```
**********************************************************
   21193 Sun Jan 31 14:08:51 2016
new/usr/src/man/man5/resource_controls.5
6591 resource_controls(5) should talk about project.cpu-cap not project.cpu-caps
**********************************************************
'\" te
.\" Copyright (c) 2007, Sun Microsystems, Inc. All Rights Reserved.
.\" The contents of this file are subject to the terms of the Common Development
.\" You can obtain a copy of the license at usr/src/OPENSOLARIS.LICENSE or http:
.\" When distributing Covered Code, include this CDDL HEADER in each file and in
.TH RESOURCE_CONTROLS 5 "Jul 19, 2013"
.SH NAME
resource_controls \- resource controls available through project database
.SH DESCRIPTION
.sp
.LP
The resource controls facility is configured through the project database. See
\fBproject\fR(4). You can set and modify resource controls through the
following utilities:
.RS +4
.TP
.ie t \(bu
.el o
\fBprctl\fR(1)
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBprojadd\fR(1M)
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBprojmod\fR(1M)
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBrctladm\fR(1M)
.RE
.sp
.LP
In a program, you use \fBsetrctl\fR(2) to set resource control values.
.sp
.LP
In addition to the preceding resource controls, there are resource pools,
accessible through the \fBpooladm\fR(1M) and \fBpoolcfg\fR(1M) utilities. In a
program, resource pools can be manipulated through the \fBlibpool\fR(3LIB)
library.
.sp
.LP
The following are the resource controls are available:
.sp
.ne 2
.na
\fB\fBprocess.max-address-space\fR\fR
.ad
.sp .6
.RS 4n
Maximum amount of address space, as summed over segment sizes, that is
available to this process, expressed as a number of bytes.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-core-size\fR\fR
.ad
.sp .6
.RS 4n
Maximum size of a core file created by this process, expressed as a number of
bytes.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-cpu-time\fR\fR
.ad
.sp .6
.RS 4n
Maximum CPU time that is available to this process, expressed as a number of
seconds.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-data-size\fR\fR
.ad
.sp .6
.RS 4n
Maximum heap memory available to this process, expressed as a number of bytes.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-file-descriptor\fR\fR
.ad
.sp .6
.RS 4n
Maximum file descriptor index available to this process, expressed as an
integer.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-file-size\fR\fR
.ad
.sp .6
.RS 4n
Maximum file offset available for writing by this process, expressed as a
number of bytes.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-msg-messages\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of messages on a message queue (value copied from the resource
control at \fBmsgget()\fR time), expressed as an integer.
.RE

.sp
```

```
.ne 2
.na
\fB\fBprocess.max-msg-qbytes\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of bytes of messages on a message queue (value copied from the
resource control at \fBmsgget()\fR time), expressed as a number of bytes.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-port-events\fR\fR
.ad
.sp .6
.RS 4n
Maximum allowable number of events per event port, expressed as an integer.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-sem-nsems\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of semaphores allowed per semaphore set, expressed as an
integer.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-sem-ops\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of semaphore operations allowed per \fBsemop\fR call (value
copied from the resource control at \fBsemget()\fR time). Expressed as an
integer, specifying the number of operations.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-sigqueue-size\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of outstanding queued signals.
.RE

.sp
.ne 2
.na
\fB\fBprocess.max-stack-size\fR\fR
.ad
.sp .6
.RS 4n
Maximum stack memory segment available to this process, expressed as a number
of bytes.
.RE

.sp
.ne 2
```

```
.na
\fB\fBproject.cpu-cap\fR\fR
\fB\fBproject.cpu-caps\fR\fR
.ad
.sp .6
.RS 4n
Maximum amount of CPU resources that a project can use. The unit used is the
percentage of a single CPU that can be used by all user threads in a project.
Expressed as an integer. The cap does not apply to threads running in real-time
scheduling class. This resource control does not support the \fBsyslog\fR
action.
.RE

.sp
.ne 2
.na
\fB\fBproject.cpu-shares\fR\fR
.ad
.sp .6
.RS 4n
Number of CPU shares granted to a project for use with the fair share scheduler
(see \fBFSS\fR(7)). The unit used is the number of shares (an integer). This
resource control does not support the \fBsyslog\fR action.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-contracts\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of contracts allowed in a project, expressed as an integer.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-crypto-memory\fR\fR
.ad
.sp .6
.RS 4n
Maximum amount of kernel memory that can be used for crypto operations.
Allocations in the kernel for buffers and session-related structures are
charged against this resource control.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-locked-memory\fR\fR
.ad
.sp .6
.RS 4n
Total amount of physical memory locked by device drivers and user processes
(including D/ISM), expressed as a number of bytes.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-lwps\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of LWPs simultaneously available to a project, expressed as an
```

```
integer.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-msg-ids\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of message queue IDs allowed for a project, expressed as an
integer.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-port-ids\fR\fR
.ad
.sp .6
.RS 4n
Maximum allowable number of event ports, expressed as an integer.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-sem-ids\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of semaphore IDs allowed for a project, expressed as an integer.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-shm-ids\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of shared memory IDs allowed for a project, expressed as an
integer.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-shm-memory\fR\fR
.ad
.sp .6
.RS 4n
Total amount of shared memory allowed for a project, expressed as a number of
bytes.
.RE

.sp
.ne 2
.na
\fB\fBproject.max-tasks\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of tasks allowable in a project, expressed as an integer.
.RE
```

```
.sp
.ne 2
.na
\fB\fBproject.pool\fR\fR
.ad
.sp .6
.RS 4n
Binds a specified resource pool with a project.
.RE

.sp
.ne 2
.na
\fB\fBrcap.max-rss\fR\fR
.ad
.sp .6
.RS 4n
The total amount of physical memory, in bytes, that is available to processes
in a project.
.RE

.sp
.ne 2
.na
\fB\fBtask.max-cpu-time\fR\fR
.ad
.sp .6
.RS 4n
Maximum CPU time that is available to this task’s processes, expressed as a
number of seconds.
.RE

.sp
.ne 2
.na
\fB\fBtask.max-lwps\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of LWPs simultaneously available to this task’s processes,
expressed as an integer.
.RE

.sp
.LP
The following zone-wide resource controls are available:
.sp
.ne 2
.na
\fB\fBzone.cpu-cap\fR\fR
.ad
.sp .6
.RS 4n
Sets a limit on the amount of CPU time that can be used by a zone. The unit
used is the percentage of a single CPU that can be used by all user threads in
a zone. Expressed as an integer. When projects within the capped zone have
their own caps, the minimum value takes precedence. This resource control does
not support the \fBsyslog\fR action.
.RE

.sp
.ne 2
.na
\fB\fBzone.cpu-shares\fR\fR
.ad
.sp .6
```

```
.RS 4n
Sets a limit on the number of fair share scheduler (FSS) CPU shares for a zone.
CPU shares are first allocated to the zone, and then further subdivided among
projects within the zone as specified in the \fBproject.cpu-shares\fR entries.
Expressed as an integer. This resource control does not support the
\fBsyslog\fR action.
.RE

.sp
.ne 2
.na
\fB\fBzone.max-locked-memory\fR\fR
.ad
.sp .6
.RS 4n
Total amount of physical locked memory available to a zone.
.RE

.sp
.ne 2
.na
\fB\fBzone.max-lwps\fR\fR
.ad
.sp .6
.RS 4n
Enhances resource isolation by preventing too many LWPs in one zone from
affecting other zones. A zone’s total LWPs can be further subdivided among
projects within the zone within the zone by using \fBproject.max-lwps\fR
entries. Expressed as an integer.
.RE

.sp
.ne 2
.na
\fB\fBzone.max-msg-ids\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of message queue IDs allowed for a zone, expressed as an
integer.
.RE

.sp
.ne 2
.na
\fB\fBzone.max-sem-ids\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of semaphore IDs allowed for a zone, expressed as an integer.
.RE

.sp
.ne 2
.na
\fB\fBzone.max-shm-ids\fR\fR
.ad
.sp .6
.RS 4n
Maximum number of shared memory IDs allowed for a zone, expressed as an
integer.
.RE

.sp
.ne 2
.na
```

```
\fB\fBzone.max-shm-memory\fR\fR
.ad
.sp .6
.RS 4n
Total amount of shared memory allowed for a zone, expressed as a number of
bytes.
.RE

.sp
.ne 2
.na
\fB\fBzone.max-swap\fR\fR
.ad
.sp .6
.RS 4n
Total amount of swap that can be consumed by user process address space
mappings and \fBtmpfs\fR mounts for this zone.
.RE

.sp
.LP
See \fBzones\fR(5).
.SS "Units Used in Resource Controls"
.sp
.LP
Resource controls can be expressed as in units of size (bytes), time (seconds),
or as a count (integer). These units use the strings specified below.
.sp
.in +2
.nf
Category             Res Ctrl      Modifier  Scale
                     Type String
-----------          -----------   --------  -----
Size                 bytes         B         1
                                   KB        2^10
                                   MB        2^20
                                   GB        2^30
                                   TB        2^40
                                   PB        2^50
                                   EB        2^60

Time                 seconds       s         1
                                   Ks        10^3
                                   Ms        10^6
                                   Gs        10^9
                                   Ts        10^12
                                   Ps        10^15
                                   Es        10^18

Count                integer       none      1
                                   K         10^3
                                   M         10^6
                                   G         10^9
                                   T         10^12
                                   P         10^15
                                   Es        10^18
.fi
.in -2

.sp
.LP
Scaled values can be used with resource controls. The following example shows a
scaled threshold value:
.sp
.in +2
.nf
```

```
task.max-lwps=(priv,1K,deny)
.fi
.in -2

.sp
.LP
In the \fBproject\fR file, the value \fB1K\fR is expanded to \fB1000\fR:
.sp
.in +2
.nf
task.max-lwps=(priv,1000,deny)
.fi
.in -2

.sp
.LP
A second example uses a larger scaled value:
.sp
.in +2
.nf
process.max-file-size=(priv,5G,deny)
.fi
.in -2

.sp
.LP
In the \fBproject\fR file, the value \fB5G\fR is expanded to \fB5368709120\fR:
.sp
.in +2
.nf
process.max-file-size=(priv,5368709120,deny)
.fi
.in -2

.sp
.LP
The preceding examples use the scaling factors specified in the table above.
.sp
.LP
Note that unit modifiers (for example, \fB5G\fR) are accepted by the
\fBprctl\fR(1), \fBprojadd\fR(1M), and \fBprojmod\fR(1M) commands. You cannot
use unit modifiers in the project database itself.
.SS "Resource Control Values and Privilege Levels"
.sp
.LP
A threshold value on a resource control constitutes a point at which local
actions can be triggered or global actions, such as logging, can occur.
.sp
.LP
Each threshold value on a resource control must be associated with a privilege
level. The privilege level must be one of the following three types:
.sp
.ne 2
.na
\fB\fBbasic\fR\fR
.ad
.sp .6
.RS 4n
Can be modified by the owner of the calling process.
.RE

.sp
.ne 2
.na
\fB\fBprivileged\fR\fR
.ad
```

```
.sp .6
.RS 4n
Can be modified by the current process (requiring \fBsys_resource\fR privilege)
or by \fBprctl\fR(1) (requiring \fBproc_owner\fR privilege).
.RE

.sp
.ne 2
.na
\fB\fBsystem\fR\fR
.ad
.sp .6
.RS 4n
Fixed for the duration of the operating system instance.
.RE

.sp
.LP
A resource control is guaranteed to have one \fBsystem\fR value, which is
defined by the system, or resource provider. The \fBsystem\fR value represents
how much of the resource the current implementation of the operating system is
capable of providing.
.sp
.LP
Any number of privileged values can be defined, and only one basic value is
allowed. Operations that are performed without specifying a privilege value are
assigned a basic privilege by default.
.sp
.LP
The privilege level for a resource control value is defined in the privilege
field of the resource control block as \fBRCTL_BASIC\fR, \fBRCTL_PRIVILEGED\fR,
or \fBRCTL_SYSTEM\fR. See \fBsetrctl\fR(2) for more information. You can use
the \fBprctl\fR command to modify values that are associated with basic and
privileged levels.
.sp
.LP
In specifying the privilege level of \fBprivileged\fR, you can use the
abbreviation \fBpriv\fR. For example:
.sp
.in +2
.nf
task.max-lwps=(priv,1K,deny)
.fi
.in -2

.SS "Global and Local Actions on Resource Control Values"
.sp
.LP
There are two categories of actions on resource control values: global and
local.
.sp
.LP
Global actions apply to resource control values for every resource control on
the system. You can use \fBrctladm\fR(1M) to perform the following actions:
.RS +4
.TP
.ie t \(bu
.el o
Display the global state of active system resource controls.
.RE
.RS +4
.TP
.ie t \(bu
.el o
Set global logging actions.
.RE
```

```
.sp
.LP
You can disable or enable the global logging action on resource controls. You
can set the \fBsyslog\fR action to a specific degree by assigning a severity
level, \fBsyslog=\fR\fIlevel\fR. The possible settings for \fIlevel\fR are as
follows:
.RS +4
.TP
.ie t \(bu
.el o
\fBdebug\fR
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBinfo\fR
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBnotice\fR
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBwarning\fR
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBerr\fR
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBcrit\fR
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBalert\fR
.RE
.RS +4
.TP
.ie t \(bu
.el o
\fBemerg\fR
.RE
.sp
.LP
By default, there is no global logging of resource control violations.
.sp
.LP
Local actions are taken on a process that attempts to exceed the control value.
For each threshold value that is placed on a resource control, you can
associate one or more actions. There are three types of local actions:
\fBnone\fR, \fBdeny\fR, and \fBsignal=\fR. These three actions are used as
follows:
.sp
.ne 2
```

```
.na
\fB\fBnone\fR\fR
.ad
.sp .6
.RS 4n
No action is taken on resource requests for an amount that is greater than the
threshold. This action is useful for monitoring resource usage without
affecting the progress of applications. You can also enable a global message
that displays when the resource control is exceeded, while, at the same time,
the process exceeding the threshhold is not affected.
.RE

.sp
.ne 2
.na
\fB\fBdeny\fR\fR
.ad
.sp .6
.RS 4n
You can deny resource requests for an amount that is greater than the
threshold. For example, a \fBtask.max-lwps\fR resource control with action deny
causes a \fBfork()\fR system call to fail if the new process would exceed the
control value. See the \fBfork\fR(2).
.RE

.sp
.ne 2
.na
\fB\fBsignal=\fR\fR
.ad
.sp .6
.RS 4n
You can enable a global signal message action when the resource control is
exceeded. A signal is sent to the process when the threshold value is exceeded.
Additional signals are not sent if the process consumes additional resources.
Available signals are listed below.
.RE

.sp
.LP
Not all of the actions can be applied to every resource control. For example, a
process cannot exceed the number of CPU shares assigned to the project of which
it is a member. Therefore, a deny action is not allowed on the
\fBproject.cpu-shares\fR resource control.
.sp
.LP
Due to implementation restrictions, the global properties of each control can
restrict the range of available actions that can be set on the threshold value.
(See \fBrctladm\fR(1M).) A list of available signal actions is presented in the
following list. For additional information about signals, see
\fBsignal\fR(3HEAD).
.sp
.LP
The following are the signals available to resource control values:
.sp
.ne 2
.na
\fB\fBSIGABRT\fR\fR
.ad
.sp .6
.RS 4n
Terminate the process.
.RE

.sp
.ne 2
```

```
.na
\fB\fBSIGHUP\fR\fR
.ad
.sp .6
.RS 4n
Send a hangup signal. Occurs when carrier drops on an open line. Signal sent to
the process group that controls the terminal.
.RE

.sp
.ne 2
.na
\fB\fBSIGTERM\fR\fR
.ad
.sp .6
.RS 4n
Terminate the process. Termination signal sent by software.
.RE

.sp
.ne 2
.na
\fB\fBSIGKILL\fR\fR
.ad
.sp .6
.RS 4n
Terminate the process and kill the program.
.RE

.sp
.ne 2
.na
\fB\fBSIGSTOP\fR\fR
.ad
.sp .6
.RS 4n
Stop the process. Job control signal.
.RE

.sp
.ne 2
.na
\fB\fBSIGXRES\fR\fR
.ad
.sp .6
.RS 4n
Resource control limit exceeded. Generated by resource control facility.
.RE

.sp
.ne 2
.na
\fB\fBSIGXFSZ\fR\fR
.ad
.sp .6
.RS 4n
Terminate the process. File size limit exceeded. Available only to resource
controls with the \fBRCTL_GLOBAL_FILE_SIZE\fR property
(\fBprocess.max-file-size\fR). See \fBrctlblk_set_value\fR(3C).
.RE

.sp
.ne 2
.na
\fB\fBSIGXCPU\fR\fR
.ad
```

```
.sp .6
.RS 4n
Terminate the process. CPU time limit exceeded. Available only to resource
controls with the \fBRCTL_GLOBAL_CPUTIME\fR property
(\fBprocess.max-cpu-time\fR). See \fBrctlblk_set_value\fR(3C).
.RE

.SS "Resource Control Flags and Properties"
.sp
.LP
Each resource control on the system has a certain set of associated properties.
This set of properties is defined as a set of flags, which are associated with
all controlled instances of that resource. Global flags cannot be modified, but
the flags can be retrieved by using either \fBrctladm\fR(1M) or the
\fBsetrctl\fR(2) system call.
.sp
.LP
Local flags define the default behavior and configuration for a specific
threshold value of that resource control on a specific process or process
collective. The local flags for one threshold value do not affect the behavior
of other defined threshold values for the same resource control. However, the
global flags affect the behavior for every value associated with a particular
control. Local flags can be modified, within the constraints supplied by their
corresponding global flags, by the \fBprctl\fR command or the \fBsetrctl\fR
system call. See \fBsetrctl\fR(2).
.sp
.LP
For the complete list of local flags, global flags, and their definitions, see
\fBrctlblk_set_value\fR(3C).
.sp
.LP
To determine system behavior when a threshold value for a particular resource
control is reached, use \fBrctladm\fR to display the global flags for the
resource control . For example, to display the values for
\fBprocess.max-cpu-time\fR, enter:
.sp
.in +2
.nf
$ rctladm process.max-cpu-time
process.max-cpu-time  syslog=off [ lowerable no-deny cpu-time inf seconds ]
.fi
.in -2

.sp
.LP
The global flags indicate the following:
.sp
.ne 2
.na
\fB\fBlowerable\fR\fR
.ad
.sp .6
.RS 4n
Superuser privileges are not required to lower the privileged values for this
control.
.RE

.sp
.ne 2
.na
\fB\fBno-deny\fR\fR
.ad
.sp .6
.RS 4n
Even when threshold values are exceeded, access to the resource is never
denied.
```

```
.RE

.sp
.ne 2
.na
\fB\fBcpu-time\fR\fR
.ad
.sp .6
.RS 4n
\fBSIGXCPU\fR is available to be sent when threshold values of this resource
are reached.
.RE

.sp
.ne 2
.na
\fB\fBseconds\fR\fR
.ad
.sp .6
.RS 4n
The time value for the resource control.
.RE

.sp
.LP
Use the \fBprctl\fR command to display local values and actions for the
resource control. For example:
.sp
.in +2
.nf
$ prctl -n process.max-cpu-time $$
    process 353939: -ksh
    NAME    PRIVILEGE    VALUE    FLAG   ACTION              RECIPIENT
 process.max-cpu-time
         privileged   18.4Es    inf   signal=XCPU                 -
         system       18.4Es    inf   none
.fi
.in -2

.sp
.LP
The \fBmax\fR (\fBRCTL_LOCAL_MAXIMAL\fR) flag is set for both threshold values,
and the \fBinf\fR (\fBRCTL_GLOBAL_INFINITE\fR) flag is defined for this
resource control. An \fBinf\fR value has an infinite quantity. The value is
never enforced. Hence, as configured, both threshold quantities represent
infinite values that are never exceeded.
.SS "Resource Control Enforcement"
.sp
.LP
More than one resource control can exist on a resource. A resource control can
exist at each containment level in the process model. If resource controls are
active on the same resource at different container levels, the smallest
container’s control is enforced first. Thus, action is taken on
\fBprocess.max-cpu-time\fR before \fBtask.max-cpu-time\fR if both controls are
encountered simultaneously.
.SH ATTRIBUTES
.sp
.LP
See \fBattributes\fR(5) for a description of the following attributes:
.sp

.sp
.TS
box;
c | c
l | l .
```

```
ATTRIBUTE TYPE  ATTRIBUTE VALUE
_
Interface Stability     Evolving
.TE

.SH SEE ALSO
.sp
.LP
\fBprctl\fR(1), \fBpooladm\fR(1M), \fBpoolcfg\fR(1M), \fBprojadd\fR(1M),
\fBprojmod\fR(1M), \fBrctladm\fR(1M), \fBsetrctl\fR(2),
\fBrctlblk_set_value\fR(3C), \fBlibpool\fR(3LIB), \fBproject\fR(4),
\fBattributes\fR(5), \fBFSS\fR(7)
.sp
.LP
\fISystem Administration Guide:  Virtualization Using the Solaris Operating
System\fR
```