IDC LINK
Real-time IDC Research® opinion on industry news, trends and events

SYSTAP's Blazegraph Targets New Levels of Price/Performance for Graph Analytics with GPUs

December 18, 2015

By: Robert Sorensen

IDC's Quick Take

SYSTAP has announced that it has developed a version of its Blazegraph high-performance graph database package that uses NVIDIA Tesla GPUs capable of 32 GTEPS (billions of traversed edges per second) that targets performance improvements of 200-300 times over existing CPU-based graph database implementations with no required code changes. SYSTAP reports that this new Blazegraph GPU offers a 40 times price/performance advantage over counterpart CPU-based graph implementations for a wide range of analytics applications including real-time fraud detection, analytics for social and communication networks, and drug discovery and testing, and SYSTAP also states that this product is being eagerly anticipated by a number of organizations, including the British Museum that hopes to boost the power of its ResearchSpace program, a collaborative environment for humanities and cultural heritage research.

News Highlights

SYSTAP, which was founded in 2006, has long offered traditional CPU-based versions of its Blazegraph product for both server CPUs and embedded processors that are currently in production use across the private and public sector for customers such as EMC, Autodesk, DARPA, and Yahoo! With this new development, SYSTAP is looking to tackle some of the vexing memory issues that graph analytics present to traditional cache-based CPU architectures by tapping into the high core counts and high memory bandwidth potential of GPUs to realize significant gains in performance on graph applications:

- This newest GPU-based version of Blazegraph, expected for delivery in 1Q16, traces its origins back to early funding to SYSTAP by DARPA that included collaboration with the University of Utah SCI Institute for running GPU-based parallel graph algorithms on the Titan supercomputer installed at Oak Ridge National Laboratory.

Drawing on those research efforts, Blazegraph GPU, running on a cluster of 64 NVIDIA K40 GPUs, reportedly demonstrated a throughput of 32 billion traversed edges per second, traversing a scale-free graph of 4.3 billion directed edges in 0.15 seconds. In addition:

- GTEPS is a standard industry benchmark used to produce the Graph 500 list, a rating of computer systems focused on data-intensive workloads that stresses the communication subsystem of the system, instead of counting the double-precision floating point and is based on a breadth-first search in a large undirected graph, a fundamental process in graph analytics.
- The Blazegraph GPU system, if listed, would have placed at about number 72 on the Graph 500, well below the largest systems, such as the number 1 K computer at RIKEN's Advanced Institute of Computational Science in Japan, which with a 38,621 GTEPS rating is over 1,200 times more

powerful than the Blazegraph GPU system but at a significantly higher price/performance ratio on this test.

Indeed, according to SYSTAP, an Accumulo graph database implemented on Hadoop costs about $18 million per GTEP, while it costs only about $16,000 per GTEP for a Blazegraph GPU cluster of NVIDIA Kepler K40s. And, the firm claims that, next year, when NVIDIA ships its next-generation Pascal devices, the price per GTEP could drop to $4,000.

IDC's Point of View

Blazegraph's long successful history in graph analytics for the public and private sector combined with SYSTAP's notable record of collaborative R&D in GPU-based graph development bodes well for the Blazegraph GPU. In addition, SYSTAP is one of first to offer GPU-based systems to the graph analytics sector and the firm has the experience to help existing customers move their workloads to this new computational platform with a minimum of software rewrites. Finally, by offering systems with a reported impressive low cost per GTEPS but with the computational heft to support interesting and innovative graph analytics, the firm could attract a significant number of new users that have compelling graph problems but not the budget required to justify buying the kinds of HPCs that populate the high end of the Graph 500 list.

Subscriptions Covered:
Technical Computing