

Kirill,

Thanks for posting the patches to ckrm-tech. I 'll look into it and post my comments tomorrow.

Some documentation (or pointer to the documentation) on how to use this feature and high level design would really help.

How does the hierarchy work ? (May be reading the code would clear it up :).

few comments below..

On Wed, 2006-08-16 at 19:24 +0400, Kirill Korotaev wrote:

<snip>

- > The patches in these series are:
- > diff-ubc-kconfig.patch:
- > Adds kernel/ub/Kconfig file with UBC options and
- > includes it into arch Kconfigs

Since the core functionality is arch independent, why not have the Kconfig stuff in some generic place like init/Kconfig ?

- >
- > diff-ubc-core.patch:
- > Contains core functionality and interfaces of UBC:
- > find/create beancounter, initialization,
- > charge/uncharge of resource, core objects' declarations.
- >
- > diff-ubc-task.patch:
- > Contains code responsible for setting UB on task,
- > it's inheriting and setting host context in interrupts.
- >
- > Task contains three beancounters:
- > 1. exec_ub - current context. all resources are charged
- > to this beancounter.
- > 2. task_ub - beancounter to which task_struct is charged
- > itself.
- > 3. fork_sub - beancounter which is inherited by
- > task's children on fork

wondering why we need three of these ?

- >
- > diff-ubc-syscalls.patch:

- > Patch adds system calls for UB management:
- > 1. sys_getluid - get current UB id
- > 2. sys_setluid - changes exec_ and fork_ UBs on current
- > 3. sys_setublimit - set limits for resources consumptions

I agree with Rohit that configfs based interface would be more easy to use (you will not get into the system call number issue that Greg has pointed too).

- >
- > diff-ubc-kmem-core.patch:
- > Introduces UB_KMEMSIZE resource which accounts kernel
- > objects allocated by task's request.
- >
- > Objects are accounted via struct page and slab objects.
- > For the latter ones each slab contains a set of pointers
- > corresponding object is charged to.
- >
- > Allocation charge rules:
- > 1. Pages - if allocation is performed with __GFP_UBC flag - page
- > is charged to current's exec_ub.
- > 2. Slabs - kmem_cache may be created with SLAB_UBC flag - in this
- > case each allocation is charged. Caches used by kmalloc are
- > created with SLAB_UBC | SLAB_UBC_NOCHARGE flags. In this case
- > only __GFP_UBC allocations are charged.
- >
- > diff-ubc-kmem-charge.patch:
- > Adds SLAB_UBC and __GFP_UBC flags in appropriate places
- > to cause charging/limiting of specified resources.
- >
- > diff-ubc-proc.patch:
- > Adds two proc entries user_beancounters and user_beancounters_sub
- > allowing to see current state (usage/limits/fails for each UB).
- > Implemented via seq files.

again, configfs would be easier.

- >
- > Patch set is applicable to 2.6.18-rc4-mm1
- >
- > Thanks,
- > Kirill
- >
- >
- > -----
- > Using Tomcat but need to do more? Need to support web services, security?
- > Get stuff done quickly with pre-integrated technology to make your job easier
- > Download IBM WebSphere Application Server v.1.0.1 based on Apache Geronimo
- > <http://sel.as-us.falkag.net/sel?cmd=lnk&kid=120709&b id=263057&dat=121642>

> _____
> ckrm-tech mailing list
> <https://lists.sourceforge.net/lists/listinfo/ckrm-tech>
--

Chandra Seetharaman | Be careful what you choose....
- sekharan@us.ibm.com |you may get it.
