STAC Update for STAC-A2

Peter Lankford
Founder and Director, STAC

peter.lankford@STACresearch.com

STAC-A2

- Non-trivial Monte Carlo
  - Heston-based Greeks for multi-asset, path-dependent options with early exercise
  - Metrics: Speed, capacity, quality, efficiency
- Numerous reports
  - Some public, some in the STAC Vault
- Premium STAC members get:
  - Reports in STAC Vault
  - Detailed config info on public and private reports
  - Code from vendor implementations of the benchmarks

www.STACresearch.com/**a2**

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Intel Skylake

- SUT ID: INTC170920
- SUT:
  - STAC-A2 Pack for Intel Composer (Rev K)
  - 2 x 28-core Intel Xeon Platinum 8180 @ 2.5 GHz
  - White box (Purley SDP)
  - RHEL 7.4

www.STACresearch.com/**INTC170920**

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Highlights

- Compared to a similar 2-socket system using Broadwell EP (SUT ID INTC160314), performance was:
  - 2.89x in warm runs of the baseline GREEKS benchmark (STAC-A2.β2.GREEKS.TIME.WARM)
  - 4.62x in warm runs of the large problem size (STAC-A2.β2.GREEKS.10-100k-1260.TIME.WARM)
  - 27% higher in max assets (STAC-A2.β2.GREEKS.MAX_ASSETS)

8 x NVIDIA Volta (V100) in HPE ProLiant XL270d Gen9 Server

- SUT ID: NVDA171020
- SUT:
  - STAC-A2 Pack for CUDA (Rev D)
  - NVIDIA CUDA 9.0
  - 8 x NVIDIA Volta V100 GPUs
  - 2 x E5-2680 v4 processors at 2.4GHz
  - HPE ProLiant XL270d Gen9
  - 512 GB DRAM
  - RHEL 7.4
- STAC Report released yesterday

www.STACresearch.com/**NVDA171020**

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Highlights

- Set new records in all performance benchmarks and energy efficiency
- Compared to all publicly reported results on non-NVIDIA based architectures:
  - 8.9x the next best throughput (STAC-A2.β2.HPORTFOLIO.SPEED)
  - 6.2x the next best time in warm runs of the baseline Greeks benchmark (STAC-A2.β2.GREEKS.TIME.WARM)
  - 2.7x the next best energy efficiency (STAC-A2.β2.HPORTFOLIO.ENERG_EFF)
  - 1.9x the maximum basket size (STAC-A2.β2.GREEKS.MAX_ASSETS)
  - 1.5x the next best space efficiency (STAC-A2.β2.HPORTFOLIO.SPACE_EFF)
- Compared to the best performing solution to date using 4 previous-generation NVIDIA Tesla P100 GPUs (SUT ID NVDA170718):
  - 2.7x in throughput (STAC-A2.β2.HPORTFOLIO.SPEED)
  - 2.4x in warm runs of the baseline Greeks benchmark (STAC-A2.β2.GREEKS.TIME.WARM)
  - 1.5x in the maximum basket size (STAC-A2.β2.GREEKS.MAX_ASSETS)
  - 1.2x in energy efficiency (STAC-A2.β2.HPORTFOLIO.ENERG_EFF)

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER

Google Cloud Platform – Skylake vs Broadwell

- Google is the only cloud provider currently offering Skylake
- SUT ID INTC171002
  - 64 vCPU Broadwell + 240GB memory
- SUT ID: INTC171003
  - 92 vCPU Skylake + 360GB memory
- Compared to Broadwell instance, Skylake instance was:
  - 87% faster in the portfolio test (STAC-A2.ß2.HPORTFOLIO.SPEED)
  - 84% faster in warm runs of the baseline Greeks benchmark (STAC-A2.ß2.GREEKS.TIME)
  - 95% faster in warm runs of the large problem size (STAC-A2.ß2.GREEKS.10-100k-1260.TIME)
- Report coming shortly

Kubernetes / 8 x NVIDIA Volta (V100) / HPE ProLiant

- SUT ID: NVDA171018
- SUT:
  - Same as NVDA171020, plus
  - Docker 1.2.16
  - Kubernetes 1.8 with device plugin and accelerator gate (alpha state)
- Red Hat submitted unaudited report to the STAC Vault today

www.STACresearch.com/**NVDA171018**

STAC®
SECURITIES TECHNOLOGY ANALYSIS CENTER