HPC/HTC vs. Cloud Benchmarking

An empirical evaluation of the performance and cost metrics

Kashif Iqbal - PhD
Kashif.iqbal@ichec.ie
ICHEC, NUI Galway, Ireland

With acknowledgment to Michele Michelotto – INFN for their kind support for the HTC Benchmarking + Resources

ICHEC - in a nutshell

- **I**rish **C**entre for **H**igh-**E**nd **C**omputing
  - National Tier-1 Centre
  - Run Irish National HPC service for Academics
  - PRACE partner
- Interest in understanding the competitive costs
  - Understanding various infrastructures & workloads
    - HPC, HTC, HPC Cloud, HTC Cloud
- What is the most effective means to address our customers (Academics) needs?

Outline

- Benchmarking - Why, which benchmark?
- HPC and HTC Benchmarking
  - Benchmarks (NPB, HEPSPEC06)
  - Environment Setup
  - Results
- Next Steps

MOTIVATION

If there is a better reason to paddle, I don't know what it is.

Overview

- Diversity
  - Diverse computing infrastructures (HPC. HTC, Cloud)
  - Diverse workloads for various academic communities
- Cost analysis and performance metrics
  - Performance and configuration overhead as indirect costs
- System benchmarking for:
  - Comparison of HPC and HTC systems vs. Cloud offerings
  - Comparison of parallelism techniques (e.g. MPI/OMP)

HPC/HTC Benchmarks

- LINPACK – Top 500
- SPEC06 – CPU intensive benchmark
  - HEP-SPEC06
- HPC Challenge (HPCC)
- Graph 500
- STREAM – for memory bandwidth
- MPPtest – MPI performance
- NAS Parallel Benchmark (NPB)
- …

NAS Parallel Benchmark

- Open-source and free CFD benchmark
- Performance evaluation of commonly used parallelism techniques
  - Serial, MPI, OpenMP, OpenMP+MPI, Java, HPF
- Customisable for different problem sizes
  - Classes S: small for quick tests
  - Class W: workstation size
  - Classes A, B, C: standard test problems
  - Classes D, E, F: large test problems

NPB Kernels

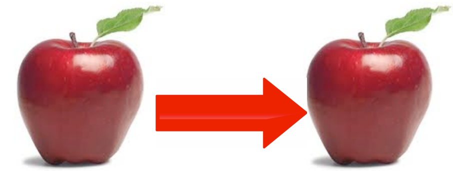
Kernel	Description	Problem Size	Memory (MW)
EP	Monte Carlo kernel to compute the solution of an integral – Embarrassingly parallel	2^{30}	18
MG	Multi-grid kernel to compute the solution of the 3D Poisson equation	256^3	59
CG	Kernel to compute the smallest eigenvalue of a symmetric positive definite matrix	75000	97
FT	Kernel to solve a 3D partial difference equation using an FFT based method	512x256x256	162
IS	Parallel sort kernel based on bucket sort	2^{25}	114
LU	Computational Fluid Dynamics (CFD) application using symmetric successive over relaxation	102^3	122
SP	CFD application using the Beam-Warming approximate factorisation method	102^3	22
BT	CFD application using an implicit solution method	102^3	96

Cloud Cluster Setup

- EC2 instance management
  - StarCluster Toolkit
    - http://web.mit.edu/star/cluster/
    - StarCluster AMIs – Amazon Machine Image
      - Resource manager plugin
- Login vs. compute instances
  - EC2 small instance as login node
  - File system shared via NFS across nodes

e-FISCAL

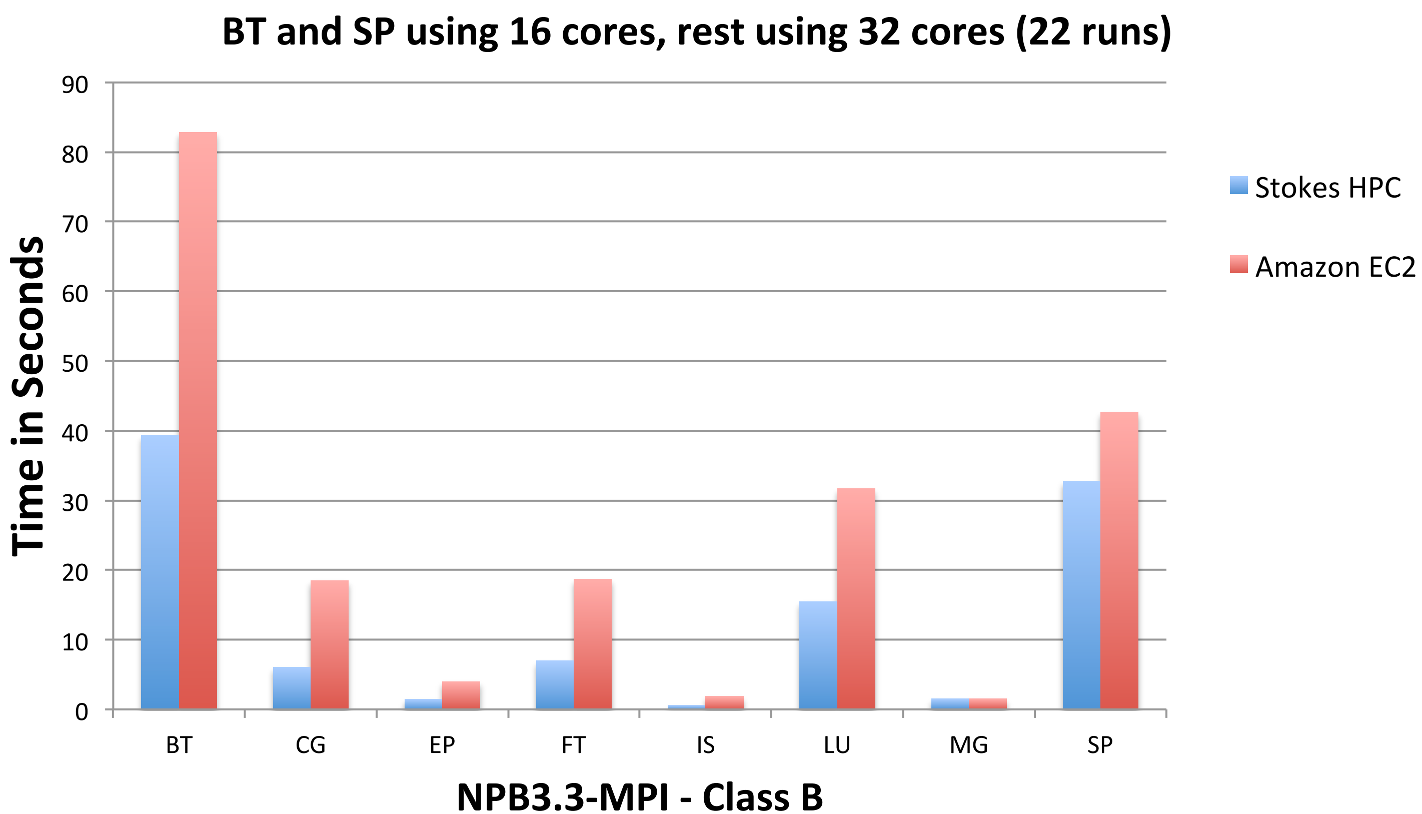
Cloud vs. HPC



	Amazon EC2	Stokes HPC
Compute Node	23 GB of memory, 2 x Intel Xeon X5570, quad-core “Nehalem” (8 cores X 4 Nodes)	24 GB memory, 2 x Intel Xeon E5650, hex-core “Westmere” (12 cores X 3 Nodes)
Connectivity	10 Gigabit Ethernet	ConnectX Infiniband (DDR)
OS	Ubuntu, 64-bit platform	Open-SUSE, 64-bit platform
Resource manager	Sun Grid Engine	Torque
Compilers & libraries	Intel C, Intel Fortran, Intel MKL, Intel MVAPICH2	Intel C, Intel Fortran, Intel MKL, Intel MVAPICH2

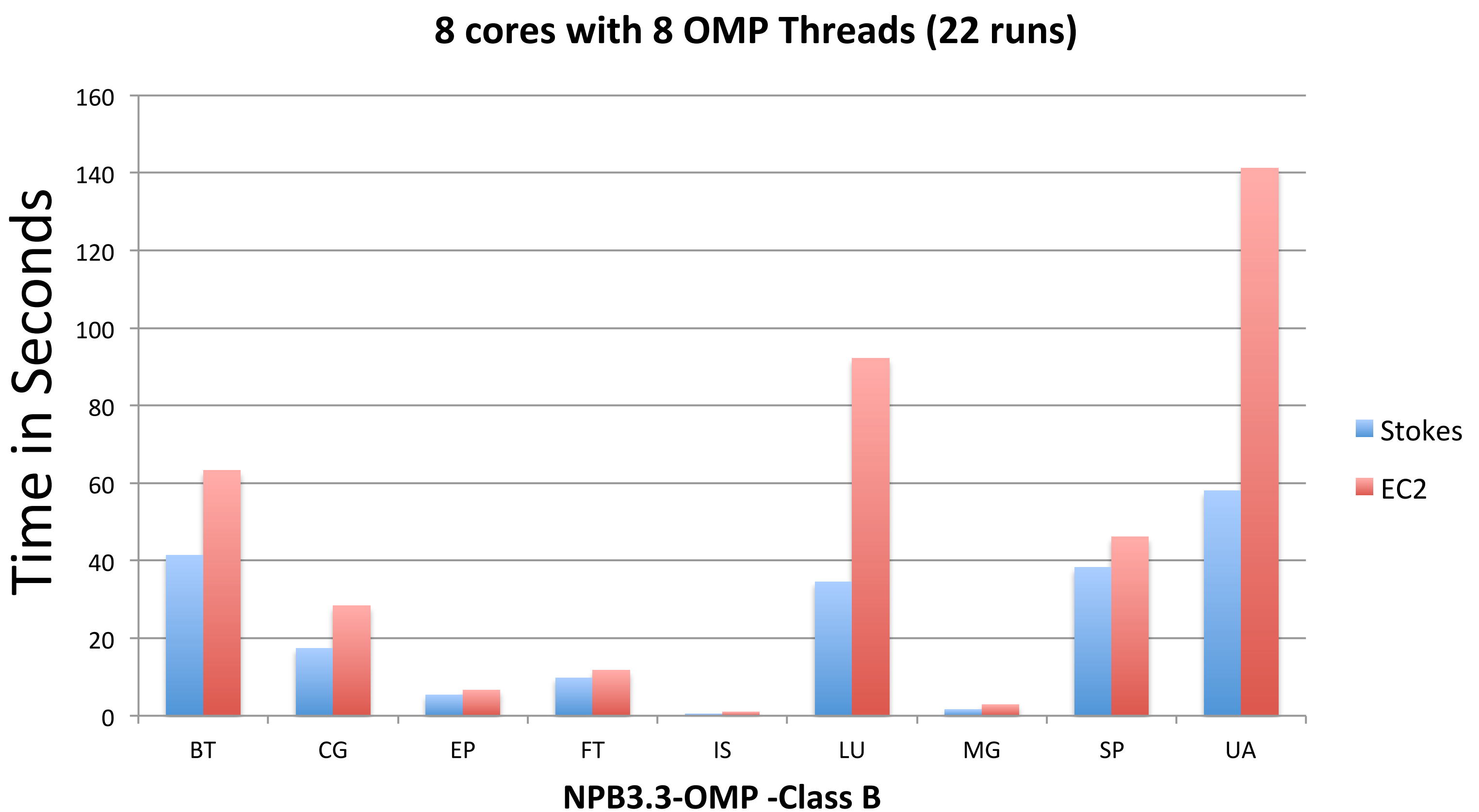
- Non-trivial to replicate runtime environments
- Large variations in performance possible
- Logical vs. Physical cores
  - HT/SMT – Hyper or Simultaneous Multi-Threading (i.e. 2 X Physical Cores)

NPB – MPI



The average performance loss ~ **48.42%** (ranging from 1.02% to 67.76%).

NPB - OpenMP



The average performance loss ~ **37.26**%
(ranging from 16.18 - 58.93%

Cost

- 720 hours @ 99.29 USD ☺
  - ~100 % utilisation
  - Compute cluster instance @ $1.300 per Hour
  - Small instance @ $0.080 per Hour
- Other useful insights:
  - Spot instances
  - Overheads (performance, I/O, setup)
  - Data transfer costs and time

HEPSPEC Benchmark

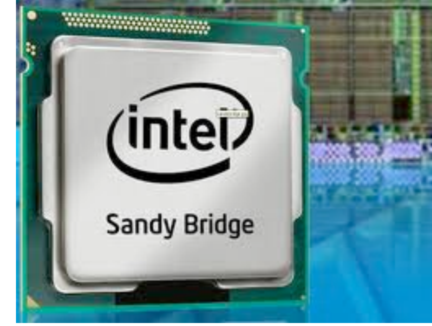
- HEP Benchmark to measure CPU performance
  – Based on all_cpp bset of SPEC CPU2006
  – Fair distribution of SPECint and SPECfp
  – Real workload

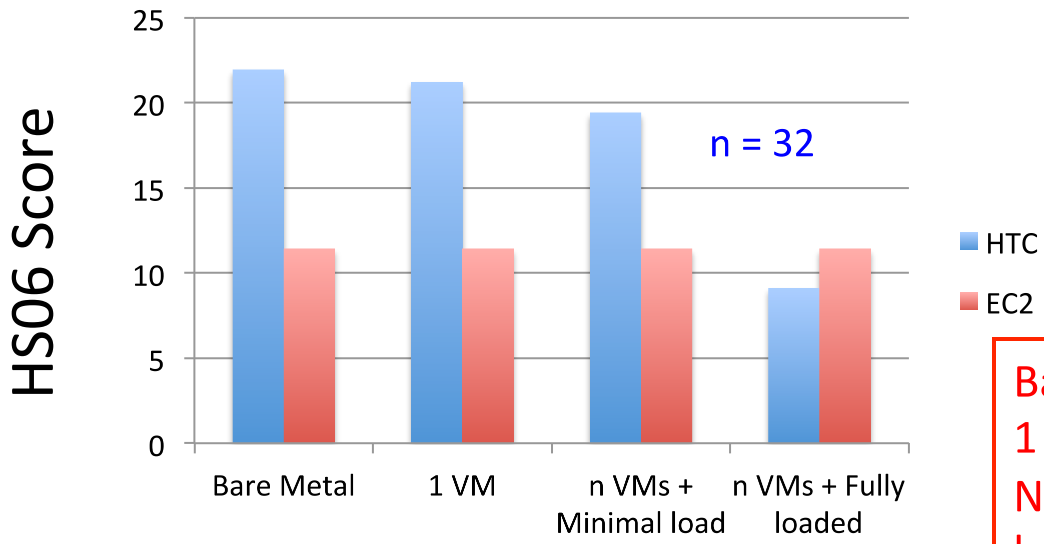
- 32-bit binaries
  – Can be compiled using 64-bit mode ~ for better results

Benchmark Environment

	Amazon EC2	HTC resource at INFN
Compute Nodes	Medium: 2 ECU Large: 4 ECU Xlarge: 8 ECU **1 ECU = 1.0-1.2 GHz**	**Intel(R) Xeon(R)** CPU E5-2660 @ 2.2 GHz, 2 X 8 cores **AMD Opteron** 6272 (aka Interlagos) @ 2.1 GHz, 2 X 16 cores **M instance Single-core VM** **L instance Dual-core VM** **XL Instance Quad-core VM**
OS	SL6.2, 64-bit platform	SL6.2, 64-bit platform
Memory	3.75 GB, 7.5 GB and 15 GB	64 GB for both Intel and AMD
Hyper-Threading	Enabled	Enabled (for Intel) ~ 32 logical cores
Compilers	GCC	GCC

HS06 for Medium

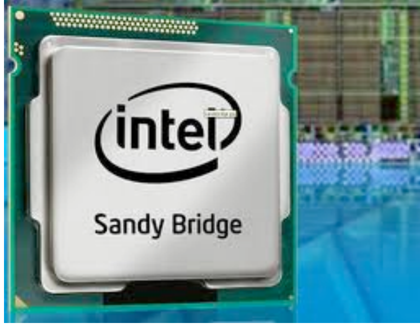




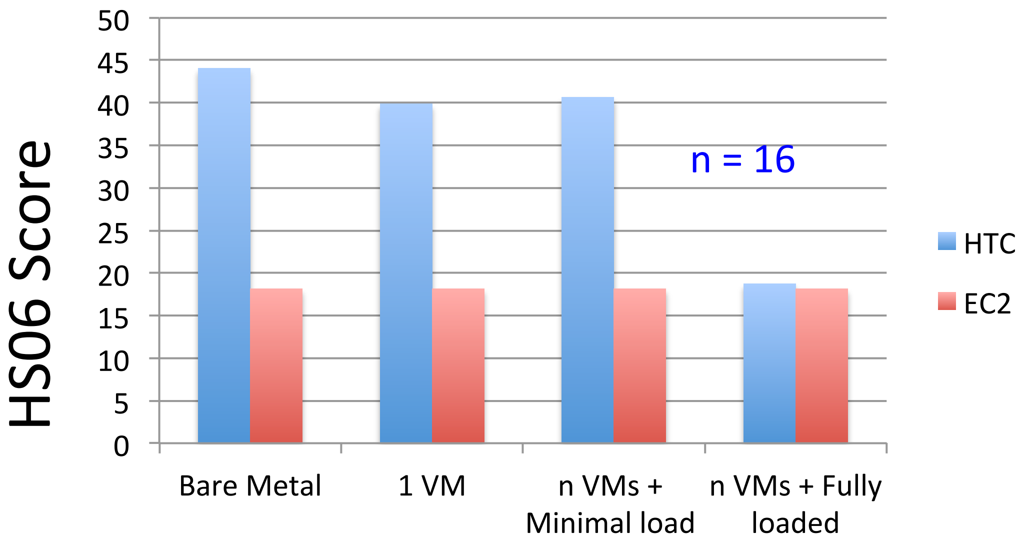
Bare Metal – no virtualisation
1 VM + idle
N VMs + minimal load
N VMs + Fully loaded

Bare Metal – NO!!!!!
1 VM + idle – UNLIKELY!
N VMs + minimal load – Possible but Unknown!
N VMs + Fully loaded – Possible but Unknown!

- SPEC score < with the > no. of VMs
- Virtualisation + Multi-Tenancy (MT) effect on performance ~ 3.28% to 58.48%
- More realistic figure ~ 11.53 to 58.48



HS06 for Large

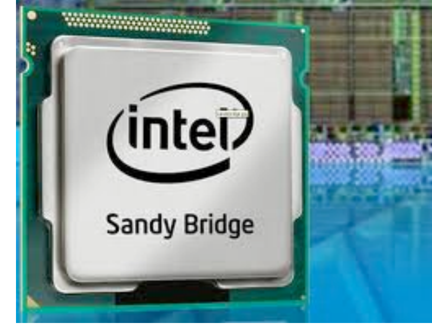


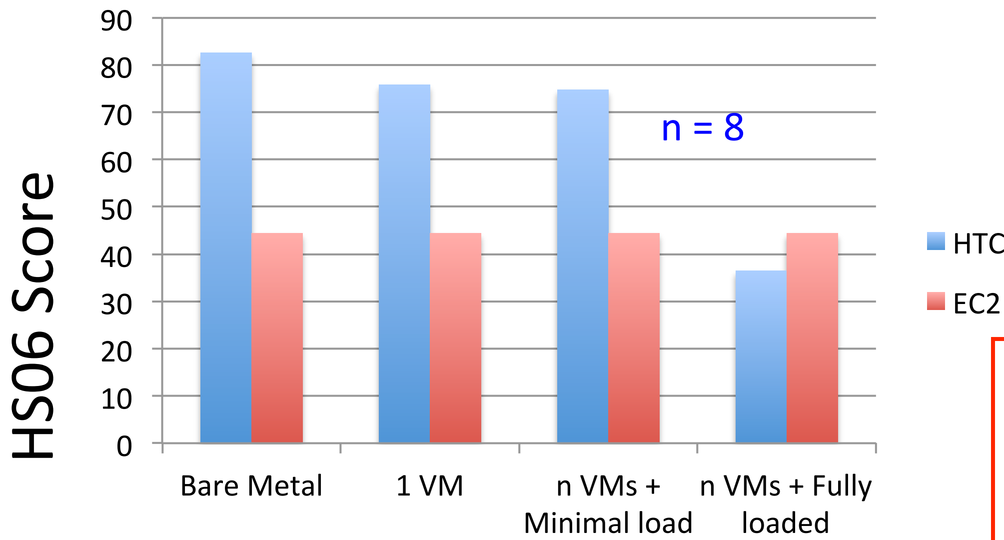
Bare Metal – no virtualisation
1 VM + idle
N VMs + minimal load
N VMs + Fully loaded

Bare Metal – NO!!!!!
1 VM + idle – UNLIKELY!
N VMs + minimal load – Possible but Unknown!
N VMs + Fully loaded – Possible but Unknown!

- Virtualisation + MT effect on performance ~ 9.49% to 57.47%
- Note the minimal effect of > no. of VMs

HS06 for Xlarge



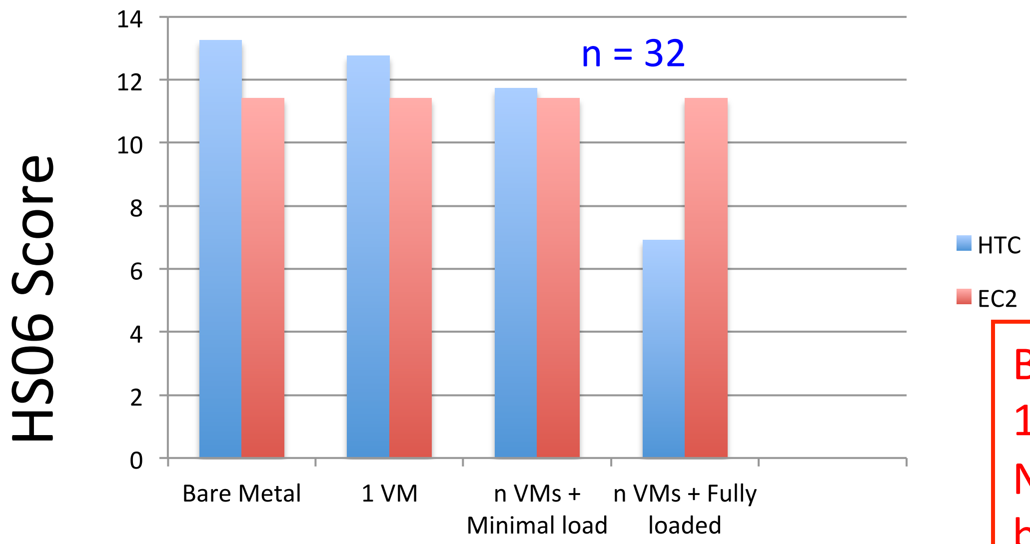


Bare Metal – no virtualisation
1 VM + idle
N VMs + minimal load
N VMs + Fully loaded

Bare Metal – NO!!!!!
1 VM + idle – UNLIKELY!
N VMs + minimal load – Possible but Unknown!
N VMs + Fully loaded – Possible but Unknown!

- Virtualisation + MT effect on performance ~ 8.14% to 55.84%
- Note the minimal effect of > no. of VMs

HS06 for Medium

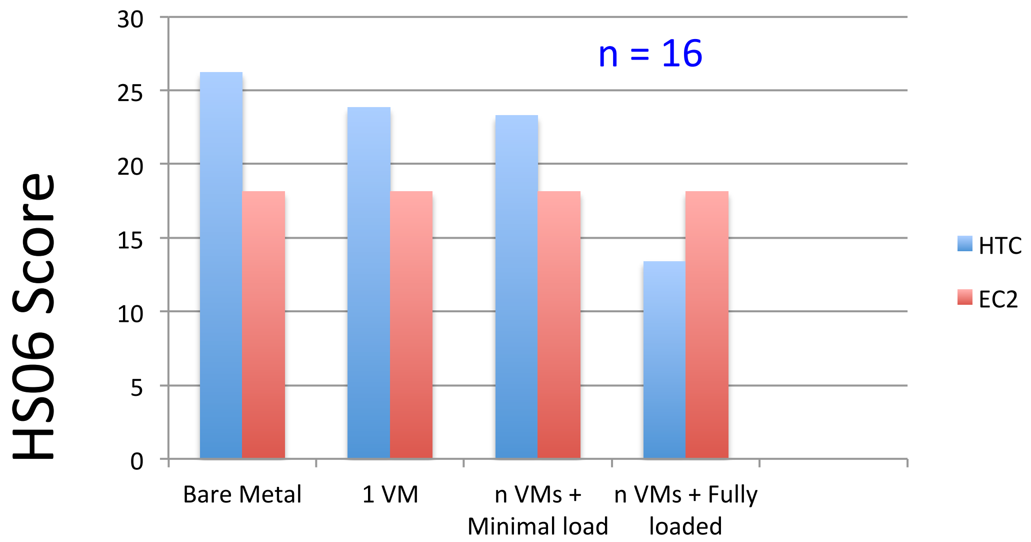


Bare Metal – no virtualisation
1 VM + idle
N VMs + minimal load
N VMs + Fully loaded

Bare Metal – NO!!!!!
1 VM + idle – UNLIKELY!
N VMs + minimal load – Possible but Unknown!
N VMs + Fully loaded – Possible but Unknown!

- Virtualisation + MT effect on performance ~ 3.77% to 47.89%

HS06 for Large



Bare Metal – no virtualisation
1 VM + idle
N VMs + minimal load
N VMs + Fully loaded

Bare Metal – NO!!!!!
1 VM + idle – UNLIKELY!
N VMs + minimal load – Possible but Unknown!
N VMs + Fully loaded – Possible but Unknown!

- Virtualisation + MT effect on performance ~ 9.04% to 48.88%

Conclusions - HPC

- As expected a purpose built HPC cluster outperforms EC2 cluster for same number of OMP threads
  – Average performance loss over all NPB tests: **~37%**

- Similarly so for when comparing 10GigE versus Infiniband networking fabrics
  – Average performance loss over all NPB test: **~48%**

- Even at a modest problem size the differences in performances between systems is highlighted.

e-FISCAL

Conclusions - HTC

- Virtualisation overhead is much less than the Multi-Tenancy effect
  - What others are running will have a direct effect!

- Standard deviation with pre-launched VMs in EC2 is significantly low!
  - Hypothesis: Variations will possibly be there!

- HS06 Scores variations on the order of 40-48%

Next steps

- HTC vs. Cloud Benchmarking
  - Cluster Compute and High-CPU Instances
  - Study pre-launch vs. new VM in EC2
- Benchmarking results in the cost model
  - As an extra weight in addition to monetary costs
- Publications

Thank you for your attention!

Questions??

kashif.iqbal@ichec.ie