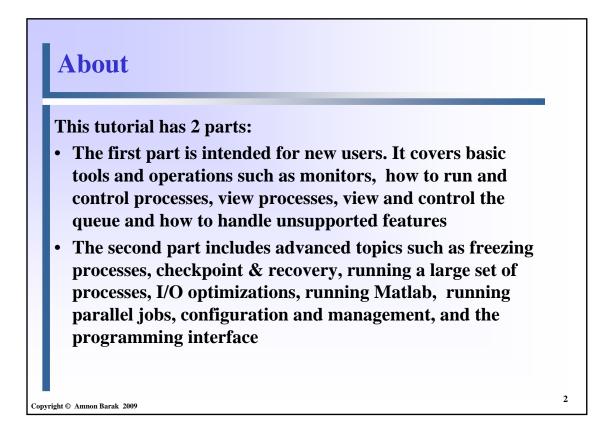
MOSIX2 Tutorial

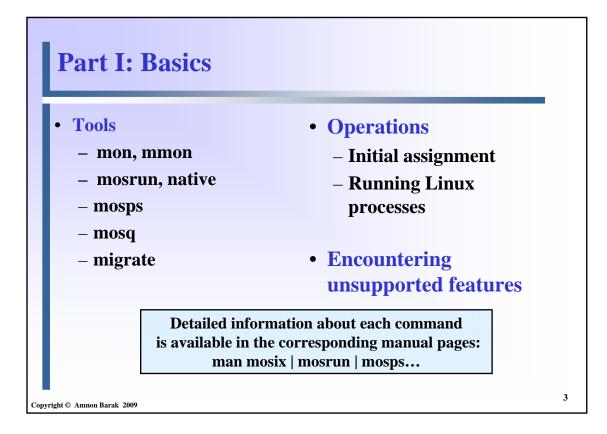
L. Amar, A. Barak, T. Maoz, E. Meiri, A. Shiloh Department of Computer Science The Hebrew University

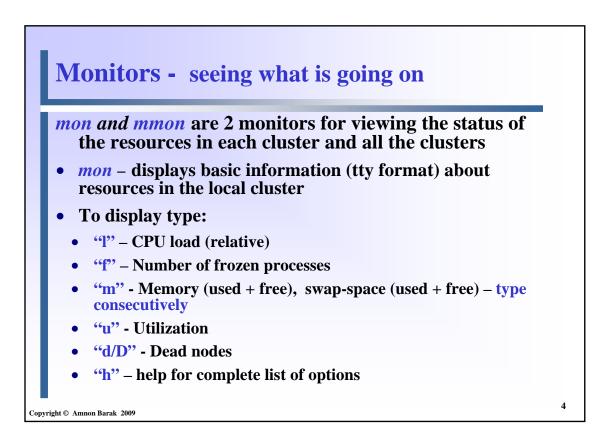
http://www.MOSIX.org

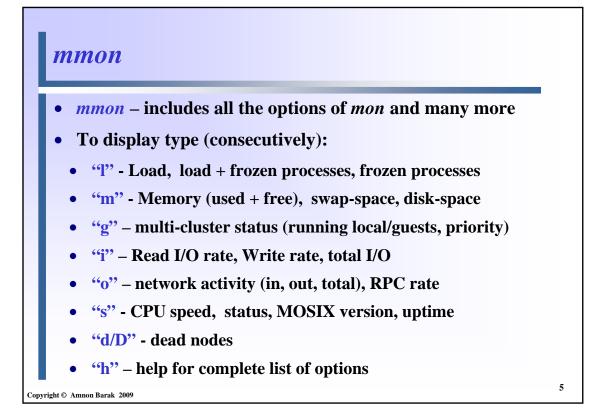
January 2009

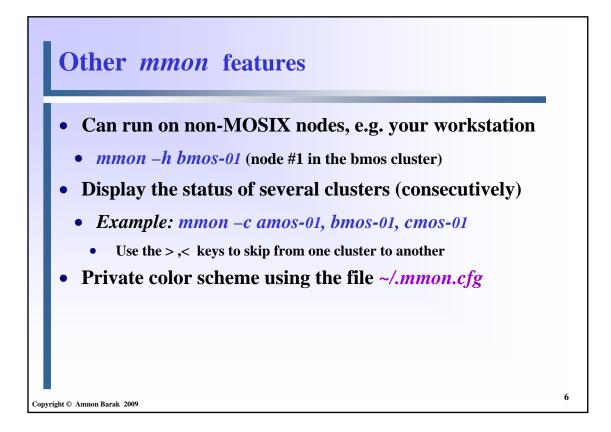
Copyright © Amnon Barak 2009

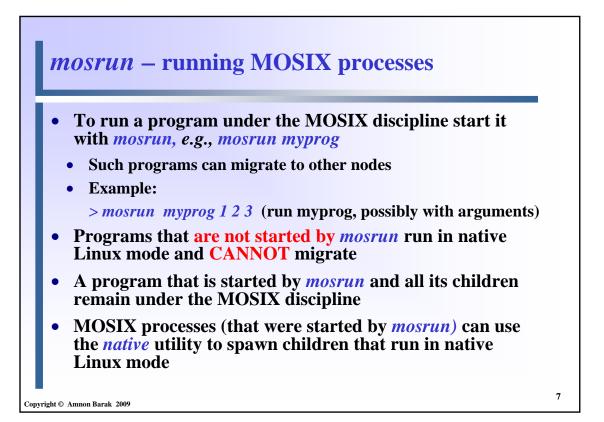


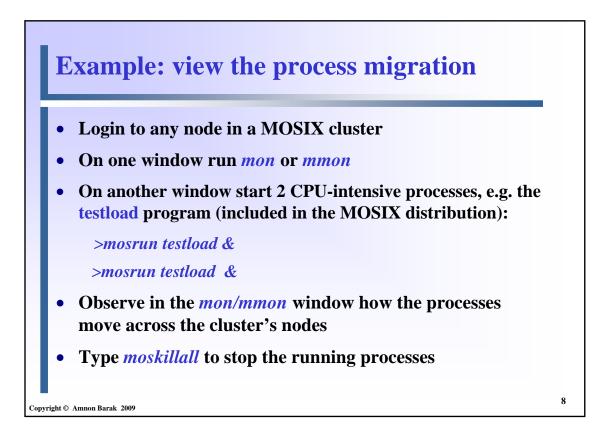


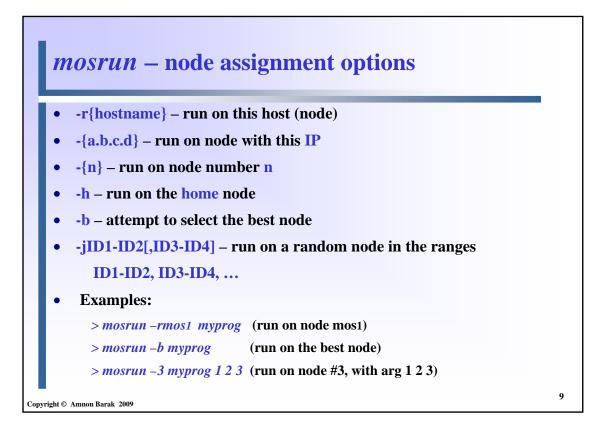


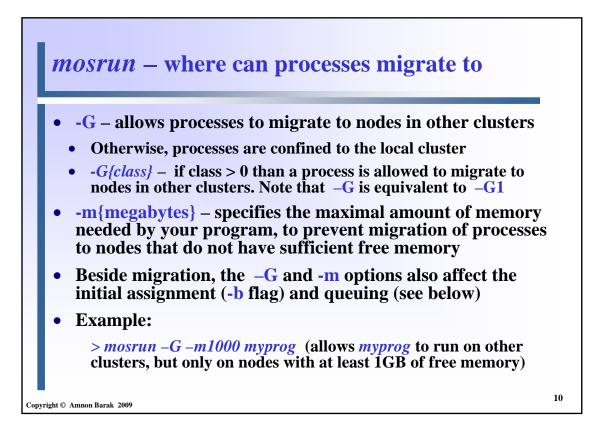


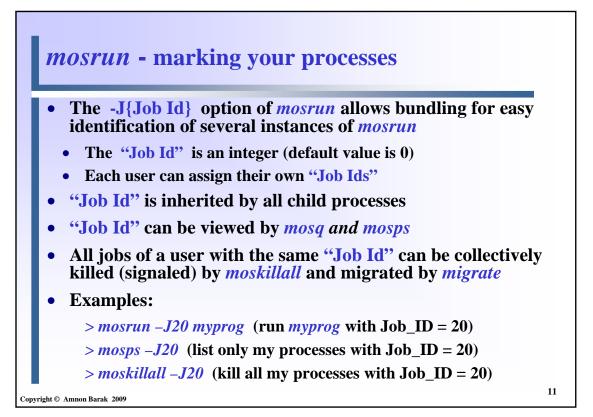


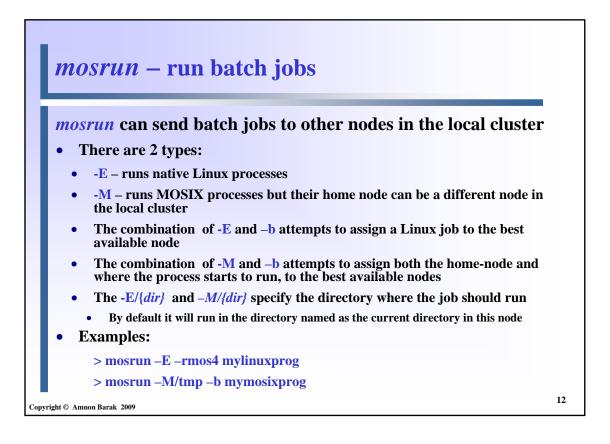












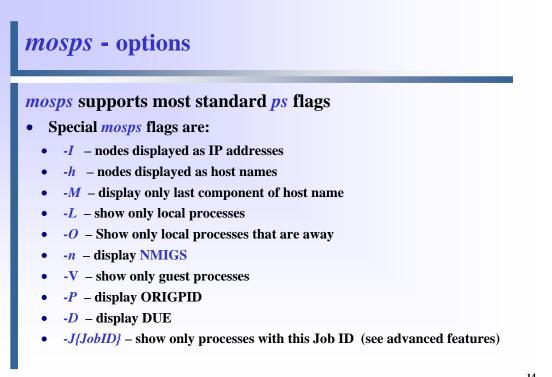


mosps (like ps) provides information about your MOSIX processes (and many standard *ps* fields, see the next slide), including:

- WHERE where (or in what special state) is your process
- from where this process came FROM
- **ORIGPID** original pid at home node
- *CLASS* class of the process
 - For MOSIX processes defined by *mosrun* –*G*{*class*) •
- Other values are *batch* and *native*
- **FRZ** if frozen why: A- auto frozen (due to load); E- expelled; *M*-manually; - NOTFROZEN; N/A – can't be frozen; **DUE** – when user expects process to complete
- **NMIGS** number of migrations so far

Copyright © Amnon Barak 2009

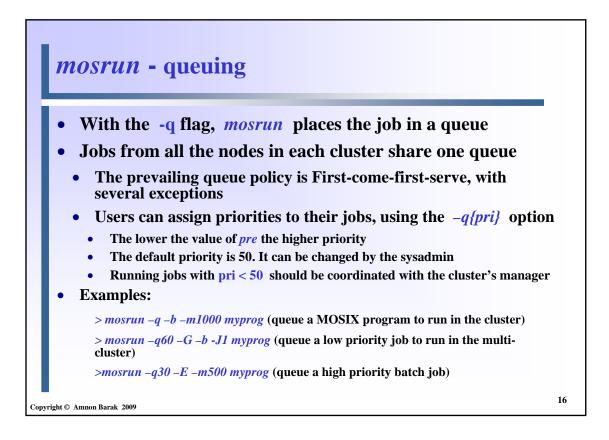
Copyright © Amnon Barak 2009

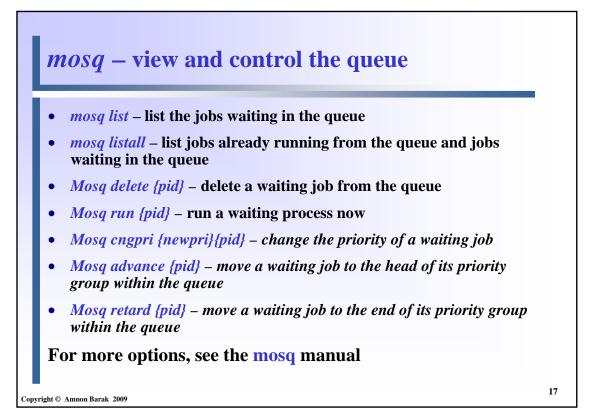


13

mosps – example

```
> mosps -AMn
    PID
          WHERE
                  FROM CLASS FRZ NMIGS TTY
                                                     CMD
   24078 cmos-18 here local
                                    1
                                        pts/1 mosrun -b testload
   24081 here
                                    0
                  here local
                                        pts/1 mosrun -b testload
                                    1
   24089 cmos-16 here local
                                        pts/1 mosrun -b testload
                              -
   30145 queue
                  here N/A N/A
                                   N/A pts/1 mosqueue -b testload
   30115 here
                  here local
                              М
                                    0
                                        pts/1 mosrun testload
   30253 here
                  mos3 local N/A
                                   N/A
                                        ?
                                               /sbin/remote
                                                                      15
Copyright © Amnon Barak 2009
```





		_	-	_	
> mosc	η lis	tal I			
PI D	USER	MEM(MB)	GRID F	PRI FROM	COMMAND
21666	lior	-	NO F	RUN here	mosrun -b testload
21667	lior	-	NO	50 here	mosrun -b testload
21155	lior	-	NO	50 cmos-17	mosrun testload
> moso	q cng	pri 20	21155		
> mosq	list	all			
PI D	USER	MEM(MB)	GRI D	PRI FROM	COMMAND
21666	lior	-	NO	RUN here	mosrun -b testload
21155	lior	-	NO	20 cmos	-17 mosrun testload
04//7	lior	-	NO	50 here	mosrun -b testload

migrate – control MOSIX processes

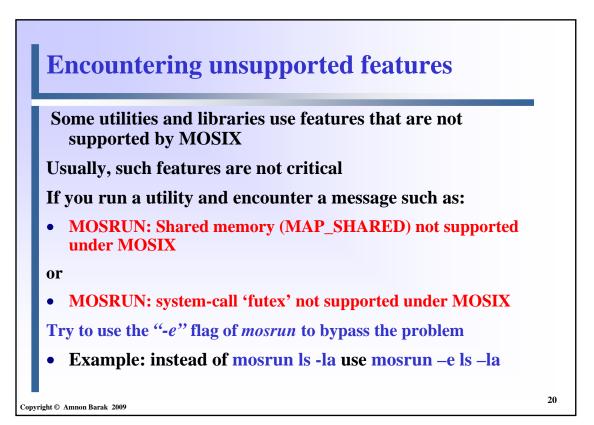
- migrate{pid} {node-number | ip-address | host-name} move your process to the given node
- *migrate{pid} home* request your process to return home
- migrate{pid} freeze request to freeze your process
- *migrate{pid} continue* unfreeze your process
- *migrate{pid} checkpoint* request your process to checkpoint
- *migrate{pid} checkstop* request your process to checkpoint and stop

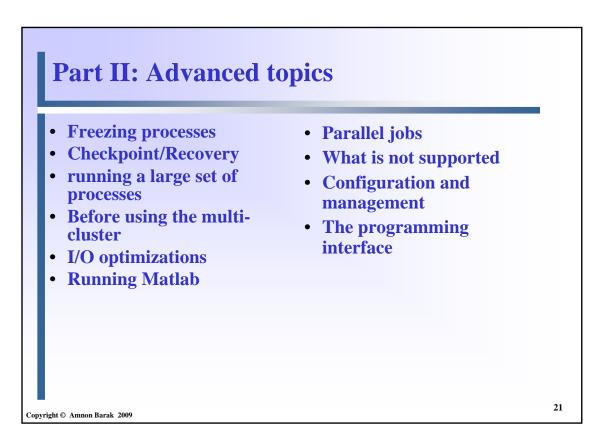
19

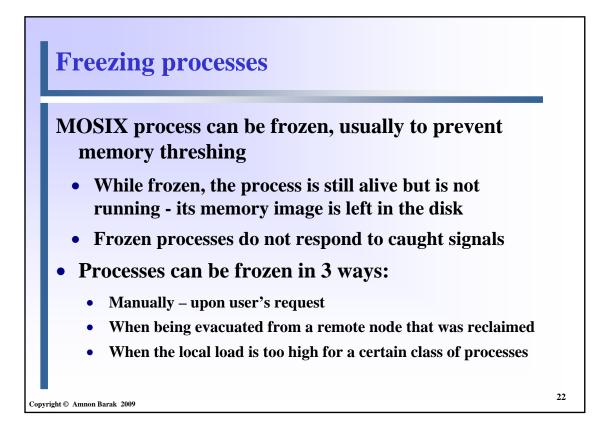
• *migrate{pid} exit* – checkpoint your process and exit

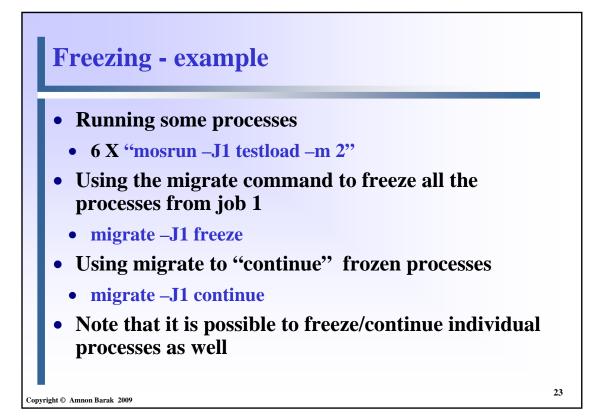
For more options, see the migrate manual

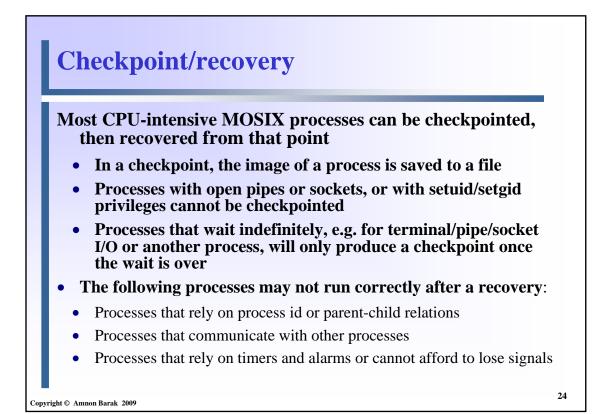
Copyright © Amnon Barak 2009

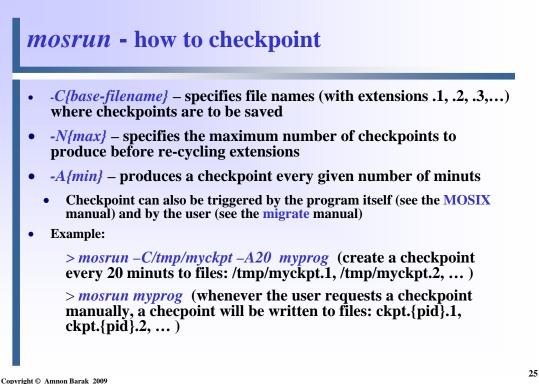


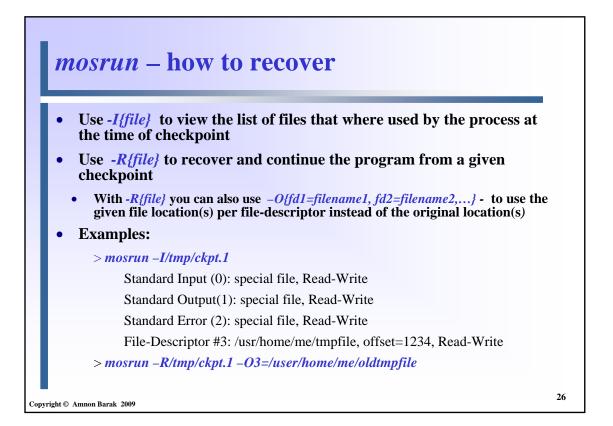


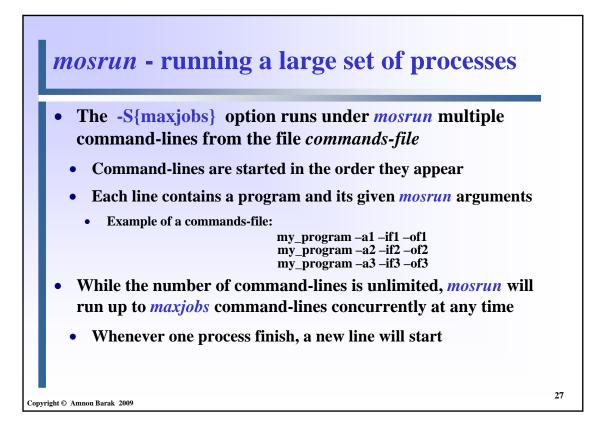


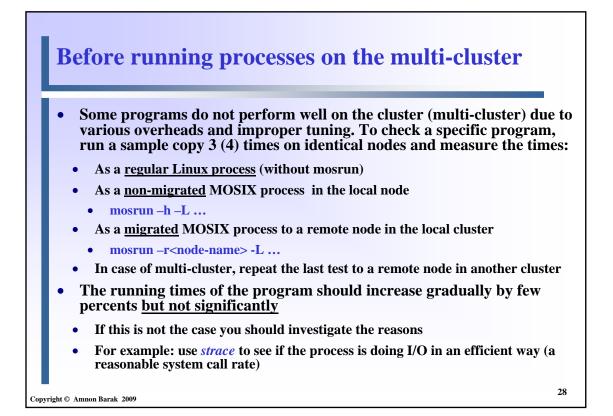


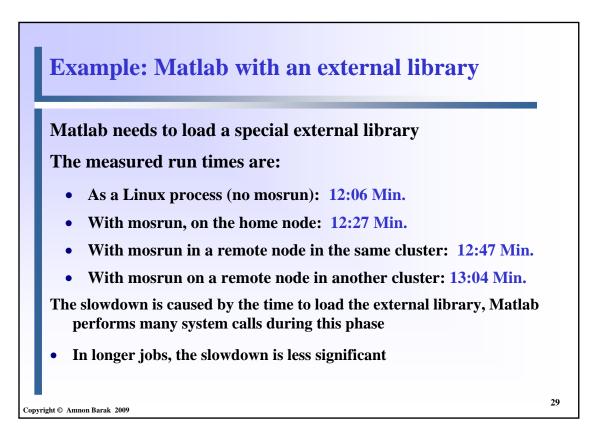


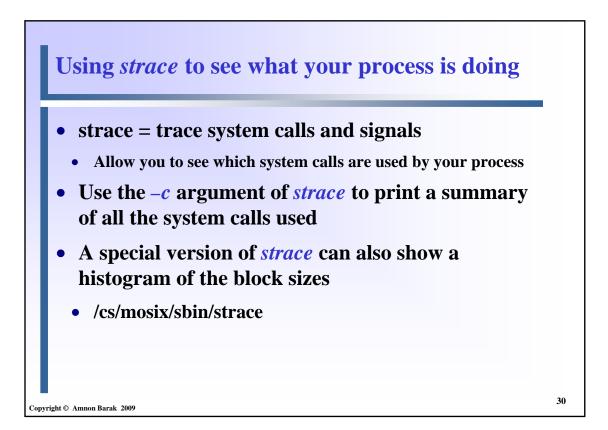












			-		
6 time	seconds	usecs/cal I	calls	errors syscall	
99. 78	16. 847495	17	1019960	2712 futex	
0. 17	0. 028892	95	303	275 open	
0.01	0. 002178	128	17	munmap	
0.01	0.002096	40	53	read	
0. 01	0. 001843	29	64	60 stat64	
0.01	0.000993	7	143	brk	
0.00	0. 000474	43	11	11 access	
0.00	0. 000370	10	38	old_mmap	
0.00	0.000300	12	25	mmap2	
0.00	0. 000236	9	27	close	
0.00	0. 000160	6	28	fstat64	
0.00	0. 000115	58	2	wri te	
0.00	0. 000087	12	7	mprotect	
0.00	0.000073	6	13	times	
0.00	0.000056	5	11	_II seek	
0.00	0. 000025	5	5	gettimeofday	
0.00	0.000020	20	1	clone	
0.00	0.000011	11	1	_sysctl	
0.00	0.000010	5	2	rt_si gacti on	

							tatio	
	on memory							
% time	seconds	usecs/cal l	calls	errors syscall				
78.32	0. 030613	99	309	281 open				
5.54	0.002167	22	98	83 stat64				
5.33	0.002083	130	16	munmap				
4.78	0.001868	37	51	read				
1.76	0.000689	5	138	brk				
0.87	0.000339	42	8	8 access				
0.84	0.000329	8	40	old_mma)			
0.68	0.000267	9	29	fstat64				
0.46	0.000178	7	27	cl ose				
0.37	0.000143	72	2	write				
0.28	0.000110	11	10	mprotec				
0.28	0.000108	5	21	mmap2				
0.13	0.000049	49	1	uname				
0.12	0.000046	4	13	times				
0.09	0.000034	3	11	_I I seek				
0.05	0.000018	18	1	set_thre	ead_area			
0.04	0.000017	3	5	gettime	ofday			
0.02	0.00008	8	1	_sysctl				
0.02	0.000007	4	2	rt_si ga	ti on			
0. 01	0.000005	5	1	futex				
	0. 039088			372 total				

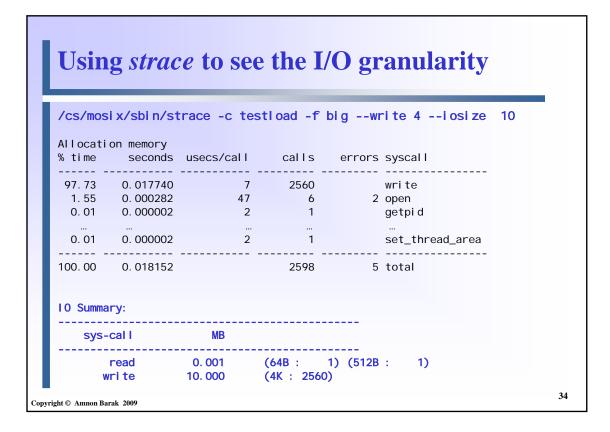


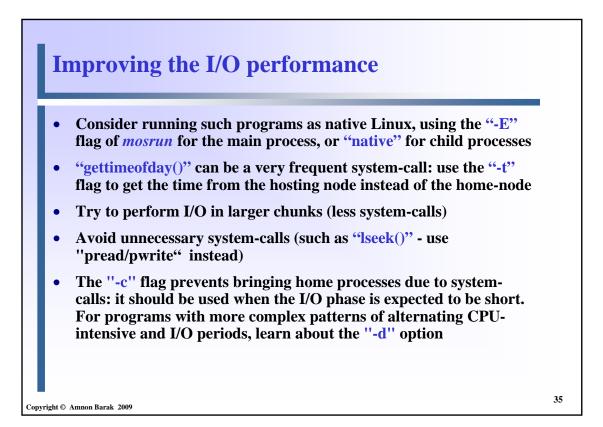
- MOSIX programs that issue a large number of system-calls or perform intensive I/O relative to the amount of computation are expensive because those operations are emulated
- When a process is running in a remote node, there is also overhead of sending those operations to the home-node over the network

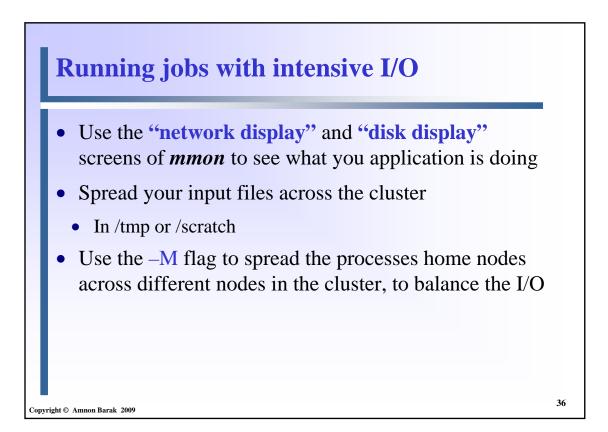
33

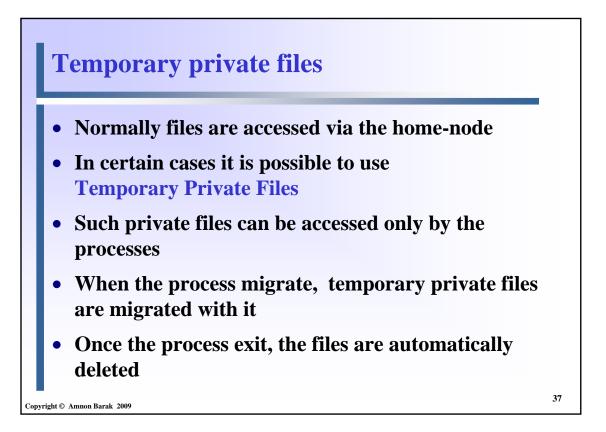
• Such processes will automatically be migrated back to the home-node, to eliminate the communication overhead

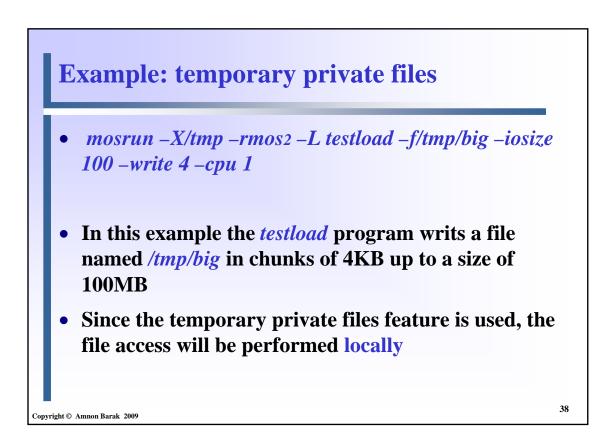
Copyright © Amnon Barak 2009

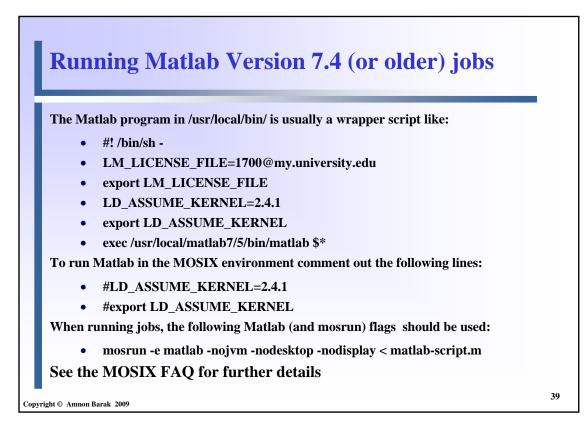


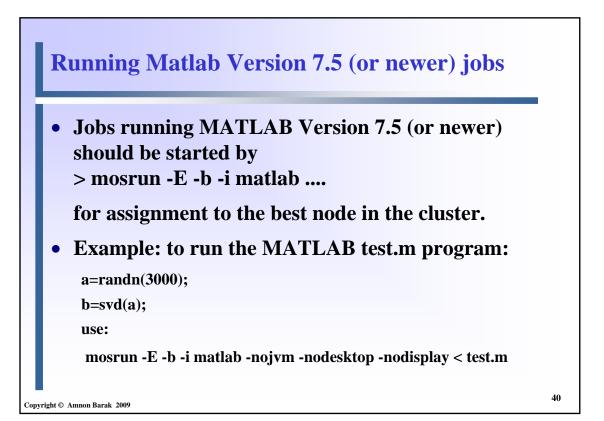


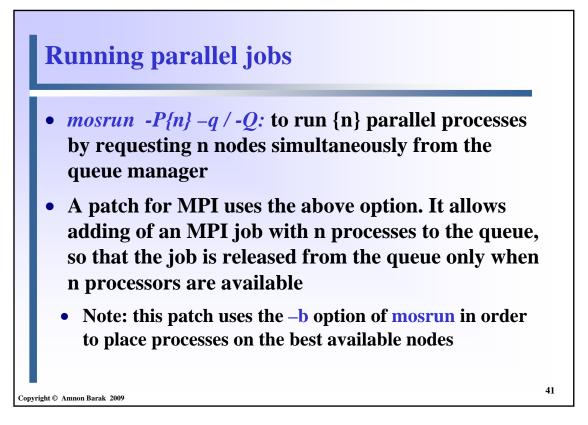


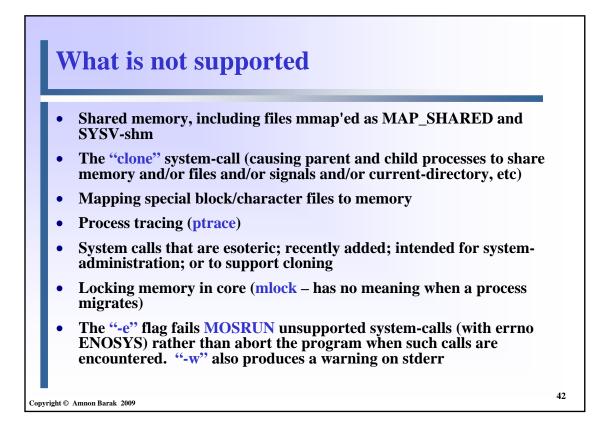


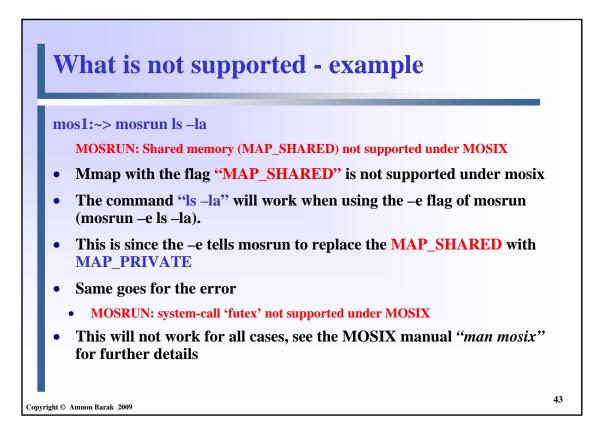


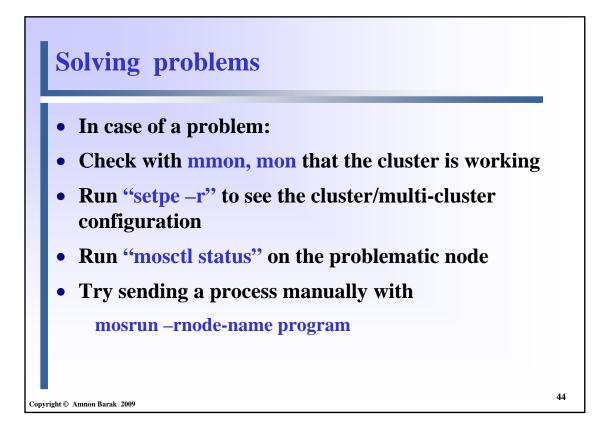


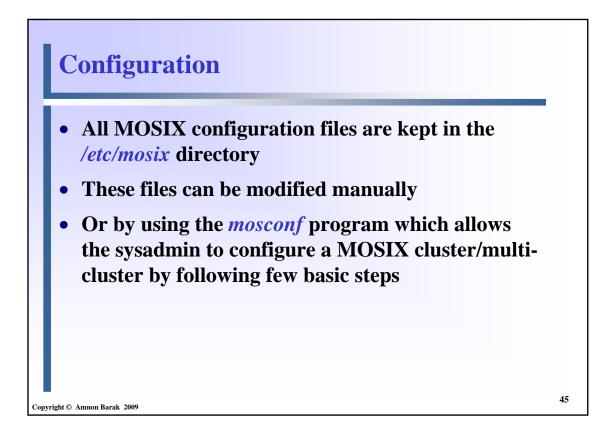


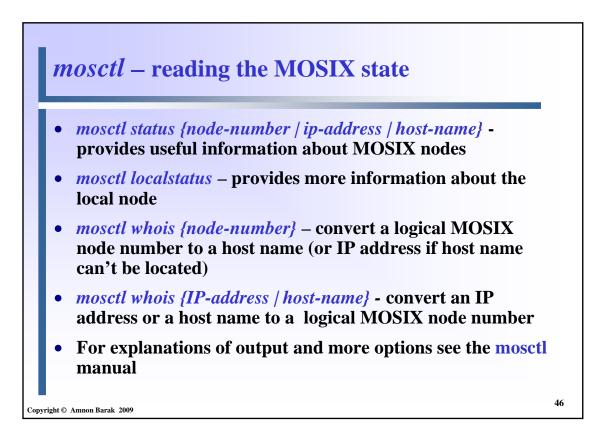








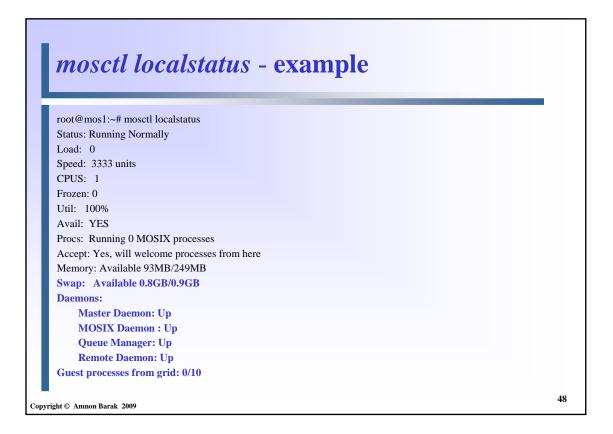




mosctl - example

> mosctl status
Status: Running Normally
Load: 0.29 (equivalent to about 0.29 CPU processes)
Speed: 10012 units
CPUS: 1
Frozen: 0
Util: 100%
Avail: YES
Procs: Running 1 MOSIX processes
Accept: Yes, will welcome processes from here
Memory: Available 903MB/1010MB

Copyright © Amnon Barak 2009



47



- setpe -r lists the nodes in the local cluster
- setpe -R lists the nodes in the local cluster and the multicluster
- Important information that may be listed:
 - pri={pri} priority we give to that cluster: the lower the better
 - proximate there is a very fast network connection to those nodes
 - outsider processes of class 0 cannot migrate there
 - dontgo local processes cannot migrate there
 - dont_take not accepting guests from there

Copyright © Amnon Barak 2009

setpe - example

mos1:~> setpe -R

This MOSIX node is: 132.65.174.1 (mos1.cs.huji.ac.il) (no features)

Nodes in this cluster:

mos1.cs.huji.ac.il - mos38.cs.huji.ac.il

Nodes from the rest of the multi-cluster:

mos51.cs.huji.ac.il - mos66.cs.huji.ac.il: pri=45,proximate,outsider

cmos-16.cs.huji.ac.il - cmos-20.cs.huji.ac.il: pri=45,proximate,outsider

vmos-02.cs.huji.ac.il - vmos-03.cs.huji.ac.il: pri=45,proximate,outsider

Copyright © Amnon Barak 2009

49

