EPV Performance University 2021

- z/OS Performance Analysis
- WLM Update
- Db2 Performance Analysis

16th June 2021

zIIP and SMT

Agenda

- Introduction
- SMT metrics
- A real case

Introduction

Introduction

Introduction

- “The CPU Activity section reports on logical core and logical processor activity. For each processor, the report provides a set of calculations that are provided at a particular granularity that depends on whether multithreading is disabled or enabled ...”
- “If multithreading is disabled for a processor type, all calculations are at logical processor granularity. If multithreading is enabled for a processor type, some calculations are provided at logical core granularity and some are provided at logical processor (thread) granularity.”

RMF Report Analysis V2R4 SC34-2665-40

Introduction

- What do you mean by zIIP if you are:
  - PR/SM ➔ PU - Physical Processor - CP - **CORE**
  - z/OS ➔ Logical Processor - LCP - **Logical CORE - Thread**
  - Application ➔ Logical Processor – **Thread**

SMT terms in red

Introduction

- Simultaneous Multi Threading (SMT) is already used on other platforms
- Currently available technologies can’t provide big additional improvements in processor speed so IBM started introducing SMT on the Mainframe with the z13 announcement
- Only for zIIP and IFL (for the moment)
- The reason of this prudent approach is that, from Capacity Management point of view, this is a very critical change

Introduction

- With SMT, multiple instruction streams can be processed simultaneously; when a thread is waiting for a cache miss the core can continue doing work on behalf of the other threads
- Unfortunately, the additional throughput from SMT does not scale very well with the number of threads
- This is because all the threads on a core share some limited resources (e.g. pipes, processor cache, TLB)

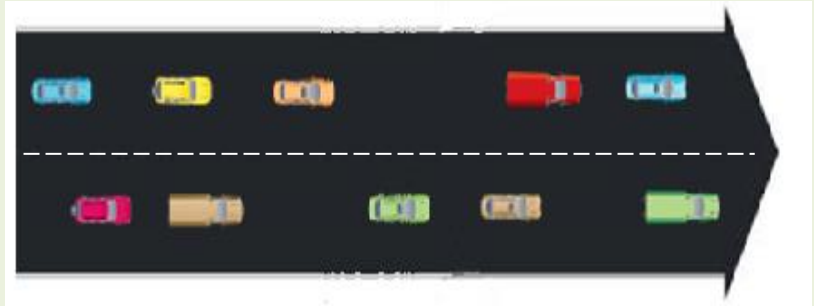
Introduction

MT-1

Faster execution

Lower throughput

MT-2



Slower execution

Higher throughput

Introduction

- Expected speed reduction when 2 threads active:
  - Similar to having more slower engines
  - In the 30-40% range
- Throughput variability:
  - Throughput depends on workload (threads) characteristics
  - On average up to 40% increase when 2 threads active
  - But it may also decrease

Introduction

- SMT can provide benefits in terms of zIIP eligible on CPU reduction (or in general more throughput)
- You can track this ratio (ideally it should be 0):

zIIP eligible on CPU time / zIIP time

- A side effect is that the with SMT the zIIP CPI will slightly increase (speed reduction which means more CP cycles per instruction)

SMT metrics

SMT metrics

- New metrics are provided for zIIP in SMF 70 with SMT:
  - Maximum Capacity Factor
  - Capacity Factor
  - Average Thread Density
- New metrics have to be calculated:
  - Core productivity
  - Core utilization

SMT metrics

- The MT-2 Maximum Capacity Factor (Max CF) is the ratio of the maximum amount of work that can be accomplished using 2 threads to the amount of work that would have been accomplished with 1 thread
- MT-1 Max Capacity Factor is 1.0
- MT-2 Max Capacity Factor is workload dependent; max theoretical value is 2

SMT metrics

- The MT-2 Capacity Factor (CF) is the ratio of the maximum amount of work that has been accomplished using 1 or 2 threads to the amount of work that would have been accomplished with multithreading disabled
- If most of the time Thread Density (TD) is 1 (only 1 thread is active), CF should be close to 1; if most of the time TD is 2 (two threads are active), CF should be close to MAX CF

SMT metrics

- IBM disclosed some info about the way they estimate the most relevant new SMT metrics
- Instructions performed and cycles used with 1 and 2 active threads are the base measurements
- Unfortunately, not all these counters are available to users yet

SMT metrics

- A new MT Diagnostic Counter set is available in SMF 113
- It only includes:
  - Counter 448 – Cycle count with one thread active
  - Counter 449 – Cycle count with two threads active
- They can be used to calculate TD

TD = Counter 448 / Counter 449

- But the key metric is MAX CF

SMT metrics

$$\text{Max CF} = \frac{IPC\_2}{IPC\_1}$$

$$\text{Instructions per cycle (1 thread)} = IPC\_1 = \frac{I\_1}{C\_1}$$

$$\text{Instructions per cycle (2 threads)} = IPC\_2 = \frac{I\_2}{C\_2}$$

SMT metrics

- In this RMF report snapshot you can note that:
  - MT-1 is used for CP; MAX CF, CF and AVG TD value is 1
  - MT-2 is used for zIIP; MAX CF is 1,804 and CF is 1,746
  - zIIP CF and MAX CF are very close because TD is almost 2

---------- MULTI-THREADING ANALYSIS ----------

CPU TYPE	MODE	MAX CF	CF	AVG TD
CP	1	1.000	1.000	1.000
IIP	2	1.804	1.746	1.928

SMT metrics

- CORE productivity is the percentage of the maximum core capacity that has been used while the logical core was dispatched to physical hardware
- If CORE productivity equals 100% all threads on the core are executing work and all core resources are being used
- It's calculated as a ratio between CF and MAX CF

CORE productivity = CF / MAX CF

SMT metrics

---CPU---		---------- TIME % ----------			--- MT % ----	
NUM	TYPE	ONLINE	LPAR BUSY	MVS BUSY	PARKED	PROD	UTIL
0	CP	100.00	4.28	4.15	0.00	100.00	4.28
1	CP	100.00	1.41	1.42	0.00	100.00	1.41
TOTAL/AVERAGE			2.84	2.79		100.00	2.84
2	IIP	100.00	76.08	75.23	0.00	96.48	73.40
				71.13	0.00		
6	IIP	100.00	67.62	66.22	0.00	97.07	65.64
				64.94	0.00		
TOTAL/AVERAGE			71.85	69.38		96.78	69.52

---------- MULTI-THREADING ANALYSIS ----------				
CPU	TYPE	MODE	MAX CF	CF	AVG TD
	CP	1	1.000	1.000	1.000
	IIP	2	1.804	1.746	1.928

MT % PROD = 1,746 / 1,804 = 96,78

SMT metrics

- zIIP LPAR busy simply tells you that the logical core is dispatched
- CORE utilization is supposed to be a more precise metric than zIIP LPAR busy; it should tell you how much work the CORE can still execute
- CORE utilization is calculated by multiplying zIIP LPAR busy and CORE productivity

  *CORE Utilization = zIIP LPAR busy * CORE productivity*

- CORE utilization is automatically used in EPV instead of zIIP LPAR busy when SMT is active

SMT metrics

---CPU---		---------- TIME % ----------			--- MT % ----	
NUM	TYPE	ONLINE	LPAR BUSY	MVS BUSY	PARKED	PROD	UTIL
0	CP	100.00	4.28	4.15	0.00	100.00	4.28
1	CP	100.00	1.41	1.42	0.00	100.00	1.41
TOTAL/AVERAGE			2.84	2.79		100.00	2.84
2	IIP	100.00	76.08	75.23	0.00	96.48	73.40
				71.13	0.00		
6	IIP	100.00	67.62	66.22	0.00	97.07	65.64
				64.94	0.00		
TOTAL/AVERAGE			71.85	69.38		96.78	69.52

---------- MULTI-THREADING ANALYSIS ----------				
CPU TYPE	MODE	MAX CF	CF	AVG TD
CP	1	1.000	1.000	1.000
IIP	2	1.804	1.746	1.928

MT % UTIL = 76,08 * 96,48 / 100 = 73,40

SMT metrics

- Exercise 1

Knowing that zIIP LPAR busy at 11 is 90%, calculate zIIP CORE utilization in that hour

SYSTEM IIP MT

SWITCH

IIP MT BY HOUR Thu, 21 Jan 2016

METRIC	8	9	10	11	12	13	14	15
IIP CF	1,12	1,14	1,15	1,15	1,14	1,14	1,14	1,14
IIP MCF	1,34	1,37	1,36	1,35	1,35	1,36	1,36	1,36
IIP TD	1,35	1,40	1,41	1,41	1,42	1,40	1,40	1,39

SMT metrics

- Exercise 1

Knowing that zIIP LPAR busy at 11 is 90%, calculate zIIP CORE utilization in that hour

  - **CF = 1,15**
  - **MCF = 1,35**
  - **Core productivity = CF /MCF = 85,2%**
  - **Core utilization = 90% * 85,2% = 76,7%**

SMT metrics

- With SMT enabled all accounting fields (SMF 30, 72, etc) reports zIIP consumption of workloads as MT-1 Equivalent Time and Service Units
- MT-1 Equivalent Time is the zIIP time that would have taken to run the same work in MT-1 mode
- MT-1 Equivalent Time is internally calculated as

  MAX CF * zIIP time
- All measurements are enflated (if Max CF is greater than 1)

SMT metrics

- Most important consequence of MT-1 Equivalent Time measurements is that when working in MT-2 you have to change the calculation of the capacity used by any workload

Example of old algorithm valid when SMT is not active:

- Workload A used 1.800 zIIP seconds in 1 hour
- 1 CORE is targeted 1.000 MIPS
- **used COREs = 1.800 / 3.600 = ,5**
- **used MIPS = 1.000 * ,5 = 500**

zIIP and SMT

- Exercise 2

Workload A uses 1800 zIIP seconds; 1 CORE is targeted 1.000 MIPS

calculate its MIPS usage if MT=2 and MAX CF is 1,25

zIIP and SMT

- Exercise 2

Workload A uses 1800 zIIP seconds; 1 CORE is targeted 1.000 MIPS

calculate its MIPS usage if MT=2 and MAX CF is 1,25

  - Workload A used 1.800 zIIP seconds in 1 hour
  - 1 CORE is targeted 1.000 MIPS
  - **used COREs = 1.800 / 1,25 / 3.600 = 0,4**
  - **used MIPS = 1.000 * 0,4 = 400**

SMT metrics

- To activate SMT on z/OS, you have to:
  - ✓ define the PROCVIEW CORE option in LOADxx; if you do not want to use SMT you can omit the PROCVIEW parameter or specify PROCVIEW CPU which is the default; IPL is needed to change it
  - ✓ set MT_ZIIP_MODE=2 in IEAOPTxx; it can be dynamically changed

SMT metrics

SYSTEMS

SYSTEMS CONFIGURATION

CEC	SYSTEM	LPARNAME	SYSNAME	SYSPLEX	OS LEVEL	GMTOFF	HDISP	CPUS	AAPS	IIPS	CORE	MT CPU	MT IIP
					ZV020200	1,0	Y	3	0	3	N	1	1
					ZV020200	1,0	Y	10	0	8	N	1	1
					ZV020100	1,0	Y	19	0	9	N	1	1
					ZV020100	1,0	Y	19	0	9	N	1	1
					ZV020100	1,0	Y	19	0	9	N	1	1
					ZV020200	1,0	Y	1	0	1	N	1	1
					ZV020200	1,0	Y	3	0	2	N	1	1
					ZV020200	1,0	Y	1	0	0	N	1	0
					ZV020200	1,0	Y	3	0	2	N	1	1
					ZV020200	1,0	Y	2	0	2	N	1	1
					ZV020100	1,0	Y	2	0	2	Y	1	2
					ZV020100	1,0	Y	2	0	2	Y	1	2
					ZV020100	1,0	Y	1	0	1	Y	1	2
					ZV020100	1,0	Y	1	0	1	Y	1	2
					ZV020100	1,0	Y	8	0	6	Y	1	1
					ZV020100	1,0	Y	8	0	6	Y	1	1
					ZV020200	1,0	Y	3	0	4	N	1	1

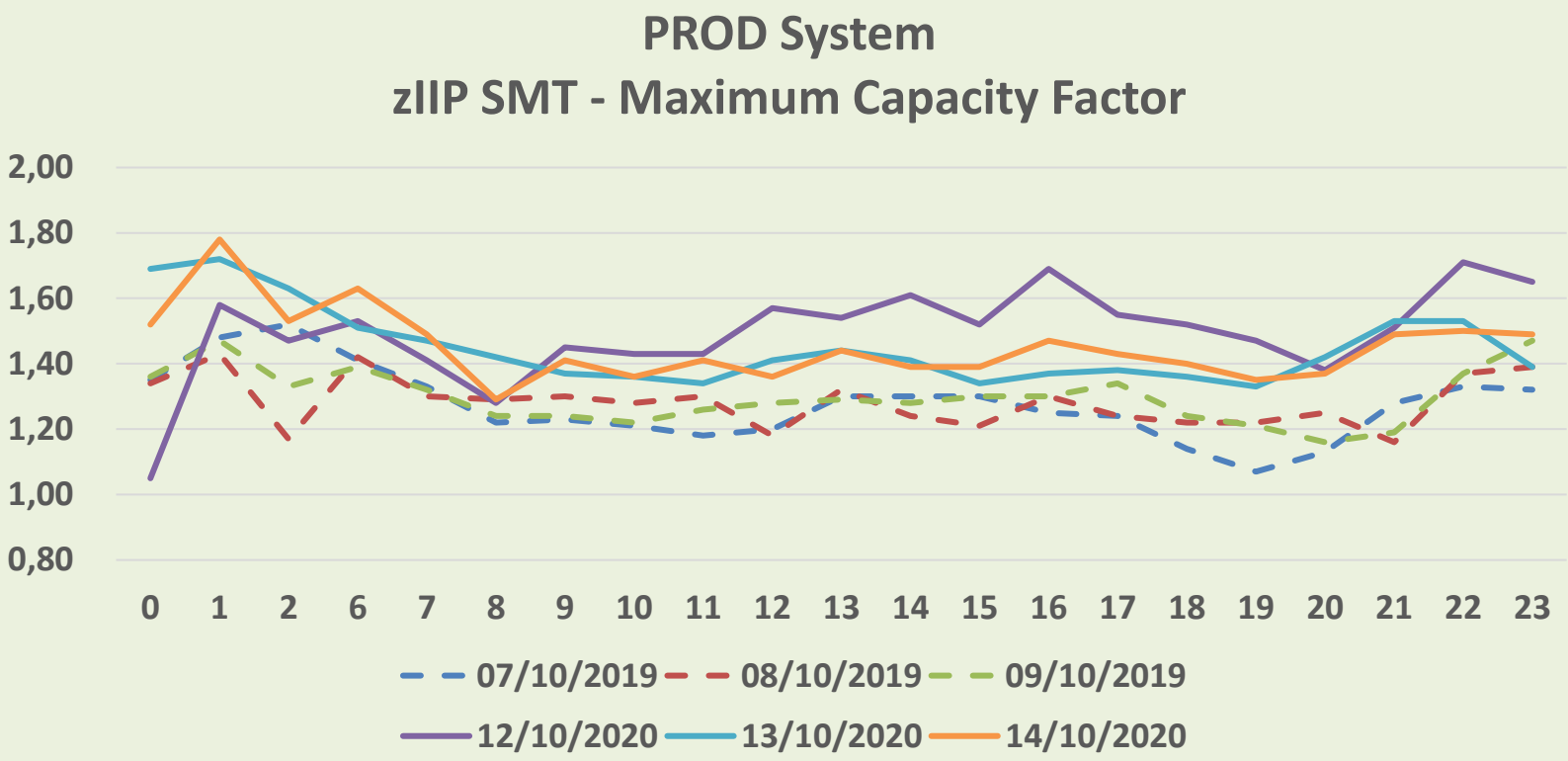
A real case

A real case

- One of our customers migrated from z13 to z15
- They greatly increased the number of zIIPs to eliminate zIIP eligible on CPU
- zIIP eligible on CPU was almost eliminated and average zIIP busy was greatly reduced
- They wanted to check the effects on SMT

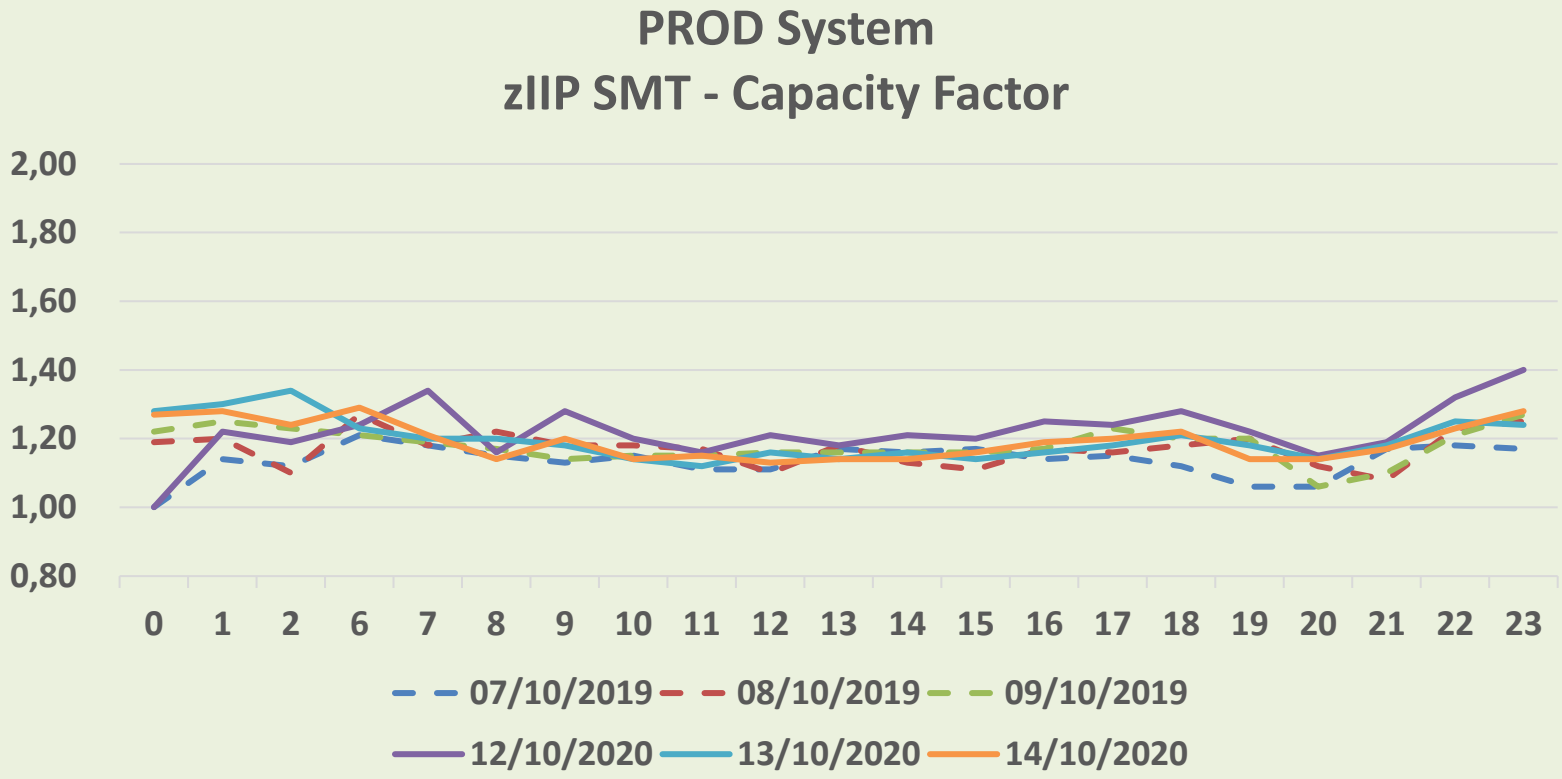
A real case

- Maximum Capacity Factor
- Dotted lines are for z13, continuous lines are for z15
- Potential SMT benefits increased



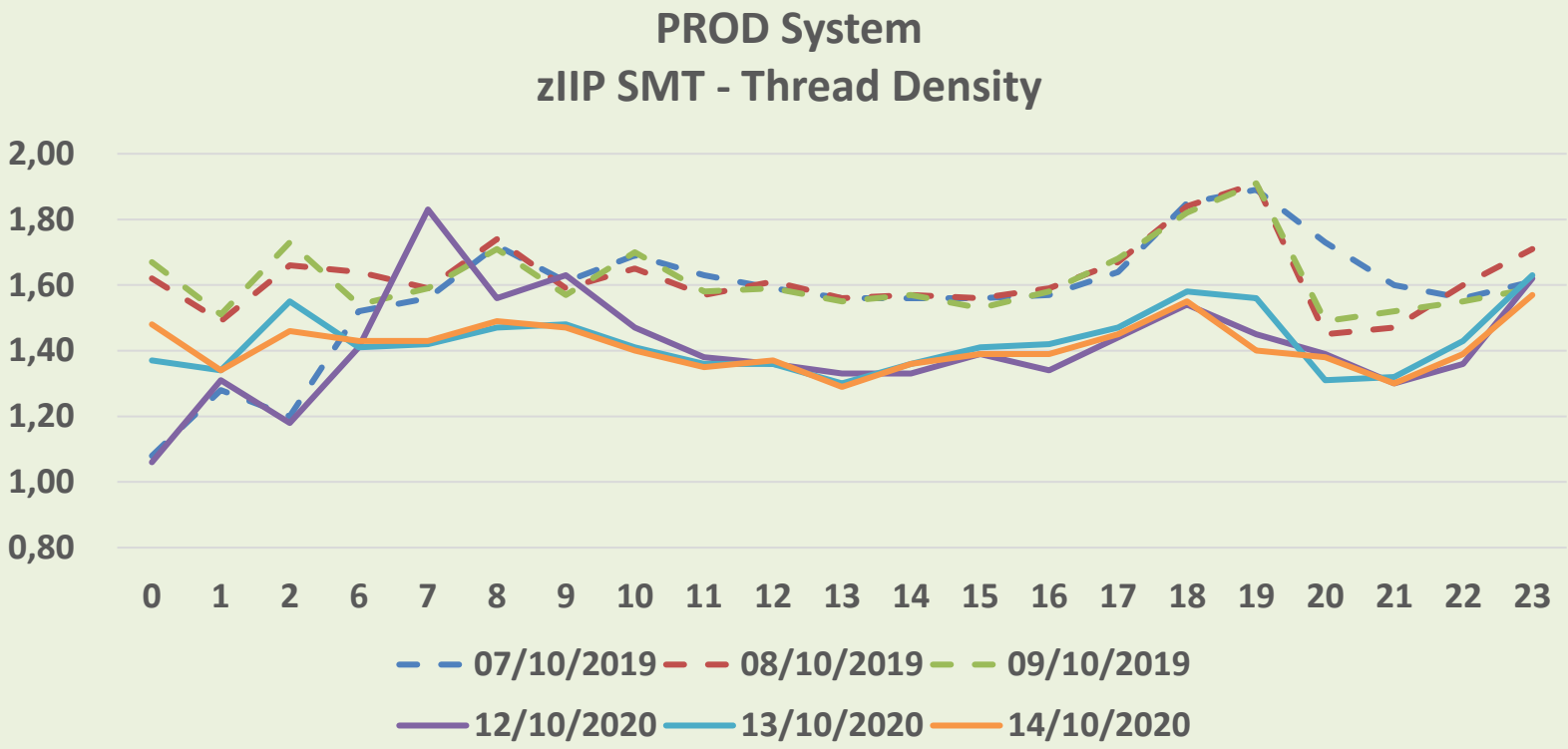
A real case

- Capacity Factor
- Real SMT benefits only slightly better



A real case

- Thread Density
- Reduced because of the increased number of zIIPs



Questions?