Subject: Re: [patch00/05]: Containers(V2)- Introduction
Posted by Christoph Lameter on Wed, 20 Sep 2006 17:15:52 GMT

View Forum Message <> Reply to Message

On Wed, 20 Sep 2006, Alan Cox wrote:

> Ar Mer, 2006-09-20 am 09:25 -0700, ysgrifennodd Christoph Lameter:
> > We already have such a functionality in the kernel its called a cpuset. A
> > container could be created simply by creating a fake node that then
> > allows constraining applications to this node. We already track the
> > types of pages per node. The statistics you want are already existing.
> > See /proc/zoneinfo and /sys/devices/system/node/node*/*.
>
> CPUsets don't appear to scale to large numbers of containers (say 5000,
> with 200-500 doing stuff at a time). They also don't appear to do any
> tracking of kernel side resource objects, which is critical to
> containment. Indeed for some of us the CPU management and user memory
> management angle is mostly uninteresting.

The scalability issues can certainly be managed. See the discussions on
linux-mm. Kernel side resource objects? slab pages? Those are tracked.

> I'm also not clear how you handle shared pages correctly under the fake
> node system, can you perhaps explain that further how this works for say
> a single apache/php/glibc shared page set across 5000 containers each a
> web site.

Cpusets can share nodes. I am not sure what the problem would be? Paul may
be able to give you more details.