



Monitoring and administration of openMosix clusters

by Matt Rechenburg

Legal openMosix is Copyright (c) 2002, 2003 by Moshe Bar. Mosix is Copyright (c) 2002 by Amnon Barak. The Mosix is a trademark of Amnon Barak. Linux is a Registered Trademark of Linus Torvalds. openMosix is licensed under version 2 of the GNU General Public License as published by the Free Software Foundation.





Matt Rechenburg

- Member of the openMosix community
- openMosix developer -> openMosixview
- Freelancer, working for Qlusters





- adminstrating openMosix
- running applications on openMosix
- the openMosixview project
- the gomd project (featuring Gian Paolo)
- usefull utilities and examples
- Time for discussion







Monitoring in general.. exactly, what is going on ?







Monitoring in detail

the openMosix commandline monitoring utilities

- mosmon
- the /proc/interface
- contributed tools -> mps + mtop





YUL/2	dem_	🗝 - Beleh si	iens er 3	- K uri=ul =
5			}	
1333	1	0 5	0:00	/sbin/mingerty ttyl
1398	14	0.5	0:00	/sbin/mingelly tty2
1399	30	0 3	0:00	/sbin/mingetty tty3
1100		0 5	0:00	/sbin/mingetty tty4
1401	5	0 5	0:00	/sbin/mingelly lly5
1402	6	0 5	0:00	/sbin/mingetty tty6
1408	2	O SN	0:00	(mfs gerver)
1404	7	0 5N	0:00	(mIs_server)
10716	7	0.5	0:01	/usr/sbin/sshd
10/18	2	0 5	0:01	bash
11.079	7	0 5	0:00	/usr/sbin/sshd
11052	7	0 3	0:00	
11228	1	0 5	0:00	/usr/sbin/sshd
11232	7	0 5	0:00	-bash
11429	7	0.5	0:00	/bin/sh /bin/mosrum -b openmosizview
11.83	12	0.5	0:03	openrosixview
11544	2	0 5	0:00	/bin/bash ./start_distkeygen.sh
11551	7	13	0:00	(distkeygen)
11552	2	0 3	1:21	./distkeygen
11553	7	1.5	0:31	(distkeygen)
11554	7	2 3	0:43	(distkeygen)
11555	2	1 5	0:31	(distkeygen)
11556	7	2 5	0:55	(distkeygen)
11557	7	1 3	0:31	(distkeygen)
11,568	12	OK	0:00	nps ax
1,251	7	0 5	0:00	(atd)
328	7	0.3	0:00	(rpc.statd)
299	12	0 5	0:00	(portnap)
Troute	demo	root]#		
_				





40-+	Zom Titte		57	nearc	1.00	d anto	an in a t		ni n	69 M		2
9210 8 977	obecces	56 01	cent	no 2	Tunni	70 0	zorb	ic.	0 sto	oned		
PTI 4	ales: 2	77 4%	inser	20	5% SVS	len	0.1%	ni	н. (1 D%	11-	
en:	255444K	AV. 1	5334	AK use	ed. 10	2096K	free		0	shrd.	272% butt	
wan:	522104K	av.		CK US	ed. 52	2104K	free				125610% cached	
PIE	USER	PRT	NT	STZE	RSS	SHARE	STAT	N#	%CPII	WVEM	TIME COVMAND	1
1552	root	18	0	972	972	784	R.	U	81.6	0.3	1:46 distkeygen	
1556	root	16	0	456	456	268	S	2	51,3	0.1	1:10 distkeygen	
1554	root	16	Э	456	456	268	3	2	49.7	0.1	0:58 distkeygen	
1555	root	18	÷	792	792	716	5	1	35.6	0.3	0:41 distkeygen	
553	root	18	0	456	456	268	S	1	34.0	0.1	0:42 distkeygen	
557	root	19	0	792	792	716	3	1	33.3	0.3	0:42 distkeygen	
564	TOOT	13	3	880	880	724	1	0	4.7	6.3	D:00 ntop	
103	root	9	9	1300	1300	1112	5	0	2.8	0.5	1:12 gsensed	
485	root	9	Ð,	7652	7652	5756	5	0	1.2	2.9	0:03 opennosixview	
228	root	9	Ð	1920	1920	1624	5	0	0.9	0.7	0:00 sshd	
1	root	8	0	480	180	-120	5	U	0.0	0.1	0:07 init	
2	root	9	Ð	0	0	0	SW	0	0.0	0.0	0:00 keventd	
9	root	9	0	0	0	0	3%	0	0.0	C.O	0:00 kapnd	
4	root	18	19	U	U	0	SWN	U	0.0	0.0	0:00 ksoftirgd CP00	
5	TOOL	9	0	0	0	0	5%	0	0.0	C.O	0:00 kswapd	
ō.	root	9	0	0	0	0	SW	0	0.0	C.0	0:00 bdflust	
1	root	9	30	11	D.	0	418	11	0.0	11.0	0:00 kupdated	
11	root	9	0	0	0	0	SW	0	0.0	C.O	0:00 khubd	
269	root	9	Ō	0	0	0	S'A'	0	0.0	0.0	0:02 rpciod	
276	TOOL	9	5	n	n	0	2.2	0	0.0	6.0	0:00 Inckd	
276	root	9.	2	0	0	0	SW	0	0.0	6.0	0:00 ovfs main server	
270	THEFT	9.	0	Ó	0	0	5'5	0	0.0	C.0	0:00 UNES BE	





The openMosix API

- user + programming interface to openMosix
- provides functions to get/set values
- Available in several programming languages





Administrating openMosix

the openMosix commandline administration utilities

- mosctl
 (basic administration)
- mosrun (running applications)
- setpe (configure cluster-map)
- migrate + miggroup (manual process migration)





User + account management

- using NIS
- home directory on oMFS/NFS
- user permissions in the cluster





Time syncronisation

- using NTP
- Using rdate





ssh/rsh access to the cluster-nodes

- not needed for openMosix at all
- needed for openMosixview
- usefull for administration
- security aspect in untrusted networks





Checkpointing on openMosix

- install chpox and add the chpox module
- register processe and its libraries
- sequential dump process state to file(s)
- restore process from dump-file
- --> full-migration of processes possible





openMosix calibration & tuning

- io and/or cpu intensive applications
- decay value / load-statistics
- speed value
- locking processes





openMosix hardware tuning

- increasing the network bandwith + latency
- increasing the processor power
- increasing the number of nodes





Running applications on openMosix

- decentralized concept -> no master-node
- applications installed once on one node
- applications can be started on any nodes





Application data / storage

- using oMFS
- using NFS
- using ogfs/GFS
- using lustre

--> openMosix and oMFSDFSA





Parallel applications on openMosix

- consists of more than one process
- starts processes on every cluster-nodes
- uses own communication mechanism
- mostly uses special parallel library





running PVM applications on openMosix

- master/slave concept of PVM applications
- static load/process assignment by PVM
- dynamic loadbalancing by openMosix
- minimizes the configuration needs for PVM





MPI applications on openMosix

- MPI standart and implementations
- similar master/slave concept as PVM
- multiple program single data (MPSD)
- limitations and common pitfalls
- openMosix patch to MPICH





openMosix and sequentiel applications

- parallelization of the problem to solve
- splitt into subtasks by splitting data
- run as MPMD (multiple program multiple data)
- forking using the openMosix API



the openMosixview project



- intuitive GUI
- supports all common used actions
- QT based user interface
- 8 usefull monitoring applications
- easy to install





the main openMosixview application window

X-⊭ openMosixview 1.3					X			
<u>F</u> ile <u>V</u> iew <u>C</u> onfig Collector <u>H</u> elp								
🚔 🔚 🍪 🗿 🛆 🍋								
id clusternodes load-balancing efficiancy o	overall Icad	overall used memory	all memory	all cpi	u			
all-nodes 90%	48%	5%	1334	MB 6				
22532 192.168.88.4 11	61%	13%	ZZ3	1				
22533 192.168.88.5	<mark>42%</mark>	3%	255	1				
22534 192.168.88.6 4 1	4 <mark>4%</mark>	3%	255	1				
22535 192.168.88.7 — Ш	4 <mark>6%</mark>		123	1	Н			
22530 192.168.88.2 — III) — III (5953	55%	4%	223	1				
	62%	3%	255	1	ŧ			
started 3dmosmon								



the configuration dialog

🗙 –¤ openmosixview 📃 💌									
🔍 openMosix-configuration									
node :		192.168.88.3							
UN	off	auto-migration on/off							
yes	no	talk to others nodes							
yes	no	local procs stay							
yes	no	send away guest procs							
start	stop	start/stop							
🛛 🔿 a	pply	🙆 cancel							
CO 🗐	console								
-display node1 💌 0€.0€									
Cle	🔞 clear Clear history 🕐 close								





advanced execution

X-¤ openmosixview								
🎲 openMosixview Advanced Execution								
/usr/bin/mybigjob								
(you can now specify add	Jitional command-line arguments)							
🔘 🔘 no migration								
🔘 run home	host-chooser —							
• run on	run job on							
🔘 🔿 cpu job	cluster-node							
🔘 🔘 io job								
🔘 🔘 no decay								
🔘 🔘 slow decay								
🔘 🔘 fast decay								
🔘 🔘 parallel	🏶 execute 🛛 🕐 close							





openMosixprocs, a graphical "top"

2	K-M proce	SSBS D	n node	4					l X		
M 🚱 remest: all 📼						esses 🛛 👘 last manag	agəd probəss				
	pic	۲Ħ	luck	nmigs	sta.	criid ine	nice	UD			
	🎲 12075	22535	U	2	S	Zdistkeygen	U	L			
	🎊 12074	22535	0	3	5	Zdistkeygen	0	C			
	\$ 12072	22535	0	0	S	Zdistkeygen	0	C			
	🏟 I 2063	22534	0	3	5	Zdistkeygen	0	С	H		
	🎊 12069	22533	0	3	S	Zdistkoyçen	0	C			
	🎇 12007	22500	0	Э	3	Zdistkeygen	0	C			
	🍣 1207 T	22531	Π	2	5	Zlislkeygen	Π	٢			
	🏟 11985	22531	U	З	S	/bin/bash	U	L			
	🎊 12073	22530	0	1	Ð	Zdistkeyçen	0	C			
	🍣 12071	22000	0	1	3	Zdistkeygen	0	C			
	🏟 12065	22530	0	Z	5	Zalistkeygen	0	С			
	🏶 583 👘	С С	1	0	S	/usi/bbin/ctd	0	С			
	🍪 C47 👘	С С	1	0	Э –	সাঁঃ	0	43			
	Sec. 2	٦	1	Π	5	cound.	N	٢			
	A 832	J	1	U	s	aom	U	L	•		
	morage p	ICCS fro)m rom:	ote	Γ	67 processes on this system		ট খ	ii _///		





migration dialog







About remote processes

X -¤ openmosixprocs ﷺ 319 ∰ 320 ∰	§ 321	?••×		
from openMosix-node	22532 with IP-adre remote identity	ss 192.168.88.4 statistics		
goto home node goto best node	pid=12795 tgid=520 uid=0 gid=0 pgrp=1 session=1 nmigs=1	utime=254 cutime=0 nice=0 state=R vsize=2613248 rss=114 nswap=0 cnswap=0		
3 remote processe	s running on this node	ОК		





the openMosixanalyzer (load)









X 🛨 Open HusisAnalyzer 1.3	
Elle Hep	
🛎 🦣 🏟 🕰 🖾 <table-cell> 🕅 subersfier:</table-cell>	MEMORY - OVERVIEW
17.10.2002-18.10.47	
si 100%	
RC 4	
S or	~
-2500	4
R 214	
Some group have been and the second	4 I
100%	
A an	
G m	·
100%	
R. arr	
S on Marine Marine	-
1008. 22535	
Sars	
A DEMONSTRATION OF THE OWNER OWNER OF THE OWNER	4
22590	
STX.	
🤤 <u>แร</u> ่งได้ให้ให้หลักพร้างการพ	t
displaying the inclusion sverview new	
and and an a strate to the strate of the str	





node overview

<u>X</u> -∺ open	imosixa	nalyzer			- ×		
Informat	tions ab	out node 3					
from :	13.7.20	002-19.16.42	to :	13.7	3.7.2002-19.45.41		
Load	ł	Memory			static data		
	I		CPUs	:	1		
			Avail.r	nem :	255 MB		
			Speed	:	15217		
min/max/	mean	min/max/mea	ın		ப் quit		





openMosixhistory

X−N openi	MinsixII	istory	1.3						×
🥐 lin =:	17.1	0.2002-	-18 25.42	2		proc	E		
					,	·			
7,20	ı	1	1	I	I	ı	ı	1	
hours ->									
pid	n#	ock.	nir gs	stat	ondine		n de	UI.	E
🎇 17880	225/5	n	7	5	/set athoma		-27	Π	
\$ 12874	22554	0	4	3	vset alhome		-22	0	
🎡 12873	225/4	n	7	5	/set athome		-77	Û.	
Sm 12077	22000	0	7	3	zset athome		-22	0	
🎇 17876	opeg (n	-	5	/set athoma		-77	n	
3 12079	22000	0	2	3	zset athome		-22	0	
4 12875	22520	n	-	5	/set athoma		77	n	
🎎 E C O)	-	2	3	/usi/sb-t/Efd		0	0	
🥸 = 4 7	1		-	5	XII		n	43	
🏩 I S 2)	-	2	3	crond		0	0	
🌼 582)		2	5	gpm		0	0	
🎡 . C1	J	-	-	د	-besh		U	U	
QE12)		2	5	sendmall		0	0	
4_14	J	-	-	ز	login		U	U	
<u>187.768</u>	<u>،</u>		-	2	in du und		0	(A D	-
									-
							ſ	d).	
(thpropen)	nosixeo	llector/	chistrius	tsecs.	dət			-	









the openMosixmigmon





3dmosmon by Johnny Cache







openMosixWebview by Ramón Pons Vivanco







- oMApplet
- KludgeKollection
- wmomload
- various bootable openMosix CD's
- openMosixLoaf
- Gentoo and Debian packages
- openMosixblaster
- and many more







summary and conclusion

- transparent clustering for users and adminstrators
- easy to control and monitor, lots of tuning options
- provides speedup for almost any application
- lots of contributed projects and tools from the community
- one of the most active linux-clustering project



Question time now









Thank you !! :)

and thanks to the community, to the Fosdem-organisators and the sponsors -> Qlusters, Myrinet, Unesco, NASA, AMD,

Johns Hopkins Medicine, Dolphin, University of Arkansas, Democritos, the university of Pisa not to mention Debian + Gentoo

Happy clustering !







