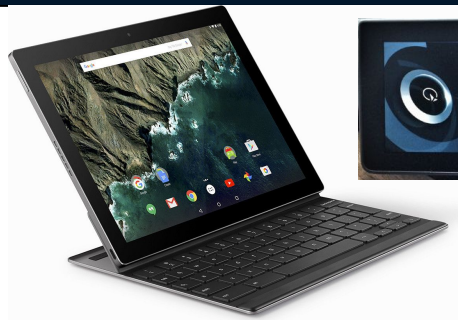
How consumer product like Google's Pixelbook benefit from NVMe Storage?

Zhiping Yang, Ph. D.
Google Inc.
Oct. 17, 2017

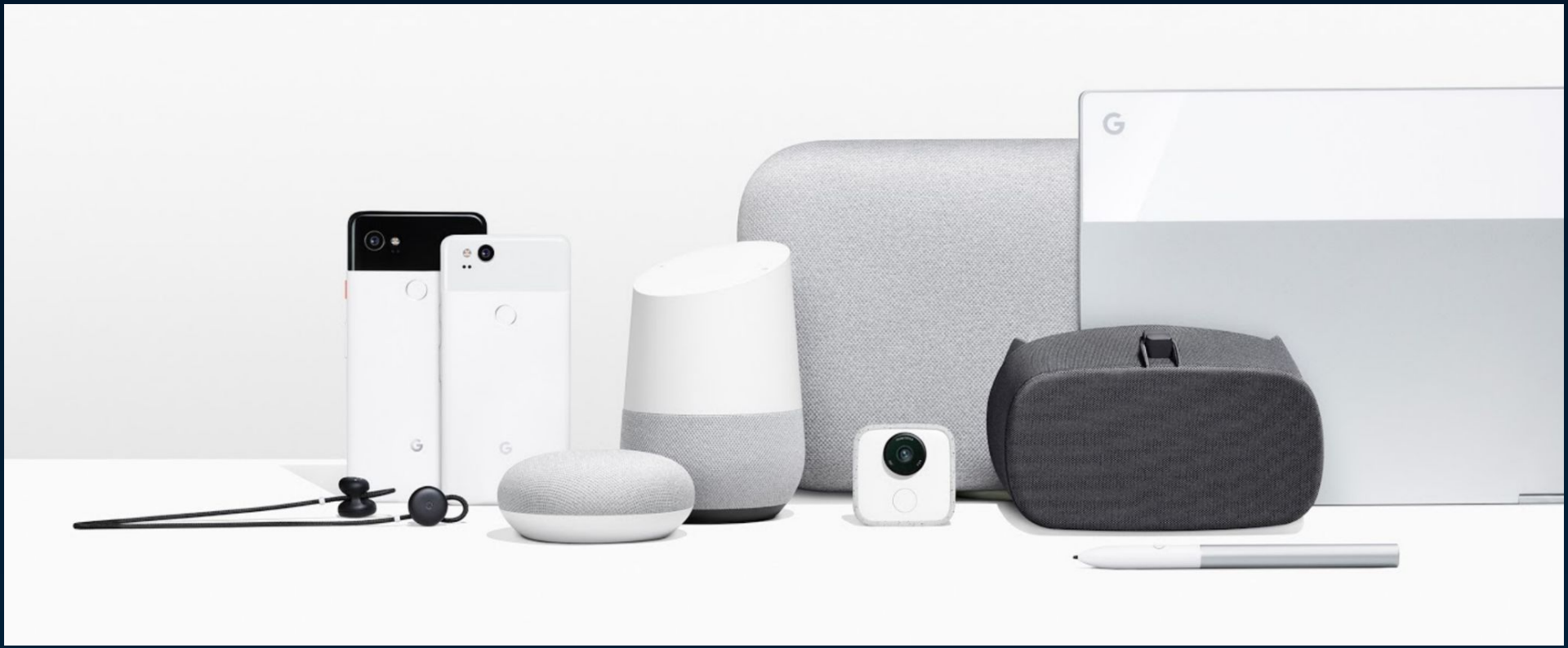
Agenda

- Consumer Hardware at Google
- Consumer Product's Storage Wishlist
- Why PCIe NVMe in Consumer Product?
  - Performance/Power/Thermal management
- Key Focus Areas for 1113 BGA SSD
  - Small size, UFS/eMMC backward compatibility
- Cost/Power Reduction, More Choices/Features
- Conclusion

Consumer Hardware at Google



Introducing a few new things made by Google



Google Pixelbook: **Designed to do everything you want to do.**

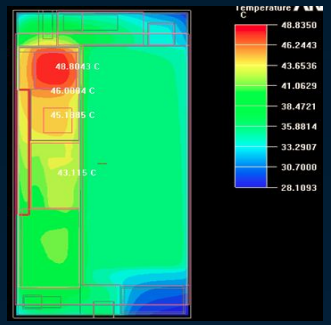
Consumer Product's Storage Wishlist

Low Cost

Small Space

Smooth Transition

High Performance



Smart Thermal/Power Management

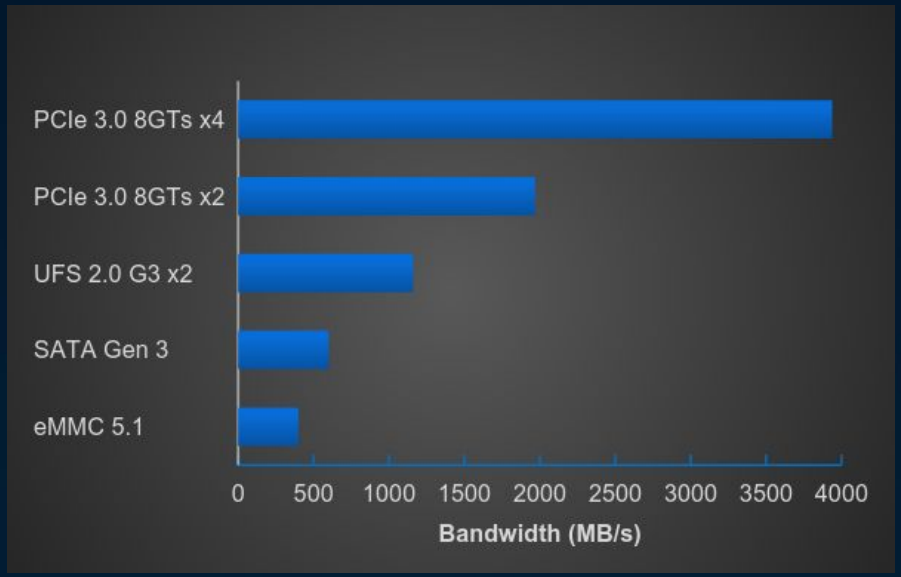
Low Power Consumption

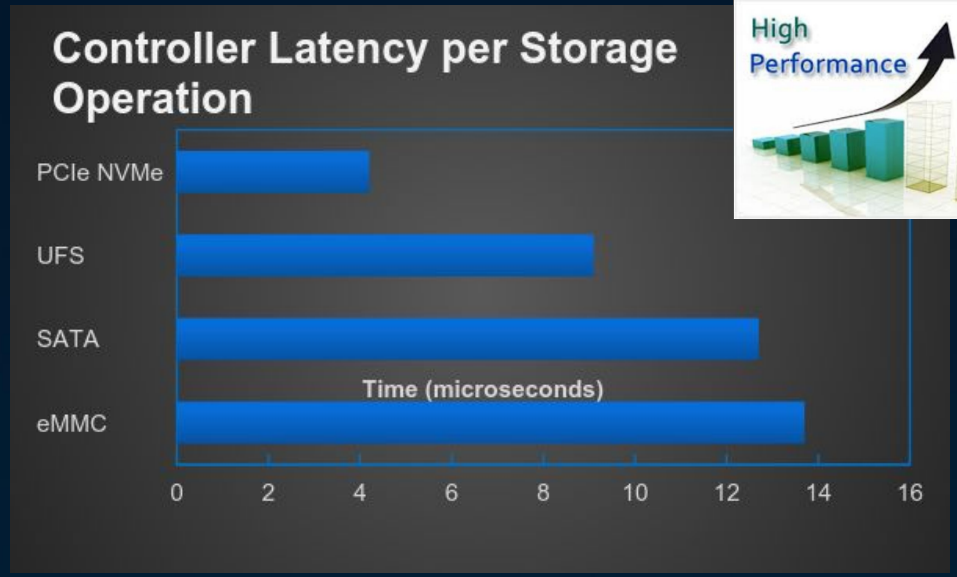
Why PCIe NVMe in Consumer Product?

- Scalable software stack across client platforms
- One less platform interface to develop and validate
- Flexibility in system design
- Low latency interface
- Drivers for all major OSes
- Host Memory Buffer provides DRAM-less solutions
- Scalability to next generation NVM
- PICe is widely adopted and well positioned for the future

PCIe and NVMe have clear benefits that are seen in the datacenter and client. These benefits can also be realized in consumer product.

Bandwidth/Latency Improvement With PCIe

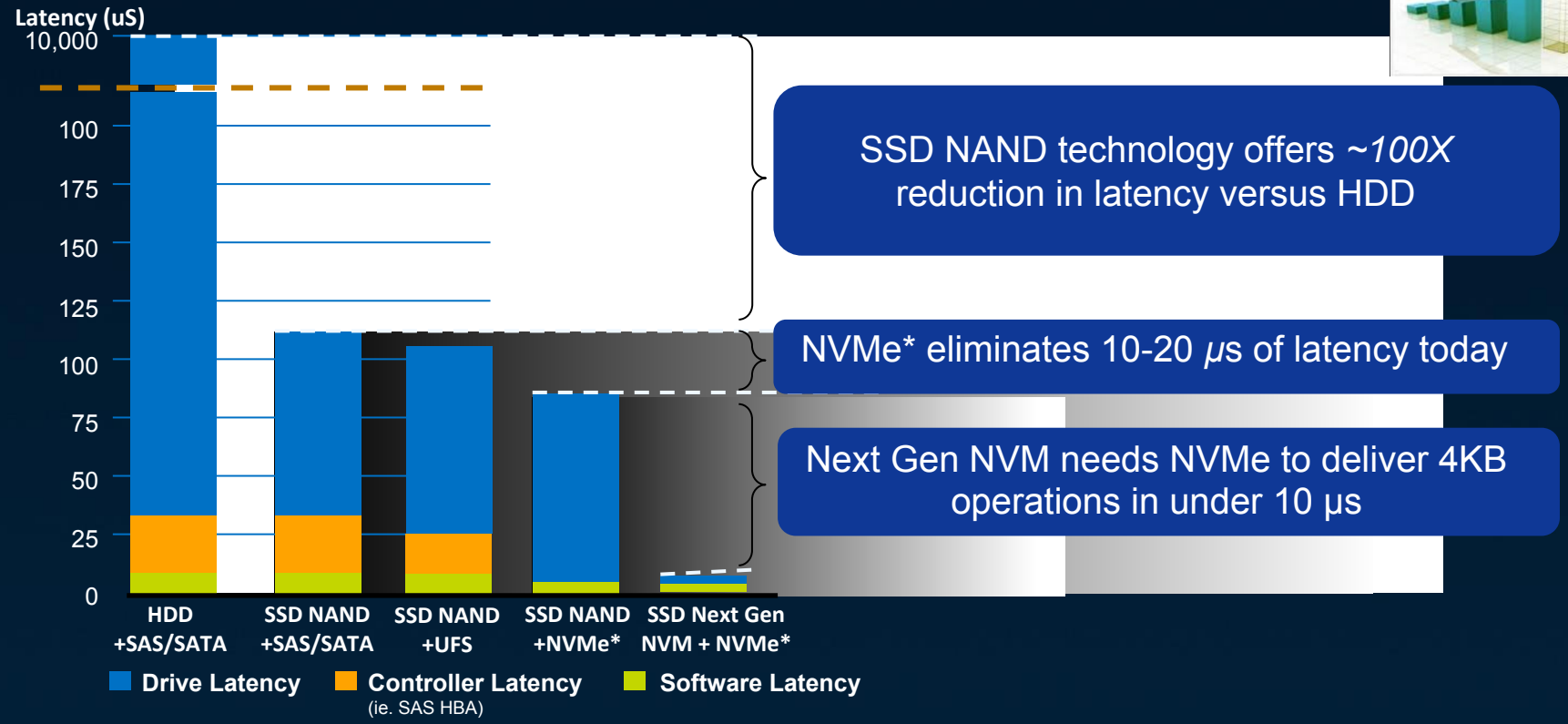




*UFS, eMMC and names associated with it are property of JEDEC.
Bandwidth from each interface are max theoretical that account for encoding overhead.
Latency values are estimated based on storage access to and from memory.

PCIe NVMe SSDs provides highest throughput and lowest latency

Latency Improvement Details



Source: Storage Technologies Group, Intel. Comparisons between memory technologies based on in-market product specifications and internal Intel specifications.

NVMe SSD Power

Two metrics of power: Active and Standby

- **Active Power:** Power consumed when doing something
- **Standby Power:** Power consumed when doing nothing

Active Power consumption is about energy consumed. PCI Express* (PCIe*) is competitive.

L1 Substates in PCIe reduce standby power to range acceptable for mobile

NVMe offers APST (Autonomous Power State Transition) to manage latency and power trade-offs.

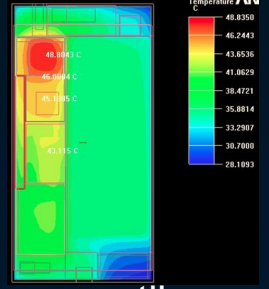
Item	PCIe Gen3	PCIe Gen2	UFS Gear3
Line Speed [Gbps]	8	5	5.83
PHY Overhead	128/130,1GB/s	8/10, 500MB/s	8/10, 583MB/s
Active Power [mW]	60 (L0)	46 (L0)	58 (HS)
Standby Power [mW]	0.11 (L1.2)	0.11 (L1.2)	0.2 (Hibern8)
MB/mJ (higher better)	14-18	8-12	8-12

UFS and names associated with it are property of JEDEC.

1. pci-sig.com: "L1 PM Substates with CLKREQ, Revision 1.0a"
2. Source: SanDisk*. Data based on PHY power estimates of PCIe vs. MPHY.

NVMe SSDs can be power competitive in Small Form Factor

NVM Express* SSD Thermals

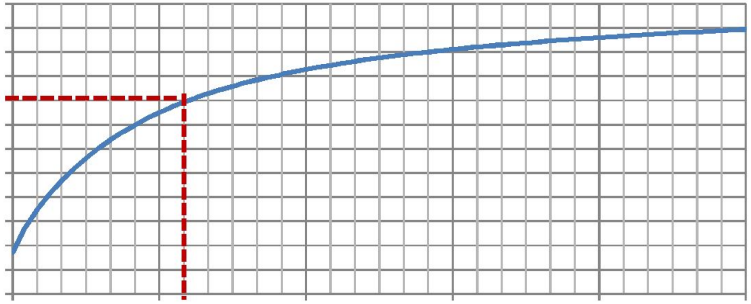


- Two ways to manage thermals in NVMe

1. Change power states: This allows the host system to run at a higher power/performance until the device or system becomes limited.
2. Host controlled Thermal Management: This allows the system to decide when the device needs to throttle from a thermal perspective.

Workload	Perf / Power	Time to Throttle
Worst case	1600MB/s @ 4W	~60s to heavy throttle
Sustained	1000MB/s @ 2W	Light throttle only
TDP	500MB/s @ <1W	Never

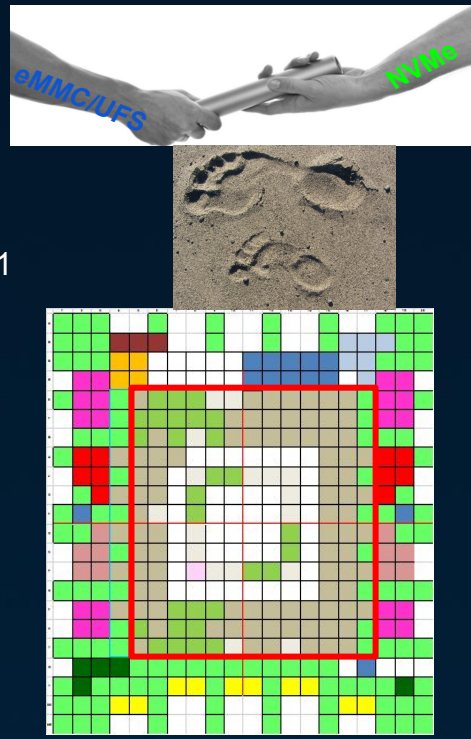
Source: NVM Solutions Group, Intel. Simulated Thermal constraints based on the PCIe* BGA 16x20mm package



NVMe SSDs provides burst performance while being able to throttle back if needed.

Key Focus Areas for 1113 BGA SSD

- Reduces the platform area in half compared to 16x20
  - 150mm2 vs. 320mm2
- 11.5x13 only supports x2 PCIe for now
- The same MLB can support either NVMe or eMMC[1]/UFS[1]
  - It provides a smooth transition from eMMC/UFS based storage to NVMe storage.
  - eMMC/UFS and NVMe together provide good performance, price, features, and capacity coverage to meet different customers' needs.
- Optional SPI interface for better system integration



1. eMMC and UFS specs are managed by JEDEC

NVMe innovations enable scaling into smaller form factors delivering new differentiated platforms.

Cost Reduction

- Controller cost has to go down
  - •Cost is a function of NAND, controller, and package/assembly/test.
  - •Higher capacity SSD means a lower percentage of cost from controller
- Cost vs Performance trade-off
  - •Lower performance & Less feature mean lower cost.
- Volume is critical to drive cost down
- Host Memory Buffer (HMB) (less cost with more performance)
  - •Support in Linux kernel has been proposed

PCIe BGA SSD cost can be competitive to eMMC depending on capacity and performance targets

Power Consumption Reduction

- Power/battery life is critical for consumer product
- PCI L1.2 power is still higher than existing eMMC/UFS solutions
- Power vs Performance trade-off
  - •Lower performance & less feature mean lower power.
- Optimize NVMe/PCIe PHY design for consumer product
  - •Shorter distance, within the same device/board
- Looking for more power saving features

PCIe BGA SSD power can be competitive to eMMC/UFS

More Suppliers & Choices

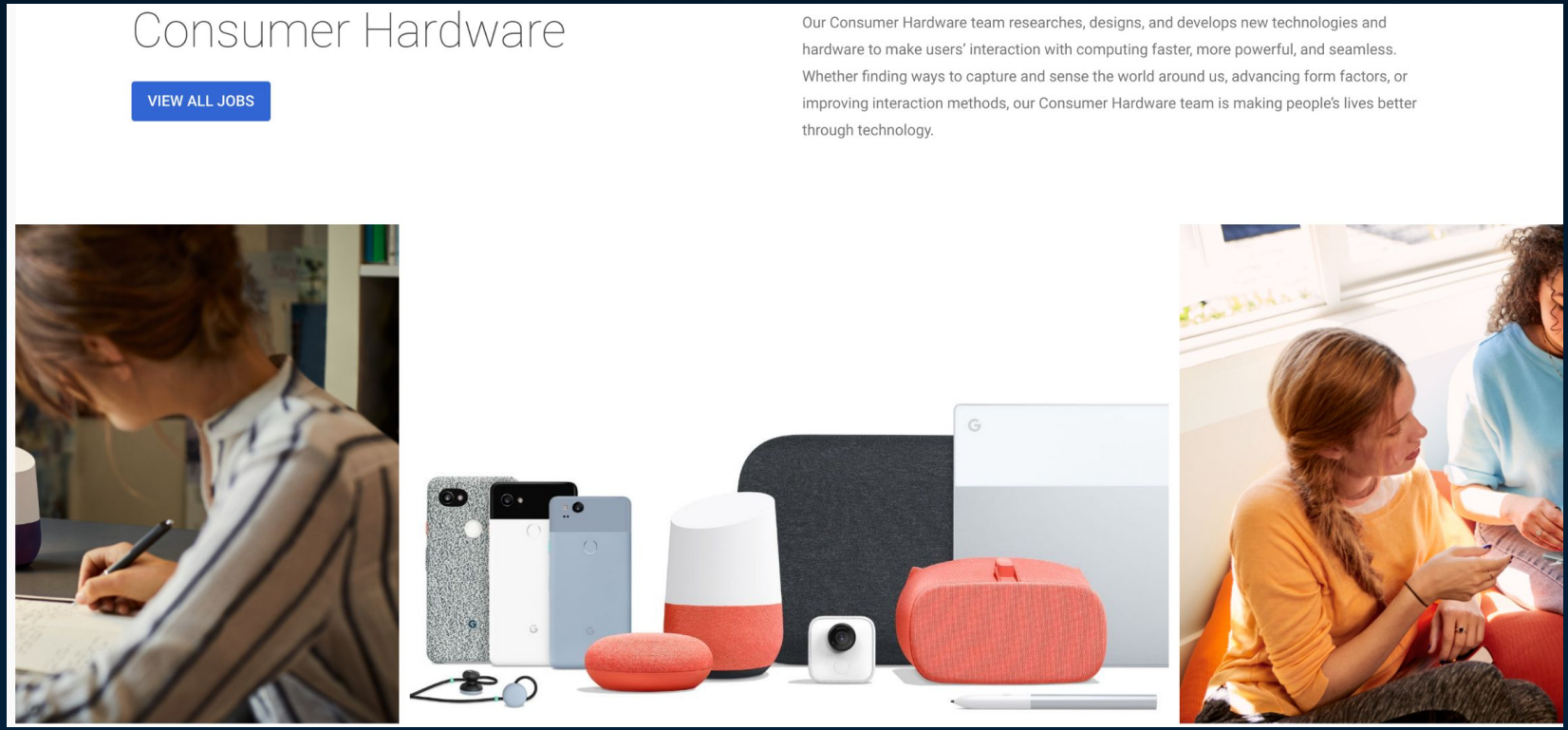
- LongSys's 1113 BGA NVMe announcement
  - Spec was approved in February 2017 by PCI-SIG
- More vendors in the pipeline for FY17/18

Conclusion

- 11x13 BGA NVMe consumer devices are coming!
  - New usages require better storage.
  - Offers many advantages over existing solutions
- Continue to push for lower cost, lower power consumption, and more choices.
- Call to action
  - Consider 11x13 PCIe BGA devices for your consumer product
  - Contribute to 11x13 development to make it more attractive
  - Support of smaller capacity (32/64GB)

Google Consumer Hardware Is Hiring!

- https://careers.google.com/hardware/



Q&A

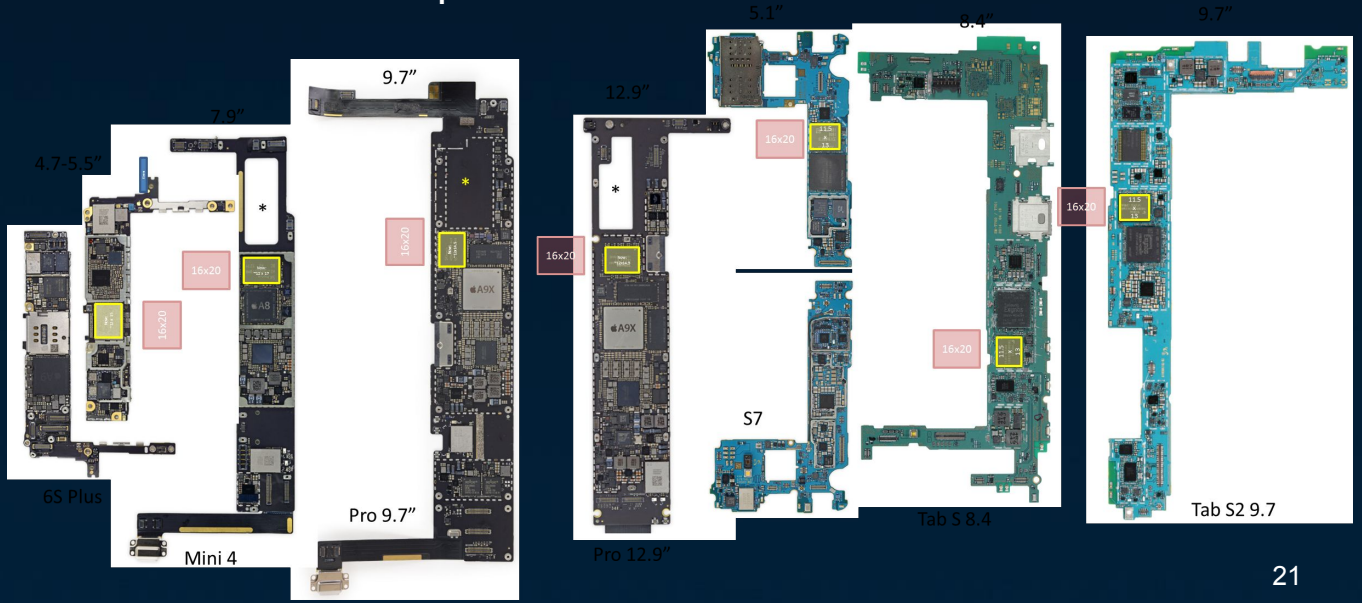
Backup

Available NVMe Form Factors

- M.2 Module: Capacity, Size, cost, backward compatibility
- 16x20mm BGA SSD: Capacity, Size, backward compatibility
- 11.5x13mm BGA SSD: **Right Solution**

Why Size Matters?

- 16x20mm does not fit in smaller form factor devices.
  - Need for a smaller performance SSD
- M.2 storage ~15% of platform area
- Allows for smaller motherboard and better product



Credit: iFixit

SPI ROM Interface for Secure Boot

Interface	Signal Name	I/O	Description	IO Voltage
SPI ROM[1]	WP_L[2]	I	Write protect signal to prevent writes from occurring to SPI NOR. Active low.	1.8 V
	SPI_CLK	I	SPI clock. Max frequency is 50MHz.	1.8 V
	SPI_MOSI	I	Master Out Slave In signal for SPI NOR.	1.8 V
	SPI_MISO	O	Master In Slave Out signal for SPI NOR	1.8 V
	SPI_CS_L	I	Chip select for SPI NOR. Active low.	1.8 V
	SPI_18	I	+1.8 V supply. Optional voltage supply if SPI NOR included in package.	1.8 V

[1] Optional
[2] WP_L can be used for NVMe hardware write protection as well.