


ARGONNE NATIONAL LAB IPU WORKSHOP - HARDWARE



GRAPHCORE

 Graphcore Confidential

GRAPHCORE'S SOLUTION

Hardware



**IPU processor
designed for AI**

Software



**Poplar SDK and
development tools**

Platform



**IPU platforms
Available in the cloud**

THE INTELLIGENCE PROCESSING UNIT (IPU) WHAT MAKES IT DIFFERENT?

CPU

GPU



IPU

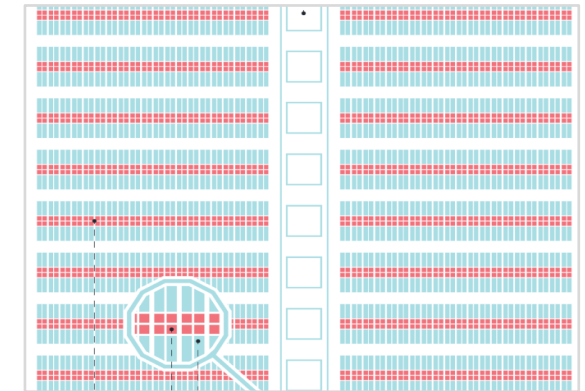
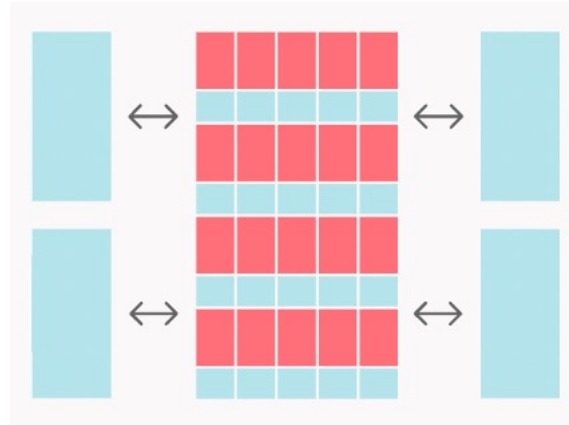
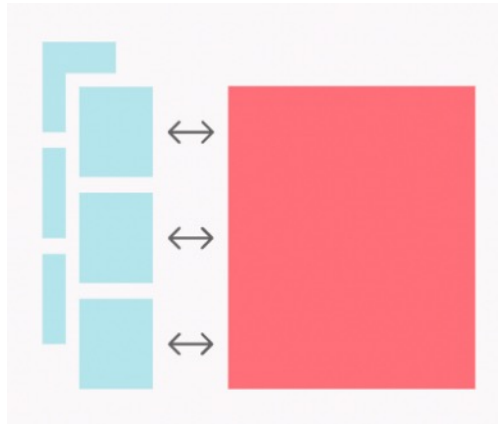
Parallelism

Designed for scalar processing

SIMD/SIMT architecture. Designed for large blocks of dense contiguous data

Massively parallel MIMD architecture. High performance/efficiency for future ML trends

Processor 
Memory 



Memory Bandwidth

Off-chip memory

Model and Data spread across off-chip and small on-chip cache and shared memory

Main Model & Data in tightly coupled large locally distributed SRAM

(2TB/s for A100 HBM)

(~65 TB/s for Bow IPU)

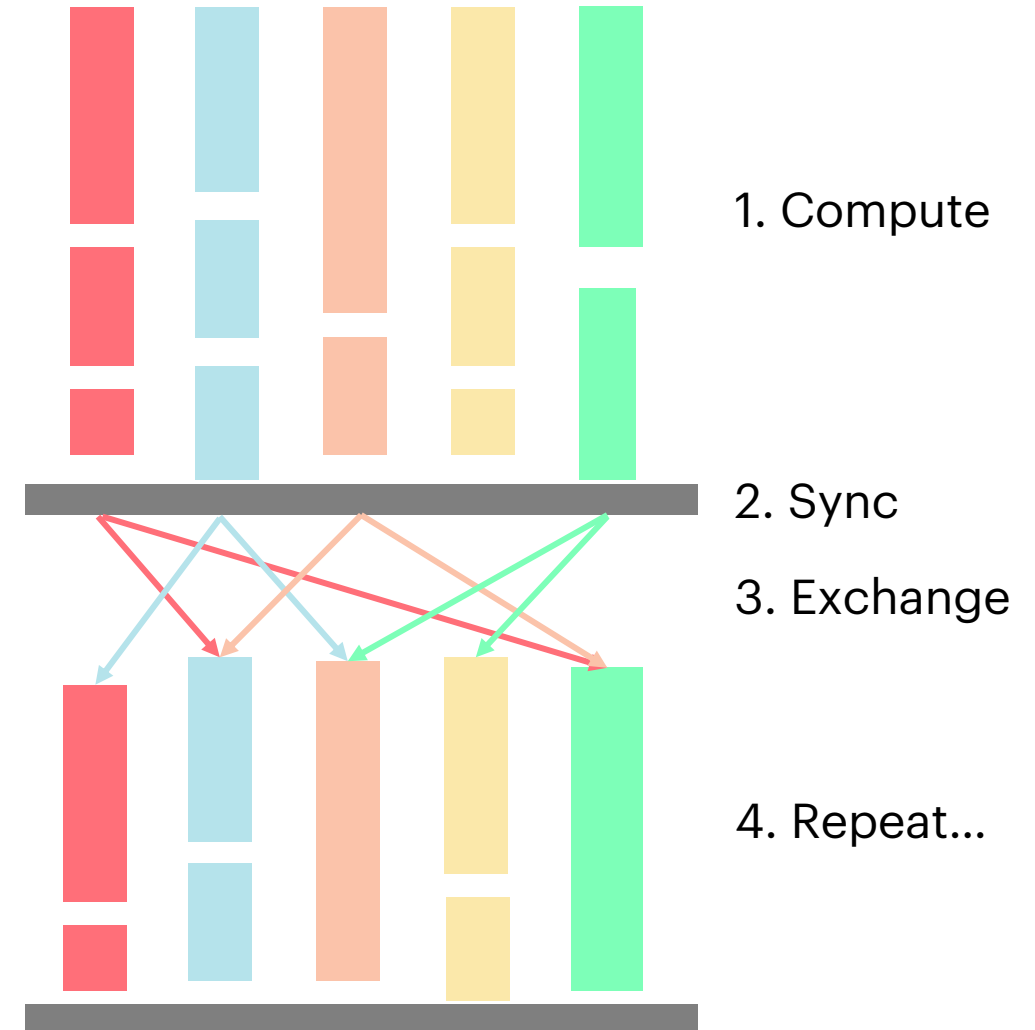
BRIDGING SOFTWARE TO PARALLEL HARDWARE

Bulk Synchronous Parallel (BSP) is the computer science solution that bridges software to parallel processors.

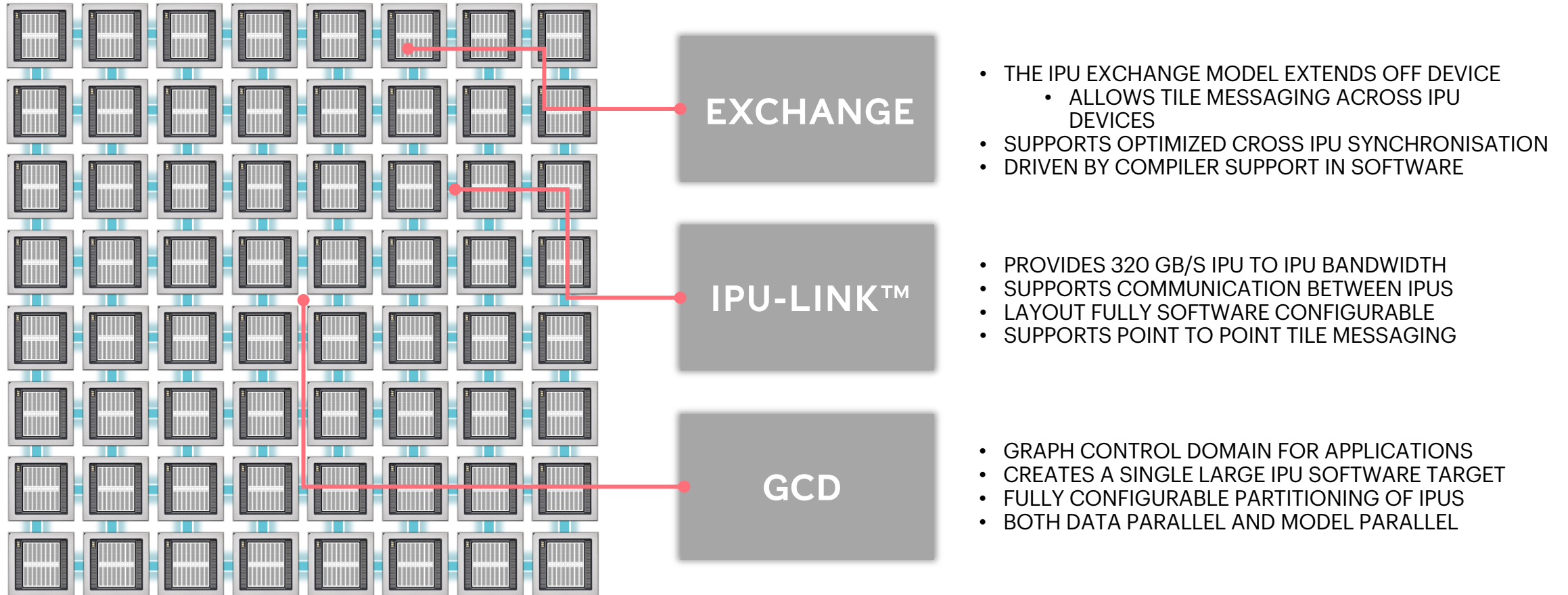
- IPU implements BSP in hardware.

BSP solves parallel software issues:

- free of concurrency hazards
- efficient for highly parallel processors
- Poplar-SDK builds BSP compute on IPU hardware



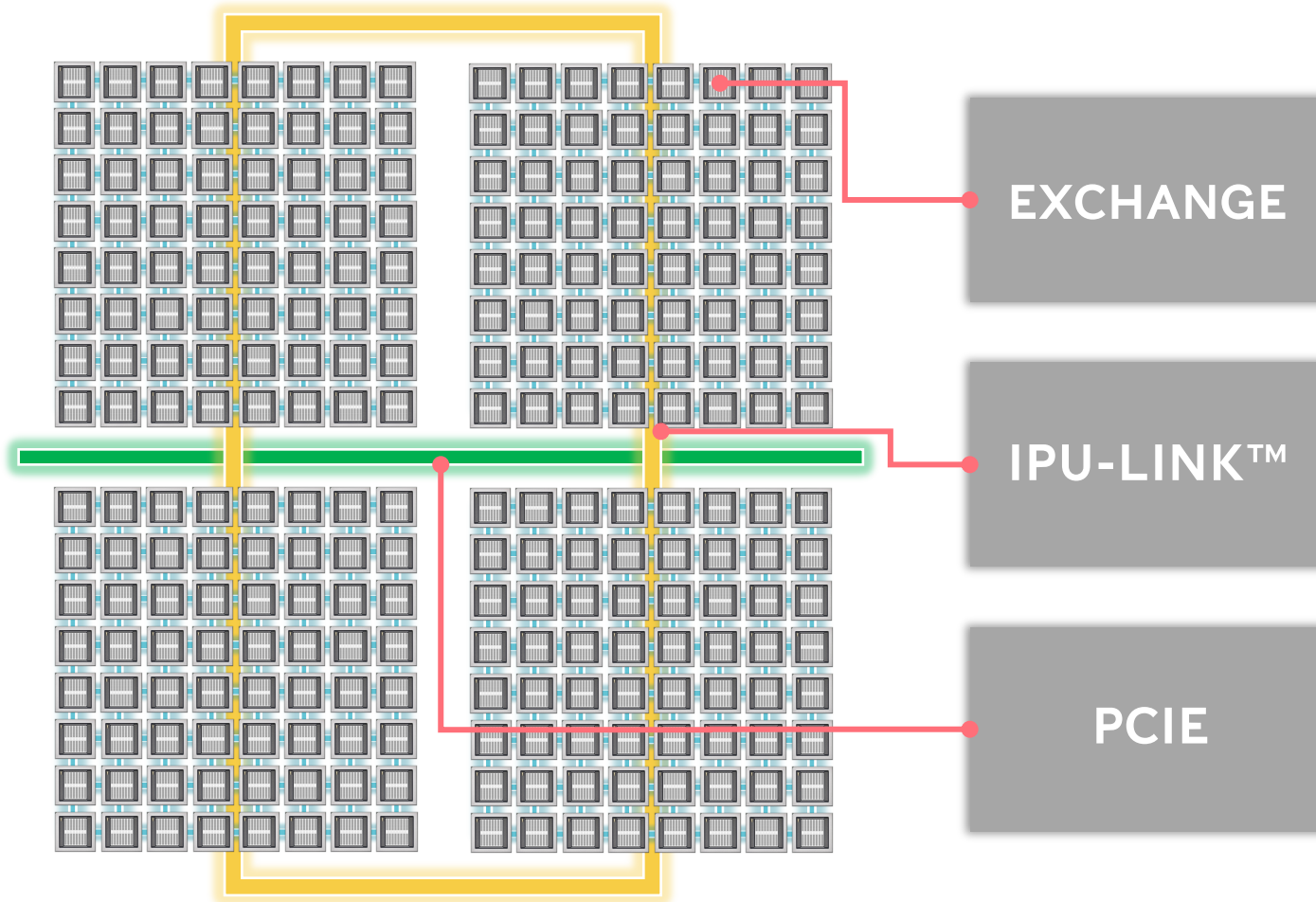
SCALING ACROSS DEVICES



UP TO 64 IPU DEVICES USABLE AS A SINGLE LARGE IPU FROM APPLICATIONS

565248 FULLY INDEPENDENT WORKERS, 57.6GB IN-PROCESSOR MEMORY™, LEVERAGING OVER 3.8 TRILLION TRANSISTORS

SCALING ACROSS SYSTEMS



- IPU EXCHANGE SUPPORT ACROSS DOMAINS
 - DRIVEN BY COMPILER SUPPORT IN SOFTWARE
- ENABLES APPLICATION COLLECTIVES SUPPORT
- ALLOWS SCALING UP TO 64000 IPU DEVICES

- IPU LINK™ CAN BE EXTENDED ACROSS DOMAINS
- SUPPORTS OPTIMIZED IPU LINK™ COLLECTIVES
- ALLOWS REPLICATION ACROSS SYSTEMS
- SUPPORTS A STANDARD IPU SOFTWARE MODEL

- IPUS CAN ACCESS MEMORY AND DEVICES OVER PCIE
- ALLOWS INTERFACING WITH HOST BASED SOFTWARE
- APPLICATIONS CAN BUILD ON HOST NETWORKING
- ALLOWS SCALING IN STANDARD SERVER PLATFORMS

256 IPU APPLICATION TARGET BUILT FROM INTERCONNECTED 64 IPU DOMAINS

PRODUCT DETAILS



GRAPHCORE

 Graphcore Confidential

BOW IPU PROCESSOR

Deep Trench Capacitor

Efficient power delivery
Enables increase in operational performance

Wafer-On-Wafer

Advanced silicon 3D
stacking technology
Closely coupled power
delivery die
Higher operating frequency
and enhanced overall
performance

IPU-Tiles™

1472 independent IPU-Tiles™ each with an
IPU-Core™ and In-Processor-Memory™

IPU-Core™

1472 independent IPU-Core™
8832 independent program threads
executing in parallel

In-Processor-Memory™

900MB In-Processor-Memory™ per IPU
65.4TB/s memory bandwidth per IPU

IPU-Links™

10x IPU-Links,
320GB/s chip to chip bandwidth

IPU-Exchange™

11 TB/s all to all IPU-Exchange™
Non-blocking, any communication pattern

PCIe

PCI Gen4 x16
64 GB/s bidirectional bandwidth to host



BOW-2000 IPU MACHINE

IPU blade form factor delivering 1.4 PetaFLOPS AI Compute

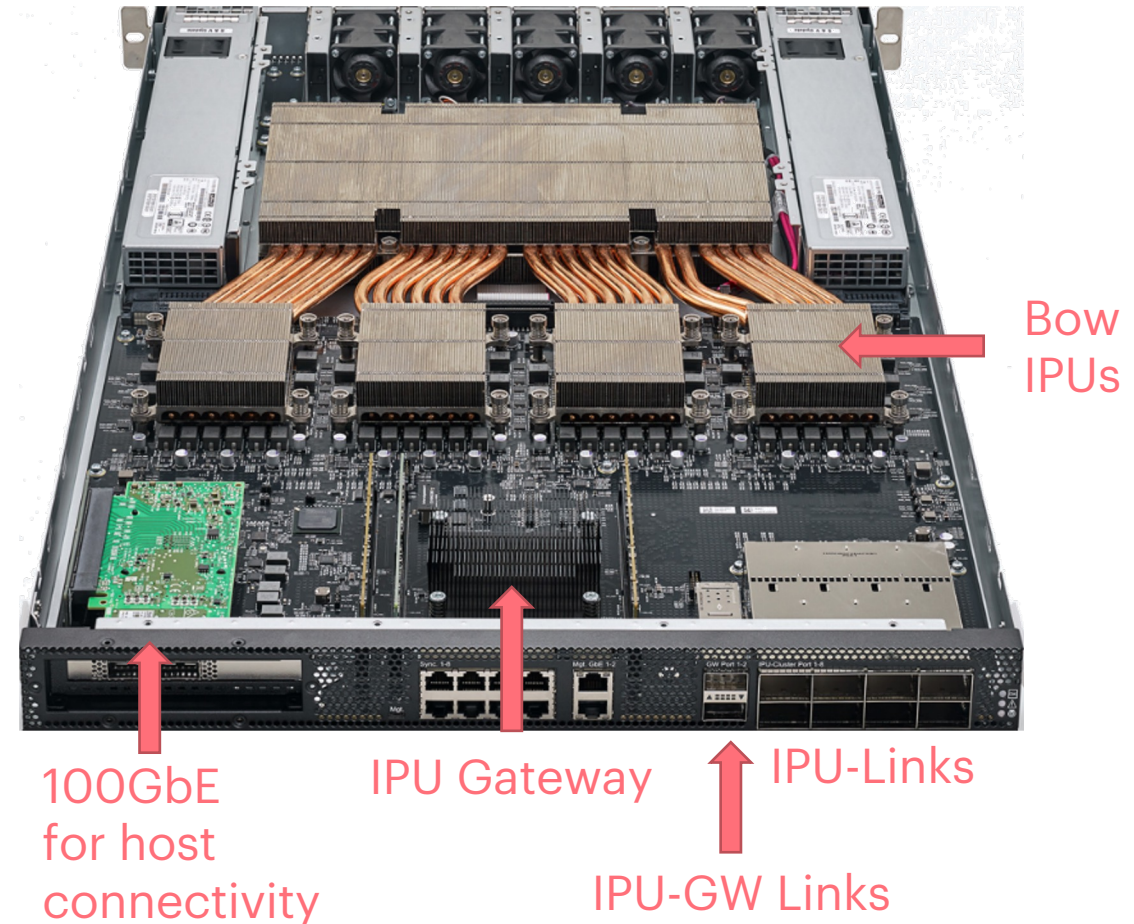
Disaggregated AI/ML accelerator platform

Excellent performance & TCO leveraging
In-Processor memory & IPU-Exchange

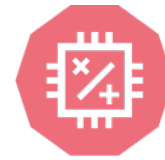
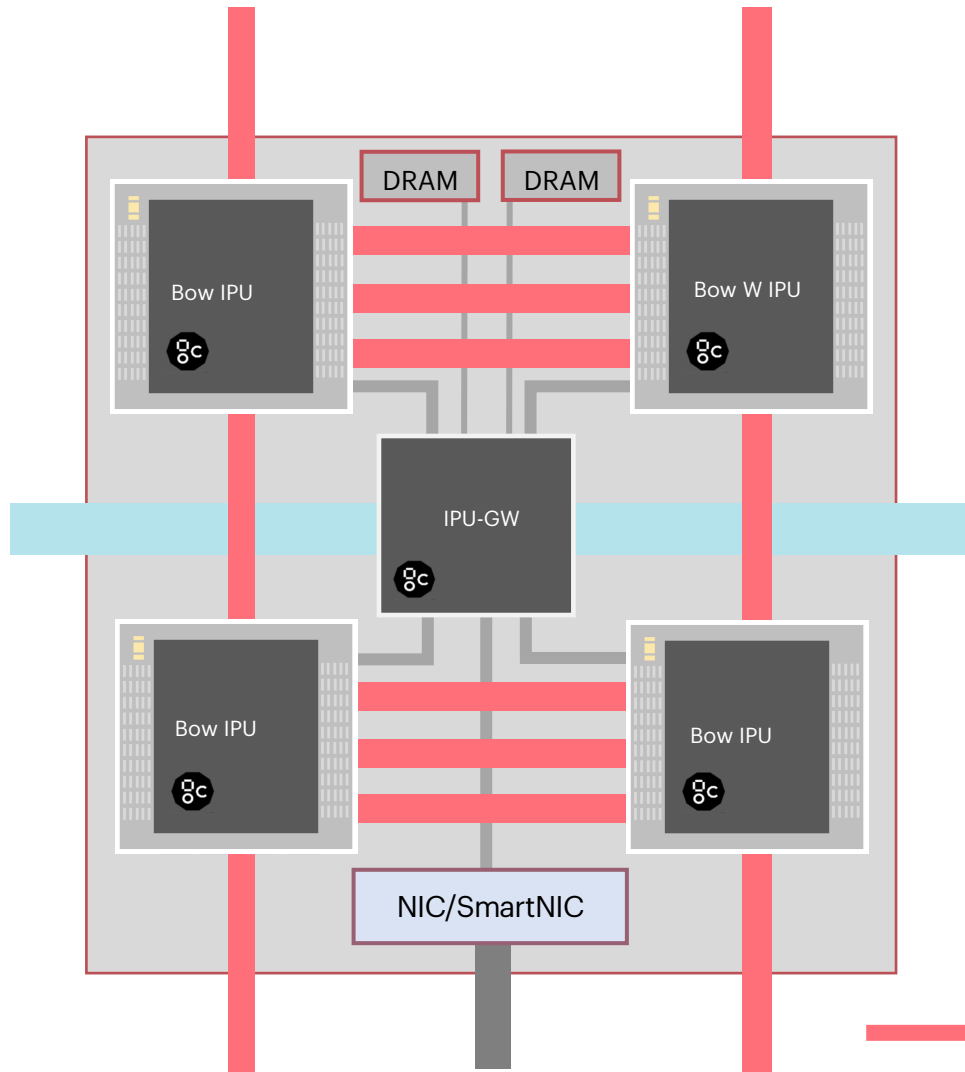
IPU-Links scale to Bow Pod64

Expansion to Bow Pod256 and beyond
with IPU-GW Links

BOW IPU-2000



BOW-2000: THE BUILDING BLOCK OF LARGE PODS



COMPUTE

4x Bow IPUs

- 1.4 PFLOP₁₆ compute
- 5,888 processor cores
- > 35,000 independent parallel threads



DATA

Exchange Memory





- 3.6GB In-Processor-Memory @ 260 TB/s
- 128GB Streaming Memory DRAM (up to 256GB) @ 20 GB/s



COMMUNICATIONS

IPU-Fabric managed by IPU-GW

- Host-Link – 100GE to Poplar Server for standard data center networking
- IPU-Link – 2D Torus for intra-POD64 communication
- GW-Link - 2x 100Gbps Gateway-Links for rack-to-rack – flexible topology

-  x16 IPU-Link [64GB/s]
-  Host-Link Network I/F [100Gbps]
-  IPU-GW Link [100Gbps]
-  x8 PCIe G4 [32GB/s]



BOW: 3RD GENERATION IPU SYSTEMS

SHIPPING TO CUSTOMERS TODAY



BOW POD₁₆

4x Bow-2000
5.6 PetaFLOPS
1 CPU server



BOW POD₆₄

16x Bow-2000
22.4 PetaFLOPS
1-4 CPU server(s)

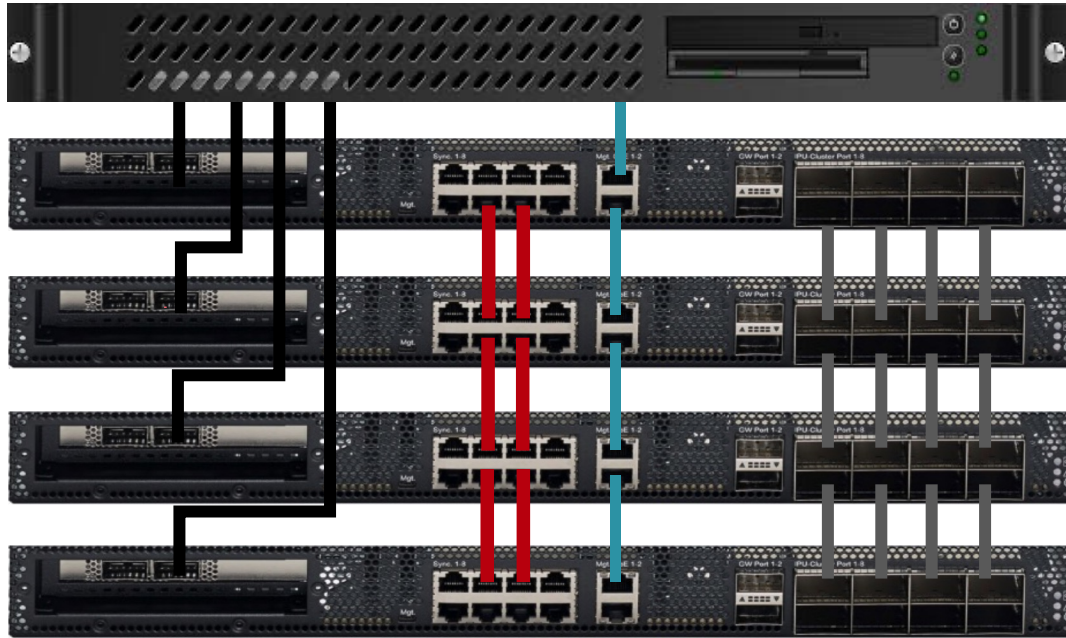


BOW POD₂₅₆

64x Bow-2000
89.6 PetaFLOPS
4-16 CPU server(s)

BOW POD16 DIRECT ATTACH

Server with x4 Bow-2000



- Host-Link 100GE network interface (QSFP, 1.0m)
- 1GbE Management (Cat5, 1.5m)
- Sync-Link (Cat5, 0.15m)
- IPU-Link (OSFP, 0.3m)

- Convenient cost effective evaluation platform
- Available through Graphcore channel for on-premise or Graphcloud
- Wide range of benchmarks and examples for Bow Pod₁₆ performance evaluation
- Scale-out with Bow Pod₆₄ and beyond

BOW POD64 REFERENCE DESIGN

Pre-Qualified 64-IPU Design with Reference Server and Switches

- Up to 16 Bow-2000 platforms
- Reference architecture supports different server requirements based on workload
- **Bow Pod₆₄ Configuration:**
 - 64 IPUs
 - 22.4 PFLOPs @ FP16.16
 - ~58GB IPU In-Processor memory
 - ~7TB Streaming Memory
- **Bow Pod Host disaggregation**
 - Flexibly connect required host server compute over fabric
- **2D-Torus topology**
 - Maximizes bandwidth across IPU-Links
 - All-Reduce 2x faster than mesh topology
- **Scalable to 64K Bow IPUs**

