

## **Tokyo Institute of Technology Submission to the CUDA CCoE Achievement Award Submission 2014 – TSUBAME-KFC**

Principal Investigator: Satoshi Matsuoka (GSIC Center, Tokyo Tech.)

**TSUBAME-KFC, a new prototype supercomputer for future power efficient supercomputers leading to exascale, designed and built by GSIC, Tokyo Institute of Technology along with partners such as NVIDIA, was ranked as No.1 supercomputer on the November 2013 edition of the Green 500 list as well as the Green Graph 500 list. This marks the first time that a single supercomputer was crowned #1 in both compute as well as data intensive applications.**

TSUBAME-KFC (Kepler Fluid Cooling), a prototype supercomputer developed by GSIC (Global Scientific Information and Computing Center, Tokyo Institute of Technology) along with partner vendors NVIDIA Corporation, NEC Cooperation and Green Revolution Cooling, as well as others, has been recognized as the greenest supercomputer in the world in both compute- and data intensive rankings for the first time. TSUBAME-KFC achieved power-efficiency of 4,503.17 MFLOPS per Watt and was ranked as No.1 on the Nov. 2013 edition of the Green 500 List announced on 20<sup>th</sup> of November at the Supercomputing 2013 conference, held in Denver, CO, USA. This bests the second ranked machine, which is also NVIDIA-driven, by 24% margin, and improves over the June 2013 #1 machine power-performance by more than 40%.

TSUBAME-KFC was also ranked as No.1 on the Green Graph 500 list “Big Data” category with 6.72 MTEPS (Mega Traversed Edges Per Second) per Watt, besting the previous record held by the JUQUEEN BlueGene/Q, by 24% margin.

This is the first time a supercomputer achieved the top spot on both lists simultaneously, and proved the value of the technological collaboration between the academia and the industry, in particular the CUDA CCoE program. Since power efficiency is the most important factor in system performance, this result is a sound demonstration that GPU technology is the forerunner for IT societies of tomorrow, not just limited to supercomputing.

By all means, the #1 statue of TSUBAME-KFC was achieved not merely with the use of GPU alone. In fact, the top 10 systems on the Green 500 list is occupied by GPU systems. TSUBAME-KFC was developed as a prototype system for TSUBAME 3.0 and onwards, supported by the “Ultra-Green Project” (2010-2015) under the Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT), with research results also applied from the precursor JST-CREST “Ultra Low Power HPC” project conducted during 2007-2012. Both projects aim at investigating every aspect of modern supercomputers so that power consumption would be minimized, including processors (CPU/GPU), memory, interconnect, etc., as well as to devise novel power conscious algorithms and software systems, with auto-tuning capabilities. Also, novel cooling and energy recovery methods are another point-of-focus especially for the former project, experimenting in efficient methods such as oil-submersion cooling with extreme architectural density, chiller-less cooling towers, as well as intricate monitoring and control of power with cooling loop in mind. Along with advances in process technology and GPU (many-core) architectural evolutions, the goal is to achieve x1000 power performance improvement in 10 years, or about x2

every year on the average. Compared do TSUBAME1.0 in 2006, which achieved approximately 40 Megaflops/W was ranked seventh in the world on the Top500, we have achieved over 100 times improvement in 7 years, largely meeting this goal.

TSUBAME-KFC consists of 40 compute nodes in a single oil-submerged rack, interconnected FDR InfiniBand network. Each compute node has 2 low power Intel Xeon E5-2620 v2 processors (Ivy Bridge EP) and 4 NVIDIA Tesla K20X GPUs, squeezed into a 1U chassis. The total number of GPUs in a single rack is 160, or if one includes the CPUs, the number of processors are 240 in total. The rack is cooled using oil submersion technology developed by Green Revolution Cooling, with appropriate oil selection made in collaboration with us to meet the stringent fire regulations of Japan, as well as the nodes being customized for oil submersion. There are numerous power and temperature sensors embedded in the system that allows non-intrusive measurement at very fast data rates of 1 second per measurement or less and displayed and archived in real time. The rack is placed in a small freight container housing and placed outside with minimum connection to the outside except for electricity and the network. The secondary water loop is connected to a evaporative cooling tower that is uses substantially little power compared to chillers, allowing KFC to achieve 1.07 effective PUE as well as plans for energy recovery with warm-water (and oil) cooling.

The theoretical peak performance is 210 Teraflops in double precision, and over 600 Teraflops in single, almost achieving petaflop per rack, despite that the rack itself is not at all oversized grossly as is with competing water-cooling machines such as BG/Q.

For the Green 500 benchmark, we employed the in-core Linpack routine supplied by NVIDIA, but conducted extensive parameter space temperature in software and machine setting as well as cooling, and even network routing using our recently developed new routing algorithm, to achieve the optimal power efficiency. In fact, the best Top 500 performance, tuned for maximum performance, turned out to be less than 80% power efficient compared to the best power-tuned configuration. For the Green Graph 500, we employed our in-house developed efficient graph BSP code that maximally utilizes the GPU memory bandwidth, and again tuned for power efficiency. We are currently preparing to publish each result at top-notch conferences.

TSUBAME-KFC received immediate attention by the media worldwide. It made national TV news on the Japanese NHK, as well as was covered by all the national newspapers in Japan. It was also covered by a recent New York Times article, not to mention every on-line new media specializing in IT.

[http://www.nytimes.com/2014/02/12/business/international/improving-energy-efficiency-in-supercomputers.html?\\_r=0](http://www.nytimes.com/2014/02/12/business/international/improving-energy-efficiency-in-supercomputers.html?_r=0)

We hope to apply our result to production TSUBAME3.0, to be commissioned early 2016. TSUBAME3.0 is expected to achieve over 10 GFlops/W as well as extremely high score on the Green Graph500.

## TSUBAME-KFC: Ultra-Green Supercomputer Testbed and TSUBAME3.0 Prototype [2011-2015]

**Fluid Submersion Cooling + Outdoor Air Cooling + High Density GPU Supercomputing in a 20-foot container**



- Target
- World's top power efficiency (>4 GFlops/Watt)
  - Average PUE 1.05, lower component power
  - Field test ULP-HPC results

## TSUBAME-KFC system



A TSUBAME3.0 prototype system with advanced cooling for next-gen. supercomputers.

40 compute nodes are oil-submerged.  
160 NVIDIA Kepler K20x  
80 Ivy Bridge Xeon  
FDR Infiniband

Peak Performance

Single Node	5.26 TFLOPS
System (40 nodes)	210.61 TFLOPS

## Nov. 2013 Green500 Ranking Announced at SC13 (Denver, CO)

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	4,503.17	OSGC Center, Tokyo Institute of Technology	TSUBAME-KFC - LX 1U-4GPU/16Rr-1G Cluster, Intel Xeon ES-2620v2 4C 2.10GHz, InfiniBand FDR, NVIDIA K20x	27.78
2	3,631.86	Cambr...		
3	3,517.84	Center Univer...		
4	3,185.91	Seres Centri...		
5	3,130.95	ROME Aidera...		
6	3,068.71	OSGC Center, Tokyo Institute of Technology		
7	2,702.16	Univer...		
8	2,625.10	Max P...		
9	2,629.10	Financ...		
10	2,358.69	CSIRO		



## 2013/11 Green Graph500 Ranking

- **TEPS (Traversal Edges Per Watt)**
- <http://green.graph500.org>

In the **Big Data** category:

Rank	MTEPS/W	Site	Machine	G500 rank	Scale	GTEPS	Nodes
1	6.72	Tokyo Institute of Technology	TSUBAME KFC	47	32	44.01	32
2							6384
3							2768
4							1
5							5536
6							1
7							64



**TSUBAME-KFC #1 For Big Data!**