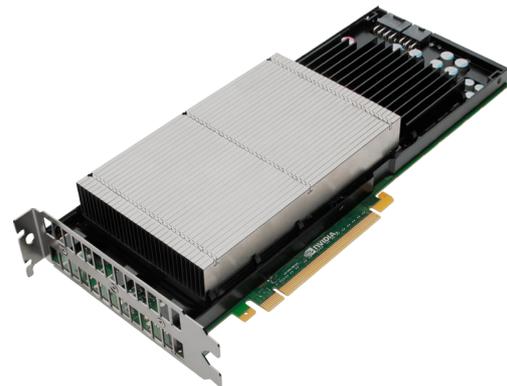




NVIDIA® TESLA® K20M
GPU COMPUTING ACCELERATOR



PART NUMBER:
TCSK20M-PB

TESLA K20M GPU COMPUTING ACCELERATOR

Designed for double precision applications and the broader supercomputing market, the Tesla K20 delivers 3x the double precision performance compared to the previous generation Fermi-based Tesla M2090, in the same power envelope. Tesla K20 features a single GK110 Kepler GPU that includes the Dynamic Parallelism and Hyper-Q features. With more than one teraflop peak double precision performance, the Tesla K20 is ideal for a wide range of high performance computing workloads including climate and weather modeling, CFD, CAE, computational physics, biochemistry simulations, and computational finance.

With teraflops of single and double precision performance, NVIDIA® Kepler GPU Computing Accelerators are the world's fastest and most efficient high performance computing (HPC) companion processors. Based on the Kepler compute architecture, which is 3 times higher performance per watt than the previous "Fermi" compute architecture¹, the Tesla Kepler GPU Computing Accelerators make hybrid computing dramatically easier, and applicable to a broader set of computing applications. NVIDIA Tesla GPUs deliver the best performance and power efficiency for seismic processing, biochemistry simulations, weather and climate modeling, image, video and signal processing, computational finance, computational physics, CAE, CFD, and data analytics.

The innovative design of the Kepler compute architecture includes:

>> SMX (streaming multiprocessor) design that delivers up to 3x more performance per watt compared to the SM in Fermi. It also delivers 1 petaflop of computing in just 10 server racks.

>> Dynamic Parallelism capability that enables GPU threads to automatically spawn new threads. By adapting to the data without going back to the CPU, it greatly simplifies parallel programming and enables GPU acceleration of a broader set of popular algorithms, like adaptive mesh refinement (AMR), fast multipole method (FMM), and multigrid methods.

>> Hyper-Q feature that enables multiple CPU cores to simultaneously utilize the CUDA cores on a single Kepler GPU, dramatically increasing GPU utilization, slashing CPU idle times, and advancing programmability. Ideal for cluster applications that use MPI.

TESLA K20M for Servers - PRODUCT SPECIFICATIONS

CUDA PARALLEL PROCESSING CORES	2496
PEAK DP IN TFLOPS	> 1 TFlop
INTERFACE	320-bit
PROCESSOR CORE CLOCK	706 MHz
FRAME BUFFER MEMORY	5 GB GDDR5
MEMORY BANDWIDTH	208 GB/s
DISPLAY CONNECTORS	None
MAX POWER CONSUMPTION	225 W
POWER CONNECTORS	(2) x 6-pin PCI Express power connectors
GRAPHICS BUS	PCI Express 2.0 x16
FORM FACTOR	110 mm (H) x 265 mm (L) Dual Slot
THERMAL SOLUTION	Active



1 - Based on DGEMM performance: Tesla M2090 (Fermi) = 410 gigaflops, Tesla K20 (expected) > 1000 gigaflops

Tesla GPU Computing Accelerator Common Features

<p>ECC MEMORY ERROR PROTECTION</p>	<p>Meets a critical requirement for computing accuracy and reliability in datacenters and supercomputing centers. External Both external and internal memories are ECC protected in Tesla K20.</p>
<p>SYSTEM MONITORING FEATURES</p>	<p>Integrates the GPU subsystem with the host system's monitoring and management capabilities such as IPMI or OEM-proprietary tools. IT staff can thus manage the GPU processors in the computing system using widely used cluster/grid management solutions.</p>
<p>L1 AND L2 CACHES</p>	<p>Accelerates algorithms such as physics solvers, ray-tracing, and sparse matrix multiplication where data addresses are not known beforehand</p>
<p>ASYNCHRONOUS TRANSFER WITH DUAL DMA ENGINES</p>	<p>Turbocharges system performance by transferring data over the PCIe bus while the computing cores are crunching other data.</p>
<p>FLEXIBLE PROGRAMMING ENVIRONMENT WITH BROAD SUPPORT OF PROGRAMMING LANGUAGES AND APIS</p>	<p>Choose OpenACC, CUDA toolkits for C, C++, or Fortran to express application parallelism and take advantage of the innovative Kepler architecture.</p>

Software and Drivers

>> **Software applications page:**
http://www.nvidia.com/object/vertical_solutions.html

>> Tesla GPU computing accelerators are supported for both Linux and Windows.

Server modules are only supported on 64-bit OSes and workstation / desktop modules are supported for 32-bit as well.

>> Drivers:

Latest drivers can be downloaded from
<http://www.nvidia.com/drivers>

- Learn more about Tesla data center management tools at
<http://www.nvidia.com/object/softwarefor-tesla-products.html>

>> Software development tools are available at
<http://developer.nvidia.com/getting-started-parallelcomputing>

