Chapter 7
Controller Architecture

Application team 김재희

CONTENTS

1. Controller Initialization and Shutdown Processing

1) Initialization

① 호스트는 컨트롤러가 CSTS.RDY가 '0'이 되는 것을 기다린다.

② Admin queue는 AQA, ASQ, ACQ를 적절한 값으로 설정한다.

③ 컨트롤러 셋팅이 이루어져야 한다. (CC.AMS, CC.MPS, CC.CSS에 대하여)

⑤. CC.EN을 '1'로 설정함 으로서 컨트롤러가 활성 된다.

⑥. 컨트롤러는 CSTS.RDY(Ready)가 '1'로 설정되면서 커맨드 처리 준비가 완료된다.

⑦. 호스트는 Set features 커맨드에 지원되는 I/O SQ, CQ 의 개수를 결정한다. 결정되면 MSI and/or MSI-X 레지스터가 설정되어야 한다.

⑧. 호스트는 컨트롤러가 지원하는 수만큼 I/O CQ/SQ 를 할당한다.

→ The controller may be used for I/O commands

1. Controller Initialization and Shutdown Processing

CSTS.RDY

00	RO	0	**Ready (RDY):** This field is set to '1' when the controller is ready to accept Submission Queue Tail doorbell writes after CC.EN is set to '1'. This field shall be cleared to '0' when CC.EN is cleared to '0'. Commands shall not be submitted to the controller until this field is set to '1' after the CC.EN bit is set to '1'. Failure to follow this requirement produces undefined results. Host software shall wait a minimum of CAP.TO seconds for this field to be set to '1' after setting CC.EN to '1' from a previous value of '0'.

CC.EN

00	RW	0	**Enable (EN):** When set to '1', then the controller shall process commands based on Submission Queue Tail doorbell writes. When cleared to '0', then the controller shall not process commands nor post completion queue entries to Completion Queues. When this field transitions from '1' to '0', the controller is reset (referred to as a Controller Reset). The reset deletes all I/O Submission Queues and I/O Completion Queues, resets the Admin Submission Queue and Completion Queue, and brings the hardware to an idle state. The reset does not affect PCI Express registers (including MMIO MSI-X registers), nor the Admin Queue registers (AQA, ASQ, or ACQ). All other controller registers defined in this section and internal controller state (e.g., Feature values defined in section 5.21.1 that are not persistent across power states) are reset to their default values. The controller shall ensure that there is no data loss for commands that have had corresponding completion queue entries posted to an I/O Completion Queue prior to the reset operation. Refer to section 7.3 for reset details. When this field is cleared to '0', the CSTS.RDY bit is cleared to '0' by the controller once the controller is ready to be re-enabled. When this field is set to '1', the controller sets CSTS.RDY to '1' when it is ready to process commands. CSTS.RDY may be set to '1' before namespace(s) are ready to be accessed. Setting this field from a '0' to a '1' when CSTS.RDY is a '1,' or setting this field from a '1' to a '0' when CSTS.RDY is a '0,' has undefined results. The Admin Queue registers (AQA, ASQ, and ACQ) shall only be modified when EN is cleared to '0'.

1) Initialization

① 호스트는 컨트롤러가 CSTS.RDY가 '0'이 되는 것을 기다린다.

② Admin queue는 AQA, ASQ, ACQ를 적절한 값으로 설정한다.

③ 컨트롤러 셋팅이 이루어져야 한다. (CC.AMS, CC.MPS, CC.CSS에 대하여)

⑤. CC.EN을 '1'로 설정함 으로서 컨트롤러가 활성 된다.

⑥. 컨트롤러는 CSTS.RDY(Ready)가 '1'로 설정되면서 커맨드 처리 준비가 완료된다.

⑦. 호스트는 Set features 커맨드에 지원되는 I/O SQ, CQ 의 개수를 결정한다. 결정되면 MSI and/or MSI-X 레지스터가 설정되어야 한다.

⑧. 호스트는 컨트롤러가 지원하는 수만큼 I/O CQ/SQ 를 할당한다.

→ The controller may be used for I/O commands

1. Controller Initialization and Shutdown Processing

AQA	Admin Queue Attributes
ASQ	Admin Submission Queue Base Address
ACQ	Admin Completion Queue Base Address

AQA

Bit	Type	Reset	Description
31:28	RO	0h	Reserved
27:16	RW	0h	**Admin Completion Queue Size (ACQS):** Defines the size of the Admin Completion Queue in entries. Refer to section 4.1.3. Enabling a controller while this field is cleared to 00h produces undefined results. The minimum size of the Admin Completion Queue is two entries. The maximum size of the Admin Completion Queue is 4096 entries. This is a 0's based value.
15:12	RO	0h	Reserved
11:00	RW	0h	**Admin Submission Queue Size (ASQS):** Defines the size of the Admin Submission Queue in entries. Refer to section 4.1.3. Enabling a controller while this field is cleared to 00h produces undefined results. The minimum size of the Admin Submission Queue is two entries. The maximum size of the Admin Submission Queue is 4096 entries. This is a 0's based value.

ASQ

Bit	Type	Reset	Description
63:12	RW	Impl Spec	**Admin Submission Queue Base (ASQB):** Indicates the 64-bit physical address for the Admin Submission Queue. This address shall be memory page aligned (based on the value in CC.MPS). All Admin commands, including creation of I/O Submission Queues and I/O Completions Queues shall be submitted to this queue. For the definition of Submission Queues, refer to section 4.1.
11:00	RO	0h	Reserved

ACQ

Bit	Type	Reset	Description
63:12	RW	Impl Spec	**Admin Completion Queue Base (ACQB):** Indicates the 64-bit physical address for the Admin Completion Queue. This address shall be memory page aligned (based on the value in CC.MPS). All completion queue entries for the commands submitted to the Admin Submission Queue shall be posted to this Completion Queue. This queue is always associated with interrupt vector 0. For the definition of Completion Queues, refer to section 4.1.
11:00	RO	0h	Reserved

1) Initialization

① 호스트는 컨트롤러가 CSTS.RDY가 '0'이 되는 것을 기다린다.

② Admin queue는 AQA, ASQ, ACQ를 적절한 값으로 설정한다.

③ 컨트롤러 셋팅이 이루어져야 한다. (CC.AMS, CC.MPS, CC.CSS에 대하여)

⑤. CC.EN을 '1'로 설정함 으로서 컨트롤러가 활성 된다.

⑥. 컨트롤러는 CSTS.RDY(Ready)가 '1'로 설정되면서 커맨드 처리 준비가 완료된다.

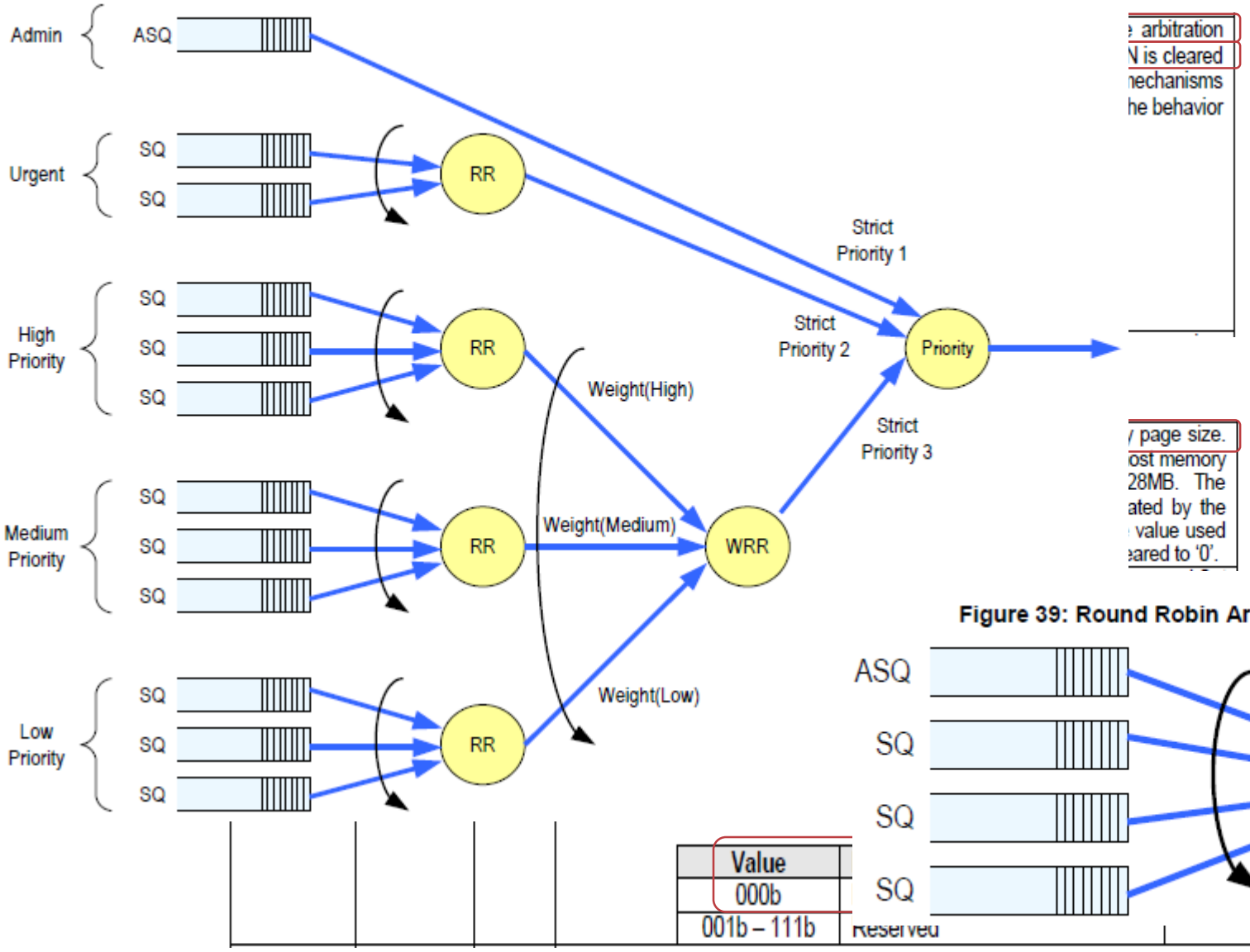
⑦. 호스트는 Set features 커맨드에 지원되는 I/O SQ, CQ 의 개수를 결정한다. 결정되면 MSI and/or MSI-X 레지스터가 설정되어야 한다.

⑧. 호스트는 컨트롤러가 지원하는 수만큼 I/O CQ/SQ 를 할당한다.

→ The controller may be used for I/O commands

1. Controller Initialization and Shutdown Processing

Figure 40: Weighted Round Robin with Urgent Priority Class Arbitration



arbitration
N is cleared
mechanisms
he behavior

page size.
ost memory
28MB. The
ated by the
value used
eared to '0'.

Figure 39: Round Robin Arbitration

Value	
000b	
001b – 111b	Reserved

1) Initialization

① 호스트는 컨트롤러가 CSTS.RDY가 '0'이 되는 것을 기다린다.

② Admin queue는 AQA, ASQ, ACQ를 적절한 값으로 설정한다.

③ 컨트롤러 셋팅이 이루어져야 한다. (CC.AMS, CC.MPS, CC.CSS에 대하여)

⑤. CC.EN을 '1'로 설정함 으로서 컨트롤러가 활성 된다.

⑥. 컨트롤러는 CSTS.RDY(Ready)가 '1'로 설정되면서 커맨드 처리 준비가 완료된다.

⑦. 호스트는 Set features 커맨드에 지원되는 I/O SQ, CQ 의 개수를 결정한다. 결정되면 MSI and/or MSI-X 레지스터가 설정되어야 한다.

⑧. 호스트는 컨트롤러가 지원하는 수만큼 I/O CQ/SQ 를 할당한다.

→ The controller may be used for I/O commands

1. Controller Initialization and Shutdown Processing

CSTS.RDY

00	RO	0	**Ready (RDY):** This field is set to '1' when the controller is ready to accept Submission Queue Tail doorbell writes after CC.EN is set to '1'. This field shall be cleared to '0' when CC.EN is cleared to '0'. Commands shall not be submitted to the controller until this field is set to '1' after the CC.EN bit is set to '1'. Failure to follow this requirement produces undefined results. Host software shall wait a minimum of CAP.TO seconds for this field to be set to '1' after setting CC.EN to '1' from a previous value of '0'.

CC.EN

00	RW	0	**Enable (EN):** When set to '1', then the controller shall process commands based on Submission Queue Tail doorbell writes. When cleared to '0', then the controller shall not process commands nor post completion queue entries to Completion Queues. When this field transitions from '1' to '0', the controller is reset (referred to as a Controller Reset). The reset deletes all I/O Submission Queues and I/O Completion Queues, resets the Admin Submission Queue and Completion Queue, and brings the hardware to an idle state. The reset does not affect PCI Express registers (including MMIO MSI-X registers), nor the Admin Queue registers (AQA, ASQ, or ACQ). All other controller registers defined in this section and internal controller state (e.g., Feature values defined in section 5.21.1 that are not persistent across power states) are reset to their default values. The controller shall ensure that there is no data loss for commands that have had corresponding completion queue entries posted to an I/O Completion Queue prior to the reset operation. Refer to section 7.3 for reset details. When this field is cleared to '0', the CSTS.RDY bit is cleared to '0' by the controller once the controller is ready to be re-enabled. When this field is set to '1', the controller sets CSTS.RDY to '1' when it is ready to process commands. CSTS.RDY may be set to '1' before namespace(s) are ready to be accessed. Setting this field from a '0' to a '1' when CSTS.RDY is a '1,' or setting this field from a '1' to a '0' when CSTS.RDY is a '0,' has undefined results. The Admin Queue registers (AQA, ASQ, and ACQ) shall only be modified when EN is cleared to '0'.

1) Initialization

① 호스트는 컨트롤러가 CSTS.RDY가 '0'이 되는 것을 기다린다.

② Admin queue는 AQA, ASQ, ACQ를 적절한 값으로 설정한다.

③ 컨트롤러 셋팅이 이루어져야 한다. (CC.AMS, CC.MPS, CC.CSS에 대하여)

⑤. CC.EN을 '1'로 설정함 으로서 컨트롤러가 활성 된다.

⑥. 컨트롤러는 CSTS.RDY(Ready)가 '1'로 설정되면서 커맨드 처리 준비가 완료된다.

⑦. 호스트는 Set features 커맨드에 지원되는 I/O SQ, CQ 의 개수를 결정한다. 결정되면 MSI and/or MSI-X 레지스터가 설정되어야 한다.

⑧. 호스트는 컨트롤러가 지원하는 수만큼 I/O CQ/SQ 를 할당한다.

→ The controller may be used for I/O commands

Figure 143: Number of Queues – Command Dword 11

Bit	Description
31:16	**Number of I/O Completion Queues Requested (NCQR):** Indicates the number of I/O Completion Queues requested by software. This number does not include the Admin Completion Queue. A minimum of one shall be requested, reflecting that the minimum support is for one I/O Completion Queue. This is a 0's based value. The maximum value that may be specified is 65,534 (indicating 65,535 I/O Completion Queues). If the value specified is 65,535, the controller should return an error of Invalid Field in Command.
15:00	**Number of I/O Submission Queues Requested (NSQR):** Indicates the number of I/O Submission Queues requested by software. This number does not include the Admin Submission Queue. A minimum of one shall be requested, reflecting that the minimum support is for one I/O Submission Queue. This is a 0's based value. The maximum value that may be specified is 65,534 (indicating 65,535 I/O Submission Queues). If the value specified is 65,535, the controller should return an error of Invalid Field in Command.

Set Features Command Dword 11

2) Shutdown

03:02	RO	0	**Shutdown Status (SHST):** This field indicates the status of shutdown processing that is initiated by the host setting the CC.SHN field. The shutdown status values are defined as: **Value** / **Definition** 00b – Normal operation (no shutdown has been requested) 01b – Shutdown processing occurring 10b – Shutdown processing complete 11b – Reserved To start executing commands on the controller after a shutdown operation (CSTS.SHST set to 10b), a Controller Reset (CC.EN cleared to '0') is required. If host software submits commands to the controller without issuing a reset, the behavior is undefined.

Shut down operation 후에 컨트롤러에
command 실행을 시작하기 위해 reset이 필요하다.

2. Asynchronous Event

2. Asynchronous Event

15	**Retain Asynchronous Event (RAE):** This field specifies when to retain or clear an Asynchronous Event. If this bit is cleared to '0', the corresponding Asynchronous Event is cleared after the command completes successfully. If this bit is set to '1,' the corresponding Asynchronous Event is retained (i.e., not cleared) after the command completes successfully. Host software should clear this field to '0' for log pages that are not used with Asynchronous Events. Refer to section 5.2.

① 호스트는 로그 페이지를 사용하는 모든 이벤트를 비활성화하는 Set Features command를 보낸다.

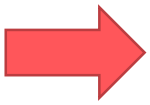
② 호스트는 Log Page를 요청하는 Get Log Page command를 보낸다.

③ - a 반환된 Log Page 상태가 not persistent이면, Log Page를 사용하는 모든 Asynchronous Event를 다시 활성화 해야 한다.

③ - b 반환된 Log Page 상태가 persistent이면, Log Page를 보고한 이벤트를 제외하고 Log Page를 사용하는 모든 Asynchronous Event를 다시 활성화 해야 한다.

④ 호스트는 Asynchronous Event request command 를 컨트롤러에 보낸다

⑤ 보고된 상태가 persistent 이면, 호스트는 이벤트 보고가 재활성화 되어야 할지 결정하기 위해, 감시를 계속한다.



Asynchronous Event Request 가 완료되었을 때 Event Type, Event Information, Log Page 정보를 제공한다.

3. Feature Values

3. Feature Values

- Optional NVM Command Support command의 비트 4 가 '1'로 설정되면,
  각 feature 별로 **<default, saveable, current>** 3 가지의 셋팅이 있다.
  Optional NVM Command Support command의 비트 4 가 '0'로 설정되면,
  각 feature 별로 **<current 와 default>** 값만 지원한다.

Bit 4 if set to '1' then the controller supports the Save field set to a non-zero value in the Set Features command and the Select field set to a non-zero value in the Get Features command. If cleared to '0' then the controller does not support the Save field set to a non-zero value in the Set Features command and the Select field set to a non-zero value in the Get Features command.

3. Feature Values

- Current value : 명시된 정보에 기초한다.
- Default value : 생산자에 의해 설정되고 변경될 수 없다.
- Saveable value :

A Select field set to 011b (i.e., supported capabilities) returns the capabilities supported for this Feature Identifier. The capabilities supported are returned in Dword 0 of the completion entry of the Get Features command.

- If Dword 0 bit 0 of the completion entry of the Get Features command is set to '1', then the Feature Identifier is saveable. If Dword 0 bit 0 of the completion entry of the Get Features command is cleared to '0', then the Feature Identifier is not saveable.

3. Feature Values

- Set Features : saveable, current 값을 수정하기 위해 사용된다.
- Get Features : default, saveable, current 값을 읽기 위해 사용된다.

07:04	**Namespace Identifier (NSID):** This field specifies the namespace that this command applies to. If the namespace is not used for the command, then this field shall be cleared to 0h. Setting this value to FFFFFFFFh causes the command to be applied to all namespaces attached to this controller, unless otherwise specified. Specifying an inactive namespace ID in a command that uses the namespace ID shall cause the controller to abort the command with status Invalid Field in Command, unless otherwise specified. Specifying an invalid namespace ID in a command that uses the namespace ID shall cause the controller to abort the command with status Invalid Namespace or Format, unless otherwise specified.

- Namespace에 명시되지 않는 feature는 CDW1.NSID에 0값을 가진다.
- Namespace에 명시된 값을 변경하기 위해서는 CDW1.NSID를 해당 namespace의 ID로 설정한다.

4. NVMe Qualified Names

4. NVMe Qualified Names

- NVMe Qualified Names은 식별/인증을 위해 호스트 또는 NVM Subsystem을 설명하는데 사용된다.

- The encoding is UTF-8 (refer to RFC 3629).
- The following characters are used in formatting:
  - dash ('-'=U+002d)
  - dot ('.'=U+002e)
  - colon (':'=U+003a)
- The maximum name is 223 bytes in length.
- The string is null terminated.

4. NVMe Qualified Names

NQN format

① 사람이 읽을 수 있는 문자열로 host / NVM subsystem 을 설명하는데 사용.

- The string “nqn.”
- A date code, in “yyyy-mm.” format. This date shall be during a time when the naming authority owned the domain name used in this format. The date code uses the Gregorian calendar. All digits and the dash shall be included.
- The reverse domain name of the naming authority that is creating the NQN.
- A colon (:) prefixed string that the owner of the domain name assigns that does not exceed the maximum length. The naming authority is responsible to ensure that the NQN is worldwide unique.

- nqn.2014-08.com.example:nvme:nvm-subsystem-sn-d78432
- nqn.2014-08.com.example:nvme.host.sys.xyz

NQN format

② Naming Authority(명명기관)이 존재하지 않을 때.

- The string “nqn.”
- The string “2014-08.org.nvmexpress:uuid:”.
- A 128-bit UUID based on the definition in RFC 4122 represented as a string formatted as “*11111111-2222-3333-4444-555555555555*”.

- nqn.2014-08.org.nvmexpress:uuid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6

5. Identifier Format and Layout

5. Identifier Format and Layout

Data : 0x01234567

Example (Value shown as ASCII characters):

- SN = "SN1"
- MN = "M2"

Big Endian

Byte	04	05	06	23 - 07	24	25	63 - 26
Value	53h ('S')	4Eh ('N')	31h ('1')	20h (' ')	4Dh ('M')	32h ('2')	20h (' ')

Example:

- VID = ABCDh
- SSVID = 1234h

Little Endian

Byte	00	01	02	03
Value	CDh	ABh	34h	12h

IEEE OUI Identifier (IEEE)

75:73	M	**IEEE OUI Identifier (IEEE):** Contains the Organization Unique Identifier (OUI) for the controller vendor. The OUI shall be a valid IEEE/RAC assigned identifier that may be registered at http://standards.ieee.org/develop/regauth/oui/public.html.

5. Identifier Format and Layout

IEEE Extended Unique Identifier (EUI64)

Byte	0	1	2	3	4	5	6	7
EUI64	OUI			Extension Identifier				

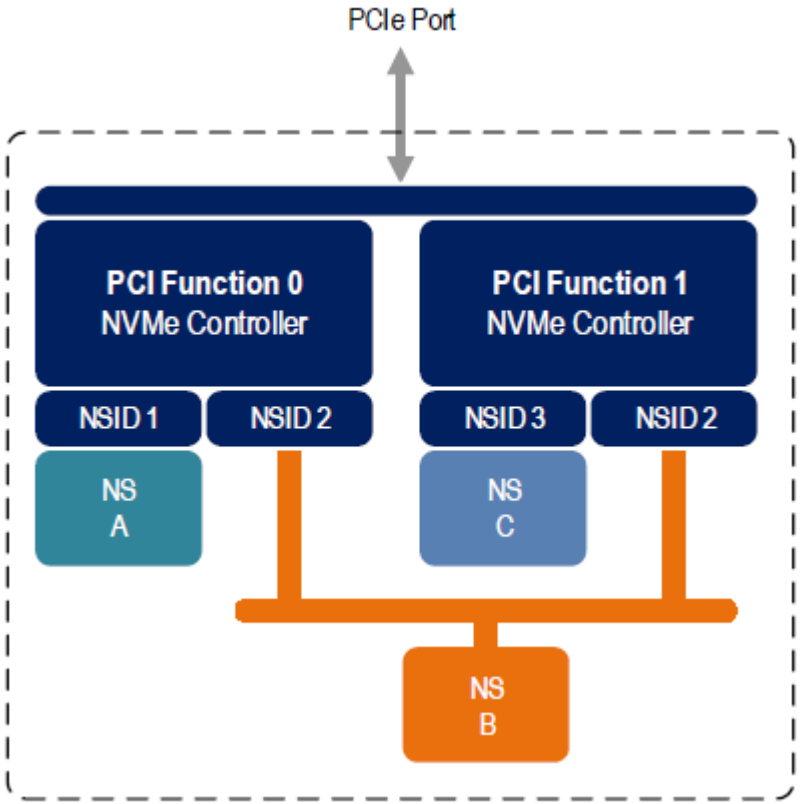
Namespace Globally Unique Identifier (NGUID)

Byte	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
NGUID	Vendor Specific Extension Identifier								OUI			Extension Identifier				

6. Unique Identifier

6. Unique Identifier

01:00	M	r identifier that is assigned by ıe ID register in section 2.1.1.
23:04	M	ır the NVM subsystem that is ection 7.10 for unique identifier quirements.
63:24	M	or the NVM subsystem that is ection 7.10 for unique identifier quirements.



NVM Su globally unique 값.

- The Contro uely identify a controller.
- When creating a namespace, the controller specifies a globally unique value in the EUI64 or NGUID field

7. Keep Alive

7. Keep Alive

Figure 161: Keep Alive Timer – Command Dword 11

Bit	Description
31:00	**Keep Alive Timeout (KATO):** This field specifies the timeout value for the Keep Alive feature in milliseconds. The controller rounds up the value specified to the granularity indicated in the KAS field in the Identify Controller data structure. If cleared to 0h then the Keep Alive Timer is disabled. The default value for this field is 0h for PCIe and fabrics that do not require use of the Keep Alive feature. For fabrics that require use of the Keep Alive feature, the default value for this field is 1D4C0h (i.e., 120,000 milliseconds or 2 minutes) rounded up to that granularity.

- CC.EN is set to '1' and CSTS.RDY is set to '1'; and
- CC.SHN is cleared to '0' and CSTS.SHST is cleared to '0'.

Thank you