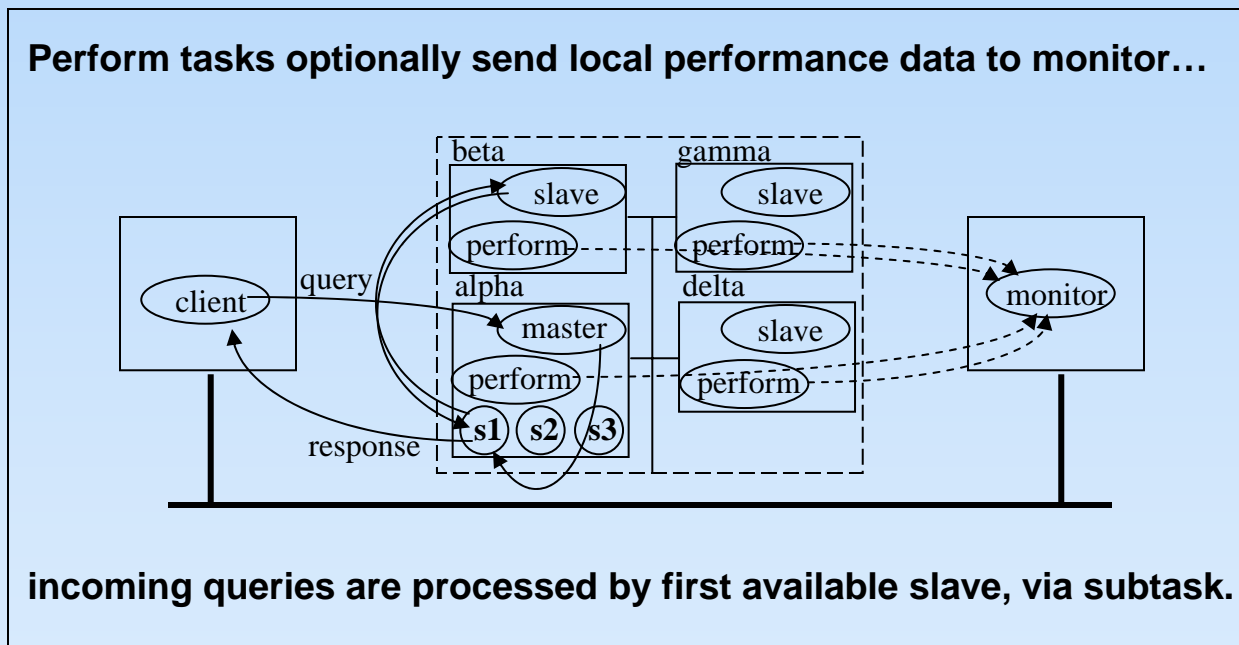
all tasks wait for work.

Slide # 21

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Master-Slave Cluster Server - Operation:

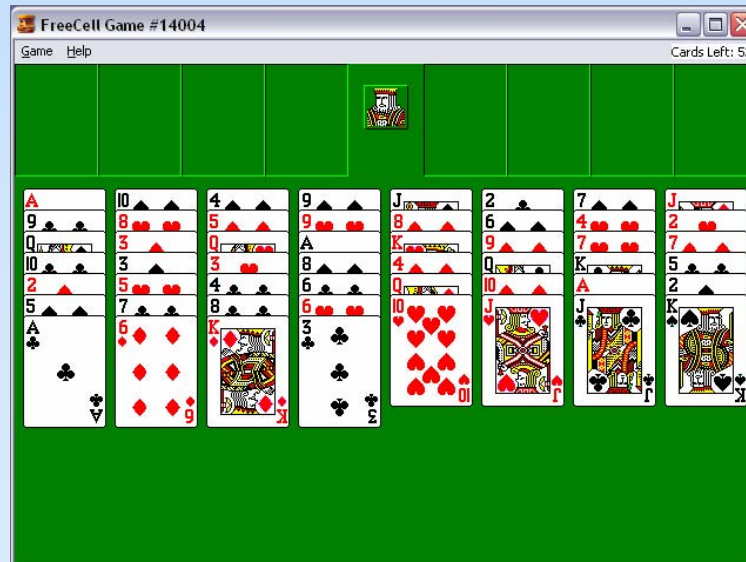


Slide # 22

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## An Example Virtual-Ring Cluster Supercomputing Architecture

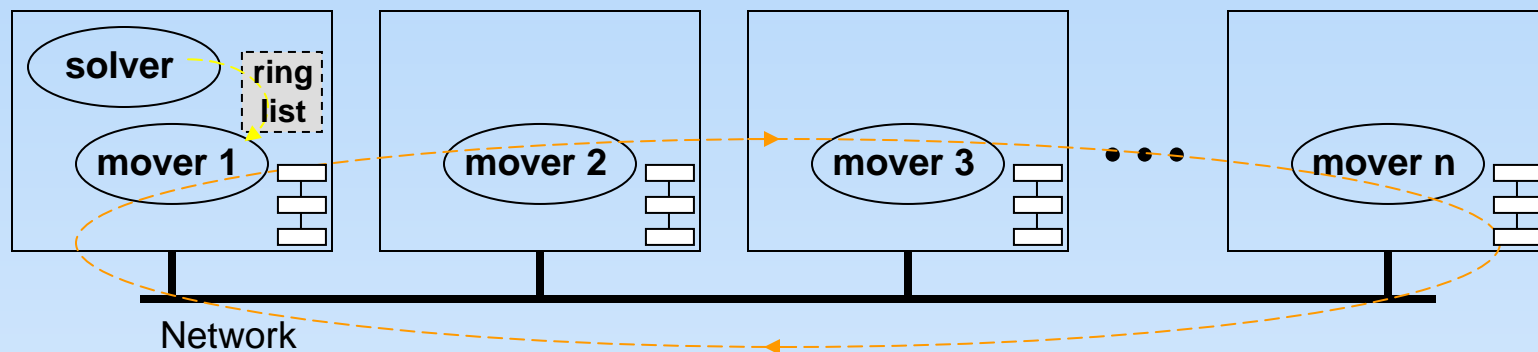


Slide # 23

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Virtual-Ring Supercomputer - Architecture:



Ring list is built from broadcast packet responses. Response order determines position on ring.

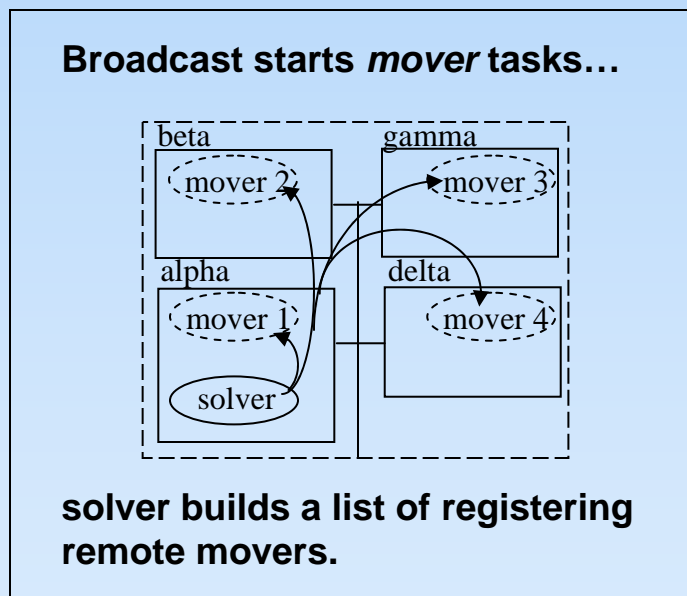
Slide # 24

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

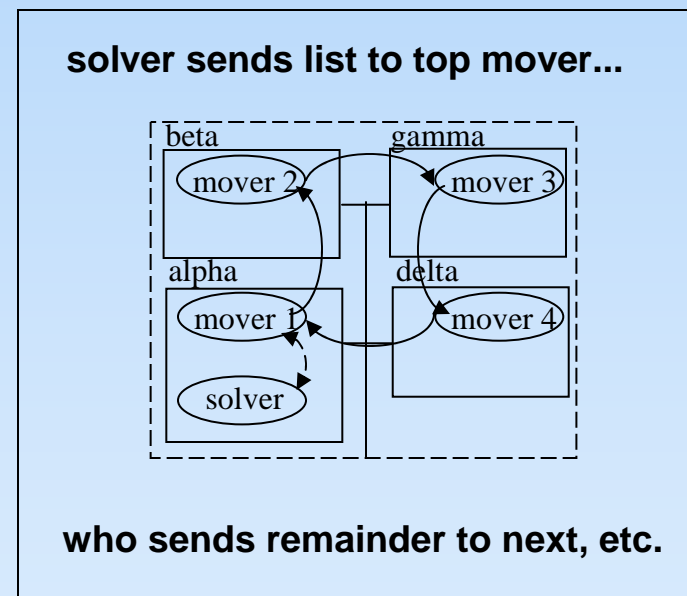
## Virtual-Ring Supercomputer - Execution:

### Initialization:



When the list comes back to the *bottom* mover, the ring is complete.

### Operation:



Solutions are returned to mover #1.

Slide # 25

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

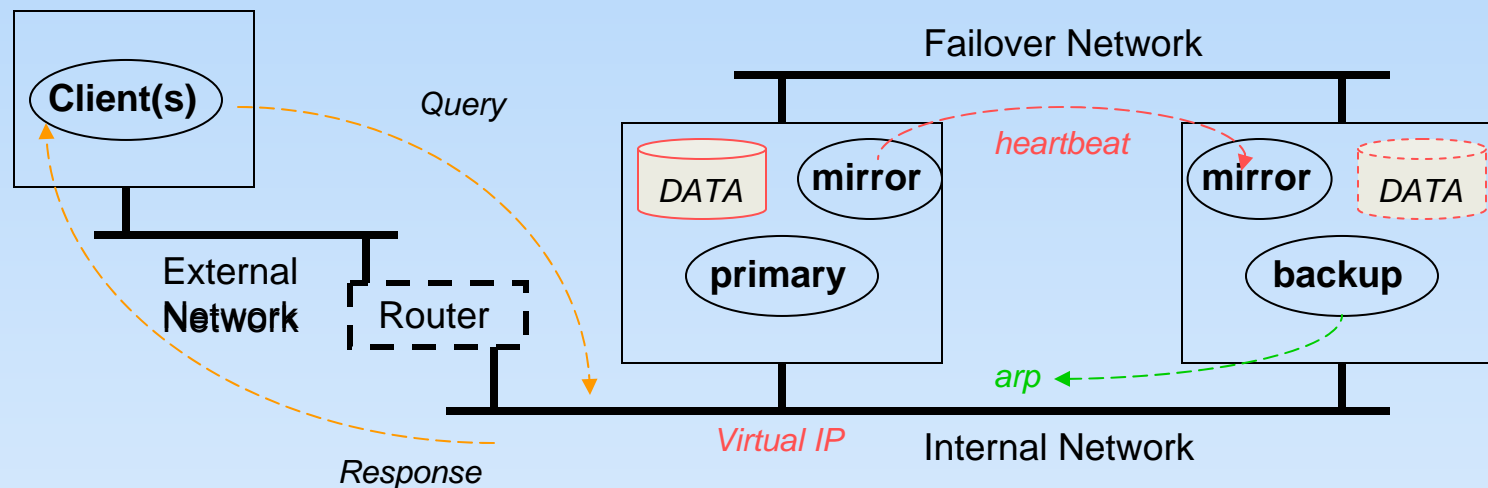
## An Example High-Availability Cluster Server Architecture

Slide # 26

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## High-Availability Server:



The *primary* server sends periodic **heartbeat** messages to the *backup* server. If the backup server detects a *missing* heartbeat, it assumes the primary's identity, responsibilities and its resources: data are continually updated; virtual IP via gratuitous arp. Primary server may be revived, or not.

Slide # 27

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Performance Considerations:

### Hardware:

- Super high speed networking hardware vastly improves *ratio*
  - Network transit time vs. execution/processing time

### Software:

- Distributed Shared Memory (DSM) operating system is better
  - Frees up software developer to write simpler programs

Slide # 28

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Recommended Material for Further Reading

Slide # 29

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## Further Details are in the Book:

- Download all the source code for free:

<http://www.sampublishing.com>

- Search on "Linux Cluster Architecture" or "Vrenios"
- Click on the "Downloads" link in the book description
  1. Individual chapter examples are in zip files
  2. A complete user *alex* environment is in a tar.gz file

- Book Signings:

Sep 8<sup>th</sup>

Borders Chandler, Sunday @ 2pm

Sep 15<sup>th</sup>

Borders Arrowhead, Sunday @ 2pm

Oct 25<sup>th</sup>

Barnes & Noble Arrowhead, Friday @ 7pm

Slide # 30

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

## References:

*Distributed Operating Systems*, Andrew S. Tanenbaum  
(of **MINIX** and **AMOEBA** fame!), Prentice Hall, 1995

*Unix Distributed Programming*, Chris Brown,  
Prentice Hall, 1994

*Advanced Programming in the UNIX Environment*,  
W. Richard Stevens, Addison-Wesley, 1992

"CHAOS Parts 1 & 2," LinuxGazette.com, Alex Vrenios, December 1998

*Linux Programming White Papers*, Rushling, et al, Coriolis Open, 1999

And of course,

*Linux Cluster Architecture*, Vrenios, A., Sams Pub., 2002

Slide # 31

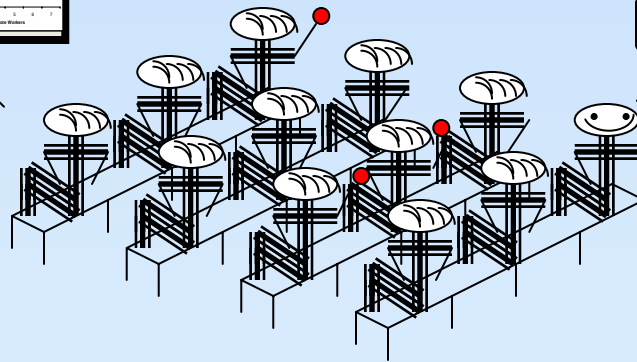
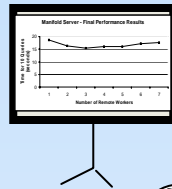
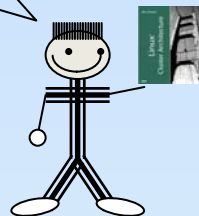
Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG

# Building a Linux Cluster

You've been a terrific audience!

Any questions?

Hurry out and  
buy this book!



¥ ¥ !



Slide # 32

Copyright © 2005 Alexander Vrenios  
Presented to the EMCC LUG