

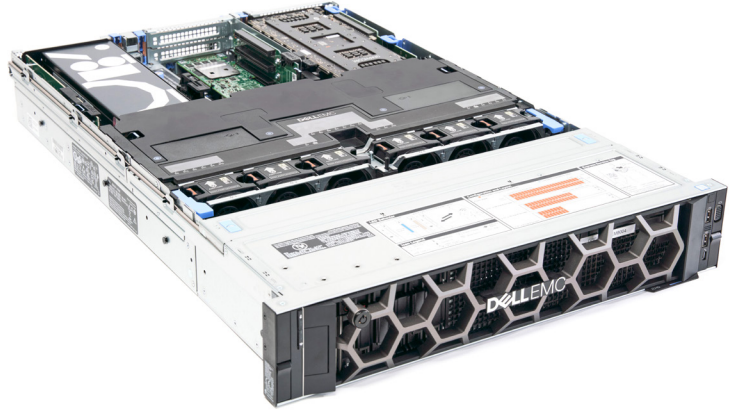
GRAPHCORE 740 IPU SERVER

This Graphcore platform based on the Dell EMC PowerEdge R740 is a high-performance server targeted for machine learning inference applications.

It is a 2U rack-mounted chassis, incorporating two Graphcore 2 IPU-Processor PCIe Cards, each of which runs two of Graphcore's Colossus GC2 IPU (Intelligence Processing Units). The 4x GC2 IPU are served by 2x Intel® Xeon Host CPU with host memory.

The IPU has been optimised to work efficiently on the extremely complex models needed for machine intelligence. It emphasises massively parallel, low-precision floating-point compute, enabling faster and more efficient deployment. With inherent IPU support for both Training and Inference, this Inference focussed platform has the additional benefit of Training support.

The platform is supported by Graphcore's Poplar® software to provide a complete platform for accelerated machine intelligence deployment.



TECHNICAL SPECIFICATIONS

IPU Accelerators	2 Graphcore C2 IPU-Processor PCIe Cards
IPU Processors	4x IPU Colossus GC2 IPU
IPU-Cores	4,864 individual IPU-Cores, with 1.2GB In-Processor Memory™
Compute	400 TeraFLOPS mixed precision
System CPU (typical)	2x Xeon Gold 6242
System Memory (typical)	24x 16GB DDR4 ECC Registered DIMM Modules
OS Storage (typical)	2x 480GB SATA 2.5in Enterprise SSD
System Storage (typical)	1.6TB NVMe Enterprise SSD
Network	10Gbps/1Gbps Ethernet
Software	Ubuntu® Linux; Graphcore Poplar®
Gross weight	TBC
Dimensions	2U rack mount
Max power	1600W (at 220-240V)