

MOSIX

Clusters voor iedereen

Geert Vernaeve & Jan De Beule

Ghent University

Clusters voor iedereen?

- Standaard huis-, tuin- en keukenhardware



Clusters voor iedereen?

- Standaard huis-, tuin- en keukenhardware (d.i. x86 hardware)



Clusters voor iedereen?

- Standaard huis-, tuin- en keukenhardware (d.i. x86 hardware)
- Open source software



Wat is mosix?

Patch voor de Linux kernel die vooral *load balancing* aanbiedt.



Wat is mosix?

Patch voor de Linux kernel die vooral *load balancing* aanbiedt.

Processen kunnen *migreren* van een zwaar belaste machine naar een minder zwaar belaste machine.



Wat is mosix?

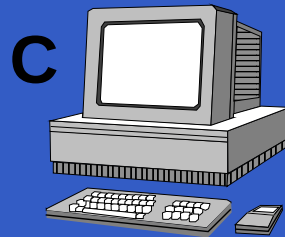
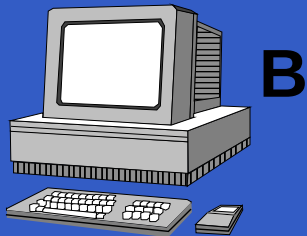
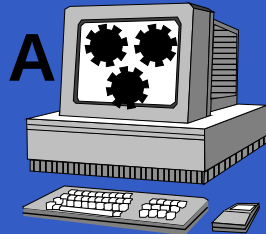
Patch voor de Linux kernel die vooral *load balancing* aanbiedt.

Processen kunnen *migreren* van een zwaar belaste machine naar een minder zwaar belaste machine.

Het geheel gedraagt zich ongeveer als een multiprocessor machine.



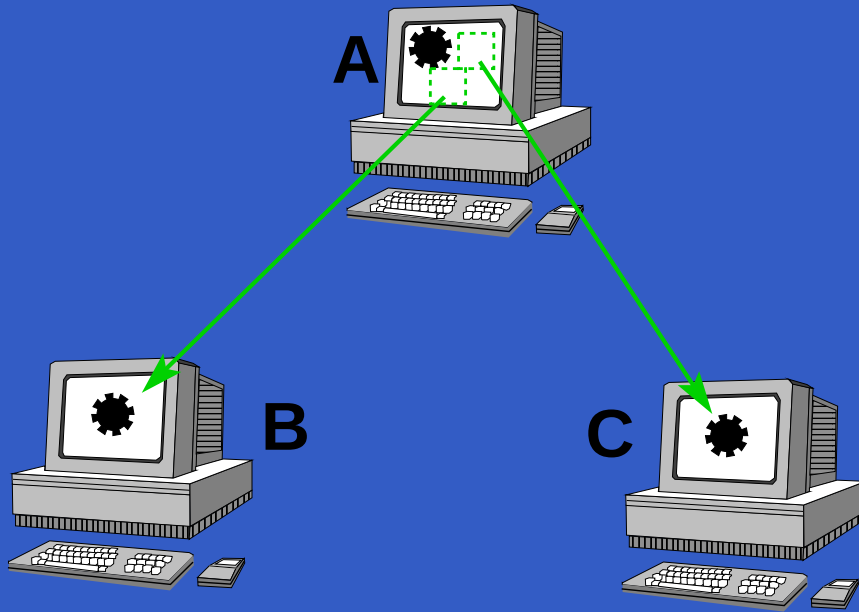
Migratie van processen



Stel dat er op A drie programma's draaien terwijl B en C niets staan te doen



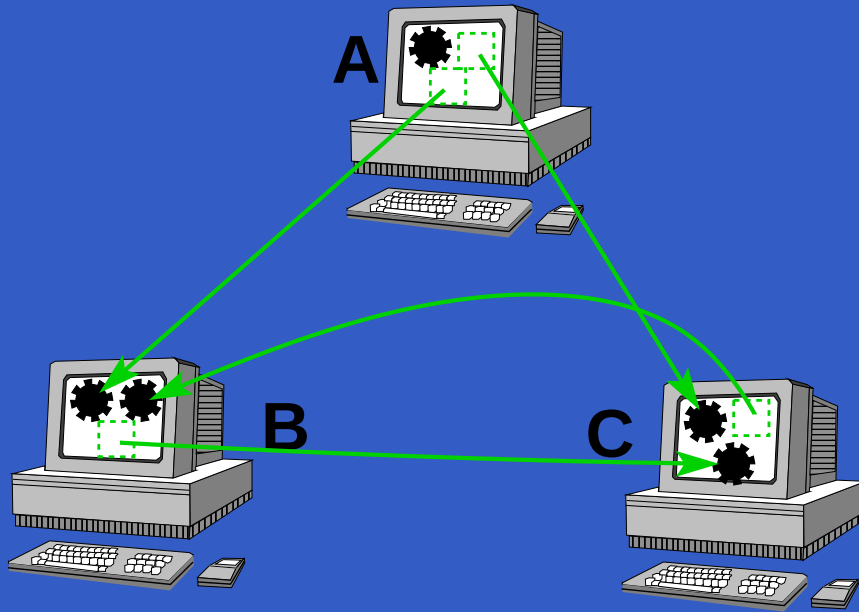
Migratie van processen



Processen zullen *migreren* naar B en C: *load balancing*. Er blijft telkens een *deputy* achter op A



Migratie van processen



Het kan altijd ingewikkelder: programma's kunnen ook op B en C gestart worden



Kenmerken

- Alle nodes zijn gelijkwaardig (geen 'master')



Kenmerken

- Alle nodes zijn gelijkwaardig (geen ‘master’)
- *Fork and forget*: programma’s hoeven niet speciaal aangepast te worden (*threads* worden evenwel niet verspreid maar blijven bij elkaar)



Kenmerken

- Alle nodes zijn gelijkwaardig (geen ‘master’)
- *Fork and forget*: programma's hoeven niet speciaal aangepast te worden
- Grote schaal (van 2 tot 2000 tot ... nodes)



Kenmerken

- Alle nodes zijn gelijkwaardig (geen ‘master’)
- *Fork and forget*: programma's hoeven niet speciaal aangepast te worden
- Grote schaal (van 2 tot 2000 tot ... nodes)
- Nodes mogen sterk verschillen (386 + dual Pentium IV; trage/snelle netwerkverbindingen)



Kenmerken

- Alle nodes zijn gelijkwaardig (geen ‘master’)
- *Fork and forget*: programma's hoeven niet speciaal aangepast te worden
- Grote schaal (van 2 tot 2000 tot ... nodes)
- Nodes mogen sterk verschillen (386 + dual Pentium IV; trage/snelle netwerkverbindingen)
- Dynamische load balancing : draaiende processen migreren over en weer als er een node meer of minder werk te doen krijgt



Kenmerken

- Alle nodes zijn gelijkwaardig (geen ‘master’)
- *Fork and forget*: programma’s hoeven niet speciaal aangepast te worden
- Grote schaal (van 2 tot 2000 tot ... nodes)
- Nodes mogen sterk verschillen (386 + dual Pentium IV; trage/snelle netwerkverbindingen)
- Dynamische load balancing
- Geen checkpointing (“niet de taak van de kernel”)



Vereisten

- x86 processor



Vereisten

- x86 processor
- identieke kernelversie



Vereisten

- x86 processor
- identieke kernelversie
- programma moet paralleliseerbaar zijn in verschillende *processen* (geen *threads*)



Vereisten

- x86 processor
- identieke kernelversie
- programma moet paralleliseerbaar zijn
- programma mag geen ‘speciale dingen’ doen (direkt I/O devices manipuleren, schrijfbaar shared memory, real-time scheduling, ...)



Vereisten

- x86 processor
- identieke kernelversie
- programma moet paralleliseerbaar zijn
- programma mag geen 'speciale dingen' doen
- voor het overige weinig beperkingen:
hardware mag sterk heterogeen zijn



Configureerbaar

- B.v. een 'parasitaire' node die wel processen naar de rest van de cluster weg migreert maar geen inkomende processen van andere nodes aanvaardt



Configureerbaar

- B.v. een 'parasitaire' node
- B.v. bureaucomputer om 6 PM volledig in de cluster schakelen, om 8 AM in 'parasitaire' stand



Configureerbaar

- B.v. een 'parasitaire' node
- B.v. bureaucomputer om 6 PM volledig in de cluster schakelen, om 8 AM in 'parasitaire' stand
- Diskless cluster: enkel één node heeft een disk; vanaf deze node worden alle processen gestart en de andere nodes staan enkel te wachten op gemigreerde processen



Configureerbaar

- B.v. een 'parasitaire' node
- B.v. bureaucomputer om 6 PM volledig in de cluster schakelen, om 8 AM in 'parasitaire' stand
- Diskless cluster: enkel één node heeft een disk; vanaf deze node worden alle processen gestart en de andere nodes staan enkel te wachten op gemigreerde processen
- . . .



-
-
-

Demo



Tools

- Mosix tools: mosmon, mosctl, mtop, ...



Tools

- Mosix tools: mosmon, mosctl, mtop, ...
- Standaard Unix tools: dsh, rsync (eventueel gecombineerd), ...



openMosix Summit 2005

openMosix Summit 2005

26–27 februari 2005

in Brussel als onderdeel van fosdem (zie
<http://www.fosdem.org/> en
<http://www.openmosix.org>)

