

Root Causing Performance Failures: A Guide to Fix Performance Issues Using QoS Knobs for the DRA74x, DRA75x, TDA2x and TDA3x Family of Devices

Prashant Karandikar, Rahul Gulati, Piyali Goswami, Prithvi Shankar Y.A., Vasant Easwaran, Chaitanya Ghone, Mihir Mody, Vinay Amancha

ABSTRACT

This application report provides a methodology through which performance issues can be identified and fixed in systems using DRA74x, DRA75x, TDA2x and TDA3x family of devices.

Contents

1	Introduction	1
2	Observability	2
3	Overview of QoS Mechanisms	3
4	Case Studies	4
5	References	11

List of Figures

1	High Level Summary of QoS Mechanisms in DRA74x/DRA75x/TDA2x/TDA3x Family of Devices	3
2	Display Bandwidth Profile for a Video Pipe @480p in DRA74x/DRA75x/TDA2x/TDA3x Family of Devices	4
3	Display Bandwidth Profile for Concurrent Display Pipelines (4 in this case – 2 video pipelines @ 720p and 1 video + graphics pipeline @ 480p) in DRA74x/DRA75x/TDA2x Family of Devices.....	4
4	Display Subsystem Profile for a Single Display Pipeline With Competing External Memory Traffic From BB2D for DRA74x/DRA75x/TDA2x/TDA3x Family of Devices	5
5	Effect of Enabling Display Subsystem Dynamic Priority Escalation for DRA74x/DRA75x Family of Devices...	6
6	BB2D Bandwidth Profile With Bandwidth Limiter Enabled for DRA74x/DRA75x Family of Devices	6
7	Concurrent Bandwidth Profile When no Underflow is Observed in Display Subsystem for DRA74x/DRA75x Family of Devices	7
8	IVA Subsystem Bandwidth Profile for 1080p Decode Concurrent With DSP EDMA and System Level EDMA traffic	8
9	IVA Subsystem Bandwidth Profile for 1080p Decode Concurrent With DSP EDMA and System Level EDMA Traffic and With IVA Subsystem Bandwidth Regulator Enabled.....	9
10	Impact of Bandwidth Regulator on DSP MDMA Latency	10
11	Write Latency With Dynamic Memory Manager Priority Extension Generator Enabled/Disabled	11

1 Introduction

Root causing performance failures due to concurrent use case scenarios in a system-on-chip (SoC) requires collaboration between hardware and software engineers. Typically, system level observability is required in order to understand any performance issues that may arise due to concurrent traffic scenarios. Having such an observability capability in the SoC eases identification of root causes related to system level performance issues. Advanced SoC family of devices like DRA74x, DRA75x, TDA2x and TDA3x provide several quality-of-service (QoS) knobs that aid to observe and control system resources in the SoC to assure the needed performance for a use case scenario.

Arteris is a registered trademark of Arteris, Inc.
 All other trademarks are the property of their respective owners.

This document covers the issues and solutions related to following different concurrent use case scenarios for DRA74x, DRA75x, TDA2x and TDA3x family of devices:

- Real-time display subsystem (DSS) underflow
- Pseudo real-time video subsystem performance (fps)
- Pseudo real-time processor subsystem performance
- External memory access latency

NOTE: The level 3 (L3) interconnect is an instantiation of the Network On Chip (NoC) interconnect from Arteris®, Inc.

All materials and trademarks are used under license from Arteris, Inc. For additional information, see the Arteris reference manuals, or contact Arteris, Inc.

2 Observability

DRA74x, DRA75x, TDA2x and TDA3x family of SoCs provide observability of performance parameters like bandwidth and latency by embedding statistics collectors within the SoC. Data from the statistics collector can be collected using the system trace subsystem (STM) or using an unused CPU in the system.

The following is a list of typical issues seen while using system trace (STM):

- Lack of availability of the STM debugger connectivity in production and customer boards.
- Fine grained observability requirements in order to identify performance issues. Typically, observability is required at finer steps of 100 μ s, whereas, the basic input/output system (BIOS) or high-level operating system (HLOS) tasks are scheduled at 2-3 ms. Hence, using BIOS or HLOS-based data collection may not be as meaningful as required. A particular IP may send high priority bandwidth burst for 500 μ s that gets averaged out while measuring in an HLOS using operating system task scheduling, which would be of order of 1 ms to 10 ms (depending on HLOS in use).

Considering the above mentioned limitations, the following setup is recommended in boards where the STM debugger connectivity is not possible:

- Availability of an unused CPU core with access to statistic collector registers within L3 interconnect
- A timer with 100 μ s interrupt capability
- A debug console universal asynchronous receiver/transmitter (UART)
- Availability of 1KB-2KB on-chip RAM usage for non-intrusive instrumentation code

For guidelines on how to configure L3 statistics collectors, timers and debug console, see the starterware APIs. For more information on starterware APIs, contact your TI sales representative.

3 Overview of QoS Mechanisms

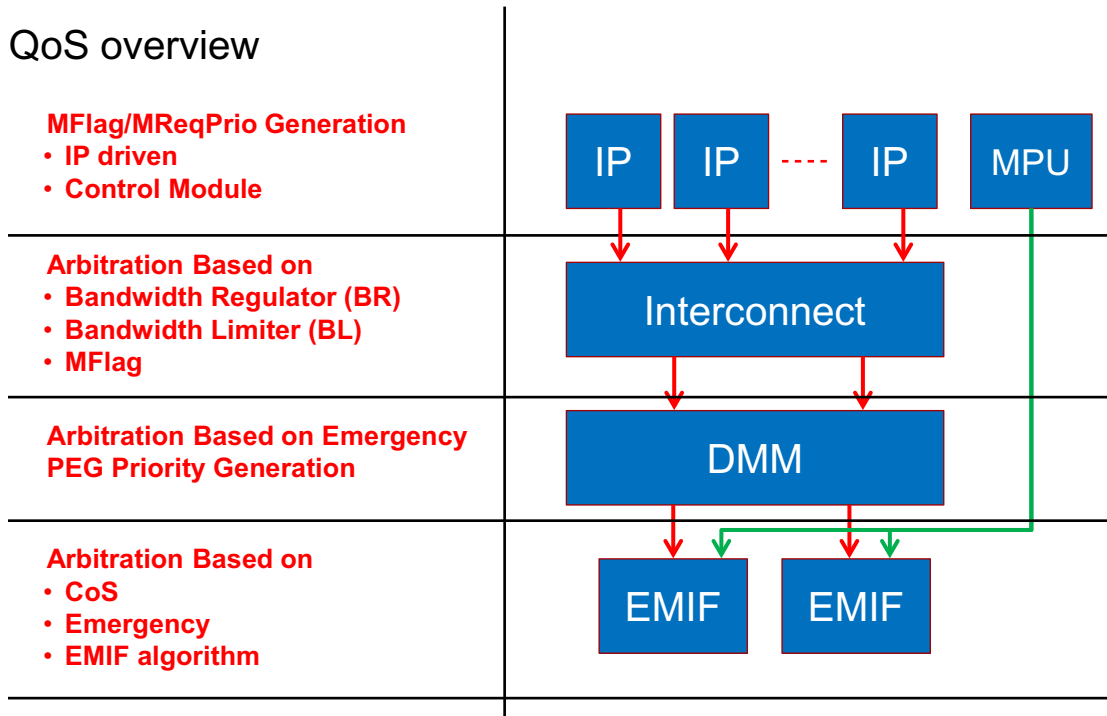


Figure 1. High Level Summary of QoS Mechanisms in DRA74x/DRA75x/TDA2x/TDA3x Family of Devices

For detailed information regarding QoS mechanisms, see the device-specific data sheets or the relevant DRA74x/DRA75x/TDA2x/TDA3x QoS application reports.

4 Case Studies

4.1 Display Underflow

In this case study, it is observed that the display subsystem exhibits underflow when high bandwidth traffic from other initiators competes with the display subsystem traffic for accessing the external memory.

4.1.1 Observability

A single DSS pipe configured for 480P ARGB display requires an average bandwidth of approximately 83 MB/s ($720 \times 480 \times 4 \text{ Bytes/Pixel} \times 60 \text{ fps}$) during frame time. Considering the blanking periods, the bandwidth requirement during active period is higher than 83 MB/s.

Figure 2 shows the display subsystem bandwidth profile for a single pipe without any competing external memory traffic.

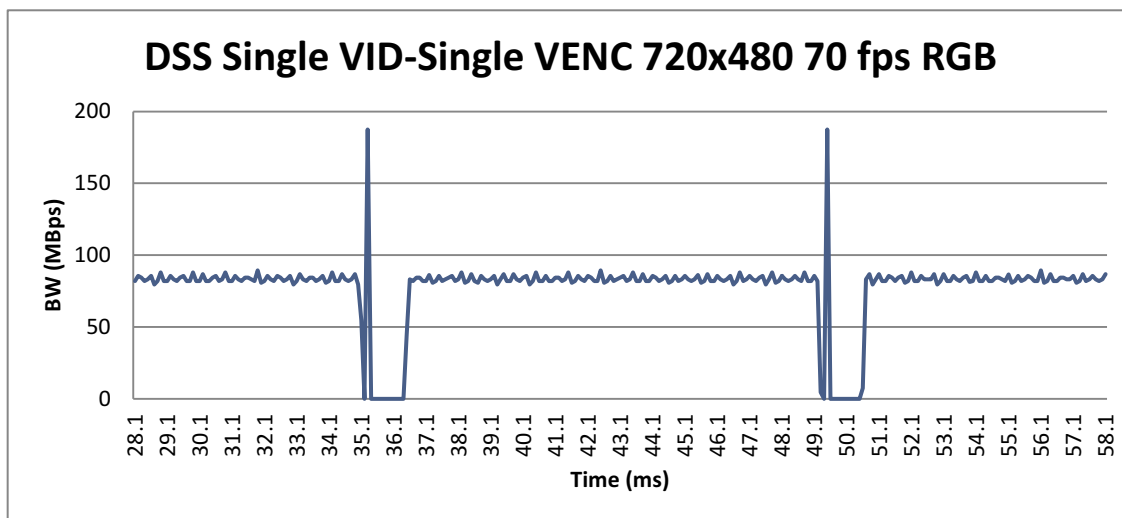


Figure 2. Display Bandwidth Profile for a Video Pipe @480p in DRA74x/DRA75x/TDA2x/TDA3x Family of Devices

Figure 3 shows the display subsystem bandwidth profile for four concurrent display pipes without any competing external memory traffic.

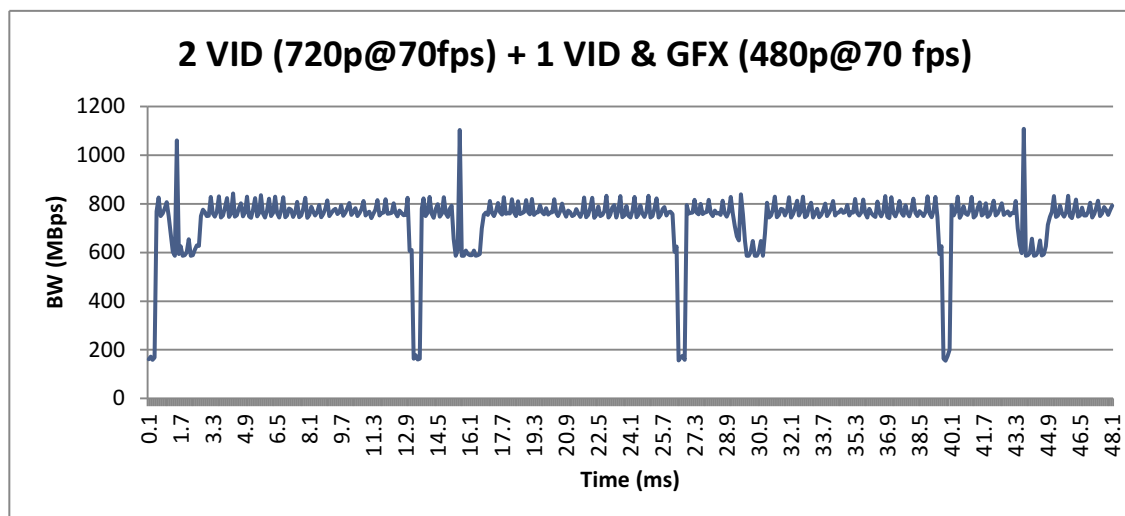


Figure 3. Display Bandwidth Profile for Concurrent Display Pipelines (4 in this case – 2 video pipelines @ 720p and 1 video + graphics pipeline @ 480p) in DRA74x/DRA75x/TDA2x Family of Devices

4.1.2 Issue: Display Subsystem not Meeting Average Bandwidth During Active Time

When concurrent traffic from BB2D is enabled, display subsystem underflow is observed.

Figure 4 depicts the bandwidth profile for display subsystem with a single pipeline and competing external memory traffic from BB2D.

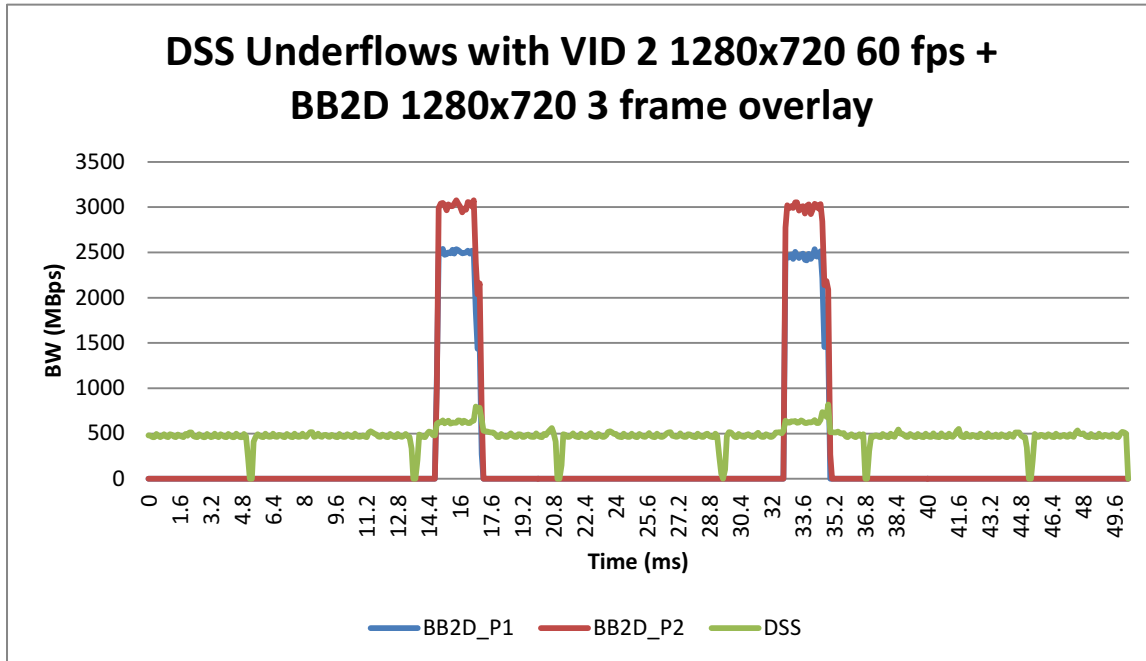


Figure 4. Display Subsystem Profile for a Single Display Pipeline With Competing External Memory Traffic From BB2D for DRA74x/DRA75x/TDA2x/TDA3x Family of Devices

Since BB2D uses a memory-to-memory processing, maximum possible traffic from BB2D is injected onto the L3 interconnect. Considering overall bandwidth available in the system, BB2D gets more than its fair share in the system.

4.1.3 Possible Solutions

There are two possible solutions to tackle this problem.

4.1.3.1 Solution 1

DSS has an in-built dynamic priority escalation and de-escalation mechanism, MFLAG, which works on FIFO levels. This mechanism can be enabled by making the pipelines that exhibit underflow issues as high priority pipelines and setting the low and high threshold at 50% and 75%, respectively. Note that the graphics pipeline within the display subsystem has a 16 KB FIFO and the three video pipelines viz. VID1/2/3 have a 32 KB FIFO each.

Use the display subsystem starterware APIs to configure the dynamic priority escalation and de-escalation as mentioned above.

Figure 5 depicts the share of BB2D bandwidth dropping when the dynamic priority escalation for display subsystem has been enabled.

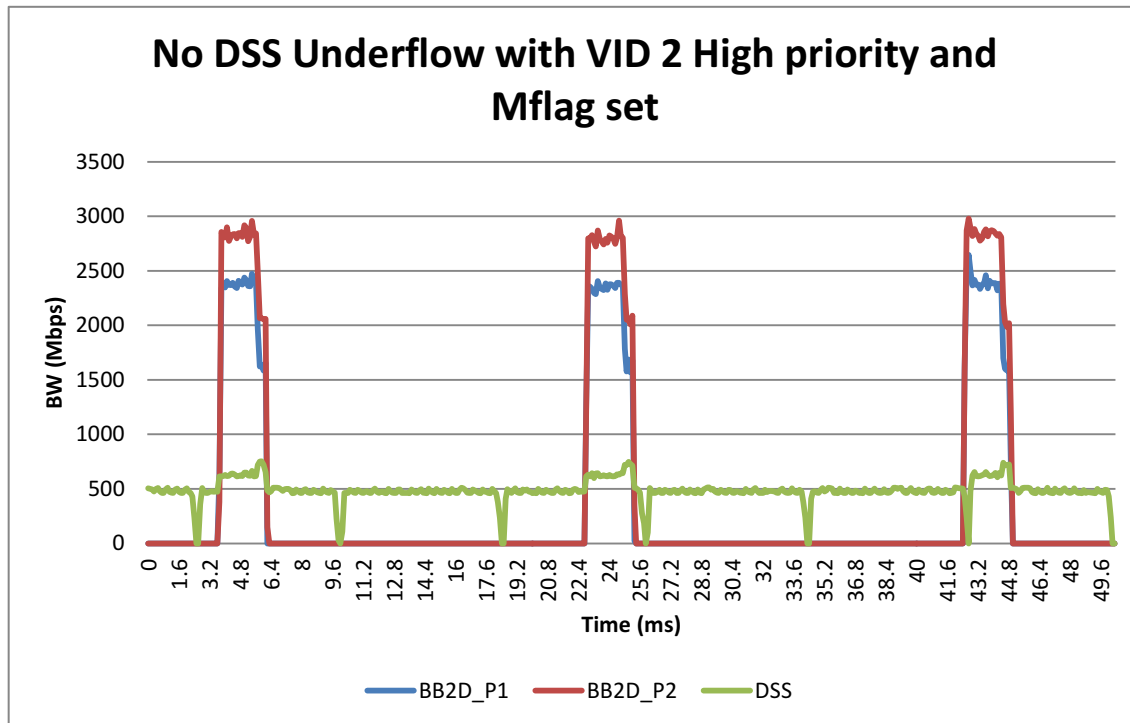


Figure 5. Effect of Enabling Display Subsystem Dynamic Priority Escalation for DRA74x/DRA75x Family of Devices

4.1.3.2 Solution 2

Since the BB2D initiator consumes more than the required bandwidth to complete the task, it finishes its task ahead of the required time, causing the display subsystem to miss its real-time deadline. The BB2D initiator bandwidth can be controlled by using a QoS mechanism known as Bandwidth Limiter, an L3 interconnect component.

Use of Bandwidth Limiter for BB2D reduces its average bandwidth consumption, at the expense of increase in its task duration. However, the Bandwidth Limiter is configured such that the BB2D initiator still completes its task with sufficient margin with respect to its deadline.

Figure 6 depicts the bandwidth profile for BB2D with the bandwidth limiter enabled.

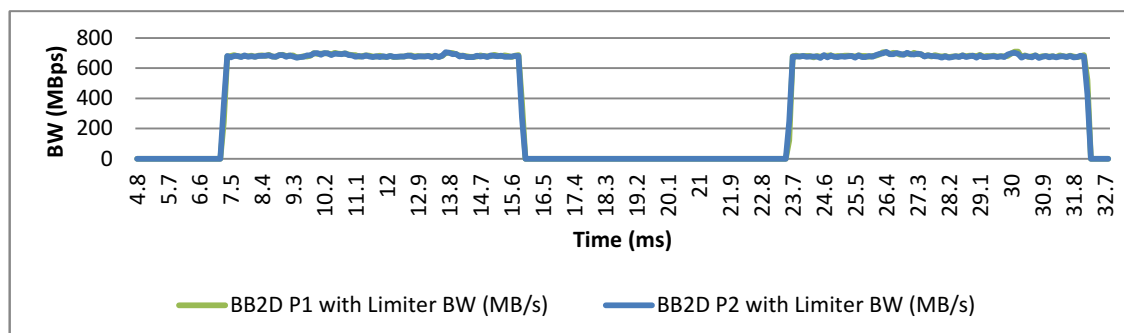


Figure 6. BB2D Bandwidth Profile With Bandwidth Limiter Enabled for DRA74x/DRA75x Family of Devices

Figure 7 depicts the concurrent bandwidth profile for the display subsystem and BB2D with the bandwidth limiter enabled for BB2D.

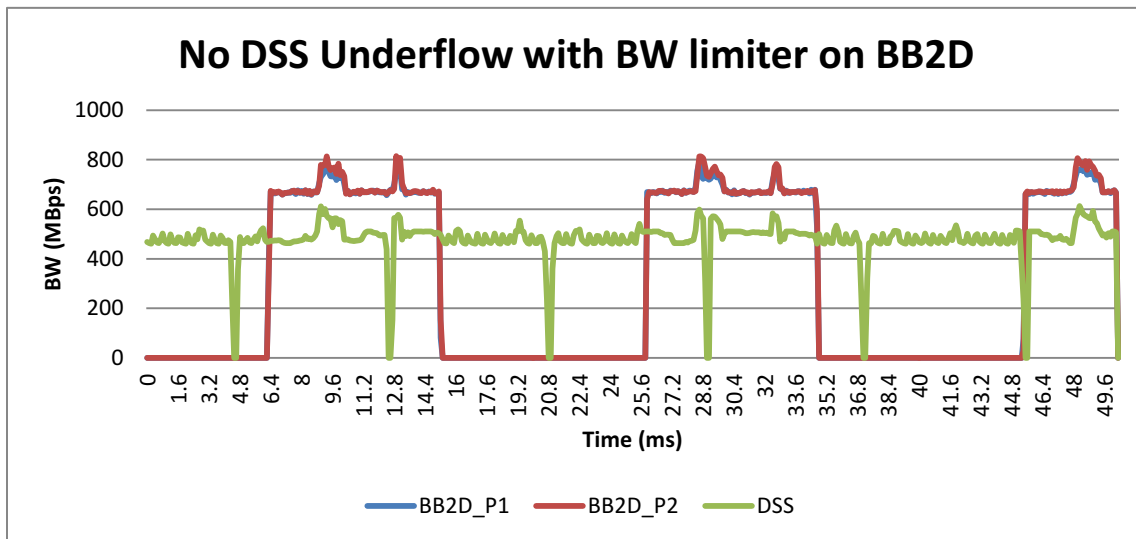


Figure 7. Concurrent Bandwidth Profile When no Underflow is Observed in Display Subsystem for DRA74x/DRA75x Family of Devices

4.2 Pseudo Real-Time Video Subsystem Performance (fps)

IVA subsystem, an engine for video compression and decompression, shows varied amounts of traffic for I/P/B kind of frames. The subsystem performance is based on both compute and data transfer time. Typically, in a concurrent traffic scenario, the IVA subsystem performance could degrade because of delays in data transfer to the subsystem. Since the IVA subsystem has a limited amount of buffering capability, it is inherently immune to certain latency expected in data transfers. In a concurrent scenario, it is observed that the IVA subsystem does not get the required average data transfer (bandwidth), causing degradation in the required fps (frames per second).

4.2.1 Issue/Observability

Figure 8 depicts the bandwidth profile for IVA subsystem traffic concurrent with DSP EDMA traffic and System EDMA traffic. The highlighted portion indicates the degradation seen by the IVA subsystem due to additional traffic. Specifically, the B-Frame takes 34.466 ms for its completion (instead of 33.33 ms) leading to an fps drop from 30 fps to 29 fps.

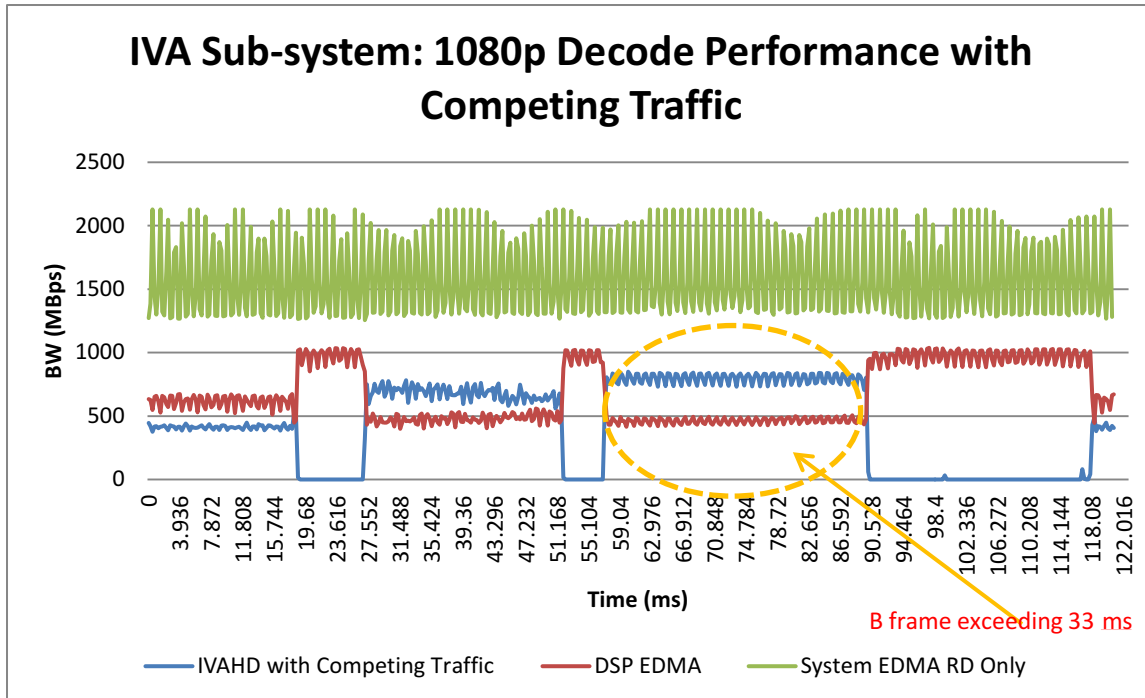


Figure 8. IVA Subsystem Bandwidth Profile for 1080p Decode Concurrent With DSP EDMA and System Level EDMA traffic

4.2.2 Possible Solution: Usage of QoS Knobs to Resolve the Degradation Seen by IVA Subsystem

Bandwidth Regulator dedicated for IVA subsystem within the L3 interconnect provides a mechanism to dynamically modify the priority of IVA subsystem in the SoC. Considering that the IVA subsystem requires about 1500 MB/s for video decode for B frames, the bandwidth regulator is configured for this requirement.

Figure 9 depicts the impact of enabling the IVA subsystem bandwidth regulator. It is evident that the performance of B-frames in particular is now well within the 33 ms requirement, thereby, enabling the overall 30 fps requirement for the IVA subsystem.

For guidelines on how to configure IVA subsystem Bandwidth Regulator within L3 interconnect, see the starterware APIs.

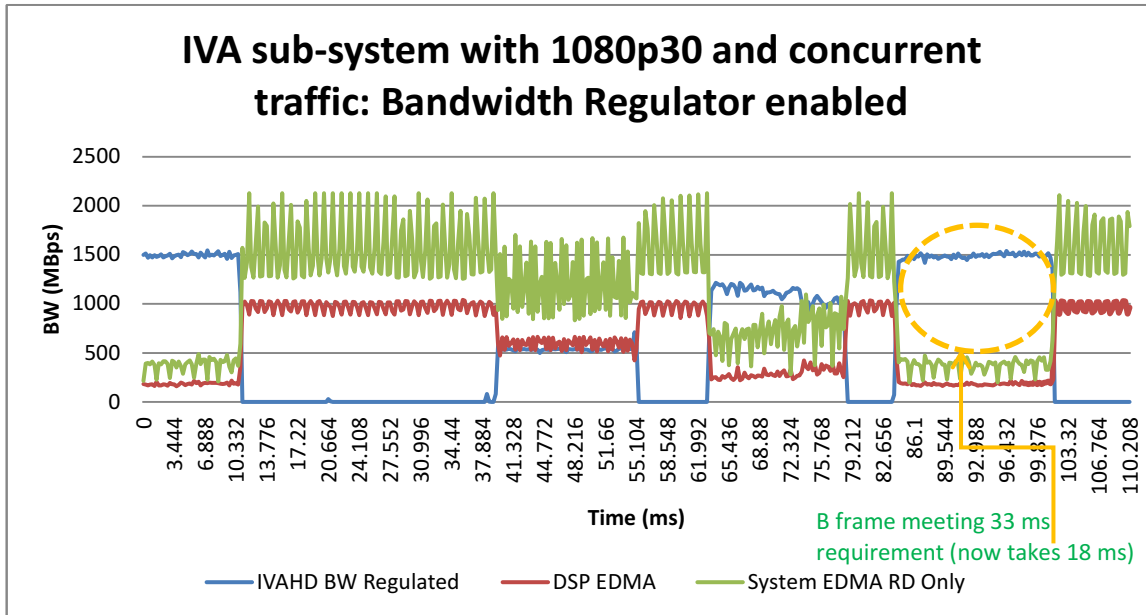


Figure 9. IVA Subsystem Bandwidth Profile for 1080p Decode Concurrent With DSP EDMA and System Level EDMA Traffic and With IVA Subsystem Bandwidth Regulator Enabled

4.3 Pseudo Real-Time Processor Subsystem Performance

In general, latency of instruction and data fetches has a significant impact on the processor subsystem performance. Figure 10 demonstrates the impact of enabling the bandwidth regulator on the processor subsystem latency. DSP1 is the processor subsystem under consideration in this case. The test case comprises of running a hand-crafted optimized memcpy routine on the C66x core within the DSP subsystem available on DRA74x, DRA75x, TDA2x and TDA3x family of devices. Such a test would lead to an average bandwidth of 2500 MBytes per second in a standalone condition. Due to the addition of competing EDMA traffic from two EDMA transfer controllers within the DSP2 subsystem, it is observed that the DSP1 subsystem bandwidth drops to 981 MBytes per second. It is possible to achieve the desired bandwidth through appropriate configuration of the DSP1 MDMA bandwidth regulator available within the L3 interconnect. The bandwidth regulator essentially reduces the latency of the DSP1 MDMA port for instruction and data fetches. It can be observed from Figure 10 that the latency for DSP1 MDMA port improves from ~180 cycles/transaction to ~110 cycles/transaction after configuring the DSP1 MDMA bandwidth regulator.

NOTE: The system integrator needs to keep in mind the bandwidth and latency requirements for all the initiators involved in the concurrent use case scenario in order to determine the appropriate configuration values for various available QoS elements, like bandwidth regulators, bandwidth limiters, and so forth. For the various QoS options available on your device, see *Quality of Service (QoS) Knobs for DRA74x, DRA75x & TDA2x Family of Devices* (SPRABX1).

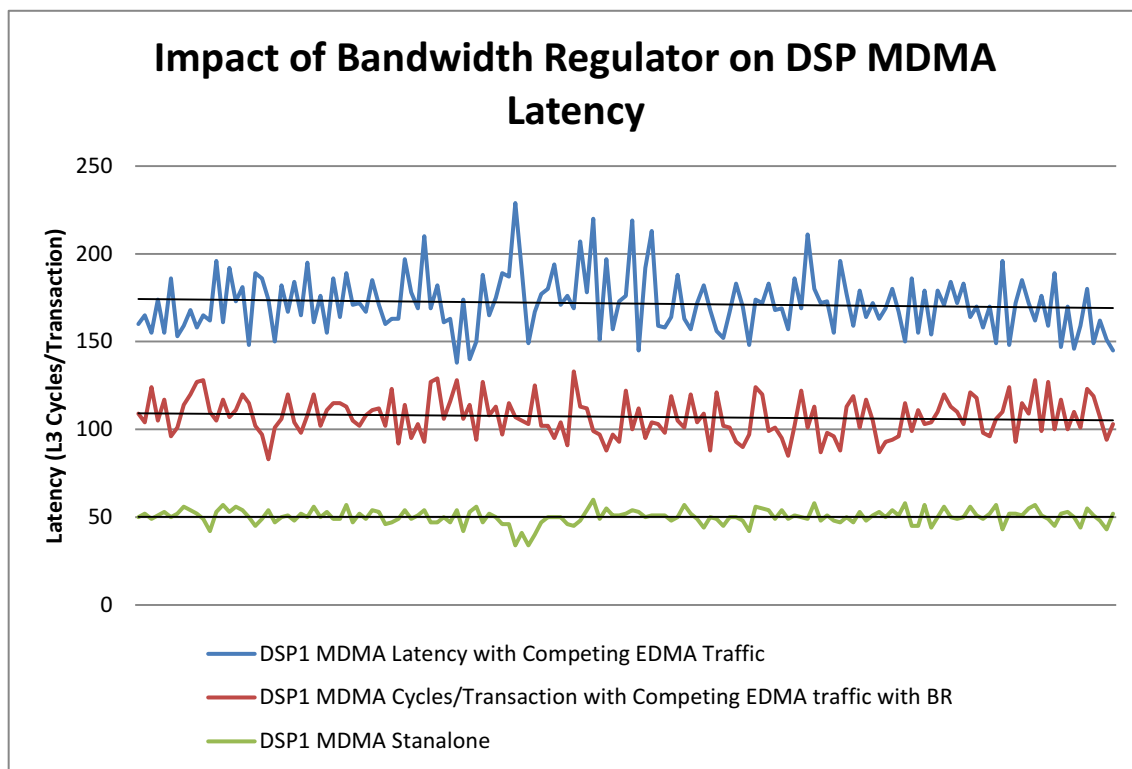


Figure 10. Impact of Bandwidth Regulator on DSP MDMA Latency

4.4 Transactions Stuck in Memory Subsystem for an Extended Time Duration

The case study chosen here comprises of concurrent traffic from 4 x EDMA TC channels. The traffic pattern is generated such that three EDMA TC are reading from the external memory and the fourth EDMA TC is writing to the external memory.

Figure 11 depicts the impact of modifying the initiator priority using the priority extension generator (PEG) feature of the dynamic memory manager (DMM) inside of the memory subsystem on latency observed by the EDMA TC channel performing write transactions. The DMM is composed of a block aimed at generating software-programmable initiator-indexed priority extensions for any of the PEG connID groups. For details of the DMM PEG configuration registers, see the device-specific technical reference manual.

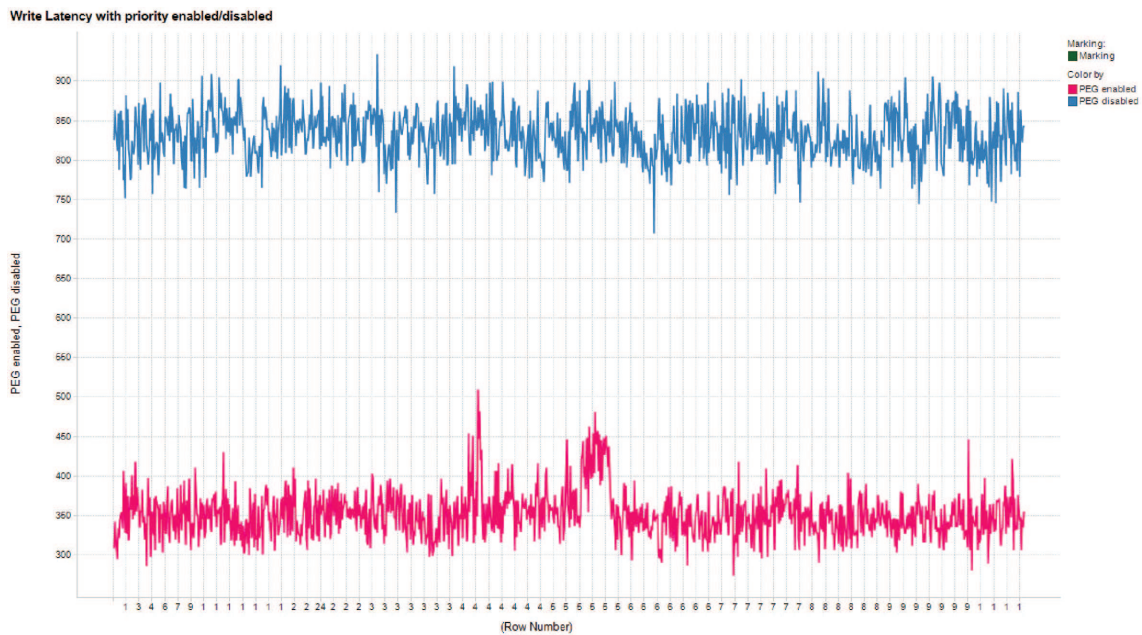


Figure 11. Write Latency With Dynamic Memory Manager Priority Extension Generator Enabled/Disabled

5 References

- Quality of Service (QoS) Knobs for DRA74x, DRA75x & TDA2x Family of Devices ([SPRABX1](#))

IMPORTANT NOTICE

Texas Instruments Incorporated and its subsidiaries (TI) reserve the right to make corrections, enhancements, improvements and other changes to its semiconductor products and services per JESD46, latest issue, and to discontinue any product or service per JESD48, latest issue. Buyers should obtain the latest relevant information before placing orders and should verify that such information is current and complete. All semiconductor products (also referred to herein as "components") are sold subject to TI's terms and conditions of sale supplied at the time of order acknowledgment.

TI warrants performance of its components to the specifications applicable at the time of sale, in accordance with the warranty in TI's terms and conditions of sale of semiconductor products. Testing and other quality control techniques are used to the extent TI deems necessary to support this warranty. Except where mandated by applicable law, testing of all parameters of each component is not necessarily performed.

TI assumes no liability for applications assistance or the design of Buyers' products. Buyers are responsible for their products and applications using TI components. To minimize the risks associated with Buyers' products and applications, Buyers should provide adequate design and operating safeguards.

TI does not warrant or represent that any license, either express or implied, is granted under any patent right, copyright, mask work right, or other intellectual property right relating to any combination, machine, or process in which TI components or services are used. Information published by TI regarding third-party products or services does not constitute a license to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property of the third party, or a license from TI under the patents or other intellectual property of TI.

Reproduction of significant portions of TI information in TI data books or data sheets is permissible only if reproduction is without alteration and is accompanied by all associated warranties, conditions, limitations, and notices. TI is not responsible or liable for such altered documentation. Information of third parties may be subject to additional restrictions.

Resale of TI components or services with statements different from or beyond the parameters stated by TI for that component or service voids all express and any implied warranties for the associated TI component or service and is an unfair and deceptive business practice. TI is not responsible or liable for any such statements.

Buyer acknowledges and agrees that it is solely responsible for compliance with all legal, regulatory and safety-related requirements concerning its products, and any use of TI components in its applications, notwithstanding any applications-related information or support that may be provided by TI. Buyer represents and agrees that it has all the necessary expertise to create and implement safeguards which anticipate dangerous consequences of failures, monitor failures and their consequences, lessen the likelihood of failures that might cause harm and take appropriate remedial actions. Buyer will fully indemnify TI and its representatives against any damages arising out of the use of any TI components in safety-critical applications.

In some cases, TI components may be promoted specifically to facilitate safety-related applications. With such components, TI's goal is to help enable customers to design and create their own end-product solutions that meet applicable functional safety standards and requirements. Nonetheless, such components are subject to these terms.

No TI components are authorized for use in FDA Class III (or similar life-critical medical equipment) unless authorized officers of the parties have executed a special agreement specifically governing such use.

Only those TI components which TI has specifically designated as military grade or "enhanced plastic" are designed and intended for use in military/aerospace applications or environments. Buyer acknowledges and agrees that any military or aerospace use of TI components which have **not** been so designated is solely at the Buyer's risk, and that Buyer is solely responsible for compliance with all legal and regulatory requirements in connection with such use.

TI has specifically designated certain components as meeting ISO/TS16949 requirements, mainly for automotive use. In any case of use of non-designated products, TI will not be responsible for any failure to meet ISO/TS16949.

Products

Audio	www.ti.com/audio
Amplifiers	amplifier.ti.com
Data Converters	dataconverter.ti.com
DLP® Products	www.dlp.com
DSP	dsp.ti.com
Clocks and Timers	www.ti.com/clocks
Interface	interface.ti.com
Logic	logic.ti.com
Power Mgmt	power.ti.com
Microcontrollers	microcontroller.ti.com
RFID	www.ti-rfid.com
OMAP Applications Processors	www.ti.com/omap
Wireless Connectivity	www.ti.com/wirelessconnectivity

Applications

Automotive and Transportation	www.ti.com/automotive
Communications and Telecom	www.ti.com/communications
Computers and Peripherals	www.ti.com/computers
Consumer Electronics	www.ti.com/consumer-apps
Energy and Lighting	www.ti.com/energy
Industrial	www.ti.com/industrial
Medical	www.ti.com/medical
Security	www.ti.com/security
Space, Avionics and Defense	www.ti.com/space-avionics-defense
Video and Imaging	www.ti.com/video

TI E2E Community

e2e.ti.com