Perf Limit Reasons TPMI Register Interface

Introduction

PLR (Performance Limit Reasons) is an observability interface specifically designed for in-field issue triaging, allowing users to quickly identify and address performance bottlenecks. The main objective is to provide information to customers on what is limiting CPU performance from a power and thermal management point of view. In power management world, frequency is loosely used as a proxy even in the case of PLR. Hence PLR provides info on what is limiting CPU frequency. In servers, each core and mesh (uncore) can run at an independent frequency. Hence the lowest granularity at which PLR are reported is per core and mesh. However, it is time consuming if a user has to read 100s of registers to figure out what is wrong with performance. Hence these per core/mesh PLR are aggregated to a higher level. PLR that is aggregated to a higher level is called Coarse Grained PLR. The per core/mesh PLR are referred to as Fine Grained PLR. User can read coarse grained PLR for initial triaging or even poll these periodically and read fine grained PLR only when additional info is needed.

The info provided over PLR might not be converged with other observability interfaces like telemetry, per-feature status registers. However, they are not expected to be completely off. There can be slight differences due to micro-architecture differences of how these interfaces are implemented.

Interface

The legacy MSR listed below are no longer supported on GNR/SRF onwards. They are replace by TPMI interface registers.

- MSR_CORE_PERF_LIMIT_REASONS (64Fh): Interface to allow software to determine what is causing resolved frequency to be clamped below the requested core frequency.
- MSR_RING_PERF_LIMIT_REASONS (6B1h) : Interface to allow software to determine what is causing the frequency to be clamped below the requested frequency for the ring (mesh) domain.

TPMI

Starting on GNR generation SOC, PLR will be exposed through TPMI with TPMI ID = 0xC. PLR interface is die-scoped.

Following are the TPMI register details:

PLR_HEADER

Name		PLR_MAILBOX_INTERFACE
Bits	**Field**	**Description**
7:0	**INTERFACE_VERSION**	Version number of this interface. Bit[7:5] denotes Major Version and bit[4:0] for Minor Version.
63:8	**RESERVED**	Reserved

PLR_MAILBOX_INTERFACE

Name		PLR_MAILBOX_INTERFACE
Bits	**Field**	**Description**
7:0	**COMMAND**	0: read, 1: write
11:8	**DOMAIN**	0: core, 1: uncore
19:12	**ID**	Logical Module ID within the domain
62:20	**RESERVED**	Reserved

63:63	**RUN_BUSY**	Set by SW, cleared by HW

PLR_MAILBOX_DATA

Name		PLR_MAILBOX_DATA
Bits	**Field**	**Description**
63:0	**DATA**	Perf Limit Reasons bit. Set by HW, cleared by SW.

Following is the bit definitions for coarse grained and fine grained perf limit reasons:

Coarse Grained PLR Bit Definition (source)

Bit #	**Coarse Grained**	
	Name	**Details**
0	FREQUENCY	Limitation due to Turbo Ratio Limit
1	CURRENT	Package ICCmax or MT-Pmax
2	POWER	Socket or Platform RAPL
3	THERMAL	Thermal Throttling
4	PLATFORM	Prochot or Hot VR
5	MCP	freq limit due to a companion die like PCH
6	RAS	freq limit due to RAS

7	MISC	Freq limit from out-of-band SW (e.g. BMC)
8	QOS	SST-CP, SST-BF, SST-TF
9	DFC	Freq limitation due to Dynamic Freq Capping
10	RESERVED	Reserved
11	RESERVED	Reserved
12	RESERVED	Reserved
13	RESERVED	Reserved
14	RESERVED	Reserved
15	RESERVED	Reserved
16	RESERVED	Reserved
17	RESERVED	Reserved
18	RESERVED	Reserved
19	RESERVED	Reserved
20	RESERVED	Reserved
21	RESERVED	Reserved

22	RESERVED	Reserved
23	RESERVED	Reserved
24	RESERVED	Reserved
25	RESERVED	Reserved
26	RESERVED	Reserved
27	RESERVED	Reserved
28	RESERVED	Reserved
29	RESERVED	Reserved
30	RESERVED	Reserved
31	RESERVED	Reserved

Fine Grained PLR Bit Definition (source)

Bit #	Fine Grained		Mapping to Coarse Grained
	Name	**Details**	
0	CDYN0	TRL index 0	FREQUENCY
1	CDYN1	TRL index 1	
2	CDYN2	TRL index 2	

3	CDYN3	TRL index 3	
4	CDYN4	TRL index 4	
5	CDYN5	TRL index 5	
6	FCT	Favored Core Turbo	
7	PCS_TRL	TRL from out-of-band (BMC)	
8	MTPMAX		CURRENT
9	FAST_RAPL		POWER
10	PKG PL1_MSR_TPMI		
11	PKG PL1_MMIO		
12	PKG PL1_PCS		
13	PKG PL2_MSR_TPMI		
14	PKG PL2_MMIO		
15	PKG PL2_PCS		
16	PLATFORM_PL1_MSR_TPMI		
17	PLATFORM_PL1_MMIO		

18	PLATFORM_PL1_PCS		
19	PLATFORM_PL2_MSR_TPMI		
20	PLATFORM_PL2_MMIO		
21	PLATFORM_PL2_PCS		
22	RSVD		
23	PER_CORE_THERMAL	Thermal Throttling	THERMAL
24	UFS_DFC	Dynamic Freq Capping	DFC
25	XXPROCHOT		PLATFORM
26	HOT_VR		
27	RSVD		
28	RSVD		
29	PCS_PSTATE		MISC
30	RSVD		RESERVED
31	RSVD		RESERVED

PLR_DIE_LEVEL

Name		PLR_DIE_LEVEL
Bits	**Field**	**Description**
63:0	**DATA**	Coarse Grained Perf Limit Reasons. Set by HW, cleared by SW.

NOTE:

- PLR_MAILBOX_DATA[31:0] = Coarse grained reasons as shown above.
- PLR_MAILBOX_DATA[63:32] = Fine grained reasons as shown above.
- PLR_DIE_LEVEL[31:0] = Coarse grained reasons as shown above.
- PLR_DIE_LEVEL[63:32] = 0x0