
Subject: Re: [PATCH v8 2/5] ipc: add sysctl to specify desired next object id
Posted by [akpm](#) on Wed, 24 Oct 2012 21:41:23 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Wed, 24 Oct 2012 19:35:09 +0400

Stanislav Kinsbursky <skinsbursky@parallels.com> wrote:

```
> This patch adds 3 new variables and sysctls to tune them (by one "next_id"
> variable for messages, semaphores and shared memory respectively).
> This variable can be used to set desired id for next allocated IPC object.
> By default it's equal to -1 and old behaviour is preserved.
> If this variable is non-negative, then desired idr will be extracted from it
> and used as a start value to search for free IDR slot.
>
> Notes:
> 1) this patch doesn't guarantee, that new object will have desired id. So it's
> up to user space how to handle new object with wrong id.
> 2) After successful id allocation attempt, "next_id" will be set back to -1
> (if it was non-negative).
>
> --- a/ipc/ipc_sysctl.c
> +++ b/ipc/ipc_sysctl.c
> @@ -158,6 +158,7 @@ static int proc_ipcauto_dointvec_minmax(ctl_table *table, int write,
>
> static int zero;
> static int one = 1;
> +static int int_max = INT_MAX;
>
> static struct ctl_table ipc_kern_table[] = {
> {
> @@ -227,6 +228,33 @@ static struct ctl_table ipc_kern_table[] = {
> .extra1 = &zero,
> .extra2 = &one,
> },
> + {
> + .procname = "sem_next_id",
> + .data = &init_ipc_ns.ids[IPC_SEM_IDS].next_id,
> + .maxlen = sizeof(init_ipc_ns.ids[IPC_SEM_IDS].next_id),
> + .mode = 0644,
> + .proc_handler = proc_ipc_dointvec_minmax,
> + .extra1 = &zero,
> + .extra2 = &int_max,
> + },
> + {
> + .procname = "msg_next_id",
> + .data = &init_ipc_ns.ids[IPC_MSG_IDS].next_id,
> + .maxlen = sizeof(init_ipc_ns.ids[IPC_MSG_IDS].next_id),
> + .mode = 0644,
```

```
> + .proc_handler = proc_ipc_dointvec_minmax,  
> + .extra1 = &zero,  
> + .extra2 = &int_max,  
> + },  
> + {  
> + .procname = "shm_next_id",  
> + .data = &init_ipc_ns.ids[IPC_SHM_IDS].next_id,  
> + .maxlen = sizeof(init_ipc_ns.ids[IPC_SHM_IDS].next_id),  
> + .mode = 0644,  
> + .proc_handler = proc_ipc_dointvec_minmax,  
> + .extra1 = &zero,  
> + .extra2 = &int_max,  
> + },  
> {}  
> };
```

ipc_kern_table[] is (badly) documented in Documentation/sysctl/kernel.txt. Can we at least mention these controls in there? Better, create a new way of properly documenting each control and document these three in that manner? Better still, document all the other ones as well ;)

The patch adds these controls to CONFIG_CHECKPOINT_RESTORE=n kernels. Why is this?
