AIX Micro-Partitioning (part 2)

Mark Cohen Austrowiek - EPV Technologies

In Part 1 an overview of Power Hypervisor, Hardware Manager Console, partition profiles, Micro-Partitioning, entitled capacity, virtual CPUs, logical CPUs, Symmetric Multi Processing and Partition Load Manager has been presented.
In Part 2 performance tools as vmstat, lparstat, mpstat and performance case examples are discussed.

3 Native Performance Utilities

3.1 lparstat

The *lparstat* utility is a new utility introduced with the AIX 5.3 operative system. This command provides a report of LPAR related information and utilization statistics. This command provides a display of current LPAR related parameters and Hypervisor information, as well as utilization statistics for the LPAR.

The various options of *lparstat* are exclusive of each other. The *lparstat* command with no options will generate two rows of statistics. The first row, included once when the command starts and again whenever there is a change in the system configuration, displays the System Configuration; the second row contains the Utilization Statistics which will be displayed in intervals and at any time the values of these metrics are deltas from pervious interval.

If the -h option is specified, the report will include summary statistics related to the Hypervisor.
If an Interval and Count are specified, a report is repeated for every Interval seconds and for Count iterations.

The following information is displayed in the system configuration row:

Statistic	Description
type	Partition Type. Can be either dedicated or shared.
mode	Indicates whether the partition processor capacity is capped or uncapped allowing it to consume idle cycles from the shared pool. Dedicated LPAR is implicitly capped.
smt	Indicates whether simultaneous multi threading is enabled or disabled in the partition.
lcpu	Number of online logical processors.
mem	Online Memory Capacity.
psize	Number of online physical processors in the pool.
ent	Entitled processing capacity in processor units. This information is displayed only if the partition type is shared.

Table 3

The following information is displayed in the utilization row:

Statistic	Description
%user	Shows the percentage of the entitled processing capacity used while executing at the user level (application).
%sys	Shows the percentage of the entitled processing capacity used while executing at the system

	level (kernel).
%idle	Shows the percentage of the entitled processing capacity unused while the partition was idle and did not have any outstanding disk I/O request.
%wait	Shows the percentage of the entitled processing capacity unused while the partition was idle and had outstanding disk I/O request(s). For dedicated partitions, the entitled processing capacity is the number of physical processors.

Table 4

The following statistics are displayed only when the partition type is shared[1]:

Statistic	Description
physc	Shows the number of physical processors consumed.
%entc	Shows the percentage of the entitled capacity consumed.
lbusy	Shows the percentage of logical processor(s) utilization that occurred while executing the user and system level.
app	Shows the available physical processors in the shared pool.
phint	Shows the number of phantom (targeted to another shared partition in this pool) interruptions received.

Table 5

The following statistics are displayed only when the -h flag is specified:

Statistic	Description
%hypv	Shows the percentage of time spent in hypervisor.
Hcalls	Shows number of hypervisor calls executed.

Table 6

More detailed information about the system configuration can be obtained using the -i option; they are listed below[2]:

Statistic	Description
Partition Name	Logical partition name as assigned at the HMC.
Partition Number	Number of this Logical Partition.
Online Virtual CPU	Number of CPUs (virtual engines) currently online.
Maximum Virtual CPUs	Maximum possible number of CPUs (virtual engines).
Online Memory	Amount of memory currently online.
Maximum Memory	Maximum possible amount of Memory.
Type	Indication whether the LPAR is using dedicated or shared CPU resource.
Mode	Indication whether the LPAR processor capacity is capped, or if it is uncapped and allowed to consume idle cycle from the shared pool. Dedicated LPAR is implicitly capped.
Entitled Capacity	The number of processing units this LPAR is entitled to receive.
Variable Capacity Weight	The priority weight assigned to this LPAR which controls how extra (idle) capacity is allocated to it. A weight of -1 indicates a soft cap is in place.
Minimum Capacity	The minimum number of processing units this LPAR was defined to ever have. Entitled capacity can be reduced down to this value.
Maximum Capacity	The maximum number of processing units this LPAR was defined to ever have.

[1] To obtain the *app* field one needs to authorize the logical partition profile to provide this information by means of the HMC.

[2] Interval and Count cannot be used with the –i option.

	Entitled capacity can be increased to this value.
Capacity Increment	The granule at which changes to Entitled Capacity can be made. A value in whole multiples indicates a Dedicated LPAR.
Maximum Dispatch Latency	The maximum number of nanoseconds between dispatches of the LPARs virtual CPUs.
Maximum Physical CPUs in System	The maximum possible number of physical CPUs in the system containing this LPAR.
Active Physical CPUs in System	The current number of active physical CPUs in the system containing this LPAR
Active CPUs in Pool	The current number of physical CPUs in the shared processor pool being used by this LPAR. (i.e. online physical processors in pool)
Unallocated Capacity	Number of processor units currently unallocated in the shared processor pool being used by this LPAR.
Physical CPU Percentage	Fractional representation relative to whole physical CPUs that these LPARs virtual CPUs equate to. This is a function of Entitled Capacity / Online CPUs. Dedicated LPARs would have 100% Physical CPU Percentage. A 4 way virtual with Entitled Capacity of 2 processor units would have a 50% physical CPU Percentage.
Minimum Memory	Minimum memory this LPAR was defined to ever have.
Minimum Virtual CPUs	Minimum number of virtual CPUs this LPAR was defined to ever have.
Unallocated Weight	Number of variable processor capacity weight units currently unallocated within the LPAR group.
Minimum Virtual Processor Required Capacity	The minimum entitled capacity required by the operating system, for each online virtual CPUs.
Partition Group ID	LPAR group that this LPAR is a member of. Shared Pool ID Identifier of Shared Pool of Physical processors that this LPAR is a member.

Table 7

The –H option provides detailed Hypervisor information. This option basically displays the statistics for each of the Hypervisor calls. The various details displayed by the –H option are listed below:

Statistic	Description
Call	Hypervisor call type
Number of calls	Number of Hypervisor calls made.
Total Time Spent	Percentage of total time spent in this type of call.
Hypervisor Time Spent	Percentage of Hypervisor time spent in this type of call.
Average Call Time	Average call time for this type of call in nanoseconds.
Maximum Call Time	Maximum call time for this type of call in nanoseconds.

Table 8

3.2 mpstat

The *mpstat* command collects and displays performance statistics for all logical CPUs in the system. Users can define both, the number of times the statistics are displayed, and the interval at which the data is updated.

When the *mpstat* command is invoked, it displays two sections of statistics. The first row, included once when the command starts and again whenever there is a change in the system configuration, displays the System Configuration; the second row contains the Utilization Statistics which will be displayed in intervals and at any time the values of these metrics are deltas from pervious interval.

When the -s flag is specified, the *mpstat* command reports simultaneous multi-threading (SMT) utilization, if it is enabled. This report displays the Virtual CPU Engines Utilization and the utilization of each thread (logical CPU) associated with the Virtual CPU engine.

Table 9 shows an example of virtual processors and logical CPUs utilization information obtained using the *mpstat –s* command for a SMT enabled LPAR. Proc0 and Proc1 are virtual CPUs while cpu0, cpu1, cpu2 and cpu3 are logical CPUs.

Proc0	cpu0	cpu1	Proc1	cpu2	cpu3
0,27%	0,17%	0,10%	49,63%	3,14%	46,49%

Table 9

All the detailed options can be located in the man utility .

3.3 vmstat

The *vmstat* utility has been on UNIX for many years. Now the utility has been enhanced to support the Micro-Partitioning technology. Now two new metrics are reported; they are the number of physical processors consumed and the percentage of entitled capacity consumed. The physical processors consumed is represented by the *pc* output variable. The percentage entitled is represented by the *ec* variable. The percentage of entitlement consumed is calculated by the following formula.

pc – number of processors used
ent – number of entitled processors
ec=(*pc*/ent)*100; percent of entitled capacity consumed.

4 Performance Case

In order to provide a better understanding of the systems behaviour we performed some tests on one of our logical partitions. A perl program was created that generates a loop 100 million times. We submitted different groups which performed parallel executions of this program in the background and timed the elapsed time of each execution and took the system performance statistics.
The name of the group represented the number of programs that were executed in parallel. For example GROUP1 executed one program while GROUP2 executed two programs in parallel.
The tests were performed with SMT enabled and disabled on an IBM 9117-570 evaluated at 54 RPERF. The operating system was AIX 5.3.
The following sections provide a description of the utilities we used and the results produced.

The *lparstat –i* command provides us with the logical partitions characteristics :

Node Name	test
Partition Name	testlpar
Partition Number	6
Type	Shared-SMT
Mode	Uncapped
Entitled Capacity	1.00
Partition Group-ID	32774
Shared Pool ID	0
Online Virtual CPUs	2
Maximum Virtual CPUs	10
Minimum Virtual CPUs	1
Online Memory	2048 MB
Maximum Memory	2560 MB

Minimum Memory	1536 MB
Variable Capacity Weight	250
Minimum Capacity	0.80
Maximum Capacity	1.20
Capacity Increment	0.01
Maximum Physical CPUs in system	12
Active Physical CPUs in system	12
Active CPUs in Pool	9
Unallocated Capacity	0.00
Physical CPU Percentage	50.00%

Table 10

As seen from the above output this logical partition has 12 physical processors while 9 physical processors are in the shared pool and 3 physical processors are dedicated. There are two virtual CPUs defined with an entitled capacity of 1.

The following *lparstat* output provides statistics for each group of processes that were executed when SMT was enabled.

GROUP#	%user	%sys	%wait	%idle	physc	%entc	lbusy	vcsw	phint
GROUP1	93,4	0,2	-	6,4	1,01	100,9	25,0	219	3
GROUP2	93,8	0,1	-	6,1	2,00	199,9	50,1	135	7
GROUP3	96,6	0,2	-	3,2	2,00	200,0	74,8	93	11
GROUP4	99,9	0,1	-	-	2,00	199,9	100,0	233	14

Table 11

As seen very clearly from the above output the *%entc* jumps immediately to 100% for GROUP1 because it already reaches the amount of its entitled capacity which is 1 CPU. This is confirmed by the *phsyc* field which says that only one physical processor is being used.

The *lbusy* fields shows the logical CPUs or threads utilization. As seen in reports the utilization always performs jumps of 25% until it uses all the available threads.

In order to understand better how the logical CPUs are working, I provided statistics for each GROUP using data obtained from the *mpstat* utility which shows the breakdown of the virtual and logical CPUs for the execution of each group.

GROUP#	Proc0	cpu0	cpu1	Proc1	cpu2	cpu3
GROUP1	12.66%	11.82%	0.84%	88.95%	80.95%	8.00%
GROUP2	100.00%	6.26%	93.74%	100.00%	6.66%	93.34%
GROUP3	100.15%	12.65%	87.50%	100.00%	43.75%	56.25%
GROUP4	99.96%	49.86%	50.10%	99.96%	50.13%	49.84%

Table 12

The following output was produced by the *vmstat* utility. The results are the same from the *lparstat* utility. Here one can see the number of processes in the run queue (*r*) each run.

GROUP#	r	b	avm	fre	re	pi	po	fr	sr	cy	in	sy	cs	us	sy	id	wa	pc	ec
GROUP1	1	0	200171	9493	0	0	0	0	0	0	13	217	212	94	0	6,0	-	1,01	100,6
GROUP2	2	0	200364	9297	0	0	0	0	0	0	0	86	185	94	0	6,0	-	2,00	200,1
GROUP3	3	0	200560	9076	0	0	0	0	0	0	1	103	175	97	0	3,0	-	2,00	200,0
GROUP4	4	0	200750	8881	0	0	0	0	0	0	1	31	180	99	0	-	-	2,00	200,0

Table 13

We performed a series of similar tests with SMT disabled[3].
The number of logical CPUs (*lcpu*) changes from 4 to 2 and the *smt* parameter changed from On to Off. The following example shows how the logical partition behaved executing the processes after the change.

GROUP#	%user	%sys	%wait	%idle	physc	%entc	lbusy	vcsw	phint
GROUP1	99,7	0,2	-	-	1,00	100,2	50,0	40	3
GROUP2	99,7	0,3	-	-	2,00	200,0	100,0	1	6

Table 14

Table 14 shows the one to one relationship between the virtual CPUs and logical CPUs. When the first process was executed the lbusy is 50%, When both of the virtual CPUs are busy the lbusy is 100%.

The following table contains a comparison of the different groups executions when SMT was enabled and disabled.
As seen from these statistics when SMT is enabled the elapsed time is much better when the number of processes increases,

Number of processes	no SMT (elapsed)	SMT (elapsed)
1	77	88
2	76	82
4	154	91
8	306	167
16	658	364
24	947	576

Table 15

5 Conclusions

Micro-Partitioning technology has introduced many new concepts which must be understood before analysing shared logical partitions.
This new technology introduced also changes in the well known native utility *vmstat* and introduced the new utility named *lparstat* which provides LPAR configuration, CPU and Hypervisor statistics.

Bibliography

Advanced POWER Virtualization on IBM System p5, IBM Redbooks, SG24-7940-01
AIX 5L Practical Performance Tools and Tuning Guide, IBM Redbooks, SG24-6478-00
AIX 5L on IBM eServer i5 Implementation Guide, IBM Redbooks, SG24-6455-00
IBM eServer p5 590 and 595 System Handbook, IBM Redbooks, SG24-9119-00
Introduction to pSeries Provisioning, IBM Redbooks, SG24-6389-00
Partitioning Implementation for IBM eServer p5 Servers, IBM Redbooks, SG24-7039-02

[3] When SMT is enabled or disabled the *mpstat* utility notifies that the change was taken place.