How to improve IMS Scheduling

GSE BeNeLux IMS User Group 2017
Antwerp, Belgium

Ulrich Dillenberger
Principal Technical Consultant, BMC Software
ulrich_dillenberger@bmc.com

Agenda

What is IMS scheduling?

Elements that affect IMS scheduling

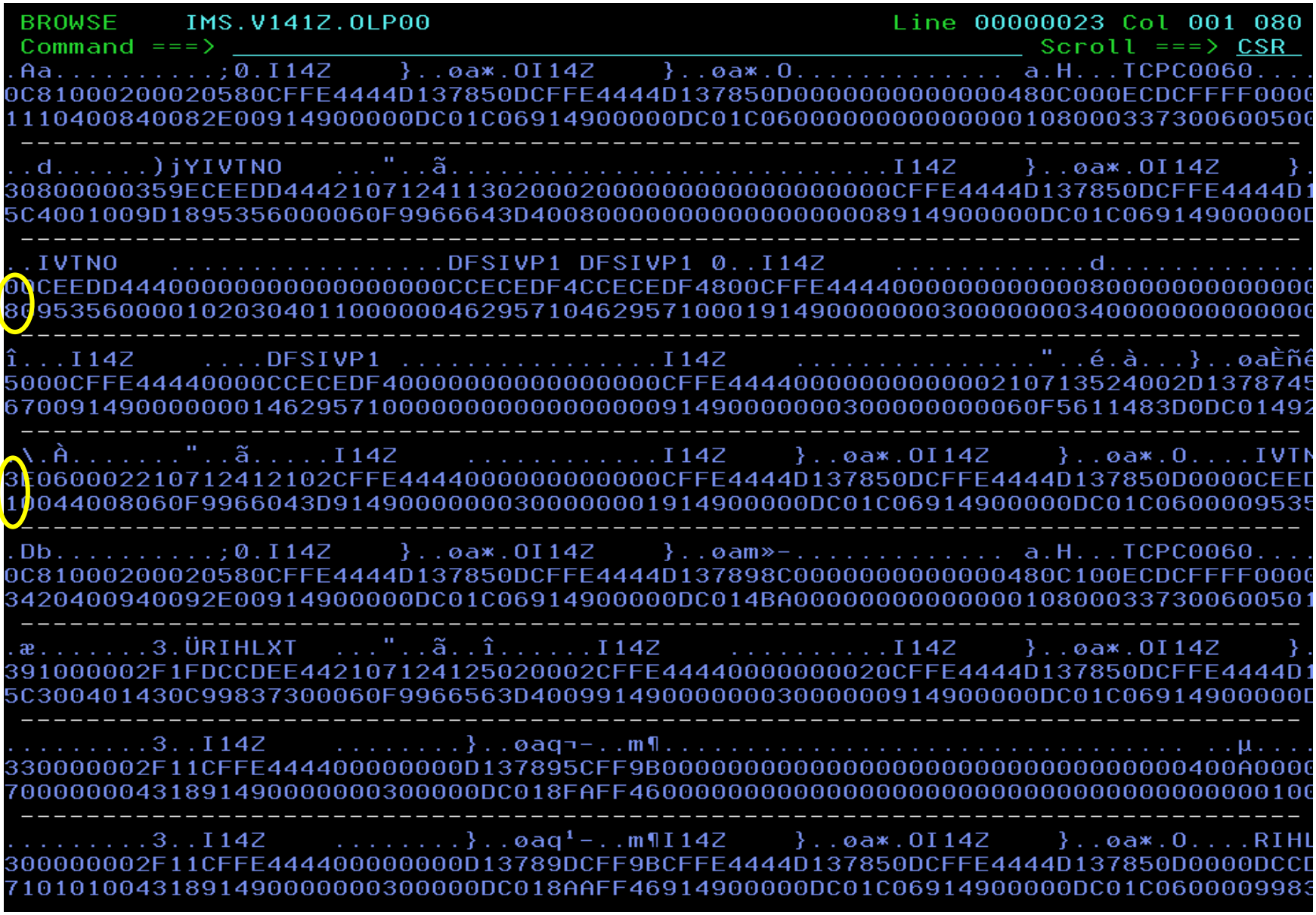
Monitoring IMS scheduling

What is IMS Scheduling ?

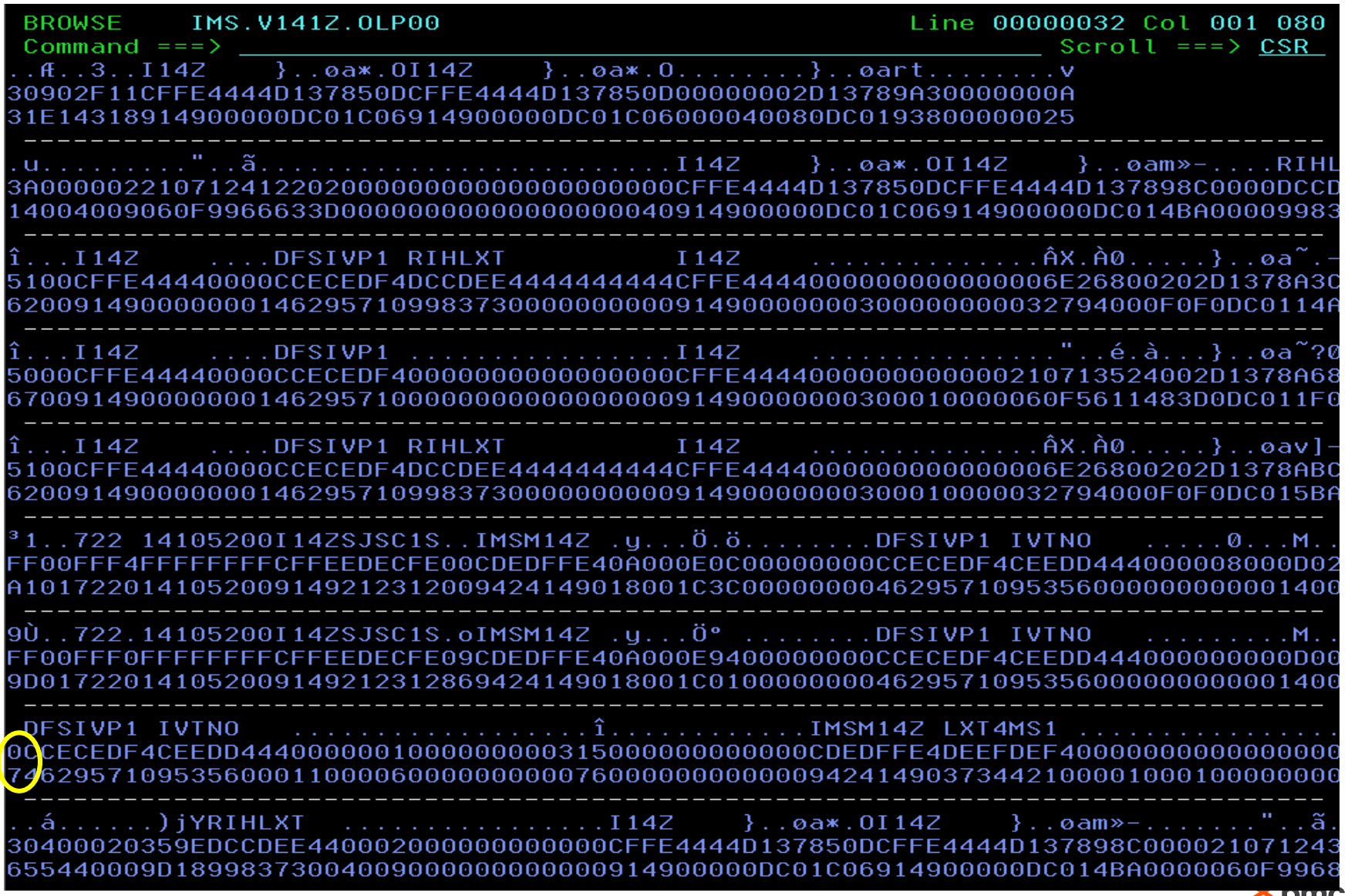
The life of a full-function IMS transaction (message)

- Message arrives in IMS
- Enqueue the message to the SMB
- Schedule the message
- Schedule-end to first DL/I call
- Program elapse
- Syncpoint
- Output message

IMS Log Records



IMS Log Records – Part 2



A more convenient way to look at the log records related to a unit of work

```
Report index file: 'RIHLXT.LUI.INDEX'
Type one or more action codes. Enter LEFT or RIGHT for additional columns.
  S=Record detail  T=List adjacent records
  D=Dump record    B=Browse record  ?=Log code help     Sorted by: TIMESTAMP
  Log
A Code     Type         Attributes (Enter ALTview for multiple lines per record)
- ---- ------------ ------------------------------------------------------------
_ 01   input msg    origims=I15Z destims=I15Z origin=RIHLXT   dest=IVTNO
_ 35   msg enqueue  dest=IVTNO    drrn=0400001D
_ 08   pgm start    recovery token=I15Z    0000000200000000 tran=IVTNO
_ 3307 recovery     (start unit-of-recovery) psb=DFSIVP1
_ 31   msg get      recovery token=I15Z    0000000200000000 dest=IVTNO
_ 36   msg return   (Input message returned to msg Q) qlriflgs=01 qlriflg2=00
_ 35   msg enqueue  dest=IVTNO    drrn=0400001D pst=00001
_ 38   msg return   (Input message returned to msg Q) qlriflgs=01 qlriflg2=00
_ FA   Mainview     tran=IVTNO    jobname=IMSM15Z  userid=RIHLXT   ver=5300
_ F9   Mainview     jobname=IMSM15Z  pst=00001 psb=DFSIVP1  tran=IVTNO
_ 07   pgm end      abendcode=U3303  psb=DFSIVP1  tran=IVTNO    pst=00001
********************************* Bottom of Data ********************************
```

What is IMS Scheduling ?

Schedule-to-1st DL/I call time

- Phase 1: Scheduling the message
  - IMS Scheduler examines the following elements
    - SERIAL/PARALLEL, MAXRGN, PARLIM, PROCLIM, PRIORITY, CLASS
  - Pool allocation
    - PSB/DMB directory (PDIR/DDIR) and PSBW pool
      - » If PSB not resident, I/Os to ACBLIB
      - » If pool space not available, intent failure
  - Schedule the message to MPR

What is IMS Scheduling ?

Schedule-to-1st DL/I call time (cont'd)

- Phase 2: Schedule-end to first DL/I call
  - Program & subroutine load from PGMLIB
  - Program initialization
  - Program does the first DL/I call to the I/O PCB

What is IMS Scheduling ?

Our goal for today's session is to:

- Reduce the message scheduling time
- Reduce or eliminate the Schedule-end to first DL/I call time

- Results
  - Improve transaction response time
  - Reduce CPU utilization
  - Increase transaction throughput
  - Meet SLA with minimum resource requirement

Elements that affect Scheduling

Define program SERIAL/PARALLEL (APPLCTN/CREATE PGM)

- Definition
  - Use SERIAL for application required serialization for a business reason
  - PARALLEL allows transaction to be scheduled in multiple MPRs
- Recommendation: Use PARALLEL
- Advantages
  - Less message queuing for high-volume transaction
  - Improve transaction response time
  - Handle high-volume transaction
- Disadvantages
  - Increase database locking contention or deadlock

Elements that affect Scheduling

Define MAXRGN=x for parallel scheduling (TRANSACTION/CREATE TRAN)

- Definition
  - Set the maximum number of regions that a transaction can run concurrently
  - Controls domination of a transaction in a class
- Recommendation: It depends on the transaction volume during peak time
- Advantages
  - Allows more concurrency
  - Handle high-volume transaction
- Disadvantage
  - Setting value x too high will cause IMS scheduling to be more complex thus causing high CPU usage for the control region
  - Increased possibility for locking conflicts

Elements that affect Scheduling

Define PSB as RESIDENT/Non-Resident

- Recommendation: Define PSB as RESIDENT
  - Use 64-bit ACB storage pool to reduce I/Os (ACBIN64=n)
- Advantage
  - Reduce I/Os to the ACBLIB
- Disadvantage
  - For non-active PSB, it occupies real/virtual storage (working set size). But in today's system, real/virtual storage is less of an issue

Elements that affect Scheduling

Define PROCLIM=(x,y) for limiting transaction scheduling or by CPU time

- Definition
  - Set the limit number of messages that can be executed in one schedule for a transaction
  - When the limit is reached, the program is ended unless it is eligible for a quick re-scheduling
  - Set PROCLIM=0 for program that does not perform storage clean up properly
- Recommendation: Set a high value for a high volume transaction
- Advantages
  - Reduce IMS scheduling time
  - Reduce program load time
  - Reduce CPU utilization
  - Reduce Db2 thread creation time
- Disadvantage
  - The transaction can dominate the MPR

Elements that affect Scheduling

Define PARLIM=x for parallel transaction scheduling

- Definition
  - Set the queue count that triggers parallel scheduling of a transaction
  - PARLIM=0 implies scheduling right away
  - In a shared-queue environment, IMS does not know the queue depth. It performs another schedule after a successful GU I/O PCB
- Recommendation: PARLIM=0 for non shared-queue and =2 for shared queue
- Advantage
  - Reduce input queue time
- Disadvantages
  - Scheduling overhead may increase if transaction has long schedule to 1st DL/I time
  - In CQS environment, PARLIM=0 may cause false scheduling

Elements that affect Scheduling

Define PRTY=(normal,limit,limit count) for setting transaction priority when transaction stays in the queue

- Definition
  - Specifies the values that determine the scheduling priority of this transaction
  - Normal: Use this value when tran in queue less than limit count (0-14, default is 1)
  - Limit: Use this value when tran in queue equals or greater than limit count
  - Use command to set PRTY to 0 to prevent scheduling
- Recommendation: Use default unless there is a good reason
- Advantages
  - Reduce input queue time
  - For a certain type of application, this maybe helpful
  - Useful to manage MSC transactions
- Disadvantage
  - Complex setting may cause scheduling overhead

Elements that affect Scheduling

Define Wait-for-Input (WFI) for transaction

- Definition
  - The transaction is scheduled just like normal. However, the next GU call to the I/O PCB will either get the next message in the queue or IMS will suspend the call (place the ITASK in the suspend queue) until the next message is available
  - The GU call will get the 'QC' status when the PROCLIM limit is reached
- Recommendation: Use WFI for transaction accessing Db2 or high volume
- Advantages
  - Eliminate the scheduling process for the transaction in the region
  - CPU, I/O, log records will be reduced (no log records x'07',x'08')
  - Response time improved
- Disadvantage
  - Region is occupied by this transaction
  - Application programs executed in this MPR needs to have correct initialization logic

Elements that affect Scheduling

Define the region (MPR) as Pseudo-Wait-for-Input (PWFI=Y)

- Definition
  - The transaction is scheduled just like normal. However, the next GU call to the I/O PCB will either get the next message in the queue or IMS will suspend the call (place the ITASK in the suspend queue) if there is no eligible message ready to be scheduled in this region
- Recommendation: Use PWFI as much as you can
  - Strongly recommended for transaction accessing Db2
- Advantages
  - Reduce the scheduling process for the transaction in the region
  - CPU, I/O, log records will be reduced (no log records X'07', X'08')
  - Response time improved
- Disadvantage
  - Application programs executed in this MPR need to have correct initialization logic

Elements that affect Scheduling

DMB, PSB (CSA, DLI), PSBW pools

- Definition
  - IMS Scheduler needs to allocate pool space for this transaction
- Recommendation: Monitor these pools and make sure they are large enough to handle the workload during peak
  - If IMS Catalog is used, IMS requires extra space for PSB (DLI & CSA) for each schedule
- Advantage
  - Reduce the scheduling failures such as block-mover wait
- Disadvantage
  - N/A

Elements that affect Scheduling

Program preload

- IMS needs to make the application program ready to run under the MPR; if the program is not in storage, it needs to load the program from the PGMLIB
- Recommendation: Use the DFSMPLxx to preload program and Cobol service routines (language environment modules)
  - Use GTF trace to find module names
- Advantage
  - Reduce/eliminate program load time
- Disadvantage
  - Occupy large storage/working set size especially with multiple regions

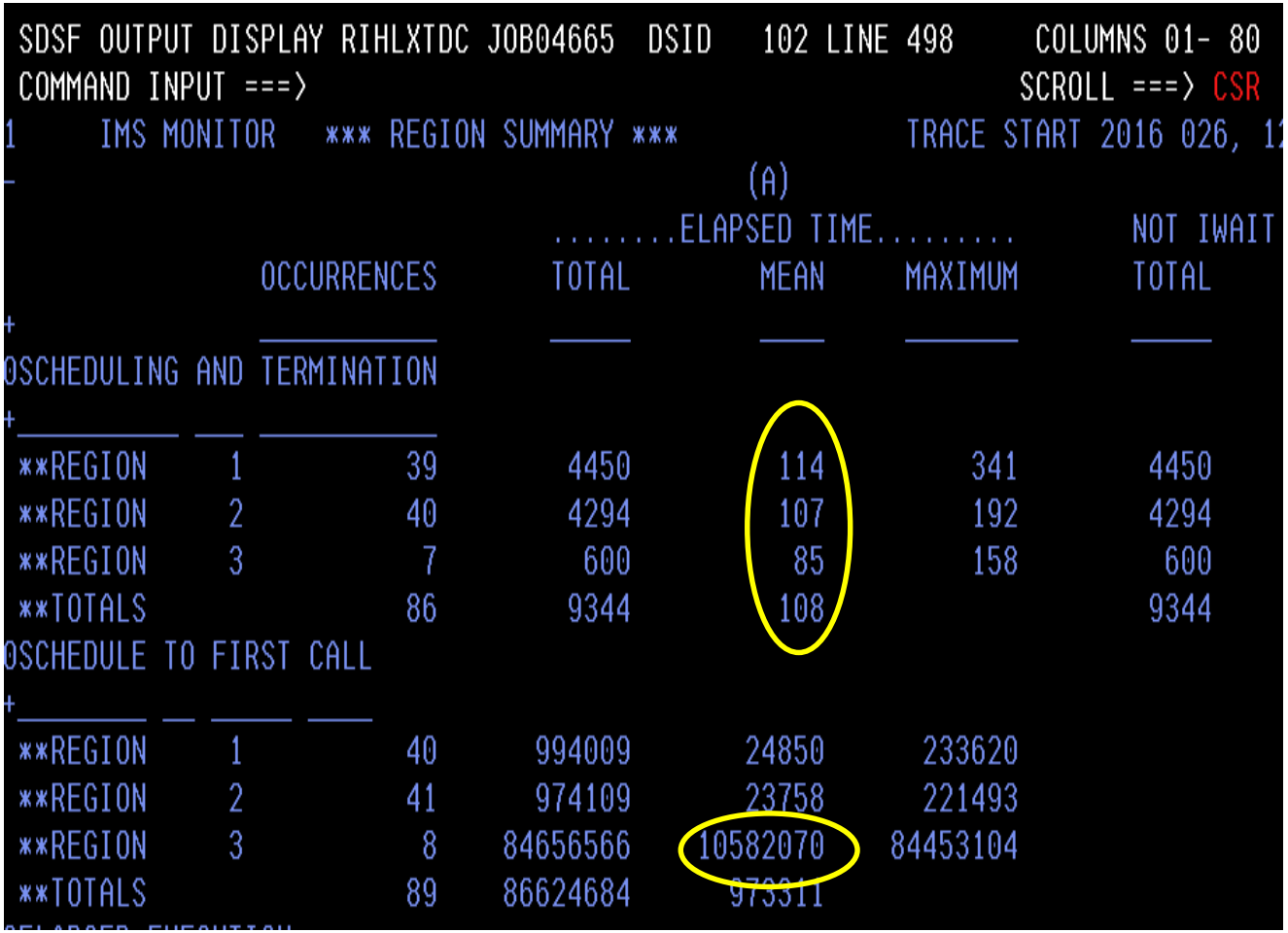
Elements that affect Scheduling

PGMLIB

- If program is not in storage, IMS loads from the PGMLIB
- Recommendation
  - Make PGMLIB under VLF or using cache/SSD device
  - Use large blocksize
  - Put RESLIB in the bottom of STEPLIB concatenation
- Advantage
  - Reduce program load time
- Disadvantage
  - N/A

Monitoring IMS Scheduling

IMS DC Monitor Region Summary Report



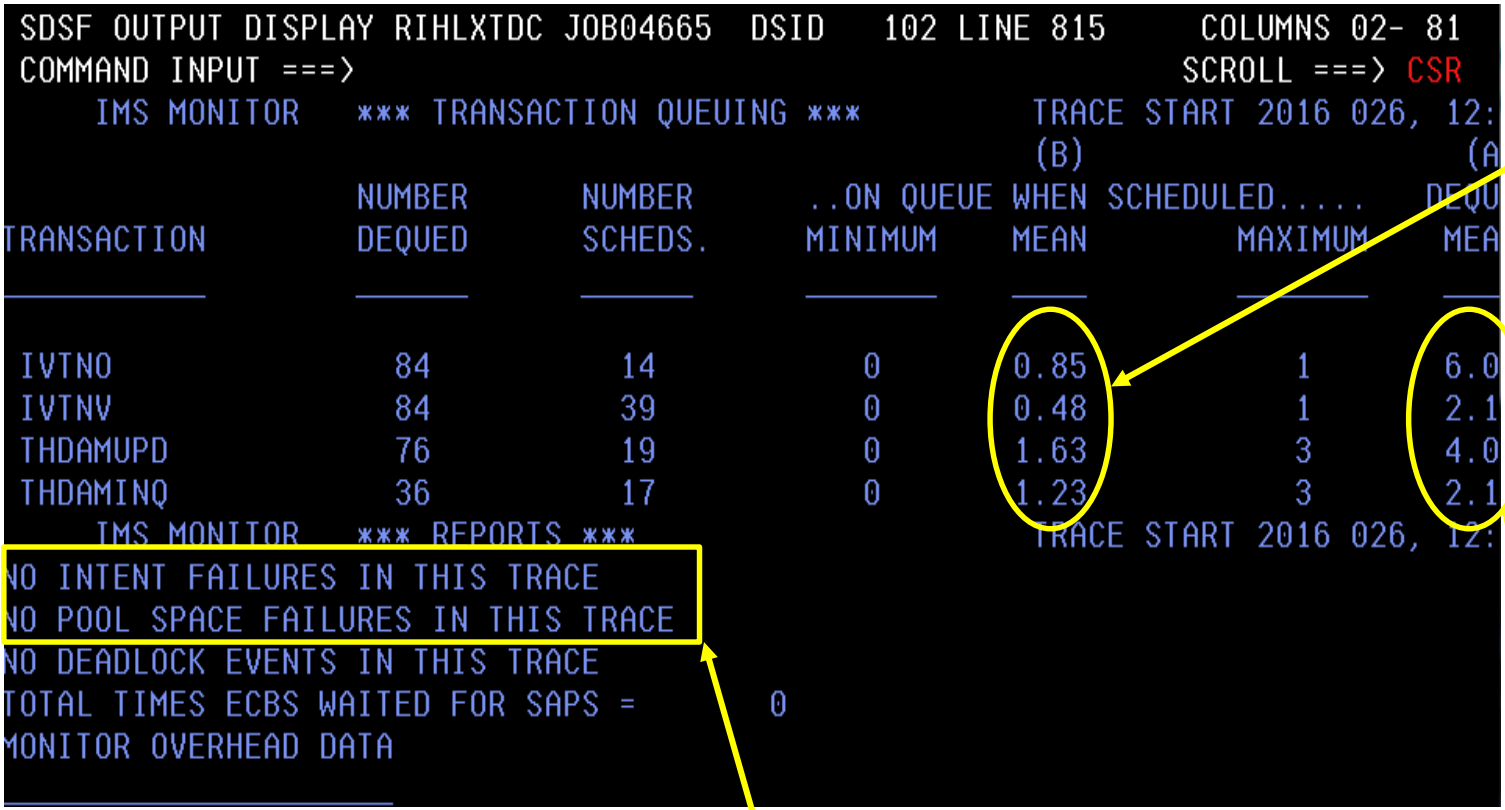
```
SDSF OUTPUT DISPLAY RIHLXTDC JOB04665  DSID   102 LINE 498     COLUMNS 01- 80
COMMAND INPUT ===>                                            SCROLL ===> CSR
1     IMS MONITOR   *** REGION SUMMARY ***            TRACE START 2016 026, 1
                                          (A)
                                 .........ELAPSED TIME.........       NOT IWAIT
                OCCURRENCES       TOTAL       MEAN     MAXIMUM       TOTAL
                ___________       _____       ____     _______       _____
OSCHEDULING AND TERMINATION
+_________ __ ___________
 **REGION     1          39        4450        114          341        4450
 **REGION     2          40        4294        107          192        4294
 **REGION     3           7         600         85          158         600
 **TOTALS                86        9344        108                     9344
OSCHEDULE TO FIRST CALL
+________ __ _____ ____
 **REGION     1          40      994009      24850       233620
 **REGION     2          41      974109      23758       221493
 **REGION     3           8    84656566   10582070     84453104
 **TOTALS                89    86624684     973311
```

- Look at the number of occurrences and the mean time (in micro secs)
- What happened to region 3 “schedule to first call” mean time?
- What is the difference between the “scheduling and termination” and “schedule to first call” time?

bmc

Monitoring IMS Scheduling

IMS DC Monitor Transaction Queuing Report



Small number is better

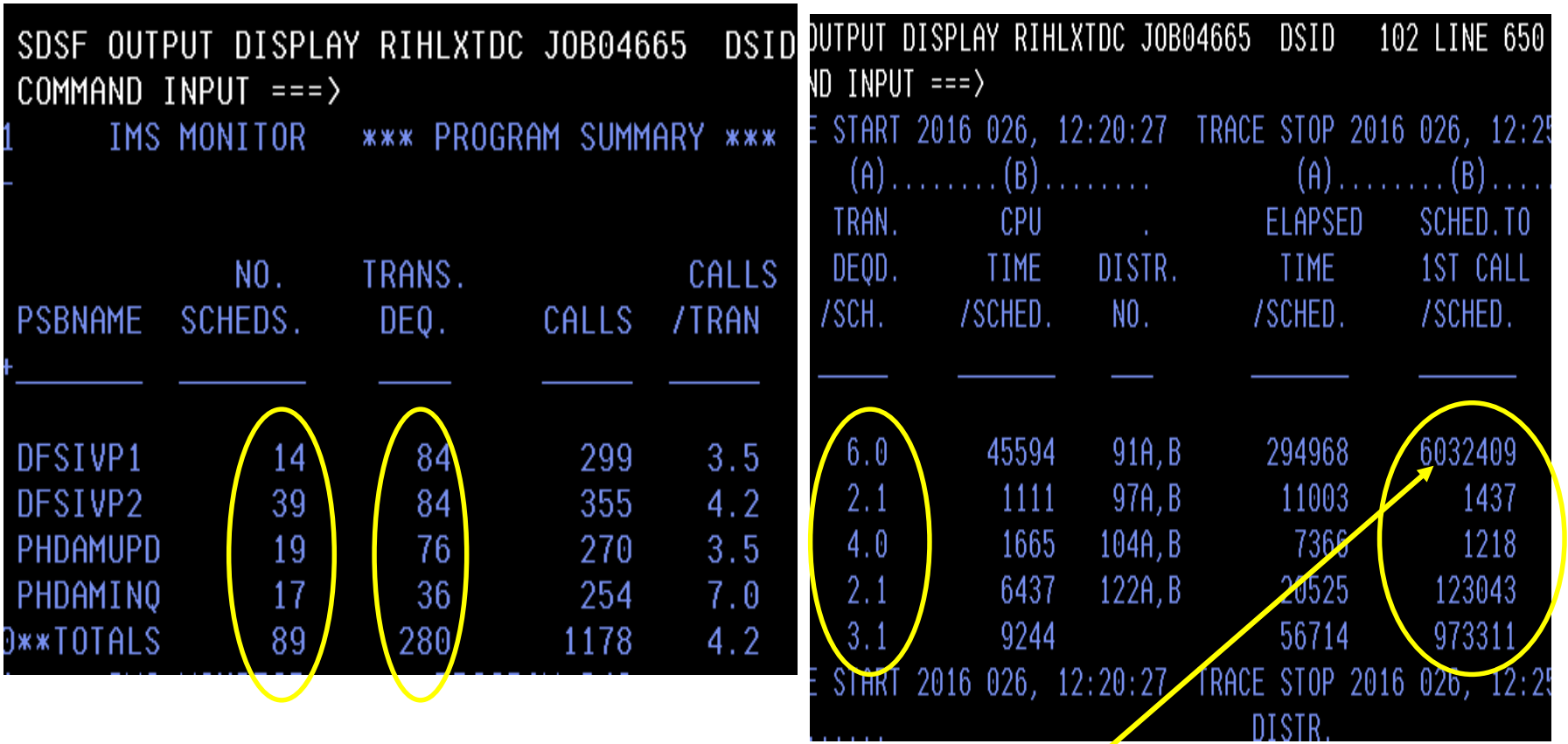
Large number is better

PSB/PSBW pools are fine

bmc

Monitoring IMS Scheduling

IMS DC Monitor Program Summary Report



```
SDSF OUTPUT DISPLAY RIHLXTDC JOB04665  DSID
COMMAND INPUT ===>
1     IMS MONITOR    *** PROGRAM SUMMARY ***
-

                NO.     TRANS.              CALLS
 PSBNAME     SCHEDS.     DEQ.      CALLS   /TRAN
+--------    -------     ----      -----   -----

 DFSIVP1         14       84        299     3.5
 DFSIVP2         39       84        355     4.2
 PHDAMUPD        19       76        270     3.5
 PHDAMINQ        17       36        254     7.0
0**TOTALS        89      280       1178     4.2
```

```
UTPUT DISPLAY RIHLXTDC JOB04665  DSID   102 LINE 650
ND INPUT ===>
E START 2016 026, 12:20:27  TRACE STOP 2016 026, 12:25
  (A).....(B).....          (A).....(B).....
 TRAN.      CPU       .      ELAPSED    SCHED.TO
 DEQD.      TIME    DISTR.     TIME      1ST CALL
/SCH.     /SCHED.    NO.     /SCHED.     /SCHED.
----      -------    ---     -------     -------

 6.0        45594   91A,B     294968     6032409
 2.1         1111   97A,B      11003        1437
 4.0         1665  104A,B       7368        1218
 2.1         6437  122A,B      20525      123043
 3.1         9244                56714      973311
E START 2016 026, 12:20:27  TRACE STOP 2016 026, 12:25
                                DISTR.
```

Why such a big number?
A possibility that GU I/O PCB calls were suspended

bmc

Monitoring IMS Scheduling

IMS DC Monitor Programs by Region Report

```
SDSF OUTPUT DISPLAY RIHLXTDC JOB04665  DSID   102 LINE 625    COLUMNS 01- 80 XTDC JOB04665  DSID   102 LINE 625
COMMAND INPUT ===>                                           SCROLL ===> CSR
1     IMS MONITOR   *** PROGRAMS BY REGION ***        TRACE START 2016 026, 12:20:27  TRACE STOP 2016 026, 12:25:17
```

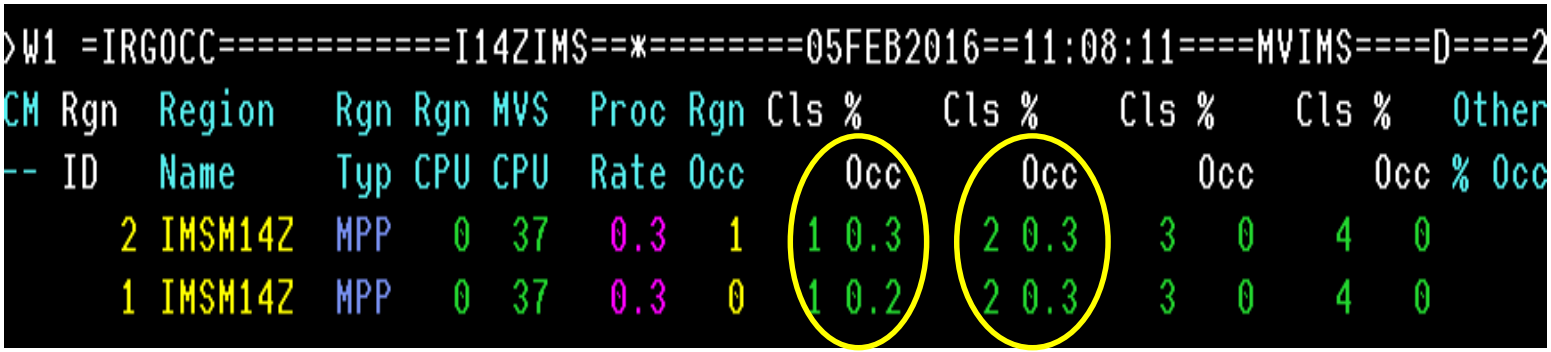
		(A) ELAPSED EXECUTION TIME			SCHEDULING	(B) END TO FIRST CALL		DISTRIBUTION
	OCCURRENCES	TOTAL	MEAN	MAXIMUM	TOTAL	MEAN	MAXIMUM	NUMBER
0**REGION 1								
PHDAMUPD	11	82649	7513	11473	13529	1229	2389	105A,B
DFSIVP2	16	139507	8719	43894	19075	1192	2283	113A,B
PHDAMINQ	8	150519	18814	47463	961095	120136	233620	125A,B
DFSIVP1	5	127956	25591	45582	310	62	96	136A,B
REGION TOTALS	40	500631	12515		994009	24850		
0**REGION 2								
DFSIVP2	20	188138	9406	24025	34714	1735	7858	98A,B
PHDAMUPD	8	57306	7163	10579	9624	1203	2015	114A,B
PHDAMINQ	7	159419	22774	47167	929497	132785	221493	123A,B
DFSIVP1	6	196262	32710	58934	274	45	66	133A,B
REGION TOTALS	41	601125	14661		974109	23758		
0**REGION 3								
DFSIVP1	3	3805346	1268448	3545819	84453145	28151048	84453104	92A,B
DFSIVP2	3	101476	33825	94677	2266	755	865	126A,B
PHDAMINQ	2	38995	19497	23052	201155	100577	145946	127A,B
REGION TOTALS	8	3945817	493227		84656566	10582070		

Was region 3 really much less efficient for program DFSIVP1 than region 1 ?
DC Monitor not helpful with PWFI !

bmc

Monitoring IMS Scheduling

Monitor IMS region occupancy percentage

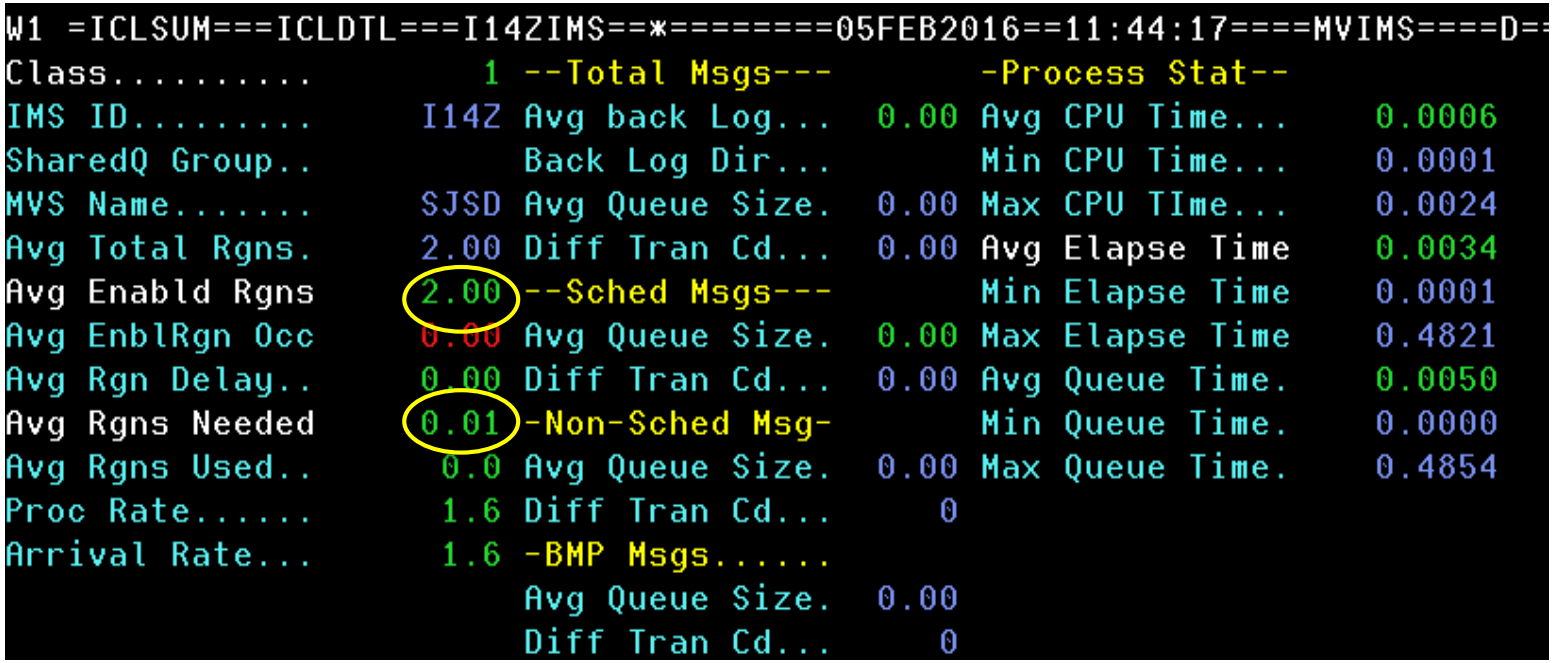


```
>W1 =IRGOCC===========I14ZIMS==*========05FEB2016==11:08:11====MVIMS====D====2
CM Rgn  Region   Rgn Rgn MVS  Proc Rgn Cls %    Cls %    Cls %    Cls %   Other
-- ID   Name     Typ CPU CPU  Rate Occ    Occ      Occ      Occ      Occ % Occ
     2 IMSM14Z  MPP   0  37   0.3   1   1 0.3    2 0.3    3   0    4   0
     1 IMSM14Z  MPP   0  37   0.3   0   1 0.2    2 0.3    3   0    4   0
```

- Good region occupancy rate under 60%
- Any class with low occupancy rate should be removed to make scheduling simple
- Group the removed class to a region (or more) so program preload is more useful

Monitoring IMS Scheduling

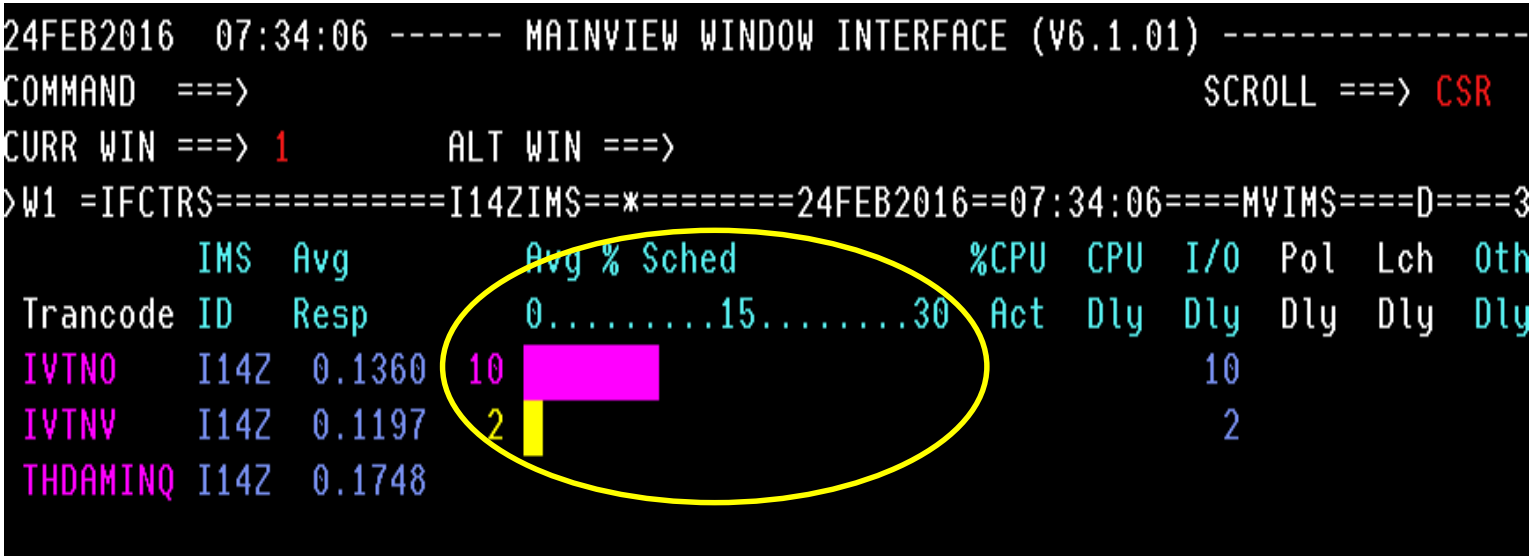
Monitor processing by class



Monitor each class to make sure there are enough regions to execute all transactions in that class

Monitoring IMS Scheduling

Monitor scheduling time by transaction



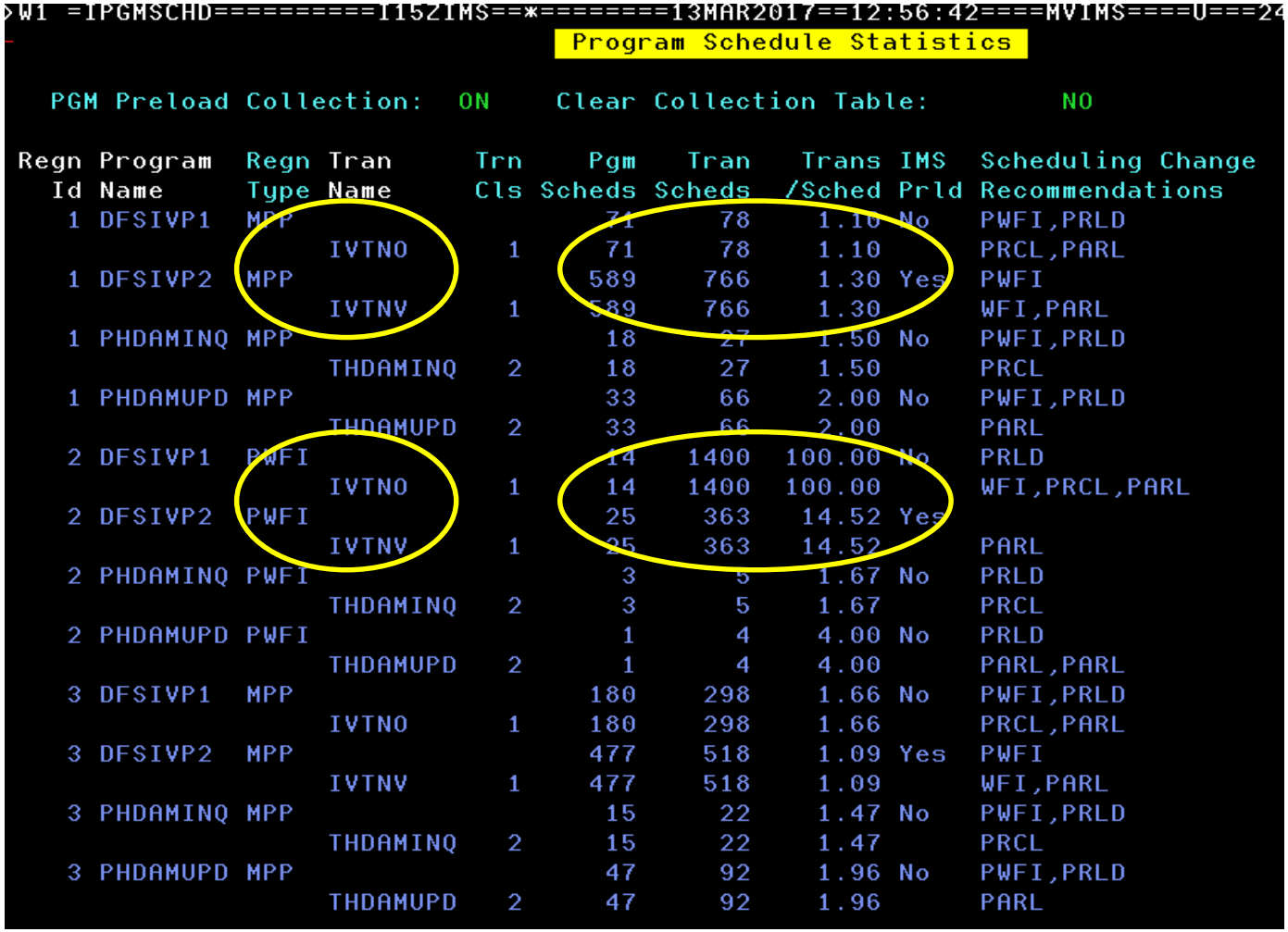
- Consider making change to the MAXRGN or PROCLIM or PARLIM
- Review your change if it improves

bmc

Monitoring IMS Scheduling

Monitor efficiency of scheduling by tran and region

- Effect of WFI transaction /PWFI region



W1 =IPGMSCHD==========I152IMS==*========13MAR2017==12:56:42====MVIMS====U===24

Program Schedule Statistics

PGM Preload Collection: ON Clear Collection Table: NO

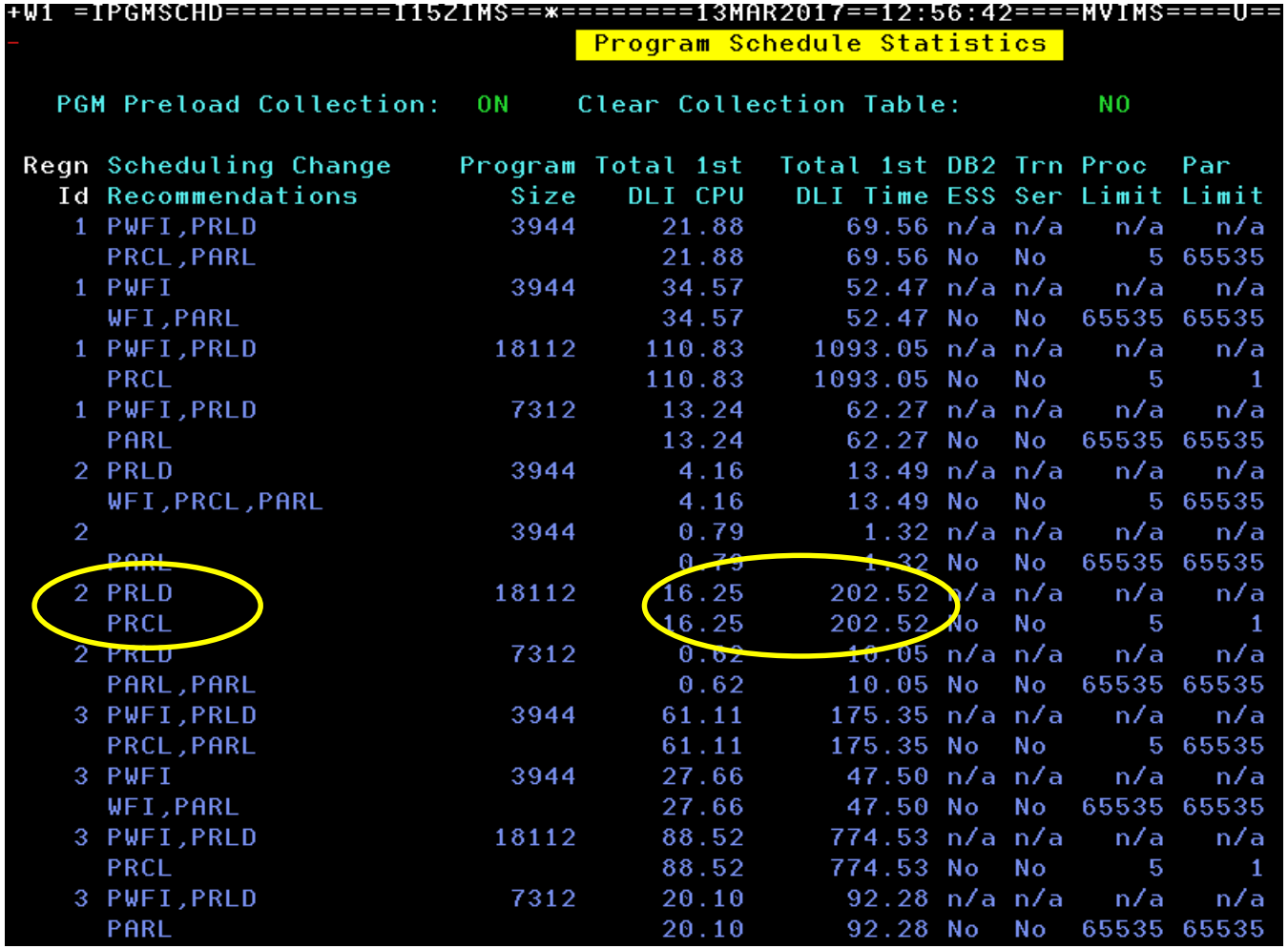
Regn Id	Program Name	Regn Type	Tran Name	Trn Cls	Pgm Scheds	Tran Scheds	Trans /Sched	IMS Prld	Scheduling Change Recommendations
1	DFSIVP1	MPP			71	78	1.10	No	PWFI,PRLD
			IVTNO	1	71	78	1.10		PRCL,PARL
1	DFSIVP2	MPP			589	766	1.30	Yes	PWFI
			IVTNV	1	589	766	1.30		WFI,PARL
1	PHDAMINQ	MPP			18	27	1.50	No	PWFI,PRLD
			THDAMINQ	2	18	27	1.50		PRCL
1	PHDAMUPD	MPP			33	66	2.00	No	PWFI,PRLD
			THDAMUPD	2	33	66	2.00		PARL
2	DFSIVP1	PWFI			14	1400	100.00	No	PRLD
			IVTNO	1	14	1400	100.00		WFI,PRCL,PARL
2	DFSIVP2	PWFI			25	363	14.52	Yes	
			IVTNV	1	25	363	14.52		PARL
2	PHDAMINQ	PWFI			3	5	1.67	No	PRLD
			THDAMINQ	2	3	5	1.67		PRCL
2	PHDAMUPD	PWFI			1	4	4.00	No	PRLD
			THDAMUPD	2	1	4	4.00		PARL,PARL
3	DFSIVP1	MPP			180	298	1.66	No	PWFI,PRLD
			IVTNO	1	180	298	1.66		PRCL,PARL
3	DFSIVP2	MPP			477	518	1.09	Yes	PWFI
			IVTNV	1	477	518	1.09		WFI,PARL
3	PHDAMINQ	MPP			15	22	1.47	No	PWFI,PRLD
			THDAMINQ	2	15	22	1.47		PRCL
3	PHDAMUPD	MPP			47	92	1.96	No	PWFI,PRLD
			THDAMUPD	2	47	92	1.96		PARL

bmc

Monitoring IMS Scheduling

Monitor 1st DLI Time

- High 1st DLI, but low CPU indicates program load ➔ consider preload



+W1 =IPGMSCHD==========I152IMS==*========13MAR2017==12:56:42====MVIMS====U==

Program Schedule Statistics

PGM Preload Collection: ON Clear Collection Table: NO

Regn Id	Scheduling Change Recommendations	Program Size	Total 1st DLI CPU	Total 1st DLI Time	DB2 ESS	Trn Ser	Proc Limit	Par Limit
1	PWFI,PRLD	3944	21.88	69.56	n/a	n/a	n/a	n/a
	PRCL,PARL		21.88	69.56	No	No	5	65535
1	PWFI	3944	34.57	52.47	n/a	n/a	n/a	n/a
	WFI,PARL		34.57	52.47	No	No	65535	65535
1	PWFI,PRLD	18112	110.83	1093.05	n/a	n/a	n/a	n/a
	PRCL		110.83	1093.05	No	No	5	1
1	PWFI,PRLD	7312	13.24	62.27	n/a	n/a	n/a	n/a
	PARL		13.24	62.27	No	No	65535	65535
2	PRLD	3944	4.16	13.49	n/a	n/a	n/a	n/a
	WFI,PRCL,PARL		4.16	13.49	No	No	5	65535
2		3944	0.79	1.32	n/a	n/a	n/a	n/a
	PARL		0.79	1.32	No	No	65535	65535
2	PRLD	18112	16.25	202.52	n/a	n/a	n/a	n/a
	PRCL		16.25	202.52	No	No	5	1
2	PRLD	7312	0.62	10.05	n/a	n/a	n/a	n/a
	PARL,PARL		0.62	10.05	No	No	65535	65535
3	PWFI,PRLD	3944	61.11	175.35	n/a	n/a	n/a	n/a
	PRCL,PARL		61.11	175.35	No	No	5	65535
3	PWFI	3944	27.66	47.50	n/a	n/a	n/a	n/a
	WFI,PARL		27.66	47.50	No	No	65535	65535
3	PWFI,PRLD	18112	88.52	774.53	n/a	n/a	n/a	n/a
	PRCL		88.52	774.53	No	No	5	1
3	PWFI,PRLD	7312	20.10	92.28	n/a	n/a	n/a	n/a
	PARL		20.10	92.28	No	No	65535	65535

Monitoring IMS Scheduling

Find candidates for preload

- Use GTF to monitor program load SVC 8 and SVC 122 for language environment
- GTF input parameters

```
****** ***************************
000001 TRACE=JOBNAMEP,SVCP
000002 *-------------------------
000003 JOBNAME=IMSI14Z
000004 *-------------------------
000005 * SVC
000006 *-------------------------
000007 SVC=(8,122)
000008 END
****** ***************************
```

- Start GTF

  /S GTF.GTF

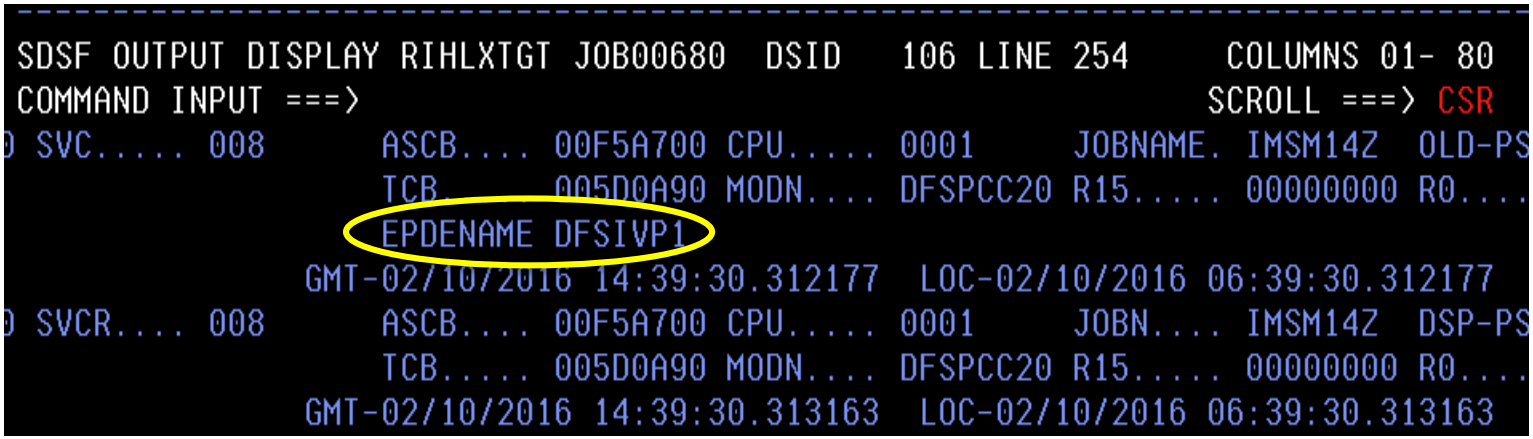
- Stop GTF

  /P GTF

Monitoring IMS Scheduling

Evaluate GTF output to find candidates for preload

- Use TSO IPCS to print the GTF trace dataset



- Find the application program(s) that has high number of loads

Conclusion

Monitor frequently

- Make sure region occupancy for all regions does not exceed 60%
- Provide enough regions to serve each class
- PSB (CSA,DLI) and PSBW pools are large enough during peak time
- Check input queue time for each transaction
- Run GTF trace for new application for preload decision
- Monitor schedule-to-1st DL/I time for each transaction
- Consider using WFI/PWFI for transaction/region

Make change and monitor again

- Look at the big picture
- Monitor virtual storage usage of region after WFI/PWFI change
- Did my change make it better or worse?

Thank you !