

Distributed Operating Systems

Overview

Ye Olde Operating Systems

OpenMOSIX

OpenSSI

Kerrighed

Quick Preview

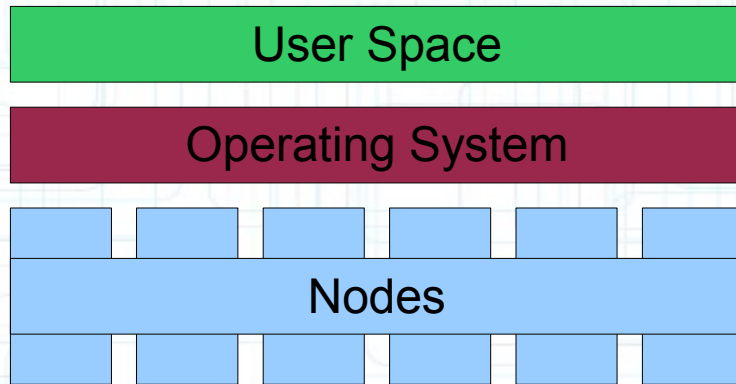


Front

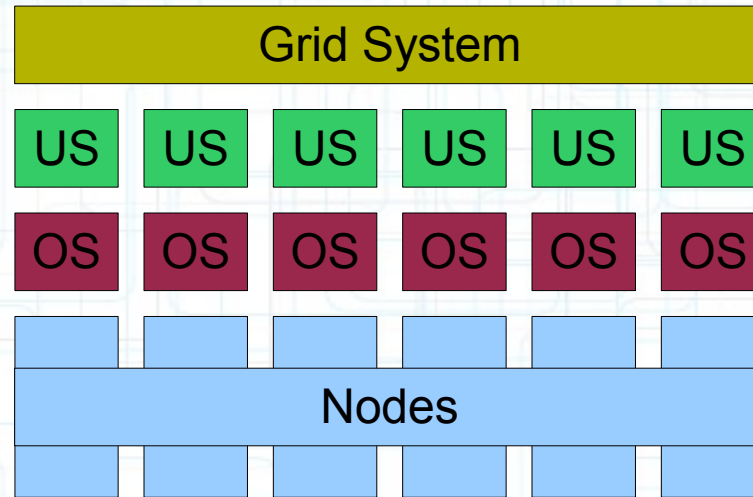


Back

Distributed Operating Systems vs Grid Computing



Amoeba, Plan9, OpenMosix,
OpenSSI, Kerrighed.



Xgrid, SGE, Condor, Distcc,
Boinc, GpuGrid.

Distributed Operating Systems vs Grid Computing

Problems with the grid.

- Programs must utilize that library system.

- Usually requiring separate programming.

- OS updates take place N times.

Problems with dist OS

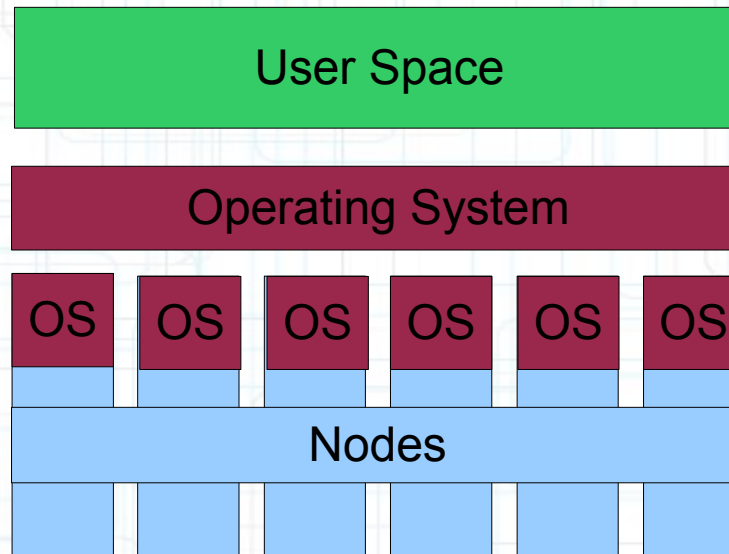
- Security issues – no SSL.

- Considered more complicated to setup.

Important Note

Each node, even with distributed operating systems, boots a kernel.

This kernel can vary depending on the role of the node and overall architecture of the system.



Amoeba

Andrew S. Tanenbaum

Earliest documentation: 1986

What modern language was originally developed for use in Amoeba?

Anyone heard of Orca?

Sun4c, Sun4m, 386/486, 68030, Sun 3/50,
Sun 3/60.

Amoeba



```
Virtual PC "Amoeba"
bootuser: using vdisk:82
bootuser: starting user process ...
bootuser: process started; wait ...
bootsvr: Reinit
bootsvr: Using disk vdisk:82 - size 2010
bootsvr: NewConf: 5 confs
BOOTSERVER INITIALIZED
bootsvr: Soap_0: cappoll failed: server not found
bootsvr: Soap_0 considered down
bootsvr: Boot(Soap_0)
SOAP 0: initializing
SOAP 0: cache_init: estimated bytes per row = 180
SOAP 0: cache_init: rowmem = 100K, Max_rows = 568
SOAP 0: cache_init: Max_dirs = 94
SOAP 0: caching 47 of 198 super blocks
SOAP 0: Bullet 0 is up
SOAP 0: super_init: global seqno in superblock is [0, 0]
SOAP 0: starting in 1 copy mode
SOAP 0: coming up in 1-copy mode
SOAP 0: 4 threads started
bootsvr: auto-switch off verbose

Welcome to standalone Amoeba
#
```

Plan9

Started development in the 1980's

Released in 1992 (universities) and 1995 (general public).

All devices are part of the filesystem.

X86, MIPS, DEC Alpha, SPARC,
PowerPC, ARM.

Union Directories, basis of UnionFS.

/proc first implemented here.

Plan9

The screenshot displays the Plan9 window manager interface. At the top, a status bar shows the date and time as "Sun Jun 11 12:45". Below this, there are several user avatars and icons, including a German flag and a "Germany" label. A terminal window in the foreground shows a directory listing for a rabbit image, with the text "Plan 9 from Bell Labs" visible below the image. To the right, another window displays a file manager view with a directory listing of files and folders, including "New Cut Paste Snarf Sort Zerox Delcol" and "New Cut Paste Snarf Sort Zerox Delcol". The bottom of the screen shows a window with a dark background and a grid of small, colorful icons, likely representing a collection of files or a specific application's interface.

Rio, the Plan9 window manager showing "faces(1), stats(8), acme(1)" and many more things.

Plan9

Split nodes into 3 distinct groupings.

Terminals

File servers

Computational servers

Uses the "9P" protocol.

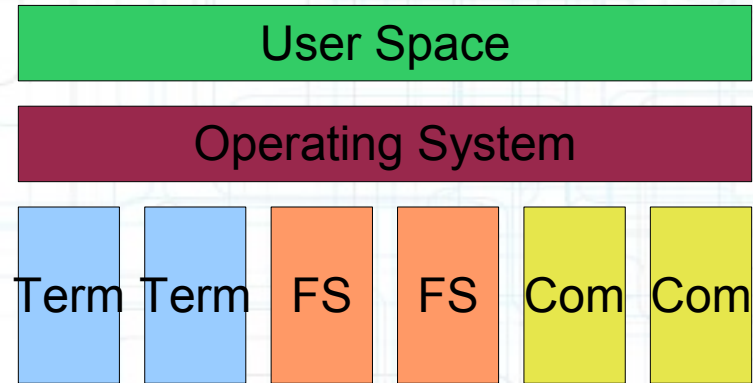
Low level, byte protocol, not block.

Used from filesystems, to printer communication.

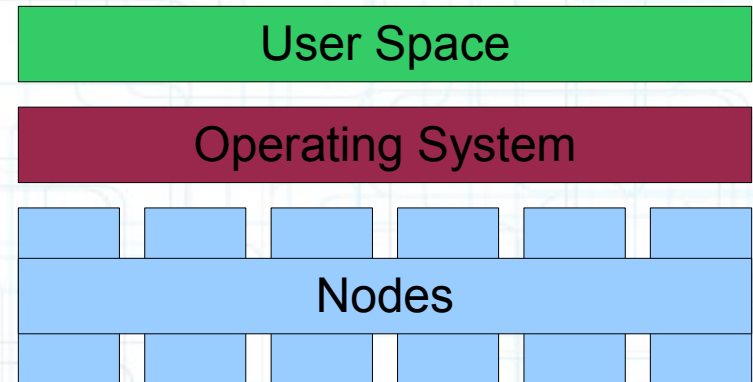
Author: Ken Thompson

Plan9 / Amoeba

Both Plan9 and Amoeba make groupings of nodes, into specific categories. This can mainly be attributed to the time period.



Starting with OpenMOSIX, there was a push to make the nodes identical, or atleast breakout from the "grouping" model.



OpenMOSIX

SSI System. (Single System Image).

Automatic load leveling. (Procs not threads).

Patch for Linux 2.4.x

EOL March 1 2008.

Linux PMI - Linux 2.6.x branch.

LiveCD autoconfiguration available.

OpenMOSIX

```
root@host118:/u2/mosix/j
18:13:35 up 22:22, 8 users, load average: 0.63, 0.44, 0.41
110 processes: 106 sleeping, 4 running, 0 zombie, 0 stopped
CPU states:  cpu  user  nice  system  irq  softirq  iowait  idle
              total 10.5%  0.0%   3.1%  0.0%   0.0%   0.0%  86.2%
Mem:  772752k av, 682460k used, 90292k free, 0k shrd, 39524k buff
      1024k active, 309628k inactive
Swap: 2088432k av, 1212k used, 2087220k free 86860k cached
```

PID	USER	PRI	NI	SIZE	RSS	SHARE	STAT	%CPU	%MEM	TIME	CPU	COMMAND
27600	root	17	0	5708	5708	172	S	99.9	0.7	56:57	22240	john
27604	root	15	0	5708	5708	172	S	99.9	0.7	56:57	22336	john
27606	root	19	0	5708	5708	172	S	99.9	0.7	56:57	22784	john
27608	root	14	0	5708	5708	172	S	99.9	0.7	56:53	22208	john
27610	root	17	0	5708	5708	172	S	99.9	0.7	56:53	22144	john
27612	root	19	0	5708	5708	172	S	99.9	0.7	56:57	22464	john
27614	root	17	0	5708	5708	172	S	99.9	0.7	56:57	22176	john
27620	root	17	0	5708	5708	172	S	99.9	0.7	56:56	22432	john
27622	root	15	0	5708	5708	172	S	99.9	0.7	56:56	22528	john
27624	root	15	0	5708	5708	172	S	99.9	0.7	56:56	22112	john
27626	root	16	0	5708	5708	172	S	99.9	0.7	56:58	22560	john
27632	root	18	0	5708	5708	172	S	99.9	0.7	56:58	22688	john
27634	root	14	0	5708	5708	172	S	99.9	0.7	56:57	22592	john
27602	root	16	0	5708	5708	172	S	99.5	0.7	56:56	22496	john
27596	root	18	0	5708	5708	172	S	99.1	0.7	56:56	22752	john
27598	root	15	0	5708	5708	172	S	98.9	0.7	56:57	22304	john
27616	root	15	0	5708	5708	172	S	98.9	0.7	56:54	22368	john
27618	root	16	0	5708	5708	172	S	98.9	0.7	56:56	22656	john
27628	root	16	0	5708	5708	172	S	98.9	0.7	56:56	22624	john
27630	root	15	0	5708	5708	172	S	98.9	0.7	56:53	22272	john

An OpenMOSIX cluster, running John The Ripper.

OpenMOSIX

Unique /mfs filesystem.

`/mfs/here` → / filesystem, current node.

`/mfs/home` → / filesystem, home node.

`/mfs/selected` → / filesystem on selected node, done by `"echo #
> /proc/self/selected"`

Added /proc support.

`/proc/hpc/nodes/[mosix ID]/(load|mem|speed)` , specific node statistics from remote /proc.

`/proc/hpc/nodes/[mosix ID]/` is not a remote /proc (only peices).

OpenMOSIX

Enable migration of sub processes:

```
"echo 0 > /proc/self/lock"
```

Useful for a shell.

Perl and Python modules available to ease programming specific applications.

Libmosix for C

Commonly used for large scale LTSP/POVRay.

OpenSSI

Last updated a year ago.

Kernel 2.6.12

<http://openssi.org/cgi-bin/view?page=docs2/1.9/Introduc>

Single process space.

Global PID's, local information.

Single root.

No specific programming required.

libcluster.so and cluster.h available.

(rexec(), rfork(), etc.)



x86_64, x86 architectures

Lenny, Etch, Sarge, FC3, FC2, RH9

Access to remote /dev

OpenSSI

Stable release is FC2

Good example on how much activity.

■ Project website

search

Go Search

toolbox

- RSS Atom
- Upload file
- Special pages

7 March 2011

- (diff) (hist) . . **N 1 Dropping a 23 98 clips in a 29 97 timeline applies a pulldown 62**; 20:56 . . (+2,759) . . [AmandaMcneal6](#) (Talk | contribs) (New page: *Image:Drop_Shipping_Companies_Drop_Ship_Companies_5647.jpg* Receiving telephone calls from unfamiliar callers can be annoying, especially if they keep calling. If you don't re...)
- (diff) (hist) . . **FC2 DRBD Root Fallover HOWTO**; 20:54 . . (+7,021) . . [Cecilkorik](#) (Talk | contribs) (Reverted to version from 2 Nov 2006 by Kristic)
- (diff) (hist) . . **N Spring collection and hoodies are dropping in a Week 6**; 17:37 . . (+2,332) . . [RaisaLuna8](#) (Talk | contribs) (New page: *Image:Drop_Shipping_Companies_Drop_Ship_Companies_2786.jpg* With folks hoping to escape the confines of exclusive office, making funds online will wide open a vast realm of op...)
- (diff) (hist) . . **N Yay baskets for free 73**; 10:31 . . (+2,558) . . [MichaelinaLovett](#) (Talk | contribs) (New page: *Image:Gift_Baskets_3391.jpg* A custom gift basket can make the best gift. Present baskets, whether giving or receiving, are really popular. According to Supermarket Information...)
- (diff) (hist) . . **N Qiero mi propio departamento y auto c 83**; 09:02 . . (+2,976) . . [RostyslavCassidy](#) (Talk | contribs) (New page: *Image:Auto_Insurance_1402.jpg* Insurance claims may seem difficult but documenting the harm caused to you and your vehicle can make the difference among any successful also a lo...)

Outdated timeline

- 2008 August - 2.0.0pre release number reserved for base kernel-2.6.16 or higher for OpenSSI stable.
- 2008 October - Preview of OpenSSI-1.9.6 (aka. 2.0.0pre3) - kernel bug fixes and performance improvements to VPROC, CFS, and PROCFS. (in CVS)
- 2009 Q1 - OpenSSI-1.9.6 for CentOS 4 - more kernel bug fixes, performance; re-enable CFS buffered I/O (**in testing**)
- 2009 Q1 - OpenSSI-1.9.6 for Debian Etch(?) (**in testing**)
- 2009 - OpenSSI-1.9 x86_64 64-bit port.
- 2009 - OpenSSI-1.9 port to kernel-2.6.18 or higher.
- TBD - OpenSSI-1.9 port to CentOS 5.
- TBD - OpenSSI-1.9 socket migration bug fixes.

OpenSSI

'localview' command

Prefix like nice.

Restricts that process to local devices, processes, and scope of ipc.

'loadlevel' command

Algorithms borrowed from openMOSIX.

Turned off by default.

Can be turned on globally, or on individual nodes.

OpenSSI

Extensive guides and tutorials

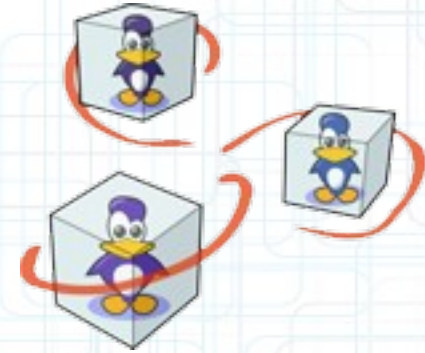
Out of date OS support.

Contrib contains xen kernel's. Possible to run massive paravirtualized guests.

Cluster virtual IP support

Similar to LVS (Linux virtual Server)

Kerrighed



Modification to the linux kernel.

2.6 branch

Current release of 2.6.30

x86, x86_64.

Single System Image

Single process space.

Checkpoint / Restart

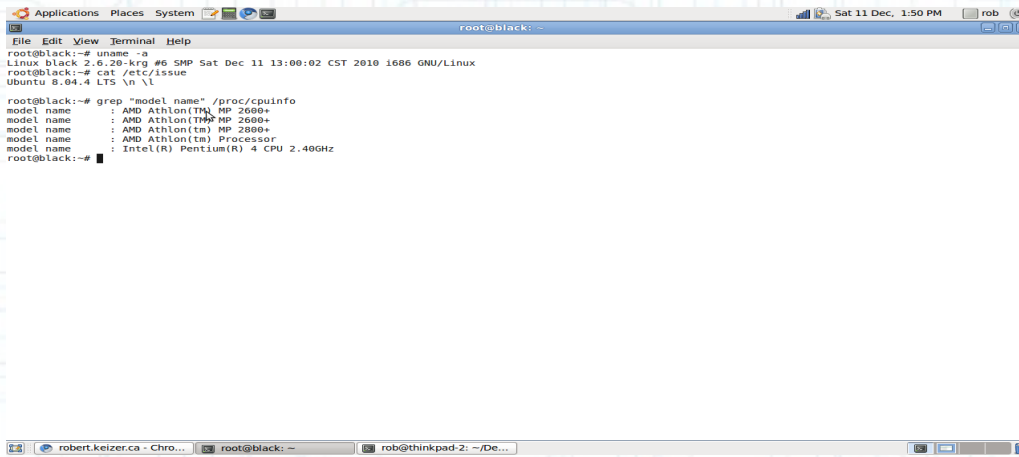
Distributed memory.

Kerrighed

Single Process Space.

Global PID's

Mashup of statistics.



The screenshot shows a terminal window with the following content:

```
Applications Places System Sat 11 Dec, 1:50 PM rob
root@black:~
File Edit View Terminal Help
root@black:~# uname -a
Linux black 2.6.20-krz #6 SMP Sat Dec 11 13:00:02 CST 2010 1686 GNU/Linux
root@black:~# cat /etc/issue
Ubuntu 8.04.4 LTS \n \l
root@black:~# grep "model name" /proc/cpuinfo
model name      : AMD Athlon(TM) MP 2600+
model name      : AMD Athlon(TM) MP 2600+
model name      : AMD Athlon(tm) MP 2600+
model name      : AMD Athlon(tm) Processor
model name      : Intel(R) Pentium(R) 4 CPU 2.40GHz
root@black:~#
```

Kerrighed

The screenshot shows a terminal window titled 'root@black: ~' with a menu bar containing 'File Edit View Terminal Help'. The terminal displays system statistics and a process list. The statistics show 47 total tasks, 1 running, a load average of 0.01 0.23 0.15, and an uptime of 00:08:06. Memory usage is 48/4035MB and swap is 0/251MB. The process list is a table with columns: PID, USER, PRI, NI, VIRT, RES, SHR, S, CPU%, MEM%, TIME+, and Command.

```
1 [          0.0%]   Tasks: 47 total, 1 running
2 [|         0.5%]   Load average: 0.01 0.23 0.15
3 [|         1.0%]   Uptime: 00:08:06
4 [|         1.0%]
5 [|         0.0%]
Mem[|||     48/4035MB]
Swp[        0/251MB]
```

PID	USER	PRI	NI	VIRT	RES	SHR	S	CPU%	MEM%	TIME+	Command
1	root	15	0	2840	1688	540	S	0.0	0.0	0:00.42	/sbin/init
231994	root	18	0	1716	504	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty1
231939	statd	15	0	4120	1148	868	S	0.0	0.0	0:00.00	~/usr/sbin/ntpd -p /var/run/ntpd.pid -u 106:108 -g
231876	root	25	0	5312	984	636	S	0.0	0.0	0:00.00	~/usr/sbin/sshd
231846	root	15	0	2184	708	480	S	0.0	0.0	0:00.01	~/sbin/syslog-ng -p /var/run/syslog-ng.pid
231816	root	18	0	1712	500	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty6
231815	root	18	0	1712	504	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty3
231814	root	18	0	1716	508	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty2
231812	root	18	0	1716	508	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty5
231811	root	18	0	1712	500	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty4
231387	sshd	25	0	1896	696	596	S	0.0	0.0	0:00.00	~/sbin/rpc.statd
231373	daemon	17	0	1836	516	420	S	0.0	0.0	0:00.00	~/sbin/portmap
230439	root	20	-4	2220	628	360	S	0.0	0.0	0:00.26	~/sbin/udevd --daemon
166389	root	18	0	1716	504	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty1
166333	statd	15	0	4120	1156	868	S	0.0	0.0	0:00.00	~/usr/sbin/ntpd -p /var/run/ntpd.pid -u 106:108 -g
166270	root	21	0	5316	1012	660	S	0.0	0.0	0:00.00	~/usr/sbin/sshd
166240	root	15	0	2188	708	480	S	0.0	0.0	0:00.02	~/sbin/syslog-ng -p /var/run/syslog-ng.pid
166210	root	17	0	1716	508	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty6
166209	root	18	0	1716	508	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty3
166208	root	17	0	1712	500	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty2
166206	root	18	0	1712	500	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty5
166205	root	17	0	1712	500	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty4
165781	sshd	16	0	1892	692	596	S	0.0	0.0	0:00.00	~/sbin/rpc.statd
165767	daemon	17	0	1832	516	420	S	0.0	0.0	0:00.00	~/sbin/portmap
164892	root	21	-4	2220	624	360	S	0.0	0.0	0:00.26	~/sbin/udevd --daemon
101243	root	18	0	1712	504	436	S	0.0	0.0	0:00.00	~/sbin/getty 38400 tty1
101216	root	19	0	2100	760	612	S	0.0	0.0	0:00.00	~/usr/sbin/cron
101205	daemon	25	0	1976	416	300	S	0.0	0.0	0:00.00	~/usr/sbin/atd
101159	dhcpd	15	0	2732	1136	640	S	0.0	0.0	0:00.00	~/usr/sbin/dhcpd3 -q -pf /var/run/dhcp3-server/dhcpd.pid -cf /etc/dhcp3/dhcpd.conf eth1
101104	root	23	0	1892	572	480	S	0.0	0.0	0:00.00	~/usr/sbin/inetd

At the bottom of the terminal, there is a footer with function key shortcuts: F1 Help, F2 Setup, F3 Search, F4 Invert, F5 Tree, F6 Sort By, F7 Nice, F8 Kill, F9 Kill, F10 Quit.

Small Kerrighed cluster running on commodity hardware.

Kerrighed

Ubuntu 8.04

Mandriva 2008.0-

Debian Lenny-

Support for x86 in $\leq 2.3.0$

Support for x86_64 $\geq 2.4.0$

Kerrighed

Kernel arguments

`session_id`

The cluster identifier. Currently 256 clusters can be on the same network.

`node_id`

Individual node id. Used in internal workings.

`Autonodeid`

If set makes `node_id=x` in `192.168.0.x`