



Resource management in Unix

Revision 1.1

Larry McVoy

BitMover, Inc.
San Francisco, California
415-821-5758
lm@bitmover.com

ABSTRACT

The management of computer resources, such as disks, networks, CPUs, and memory, is suffering from outdated interfaces. Users may wish to have close control over the use of any or all of the above resources, but Unix does not provide the interfaces to control those resources. This note describes a set of resources which managers may wish to control, and describes a model for controlling them. This is an evolving document, with the intended goal of arriving at a suitable specification for modifications to the Linux system to provide better resource control.

1. Outline

- What are the resources?
- What is the model for controlling the resources?
- How would this work?

2. What are the resources?

What resources do we wish to control? The set of interesting resources are

- CPU time.
- Memory.
- Network utilization in terms of transfers and/or bandwidth.

- Disk utilization in terms of transfers and/or bandwidth.
- File system utilization in terms of transfers and/or bandwidth.
- File system free space.

3. What is the model?

3.1. Current models

The current resource management models are the following:

- a) Hard limits as defined by `setrlimit(2)`. When these limits are exceeded the process is killed.

