**# Title:**

RAS2 Extensions

**# Status:**

Draft

**# Document:**

ACPI Specification Version 6.next

**# License:**

SPDX-License-Identifier: CC-BY-4.0

**# Submitter:**

* Jamie Iles (Qualcomm)
* Harb Abdulhamid (Ampere)
* Thanu Rangarajan (Arm)
* Samer El-Haj Mahmoud (Arm)
* TianoCore Community (<https://www.tianocore.org>)

**# Summary of the change**

This ECR is related to a proposal to extend the capabilities of the RAS2 table. The extensions are related to better control and description of patrol scrub rates. In particular:

* The current abstract format, which is of the form {minimum rate, maximum rate}, is being enhanced to {minimum, nominal, maximum}. The new nominal rate is being defined as the scrub rate that guarantees that the entire memory in the NUMA domain is scrubbed in a 24-hour period. The choice of 24 hours is based on industry best practices.
* The nominal scrub rate is associated with an actual rate value in Gb/s. This enables the OS to make informed decisions regarding RAS vs. bandwidth tradeoffs.
* An abstract cost factor is being introduced, defined in terms of percentage of maximum memory bandwidth consumed. The cost factor is assigned to each abstract scrub rate value in the range {minimum, maximum}, which then informs the OS what % of total bandwidth is lost when a particular rate in the range is chosen. Again, this enables the OS to perform informed decision making, and to optimize patrol scrubbing to best meet RAS and bandwidth requirements of the system.

These extensions are being introduced in a manner that ensures backward compatibility with the existing RAS2 table.

Additional minor errata are also fixed as part of this ECR:

* In Table 5-83, the Set RAS Capabilities Status field has a small spelling mistake in the return status called *Failed*.
* In Table 5-e, there is a vestigial sentence from the RASF table that references the PATROL\_SCRUB parameter block, which needs to be removed.

**# Benefits of the change**

Improved patrol scrubbing control by OSPM.

**# Impact of the change**

Platforms firmware will have to modified to support the extended RAS2 format. Likewise, operating systems will require new drivers/code to support the extensions.

**# Detailed description of the change [normative updates]**

Existing text

New text

Deleted Text

**5.2.X ACPI RAS2 Feature Table (RAS2)**

…

**5.2.X.1.3 RAS2 Platform Communication Channel**

…

**Table 5.83: RAS2 Platform Communication Channel Shared Memory**

**Region**

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Byte Length** | **Byte Offset** | **Description** |
| … | … | … | … |
| Set RAS Capabilities Status | 4 | 44 | Status:  0000b = Success  0001b = Not Valid  0010b = Not Supported  0011b = Busy  0100b = Failed~~F~~  0101b = Aborted  0110b = Invalid Data |

…

**Table 5.e RAS2 Platform Communication Channel Shared Memory Region**

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Byte Length** | **Byte Offset** | **Description** |
| … | … | … | … |
| Parameter Blocks | Varies  (N  Bytes) | 48 | Start address of the parameter blocks~~, the structure of which is shown in the Parameter Block Structure for PATROL\_SCRUB~~.  These parameter blocks are used as  communication mailbox between the OSPM and the  platform, and there is 1 parameter block for each RAS  feature. NOTE: There can be only on parameter block  per type. |

**5.2.X.2.1 Hardware-based Memory Scrubbing**

…

**Table 5.j Parameter Block Structure for PATROL\_SCRUB**

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Byte Length** | **Byte Offset** | **Description** |
| … | … | … | … |
| Version | 2 | 2 | Byte 0 – Minor Version  Byte 1 – Major Version  For this format of the parameter block, this field should be set to 0x000~~1~~2. |
| Length | 2 | 4 | Length, in bytes of the entire parameter block structure. The total length must include the size of the optional Extended data region, if present. OSPM must account for the optional Extended data region when allocating buffers for storing this parameter block, and then use the Length field to indicate or determine validity and presence of the extended data region. |
| … | … | … | ... |
| Scrub Parameters (OUTPUT) | 4 | 44 | The platform returns this value in response to GET\_PATROL\_PARAMETERS.  If additional information in the Extended Data region is not present, the scrub rates returned by the platform in this field must be treated as integer values in the range {Minimum, Maximum}, where:  Rate N < Rate N+1  , where each value in this range is an abstract value that represents a certain supported scrub rate. OSPM can select a rate from this abstract range based on a heuristics-based assessment of parameters such as power, bandwidth and error rates. For example, if the error rate is high, the OS can choose a higher (more aggressive) scrub rate, and vice versa. The physical scrub rates are not relevant to such schemes.  If extended information is returned in the Extended data region, the Minimum and Maximum scrub rate fields must be used as indexes into an array of scrub rate descriptors, where each descriptor provides a set of parameters related to that scrub rate. The Minimum scrub rate field must always be 0 as it points to the first descriptor of the array, and the Maximum scrub rate field represents the index of the highest scrub rate descriptor in the array. The scrub rate descriptors provide additional information about the scrub rates, including their physical values and their impact on bandwidth and power. This format enables OSPM to perform precision-based scrub control.  Bits [7:0]: Current scrub rate that is in effect on the memory region specified in “Actual Address Range”. If OSPM requested background scrubbing, then this field will reflect the current background patrol scrubbing rate.  Bits [15:8]: Minimum scrub rate supported.  Bits [23:16]: Maximum scrub rate supported.  Bits [31:24: Reserved, must be zero. |
| … | … | … | … |
| **Extended data region** |  |  |  |
| Extended Scrub Parameters  (OUTPUT) | 4 | 52 | This field is valid only for the response to GET\_PATROL\_PARAMETERS. The platform returns this value in response to GET\_PATROL\_PARAMETERS.  Additionally, OSPM must check the Length field to determine whether this field is present.  Bits[7:0]: Nominal scrub rate index.  Bits[23:8]: Nominal scrub rate in MB/s, for a maximum nominal scrub rate of 64GB/s.  Bits[31:24]: Reserved, must be zero.  The Nominal scrub rate index must satisfy the condition:  Minimum Nominal Maximum  The Nominal rate is defined as the rate at which all memory in this proximity domain is scrubbed in a 24-hour period. |
| Array of Scrub rate descriptors [N]  (OUTPUT) | 256 \* sizeof (BYTE) | 56 | This field is valid only for the response to GET\_PATROL\_PARAMETERS. The platform returns this value in response to GET\_PATROL\_PARAMETERS.  Additionally, OSPM must check the Length field to determine whether this field is present.  Each descriptor in this array is a BYTE that represents the percentage of total memory bandwidth consumed by the scrub engine when operating at that scrub rate, for a duration of 24 hours.  A maximum of 256 distinct scrub rates can thus be specified. Descriptor[0] to Descriptor[Maximum scrub rate] are valid.  The combination of the bandwidth consumed, the index of the nominal rate and the real value of the nominal scrub rate, allows the OS to make informed decisions regarding choice of scrub rates. Lower scrub rates consume less bandwidth at the cost of reliability, while higher scrub rates consume more bandwidth to offer improved reliability. |