A Survey of Form Factors for NVM Express®

Sponsored by NVM Express organization, the owner of NVMe® specifications

A short panel introduction

John “KIOXIA” Geldman (yours truly): Current Board Member of NVM Express and SNIA, Past Board Member of SD Card Association and CompactFlash Association, PCI-SIG® contributor, SNIA SFF TA contributor

Paul “HPE” Kaler: Responsible for researching and evaluating future storage and interconnect technologies and defining the server storage strategy for HPE ProLiant servers, active in SNIA SFF TA, PCI-SIG, OCP NIC and Storage

Bill “Nantero” Gervasi: So many JEDEC roles that we don’t have space

Michael “WDC” Lavrentiev: SD Card contributor, …

Dave “WDC” Landsman: Current Board Member of CompactFlash, NVM Express, SD Card Association, and SNIA, …

A Survey of Form Factors for NVM Express® Architecture

NVM Express® architecture is now ubiquitous across pretty much ‘all’ storage

Hyperscalars, Enterprise Servers, Desktops, Laptops, Digital Cameras, Cinema Cameras, Drones, Industrial, IOT

Servers & EDSFF get a lot of attention

This panel will share about EDSFF and also, the rest

- (at least what I knew to invite…)

PCI-SIG® Defined Form Factors

John Geldman, KIOXIA

CEM Add In Card (AIC)

1-16 lanes

Published for PCIe® 5.0 specification, under development for PCIe 6.0 specification

Every other function option beyond NVM Express® technology…

Size Options:

- Standard Height (111.28 mm) or Low Profile (68.09 mm)
- Length (varies: up to 312 mm, 241.30 mm recommended)
- Single, Dual, or Triple Slot

Strengths:

- The leading edge and workhorse of PCIe FFs
- Backwards compatible to PCIe Rev 1.0
- Supports up to 600 W (Air Cooling stops at 300 W!)
- Supports up 1.5 KG!

CEM NVMe® Technology Examples

SFF 8639 Connector Module (U.2)

- 4 lanes
- Published for PCIe® 4.0 specification, under development for PCIe 5.0 specification
- Size Options: 69.85 mm x 100.45 mm x (5 to 19 mm)
- Mostly used for Enterprise and Data Center Storage
- Strengths:
    - Chassis compatible with SAS and SATA 2.5 inch HDDs
    - The go-to enterprise storage since 2012 (the Enterprise SSD Form Factor Forum)
- Weaknesses?
    - Will PCIe 5.0 architecture be the most we can pull out of this connector/pinout?
    - Can SFF-TA-1001 be supported at Gen 5 or higher speeds?

U.2 SSD Examples

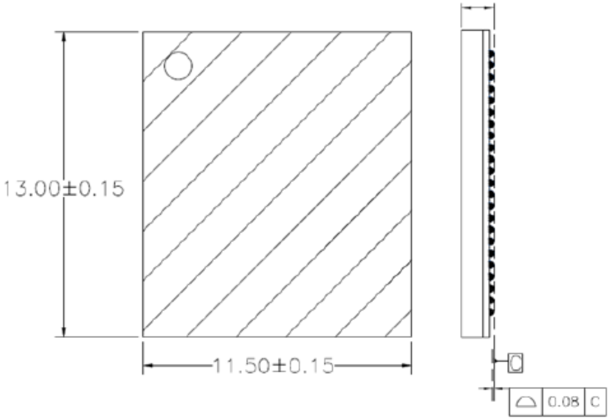
M.2 (Mini Express)

- 2-4 lanes, published for PCIe® 4.0 specification, under development for PCIe 5.0 specification
- Size Options: Too Many
  - M.2 Cards (typically 22 mm wide for SSDs, lengths include 30, 42, 80, 110 mm)
  - Two families of BGA pinouts starting at 11.5x13 mm and 16x20 mm
  - LGA options
- Developed for mobile, but consumed in:
  - Hyperscalars, Enterprise Servers, Desktops, Laptops, Digital Cameras, Cinema Cameras, Drones, Industrial, IOT
- Strengths:
  - Low power at acceptable performance (e.g., 11 W, 8 W, 4 W)
  - Wide range of functions available
  - Wide range of sizes
- Weaknesses?
  - Will PCIe 5.0 architecture be the most we can pull out of this connector/pinout?

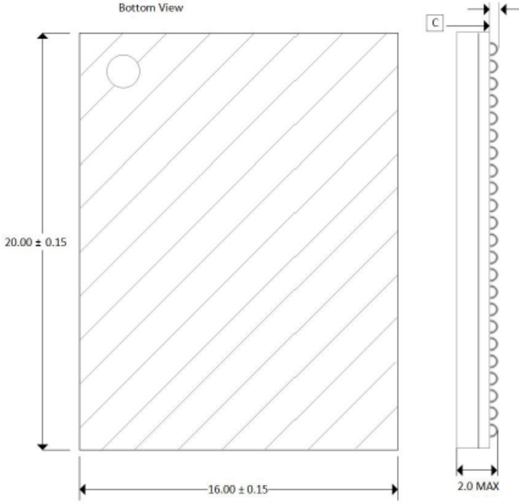
M.2 (Mini Express) Examples

Adapter Cards

BGAs

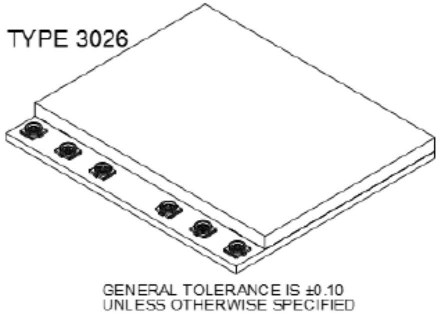


Type 1620
Bottom View



LGAs

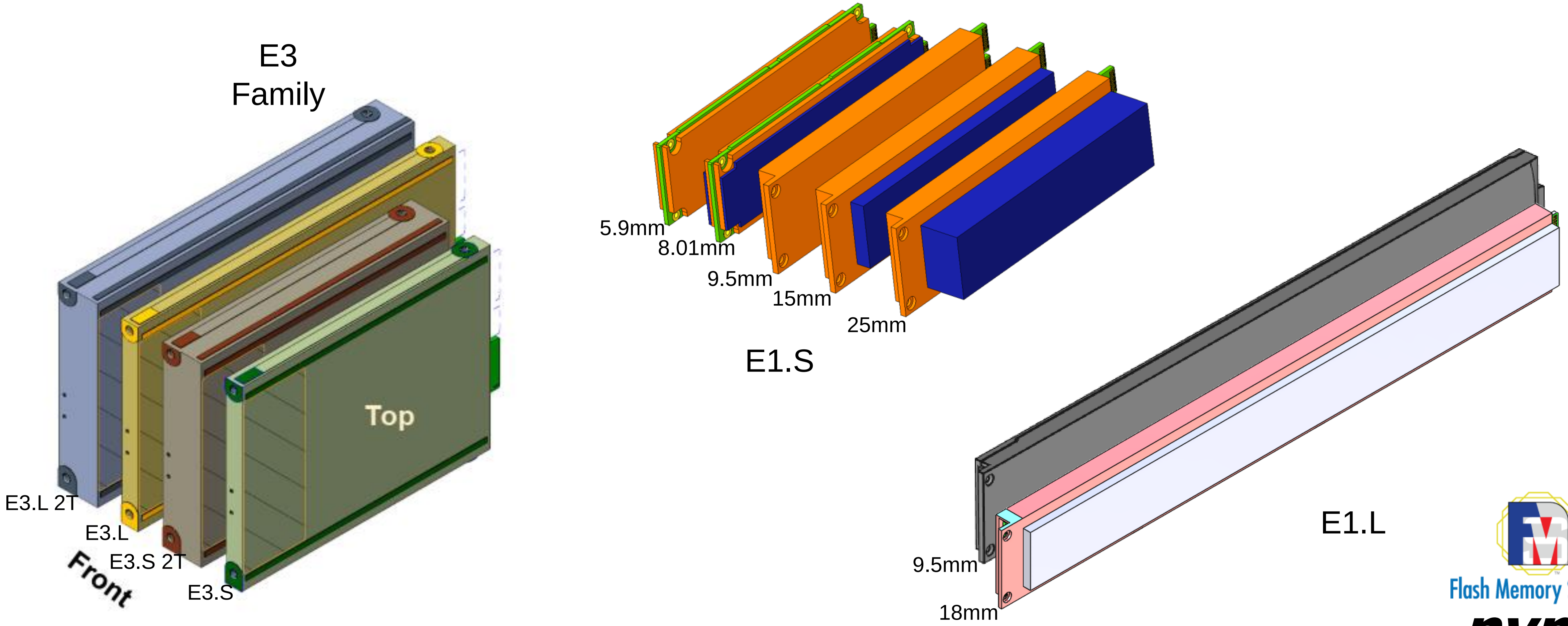
TYPE 2226-S3



SNIA SFF TA Defined Form Factors
(and SNIA Object TWG)

Paul Kaler, HPE

EDSFF (Enterprise and Datacenter Standard Form Factor)



E3 Family

- Supports up to x4/x8/x16 lane width spec'd up to PCIe® 5.0 architecture
  - Future plans for PCIe 6.0 specification and beyond
- E3.S and E3.L support 7.5/16.8 "2T" (mm) thicknesses
- 7.5mm targets mainstream capacities and performance
- 2T (16.8mm) is for higher capacity (e.g. 60TB) and/or full Gen5 performance (>25W)
- Current defined protocols– NVM Express®, CXL, Native NVMe-oF™ technologies
  - Future OCP NIC 3.0 enablement
- Predominately used in Enterprise servers and storage
  - Future—use E3.S anywhere a 2.5" drive is used today
- Strengths: Works well for both 1U and 2U servers, 2 for 1 interchange (2x 1T↔2T), scalable thermals—full PCIe 5.0 architecture and beyond performance, x16 lane width enables performance headroom

E1.S

- Supports up to x4/x8 lane width spec'd up to PCIe® 5.0 specification
  - Future plans for PCIe 6.0 architecture and beyond
- Five thickness options 5.9/8.01/9.5/15/25 (mm)
- Current defined protocols– NVM Express®, CXL, Native NVMe-oF™ technologies
- Predominately used by Hyperscalers
- Strengths: Optimized for 1U servers, wide range of thickness options provide ability to tailor for varying thermal, density, and performance requirements

E1.L

- Supports up to x4/x8 lane width spec'd up to PCIe® 5.0 architecture
- Two thickness options 9.5/18 (mm)
- Current defined protocols– NVM Express®, CXL, Native NVMe-oF™ technologies
- Predominately used by Hyperscalers for “cooler” storage tiers (e.g. QLC)
- Strengths: Very optimized to achieve high capacity (~1PB) in 1U servers and JBOFs

SFF-TA-1001 (U.3)

- Supports up to x4 lane width spec'd up to PCIe® 4.0 architecture
- Based on PCI Express® SFF-8639 Module Spec Rev 4.0 with a new pinout that enable tri-mode host drive bays
- PCIe lanes are shared with SAS/SATA enabling tri-mode RAID controllers to support NVMe®/SAS/SATA technology with one to four ports
- Predominately used in Enterprise servers and storage
- Strengths: Sharing high speed lanes enables lower cost systems for mixing SAS/SATA/NVMe architectures and makes it easier for customers to transition from SATA/SAS to NVMe technology

Native NVMe-oF™ Drive

- Supports up to x2 lane width spec'd up to Ethernet 25G
    - Future plans under development for Ethernet 50G
- Spec supports several form factor options (2.5”, 3.5”, and EDSFF)
- Predominately targeting use cases where performance scaling per drive is important
- Strengths: Enables end-to-end Ethernet connectivity eliminating potential performance bottlenecks with Ethernet to PCIe® technology conversions.

JEDEC Defined Form Factors

Bill Gervasi, Nantero

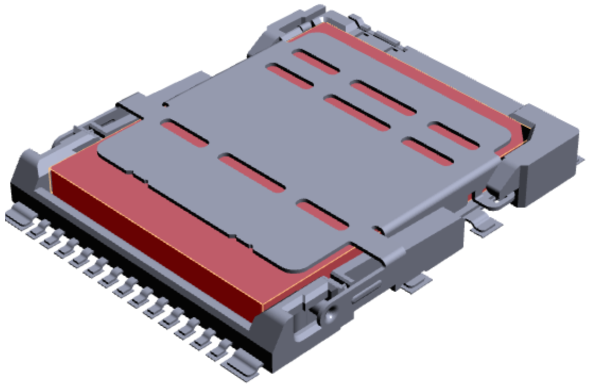
XFM (JESD233)

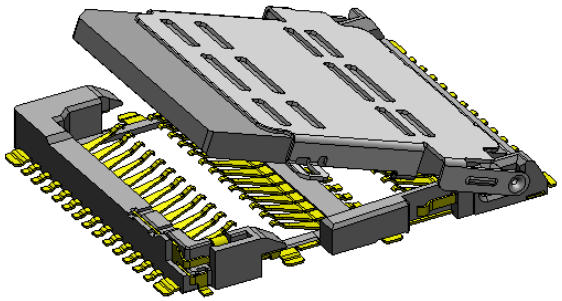
- 1-2 lanes, published for PCIe® 4.0 architecture
- Card Size
  - 14 mm wide, 18 mm long, 1.4 mm thin
- Developed as replaceable storage (not removable)
- Strengths:
  - Embedded connectorized storage
  - A balance between a small form factor and support of SSD-class components (e.g., current and future 3D flash)
  - 2.5 V and 1.2 V power inputs
  - Four current classes for targeted performance and power

XFM

Embedded connectorized storage designed for:

- Easy replacement
- Minimal real-estate (3D)
- Heat transfer mechanisms optimizable for system target





Automotive BGA

JESD312

PCIe® Shared Resource: Automotive SSD

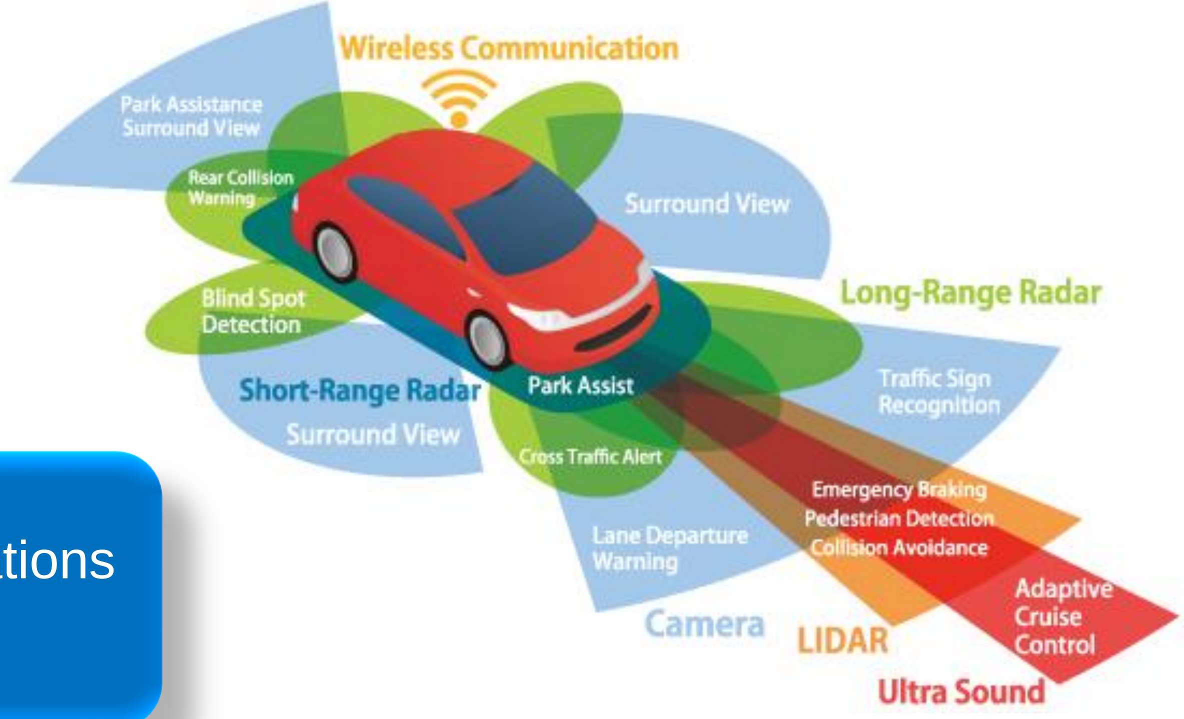
PCIe Switch

Display SoC

Communications SoC

Entertainment SoC

Vehicle Control SoC



Variety of Control Systems

JEDEC STANDARD

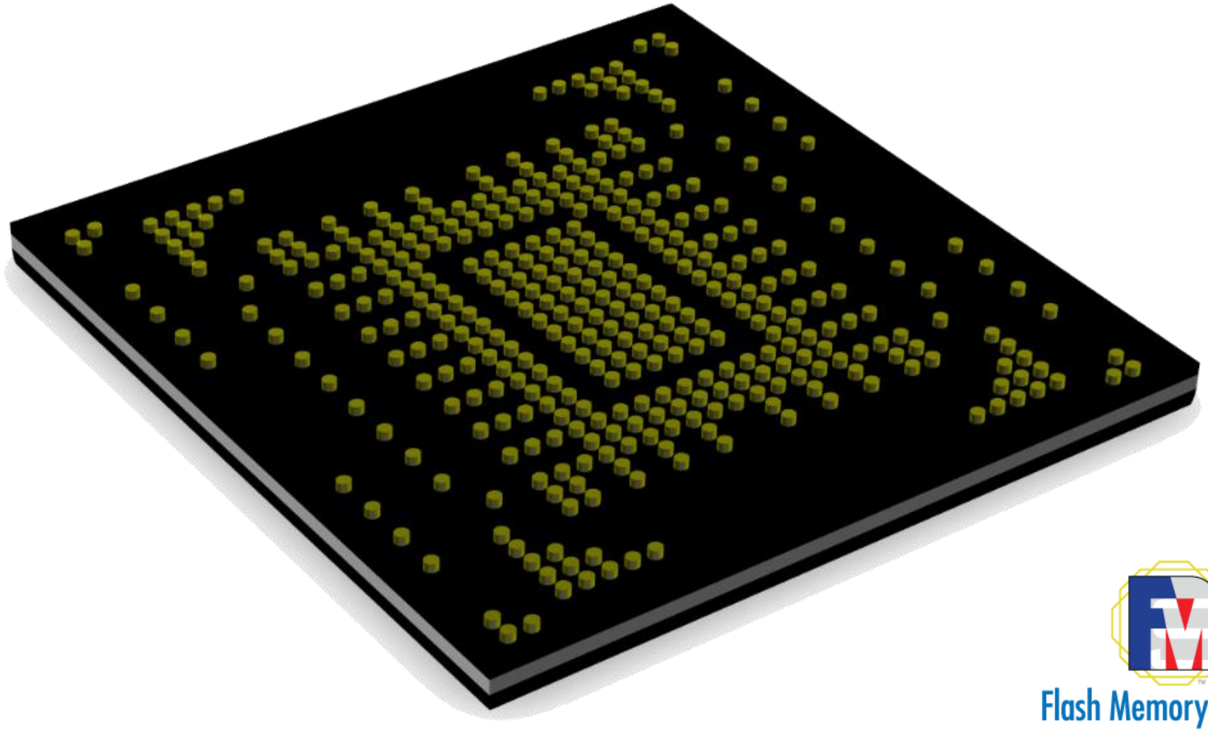
Automotive Solid State Drive (SSD) Device Specification

Rev 1.0

JESD312

Target: Year End 2022

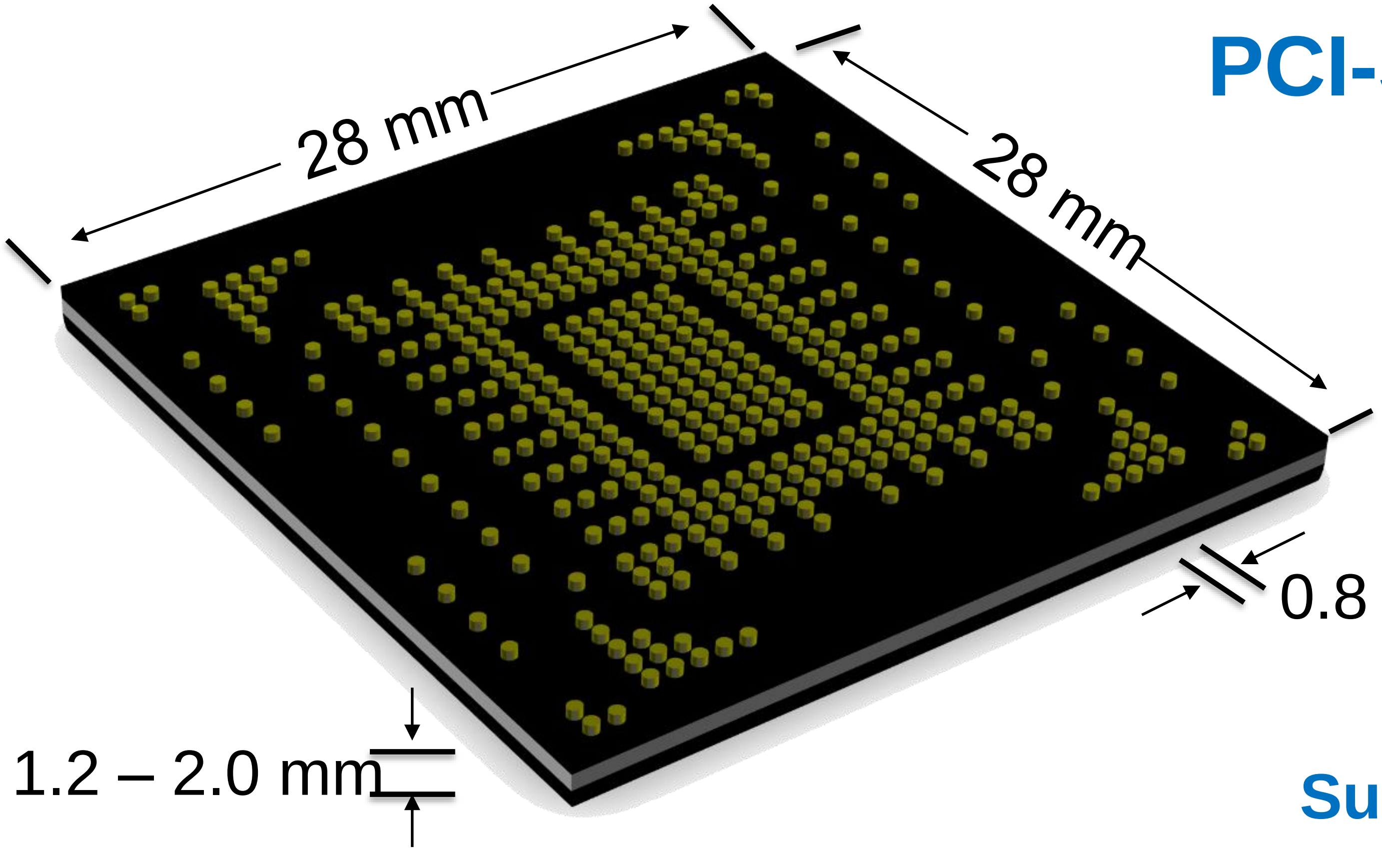
A New Standard SSD for automotive applications



Flash Memory Summit

nvm EXPRESS®

Package and Pinout

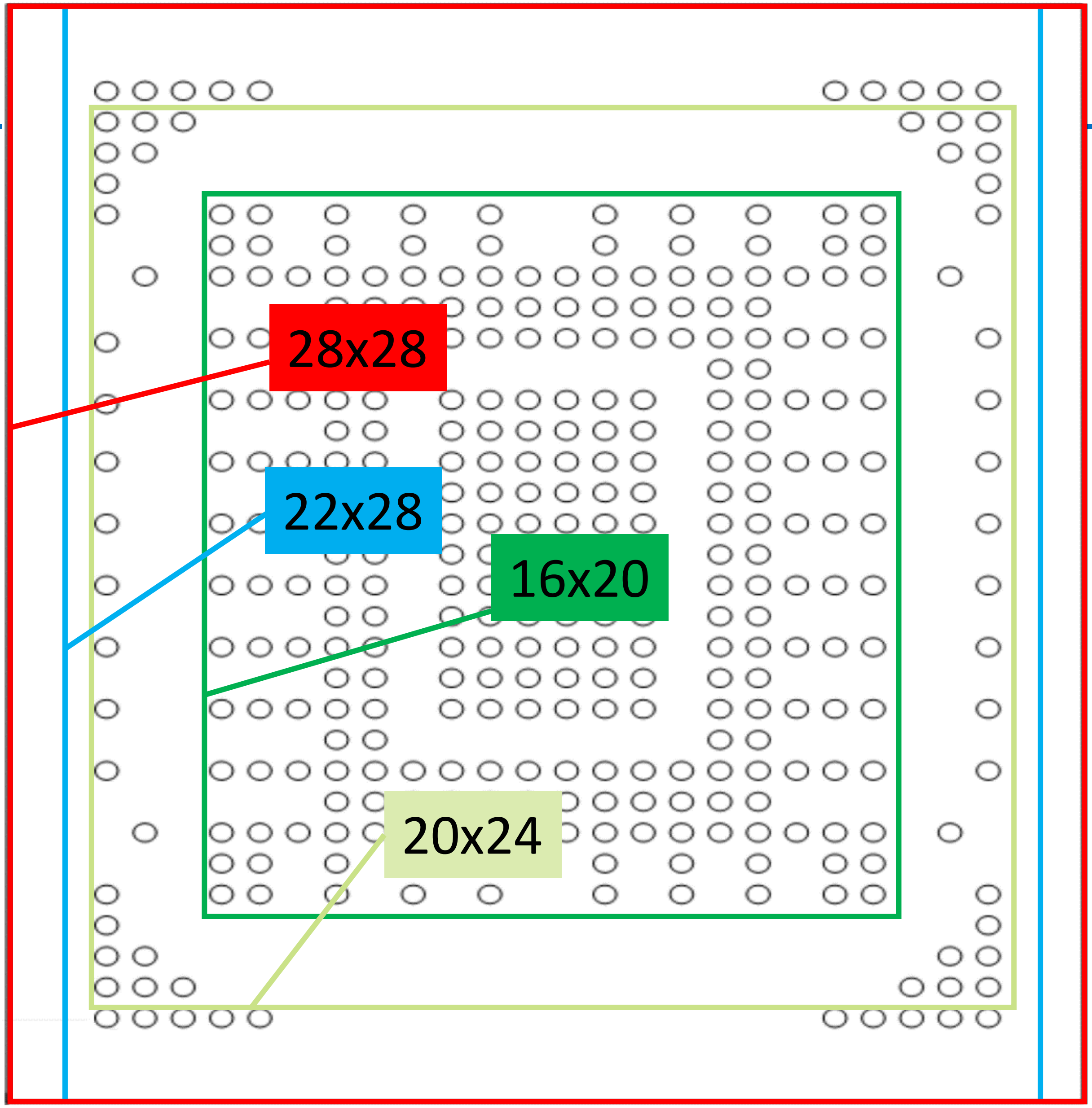


PCI-SIG® 2828 BGA

Sufficient power delivery

Case exposed for cooling

Ref: PCI Express® M.2 Specification



To assist suppliers in offering well priced options without sacrificing compatibility:

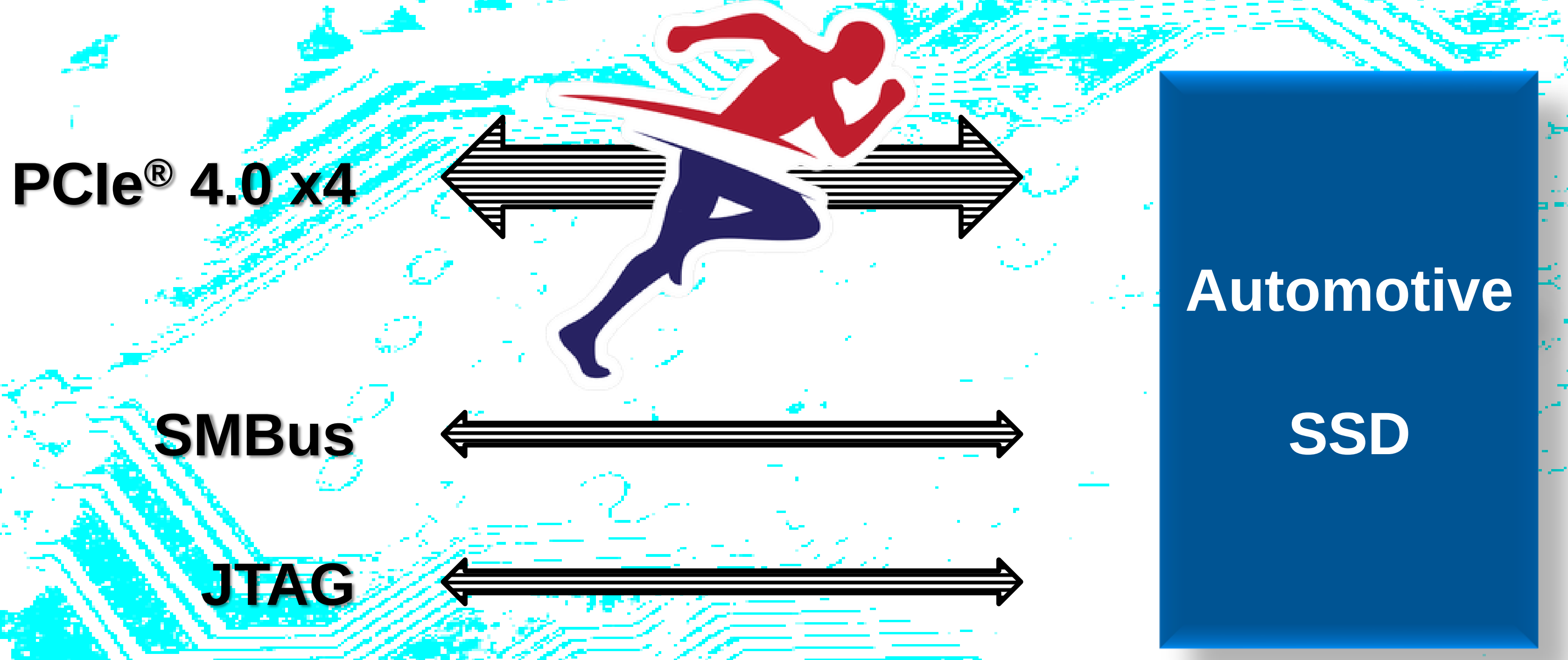
1) End users design to 2828, allow any part to drop in

2) Allows suppliers to use any of the footprint compatible options

   16 x 20 mm
   20 x 24 mm
   22 x 28 mm
   28 x 28 mm

Electrical Interface

32 GB/s peak throughput



Ref: PCI Express Base Specification 4.0

Command Protocol

Required: PCIe® 4.0	Required: NVMe® 1.4c+	Encouraged: Single Root I/O Virtualization	Required: SMBus	Required: JTAG
Physical interface	Logical interface	Optional virtualization	System and power management	Testability

Physical interface
Logical interface
Optional virtualization
System and power management
Testability

Ref: JTAG (IEEE 1149.1) Specification

Ref: System Management Bus (SMBus) Specification

Ref: NVM Express® (NVMe) protocol

Ref: PCI Express® Base Specification 4.0

Security

Requester

Responder

SPDM
Truth Boundary

Secure
Response

Firmware Resilience

384-bit
Minimum
Security

Signature: TPM_ALG_ECDSA_ECC_NIST_P384

Hash: TPM_ALG_SHA_384

Ref: Component Measurement and Authentication (CMA)

Ref: NIST Platform Firmware Resiliency Guidelines 800-193

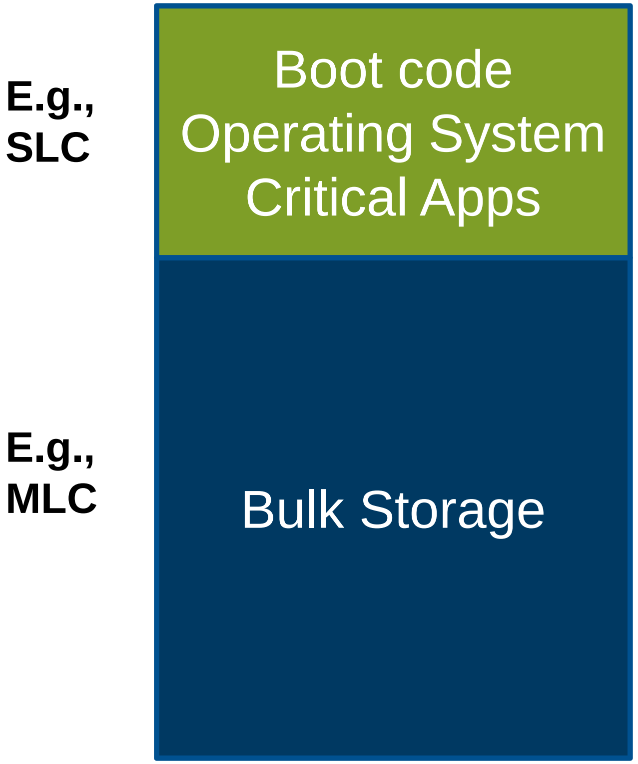
Ref: FIPS PUB 180-4 Secure Hash Standard (SHS)

Ref: Digital Signature Standard (DSS)

Ref: Security Protocol and Data Model (SPDM)

Storage Regions

Optional Feature:
High reliability system storage region



Drive Capacity Class	Minimum System Region Capacity	Bulk Region Capacity
128 GB	0	128 GB
256 GB	0	256 GB
512 GB	32 GB	512 GB
1 TB	32 GB	1 TB
2 TB	64 GB	2 TB
4 TB	64 GB	4 TB

The system and bulk regions may have distinct parameters including temperature range, retention, etc.

E.g., Terabyte Write (TBW) for -40°C to +95°C supported for system and storage regions, -40°C to +105°C for system region only

Endurance

Defined by Market Segments

Personal Auto
344 days/year
3 hours/day
15 year life

Professional Auto
365 days/year
12 hours/day
8 year life

Professional Auto market, bulk storage region capacity 1 TB class from -40 to +95 °C = 200 TBW
DWPD = 200 TBW [1 TB * 8 years * 365 days/year * (12 ÷ 24 hours)] = minimum 0.24 DWPD

Personal Auto market, system storage region capacity 64 GB from +95 to +105 °C = 12.8 TBW
DWPD = 12.8 TBW [0.064 TB * 15 years * 344 days/year * (3 ÷ 24 hours)] = minimum 0.31 DWPD

Data usage model = Enterprise model

TBW = Terabytes written
DWPD = Drive writes per day

Ref: JESD218B-01 Solid State Drive (SSD) Requirements and Endurance Test Method

Ref: JESD219 Solid-State Drive (SSD) Endurance Workloads

SD Card Defined Form Factors

Michael Lavrentiev, WDC

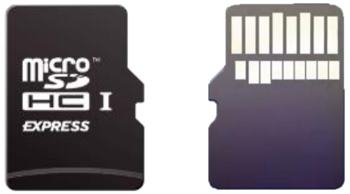
SD Express Card - Basics

The fastest SD™ and microSD™ memory cards with backward compatibility

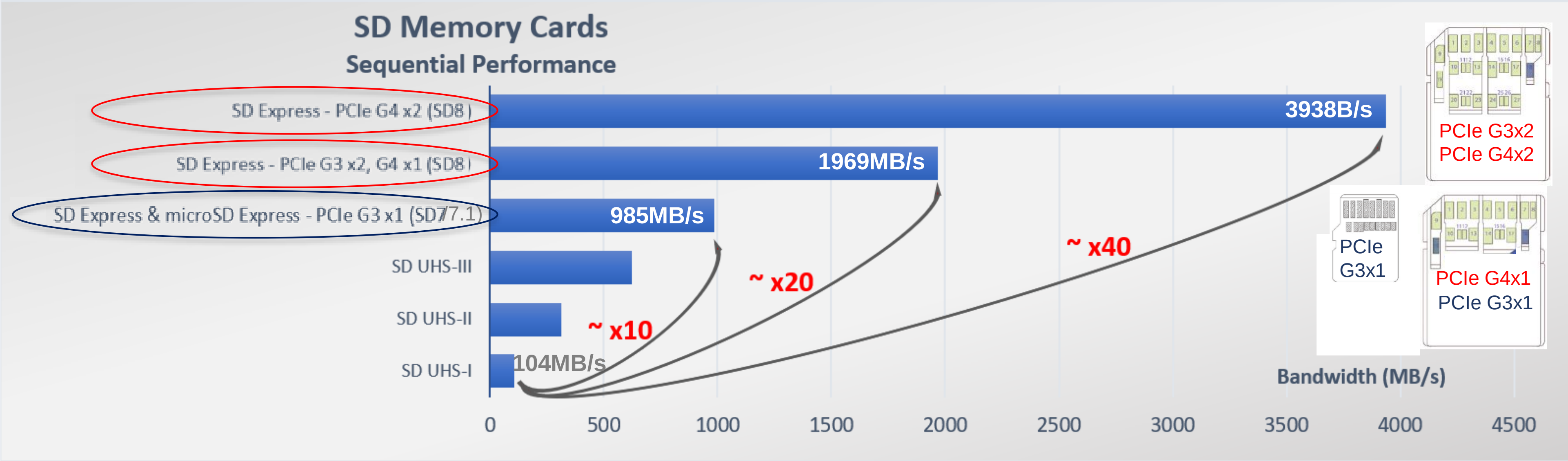
Supporting the following interfaces:

- NVMe® + PCIe® interface – up to PCIe 4.0 x2
- SD interface (UHS-I up to 105MB/s)

32.0×24.0×2.1 mm



15.0×11.0×1.0 mm



SD Express Card – Features

Initiate either directly from the PCIe®/NVMe® technology or SD

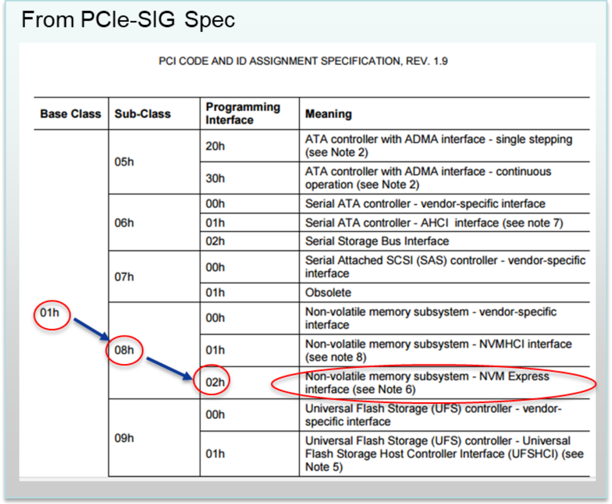
- Fully compatible to PCIe/NVMe standards. Identifies itself as a standard NVMe Memory

ESD protection up to 4KV on all pads (Same as legacy SD card requirements)

Hot Plug-In/Removal support

Boot, TCG and RPMB (SD9) may be supported by the SD interface as well

Working on New Speed Classes over NVMe technology[(1)]


From PCIe-SIG Spec

PCI CODE AND ID ASSIGNMENT SPECIFICATION, REV. 1.9

Base Class	Sub-Class	Programming Interface	Meaning
	05h	20h	ATA controller with ADMA interface - single stepping (see Note 2)
		30h	ATA controller with ADMA interface - continuous operation (see Note 2)
	06h	00h	Serial ATA controller - vendor-specific interface
		01h	Serial ATA controller - AHCI interface (see note 7)
		02h	Serial Storage Bus Interface
	07h	00h	Serial Attached SCSI (SAS) controller - vendor-specific interface
		01h	Obsolete
01h	08h	00h	Non-volatile memory subsystem - vendor-specific interface
		01h	Non-volatile memory subsystem - NVMHCI interface (see note 8)
		02h	Non-volatile memory subsystem - NVM Express interface (see Note 6)
	09h	00h	Universal Flash Storage (UFS) controller - vendor-specific interface
		01h	Universal Flash Storage (UFS) controller - Universal Flash Storage Host Controller Interface (UFSHCI) (see Note 5)

(1) Forward-looking statement: SDA undertakes no obligation to realize these forward-looking statements, which speak only as of the date hereof.

SD Express - Applications

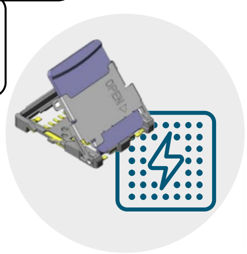
Multi Channel Video Capturing ➔ requires multi-stream high speed recording and captures large amount of data

4K cameras are everywhere
Plus growing 8K, 12K and 8k360o VR cameras with huge data/speed requirements
(8K/24fps uncompressed requires 6GB per minute or 360GB per hour!)
Off-the-shelf bridge solutions allow full support of SD-UHS-II cards as well as SD Express enabling smooth transition

Gaming with 3D high-resolution graphics ➔ requires more memory and high-speed capability for real-time usage

Multi-sensor Data Collection
And/or
Multimedia Apps running from cards

VR & AR video increasing in quality ➔ requires a high-speed real-time view of 360º



Right To Repair

Semi-embedded applications (IoT, Mobile-Compute etc)

CompactFlash Defined Form Factors

Dave Landsman, WDC

CompactFlash Association

Charter

Create a standards-based removable storage ecosystem for the professional/prosumer imaging, automotive/industrial markets.

History

Established 1995 by SanDisk and a group of Consumer Electronics manufacturers during infancy of digital photography to establish the CompactFlash into an industry standard

Standards evolved through CFA’s successful 27-year history from PCMCIA to PCIe®+NVMe® technology. Widely adopted by professional and prosumer digital cameras and camcorders

Membership

81 Corporate Members consisting of Host, Card, Peripheral and Tester manufactures

Americas = 25, Japan = 25, Asia = 20, Europe = 11

Addressable Markets

Professional Imaging

Renowned professional photo and video camera manufacturers have been loyal to CompactFlash since its inception for its performance, capacity, and reliability

Automotive and Industrial

Data acquisition, analysis, and storage is key to AI-based

industrial and autonomous transportation applications

CFexpress

Value Proposition

Highest performance, highest capacity, and highest reliability removable storage solution for the market

Form Factor	TYPE A	TYPE B	TYPE C
Dimensions width x length x thickness (mm)	20.0 x 28.0 x 2.8	38.5 x 29.6 x 3.8	54.0 x 74.0 x 4.8
PCIe Max Lanes on Card	1	2	4
CFexpress 1.0 = PCIe® Gen 3 and NVMe® 1.2	NO	YES	NO
1 Lane	N/A	8Gbps	N/A
2 Lanes	N/A	16Gbps	N/A
4 Lanes	N/A	N/A	N/A
CFexpress 2.0 = PCIe® Gen 3 and NVMe® 1.3	YES	YES	YES
1 Lane	8Gbps	8Gbps	8Gbps
2 Lanes	N/A	16Gbps	16Gbps
4 Lanes	N/A	N/A	32Gbps
Capacities in Market (as of July 15, 2022)	<=160GB	<= 4TB	N/A

Flash Memory Summit

nvm EXPRESS®

CFexpress: Why NVMe®/PCIe®?

	NVMe	PCIe
Solid foundation	• Optimized for low latency NVM • Exploits platform parallelism • Efficient SW stack	• No HBA • P2P transfers
By ~2015, NVMe & PCIe had added various features applicable to mobile and removable cards	• Host memory buffer (less/no DRAM in device) • Enhanced Power Management • Boot, Write Protect, RPMB	• L1.2 Sub-states (low power)
Continued development	• Data placement interfaces, driven by datacenter which optimize latency/QoS (ZNS, FDP, …)	• FW Attestation • Link encryption
Other advantages across NVMe/PCIe ecosystem	• Hot plug infrastructure • Better integration w/ internal storage	

Why NVMe/PCIe?
Datacenter to Removable; Investment and Innovation

Full BIO Slides

John Geldman
Director, SSD Industry Standards at KIOXIA

Member Board of Directors, NVM Express

Member Board of Directors, SNIA

Co-Chair NVMe-MI™

Currently an active contributor to the following standards organizations:

- NVM Express, INCITS T10, INCITS T13, JEDEC, PCI-SIG, SATA IO, SNIA, IEEE SISWG
- In addition, John's team members are also active in CXL, DMTF, TCG, OCP

Corporate leadership responsibility for standards for multi-billion dollar storage vendors since 2011

Involved in storage standards since 1992, with an early introduction to standards including the transition from X3T9 to ATA & SCSI, and the transition from PCMCIA to CardBus

An active FMS CAB member for at least 10 years

Paul Kaler
Future Storage Architect at HPE

Brings over 20 years of experience to his current role where he is responsible for researching and evaluating future storage and interconnect technologies and defining the server storage strategy for ProLiant servers

Actively involved in multiple standards and industry organizations, and has been a key driver of standards including U.3, EDSFF E3, and the OCP Datacenter NVMe® SSD spec.

Previously led development of SSD storage arrays, been founder and co-founder of a couple of startups, and helped develop the first dual-screen smartphone.

Bill Gervasi,
Principal Systems Architect, Nantero, Inc.

Mr. Gervasi has over 40 years of experience in high speed memory subsystem definition, design, and product development. Career highlights include 19 years at Intel where he was systems hardware designer, software designer, and strategic accounts manager. Mr. Gervasi subsequently was with S3 where he was a graphics architecture specialist and at Transmeta as memory technology analyst. Most recently he held several key positions with companies such as Netlist, SimpleTech, and US Modular driving unique memory module configurations. He is now Principal Systems Architect for Nantero, developing non-volatile RAM-class memories.

Mr. Gervasi been involved in the definition of Double Data Rate SDRAM since its earliest inception. He has served on the JEDEC Board of Directors and chaired committees for DRAM parametrics and small form factor memory modules during the development of DDR1 through DDR5. He is currently the chairman of the JEDEC Alternative Memory committee.

Michael Lavrentiev
Technologist, Systems Engineering at Western Digital

Chair of SWG at SDA since 2018

Developing SD Express since Nov 2016

Contributor to SDA since 2012

Contributed for the development of new generations of market leading SD and microSD cards.

Handled product management and product requirements for various flash memory solutions.

Before joining Western Digital, worked at KLA-Tencor, RSIP, Gyrus-ACMI and Intel.

Earned M.Sc. in Electrical Engineering from the Technion – Israel Institute of Technology.

Dave Landsman
Sr. Director, Distinguished Engineer, Western Digital

Manages storage standards across Western Digital's businesses.

Active in storage standards since 2008, representing SanDisk and then Western Digital.

Contributions to NVMe®, PCI-SIG, JEDEC, SATA-IO, T10, T13, SNIA, SFF, and others.

Currently WD's board representative for NVMe, SNIA, CFA and DNA Data Storage Alliance.

Has stopped counting years in the industry. Had "first career" at Intel, "second career" in storage at msystems/SanDisk/WD, and a startup in between.

BA in computer science from the University of California, San Diego. Aside from UCSD Pascal, most coding in ancient asm (VAX/PDP-11/misc).

Questions?

Architected for Performance