## NVIDIA Tesla K40 (Actively Cooled) Product Description

**PNY PN: TCSCK40KIT**


**Tesla K40 Overview**

Equipped with 12GB of memory, the Tesla K40 (Active) GPU accelerator is ideal for the most demanding HPC and big data problem sets. It outperforms CPUs by up to 10x and includes a Tesla GPUBoost feature that enables power headroom to be converted into a user-controlled performance boost. The innovative Kepler compute architecture, CUDA support, and active cooling system makes this board an ideal choice for demanding HPC applications.

| Technical Specifications | |
|---|---|
| **Peak Double-Precision Floating Point Performance** | 1.43 Tflops |
| **Peak Single-Precision Floating Point Performance** | 4.29 Tflops |
| **Number of GPUs** | 1 x GK110B |
| **Number of CUDA Cores** | 2,880 |
| **GPU Memory** | 12GB GDDR5 |
| **Memory Bandwidth (ECC Off)** | 288 Gbytes/sec |
| **Architecture Features** | SMX, Dynamic Parallelism, Hyper-Q |
| **System Type** | Servers and Workstations |

| Features and Benefits | |
|---|---|
| **Memory Error Protection** | Meets a critical requirement for computing accuracy and reliability. External and internal memories are ECC protected. |
| **System Management Features** | Integrates the GPU subsystem with the host system's monitoring and management capabilities such as IPMI or OEM-proprietary tools. IT staff can now manage the GPU processors in the computing system using widely available cluster/grid management solutions. |
| **L1 and L2 Caches** | Accelerates algorithms such as physics solvers, ray-tracing, and sparse matrix multiplication where data addresses are not known beforehand. |
| **Asynchronous Transfer with Dual DMA Engines** | Turbocharges system performance by transferring data over the PCIe bus while the computing cores are crunching other data. |
| **Tesla GPUBoost** | End-user can convert power headroom to higher clocks and achieve even greater acceleration for various HPC workloads. |
| **Flexible Programming and API Support** | Choose OpenACC, CUDA toolkits for C, C++, or Fortran to express application parallelism and take advantage of the innovative Kepler architecture. |