

Disaggregation for customised compute

Machine intelligence workloads have very diverse compute demands. For production deployment, optimising the ratio of AI to host compute can help maximise performance, while improving total cost of ownership. Bow Pod systems allow flexible mapping of the number of servers and switches to the requisite number of Bow-2000 machines, so deployment is better tailored to production AI workloads. Bow Pod_{v6} supports multiple server configurations.

Communication architecture built for scaling

Efficient data access and transfer can unlock greater AI performance. IPU-Fabric is an innovative communication architecture for system-wide data transfer, extending high-speed interconnect within individual Bow IPUs, across Bow-2000s, between Bow Pods and throughout the data centre. IPU-Fabric delivers high-performance low-latency communication to maximise AI application efficiency and is built to work with standard data centre communication technologies.

Platform for AI developers

TensorFlow, PyTorch, PaddlePaddle, and many other popular ML frameworks are supported and available on open source platforms with the comprehensive PopLibs™ library, for community driven

collaboration and innovation. For developers who want full control

with the comprehensive PopLibs™ library, for community driven