

OPEN.



FOR
BUSINESS.



OCP
SUMMIT



Facebook Latest server/ML hardware Overview

Whitney Zhao/Hardware Eng./Facebook Inc.

OPEN. FOR BUSINESS



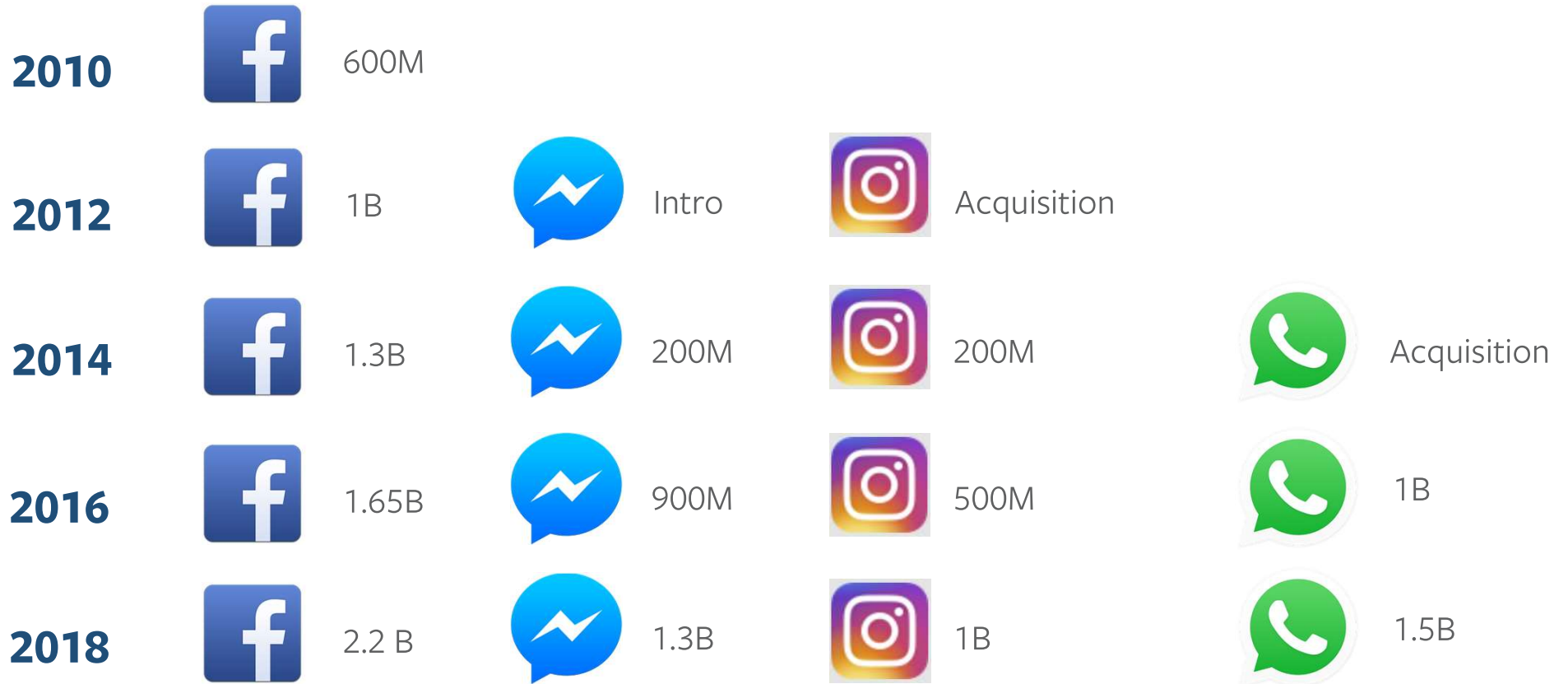


Agenda

- Overview
- 1S Server Yosemite V2
- 2S Server Tioga Pass
- AI/ML Server Big Basin V2

OPEN. FOR BUSINESS





A history of building for scale

2011



Data Center



Triplet Rack



Battery Cabinet



Freedom Servers



Spitfire Server (AMD)



Power Supply

2012



Windmill (Intel)



Watermark (AMD)



Mezzanine Card V1

2013



Knox



Winterfell



Open Rack V1



Group Hug

2014



Open Rack V2



Mezzanine Card V2



Cold Storage



Micro Server (Panther)

2015



BluRay



Leopard



Wedge



Honey Badger

2016



Wedge 100



Big Sur



Yosemite



Six Pack

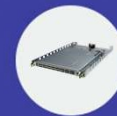


Backpack



Lightning

2017



Wedge 100S



Bryce Canyon



Yosemite V2



Tioga Pass



Big Basin

2018



Big Basin V2



Twin Lake



OCP NIC3.0

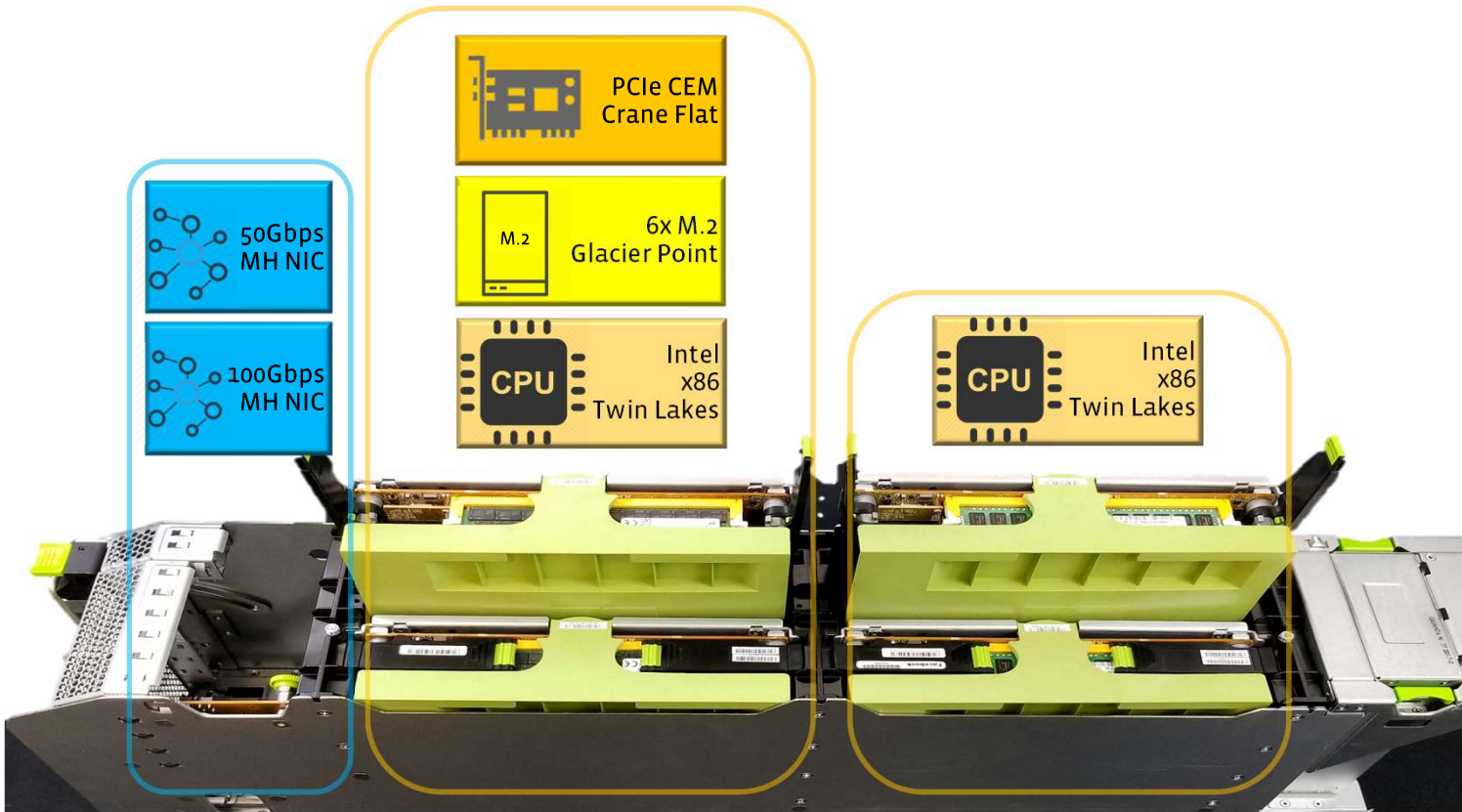


Modular Design 1S Server Yosemite V2

OPEN. FOR BUSINESS



Yosemite V2: Modular platform



OpenRack
V2

Hot
Service

Dual ODM

EMI
& Safety

<http://files.opencompute.org/oc/public.php?service=files&t=e4c2daeeb6438194ff3b2d20e61dba89>



System Overview

- Flexible Multi-Node 1S platform
- Twin Lakes 1S server
- Open Rack V2 compatible with new 40U vCubby design
- Support 4x 1S server cards or 2x 1S server + 2x Device cards
- Single smart baseboard to handle different configurations
- 100G/50G Multi-Host Network options



	Intel® Xeon® D-2100 Processor	Intel® Xeon® D-1500 Processor
CPU Cores	Up to 18 with Intel® HT	Up to 16 with Intel® HT
Cache	LLC: 1.375 MB/Core MLC: 1MB/Core	LLC: 1.5KB/Core MLC: 256K/Core
Memory	4 Channels, DDR4 1866/2133/2400/2666*	2 Channels DDR4/DDR3L 1600/1866/2133
PCIe	CPU: x32 PCIe Gen 3 <ul style="list-style-type: none">• Twin Lakes uses all 32 lanes FlexIO: x20 PCIe Gen 3 <ul style="list-style-type: none">• Twin Lakes uses 9 FlexIO PCIe lanes	CPU: x24 PCIe Gen 3 lanes FlexIO: x8 PCIe Gen 2 lanes
Acceleration Engines	AVX512 Intel® QuickAssist Technology <ul style="list-style-type: none">• Up to 100 Gbps Crypto/Compression• 100 KOps PKE 2K	AVX256 No built-in Intel® QAT

*Twin Lakes validated at 2400 MT/s

Twin Lakes is architected for high compute performance

<http://files.opencompute.org/oc/public.php?service=files&t=9d240277a11b83f9764420e99e18ac38>

OPEN. FOR BUSINESS





General Purpose 2S Server Tioga Pass

OPEN. FOR BUSINESS



Overview

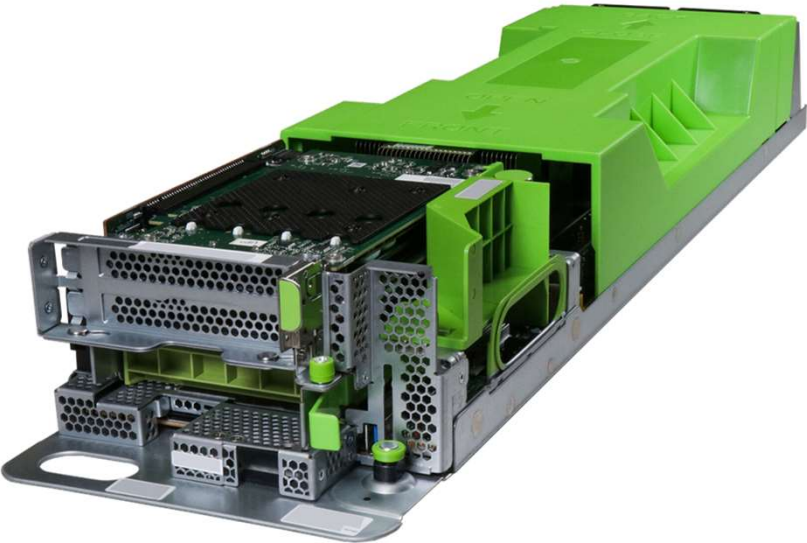
Facebook general purpose OCP 2S server

Higher performance over previous gen

Modularity

Easy to service

Open

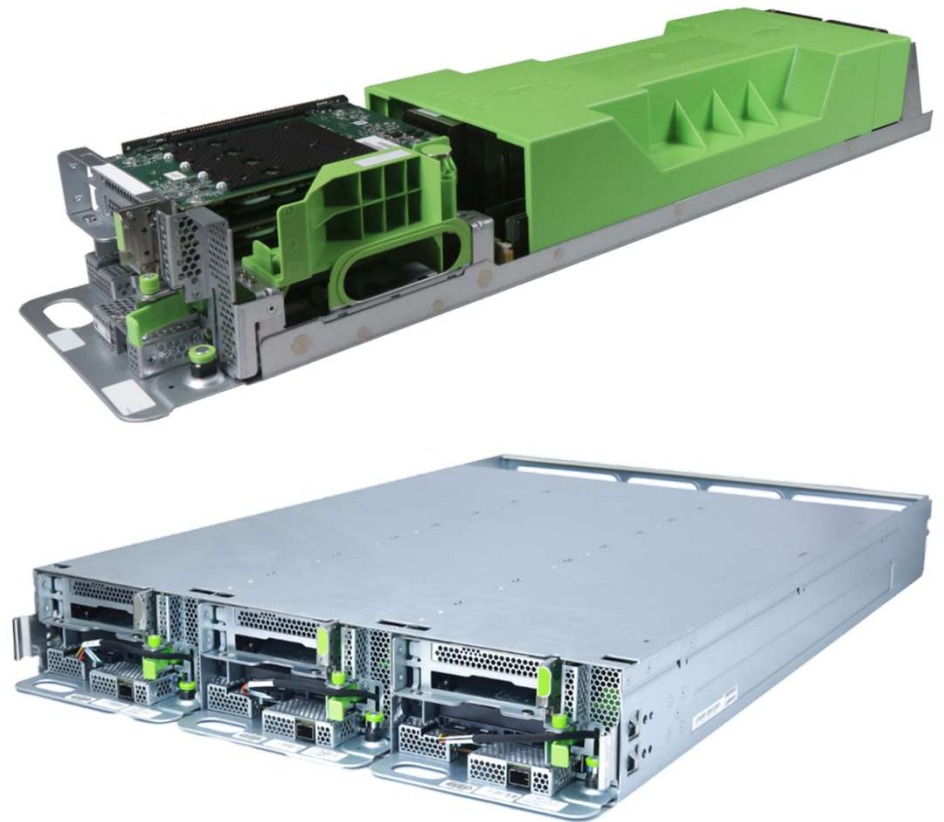


<https://www.opencompute.org/products/108/wiwynn-tioga-pass-standard-sv7220g3-s-2u-ocp-server-up-to-768gb-8gb16gb32gb-ddr4-up-to-2666mts-12-dimm-slots>
<http://files.opencompute.org/oc/public.php?service=files&t=e4c2daeeb6438194ff3b2d20e61dba89>



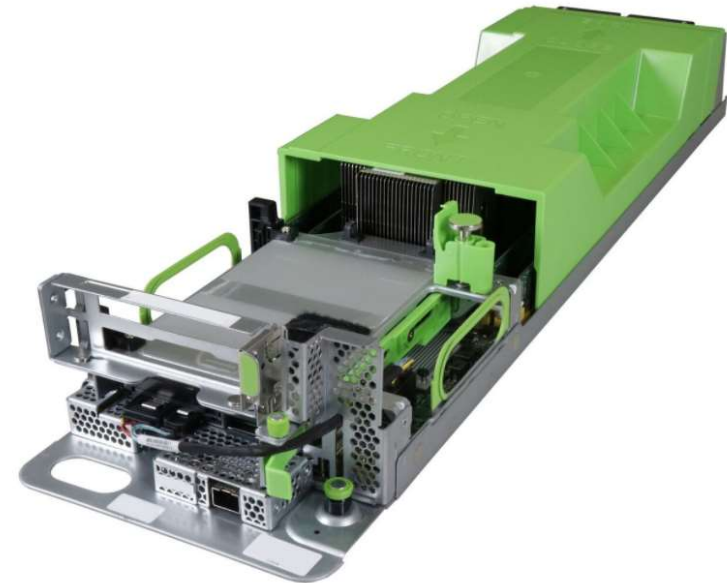
System Overview

- Intel Skylake-SP Processor up to 165w
- Single/Double Sided DIMM w/ 12 memory channels
- PCIe X32 to front I/O
- 3.5''HDD/Nvme M.2 ssd as boot drive
- Support up to 100G OCP NIC2.0
- Half width/High density
- Open Rack V2



Configuration

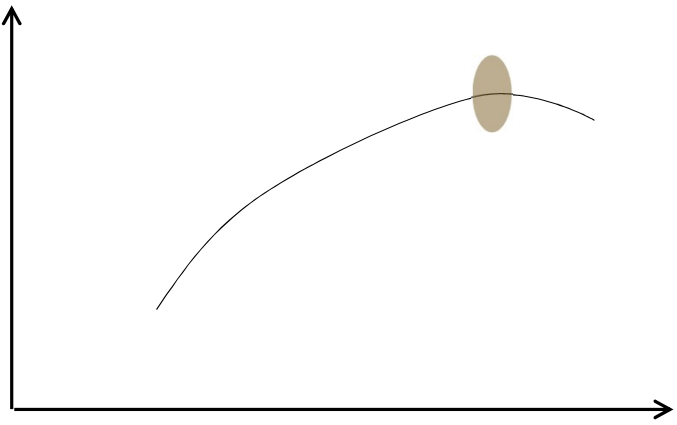
- Compute servers
- Compute server w/ PCIe AIC
- Headnode for PCIe expansion box
- Boot drive
 - 3.5" HDD
 - Nvme M.2 SSD
- NIC Card
 - OCP NIC2.0 25G/50G/100G
 - Intel KR Mezz Card



Optimize Power & Performance for FB workload

In search of best efficiency for scaling out

F(X)



$F(X) = \text{System performance} / w$
 $X = \{x_1, x_2, \dots, x_n\}$

Parameters to explore across different SKUs	
x_1	TDP
x_2	Core Count
x_3	Frequency
x_4	Tcase
x_5	Others



Big Basin Machine Learning Training Server

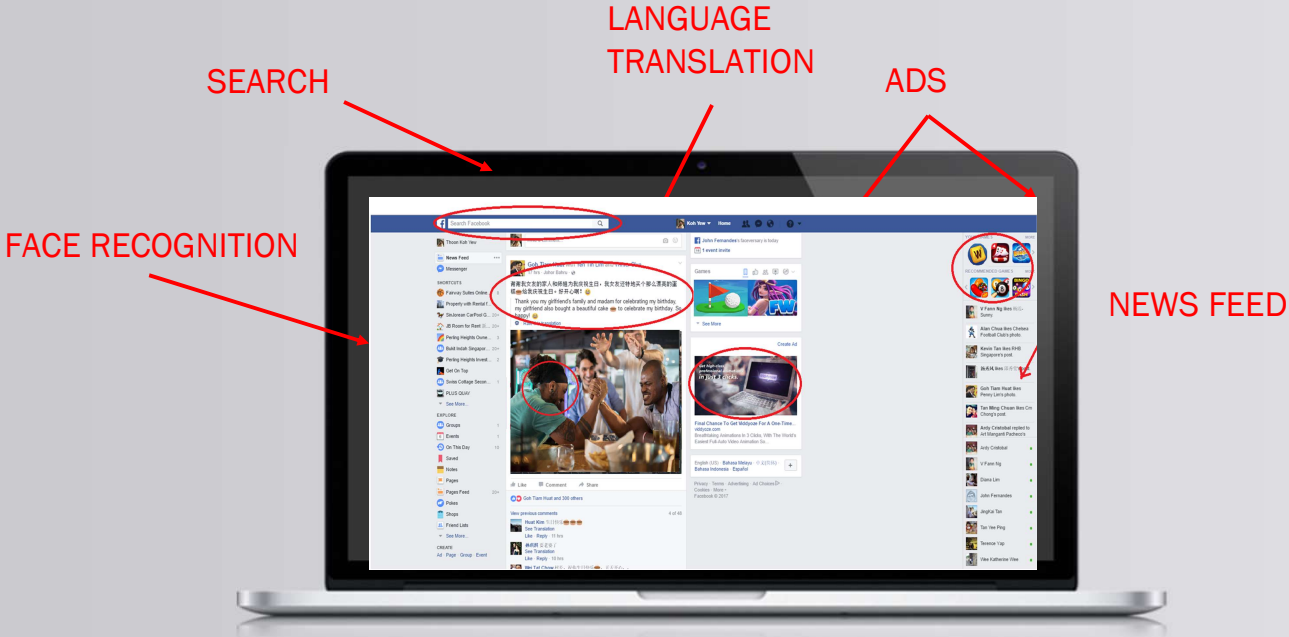
OPEN. FOR BUSINESS



Impact

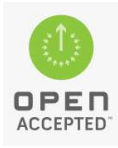
Facebook's commitment to developing AI & advancing ML

- LANGUAGE TRANSLATION
- FACE RECOGNITION
- SEARCH
- ADS
- NEWS FEED
- SIGMA
- LUMOS



Goal

- Open, full contribution to OCP
- Disaggregation/Modularity
- Serviceability



2016: Big Sur



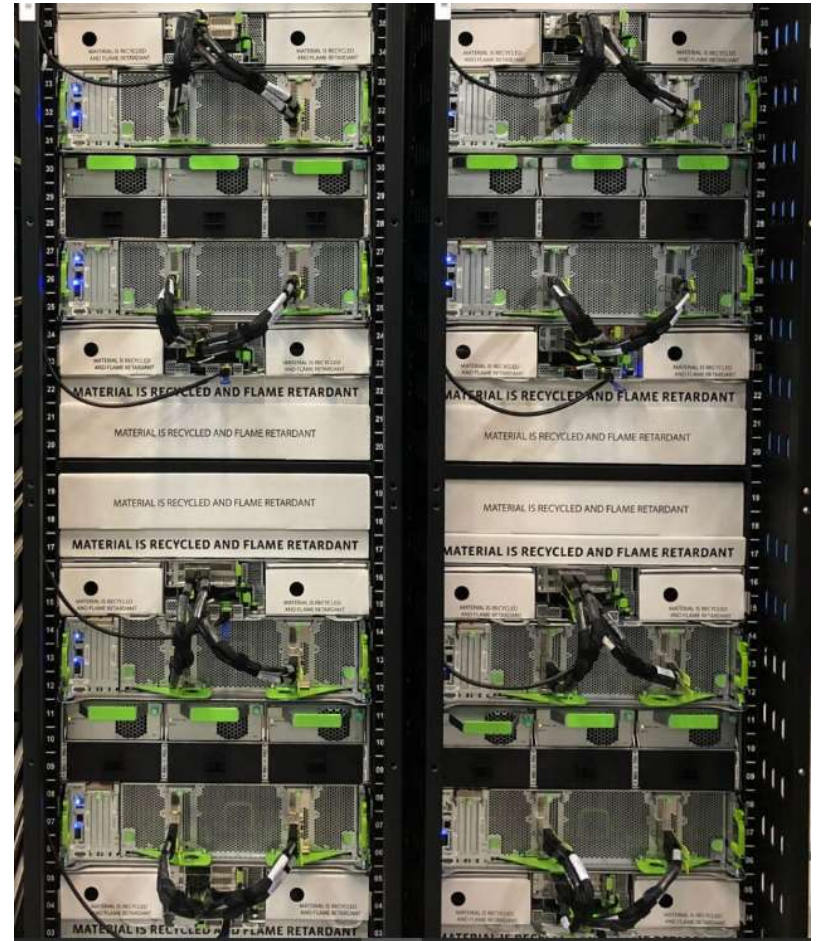
2017: Leopard + Big Basin
2018: Tioga Pass + Big Basin V2



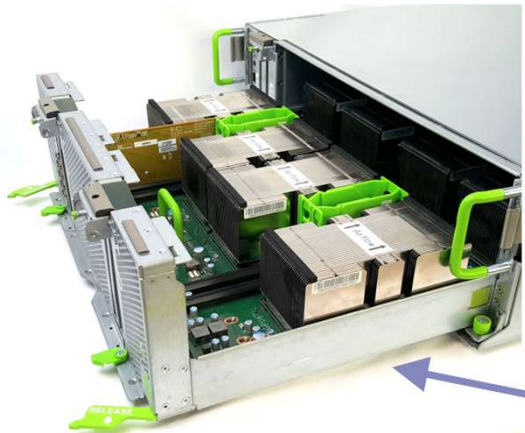
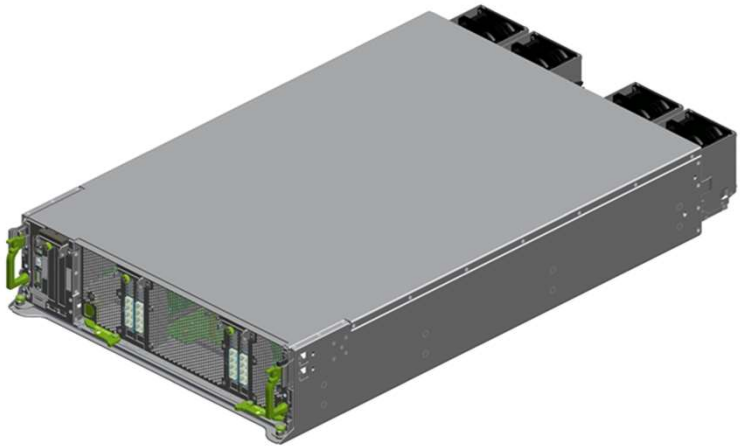
<http://files.opencompute.org/oc/public.php?service=files&t=4131a08f46af65f6f3126de39f6198d3&download>

Big Basin V2 Overview

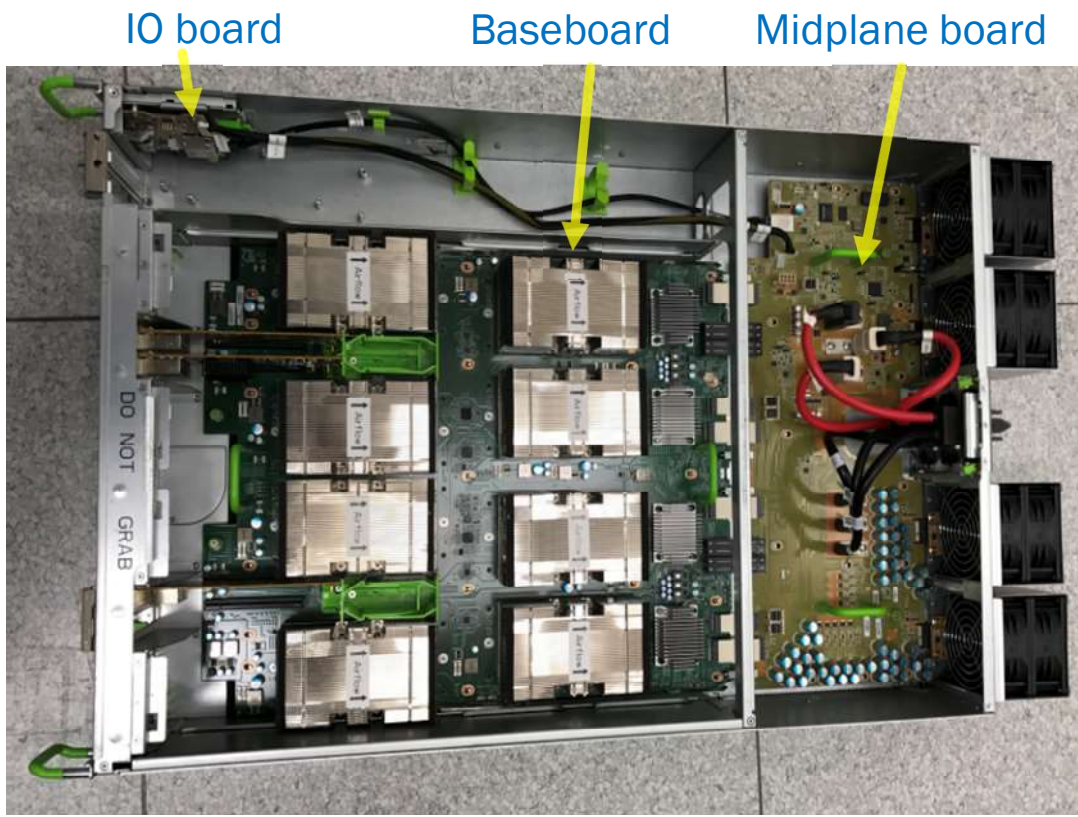
- 3 OU chassis
- Open Rack v2 compatible
- 8x Nvidia Tesla V100 GPUs; NVLink capable
- 300W TDP for each Tesla V100 GPU
- Facebook 2S Server Tioga Pass as Head node



A deeper look into Big Basin



Baseboard on sliding tray



IO board

Baseboard

Midplane board

Performance

- Comparisons of GPU Hardware

	Metrics	NVIDIA V100	NVIDIA P100	Improvement
Performance	FP-32	15 TFLOPS	10.6 TFLOPS	1.42x
	FP-16	30 TFLOPS	21.2 TFLOPS	
	TensorCore	125 TFLOPS	NA	Up to 5x
	Mem Bandwidth	900 GB/s	720 GB/s	1.25x
	NVLink	300 GB/s	160 GB/s	1.88x
Power		300 W	300 W	

Performance

- Comparisons of GPU Hardware
- Head-node upgrade: Tioga Pass
 - New CPU architecture: Broadwell-SP to Skylake-SP
 - Double PCIe bandwidth
 - Upgraded 100G NIC
- CUDA 9 + cudnn 7: faster libraries, etc.



[Learn More About Facebook Hardware](#)

OPEN. FOR BUSINESS





OCP
SUMMIT

OPEN.



**FOR
BUSINESS.**

