Scientel achieves 6 Trillion Parameter LLM run on Ohio State OSC Supercomputer

6 Trillion Parameter run achieved with DeepSeek R1 671B model on 36 Nvidia H100 GPUs

ANN ARBOR, MI, UNITED STATES, March 3, 2026 /EINPresswire.com/ -- Scientel's LLM systems are designed to operate in conjunction with its own Gensonix NewSQL AI DB. These mid-to-large scale LLM systems are more

Scientel - Gensonix AI DB

efficient than most alternatives due to the support of Relational, Document, Text and Vector data in distinctly different stores within a single AI DB. This storage of all data types in its own native form allows for the very high efficiency. Gensonix AI DB as a single data repository handles the various data storage tasks that are required by smart AI applications without the need to convert to other formats such as XML or JSON for storage. This should further enhance storage efficiency.

> We are pleased to offer a Trillion Parameter LLM solution for serious AI applications with the power of Gensonix, Nvidia H100 and DeepSeek R1 for entities aiming to host large data set AI systems"
>
> *Norman Kutemperor, CEO*

The DeepSeek R1 is a 671Billion parameter model which was loaded utilizing 4 H100 GPUs. Scientel utilized 36 H100 GPUs in a parallel mode to reach a total of 6.039 Trillion Parameters. This capability allows users to host much larger large language model (LLM) private systems whether in the cloud or on site. Data privacy and security are important factors, especially for large sets of important private data. Therefore, entities should take an interest in hosting their own private data centers rather than riskier public systems. Gensonix Large scale LLMs may be an ideal solution for this problem. DeepSeek R1 should be considered as a significant breakthrough in AI, offering high-level reasoning capabilities without having to depend on large energy-intensive data centers. With the DeepSeek model, Scientel's Gensonix AI becomes a serious challenger to current large LLM systems.

The NVIDIA H100 is a top-tier data center GPU that is based on the Hopper architecture. It is specifically designed to accelerate AI training, inference, and high-performance computing (HPC). It features 80 billion transistors, offering up to 9x faster AI training and 30x faster inference for

(LLMs) than its competitors. The H100 is equipped with up to 96 GB of high-bandwidth HBM3 memory. It is Ideal for training, deep learning, data analytics, and scientific simulations.

As LLMs become more popular for both individuals and businesses, the data centers that host these systems will likely become more important. However, they are facing major power constraints such as additional need for air conditioning, due to increased demand. Therefore, greater attention should be given to smaller and more efficient systems that can run on local networks or in the cloud. Scientel's LLM solutions are designed to address customer specific databases that may be queried locally rather than via public access.

About Scientel

Scientel, is a U.S.-based, Systems Technology Company. It designs/produces highly optimized high-end servers, bundled with its "GENSONIX™ AI" DB software, as a single-source supplier of complete systems for AI solutions. It also customizes hardware and software for specific applications resulting in higher performance. It's Elastic Scaling Servers can support up to 1000's of compute nodes in a single cluster to meet customer required speeds for virtually any data size.

Scientel's specialty is advanced NewSQL DBMS design and applications/systems integration for advanced system processes. This includes applications in AI, Big Data, and Commercial Intranets combined with IT consulting and support, for "beyond mainframe-level" Large Data Warehouse Appliances.

GENSONIX allows for very user-friendly data manipulation capabilities that are typically found in standard, SQL-based, database management systems. However, it goes beyond this, with its ultra-procedural SDBL™ development system. It is truly an "ALL-in-One SQL" -- an "All Data Management System" in the form of an ultra-flexible, NewSQL DBMS of general capabilities and application potentials. It can also function in concert with SQL systems to efficiently handle both structured and unstructured data that are stored in different data stores within a single DB as a large data warehouse repository. However, it can also handle heavy database loads by itself with the aid of the SDBL query/procedural language, involving up to trillions of rows/transactions for billions of customers. This is a huge advantage in "truly Big Data/AI" structured applications.

Business customers can take advantage of Scientel's capabilities in advanced Business Intelligence and Data Analytics to grow their business by handling AI/Big Data more cost-effectively and with greater insights so to remain competitive when compared to their peers. Scientific, government, and similar organizations can also utilize these capabilities to efficiently process AI/Big Data.

Norman Kutemperor
Scientel IT Corp
+1 248-433-4700

norm@scientel.com
Visit us on social media:
X

This press release can be viewed online at: https://www.einpresswire.com/article/896912803