e-FISCAL

HPC/HTC vs. Cloud Benchmarking

An empirical evaluation of the performance and cost implications

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

Outline

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

e-FISCAL

Overview

- Rationale
  - Diverse computing infrastructures (HPC. HTC, Cloud)
  - Diverse workloads (SMEs, academics, app. Domains)

- Methodology
  - System benchmarking for:
    - Comparison of HPC and HTC systems vs. Cloud offerings
    - Comparison of parallelism techniques (e.g. MPI/OMP)
      - ***To calculate the performance variation factor***
  - For Cost analysis and infrastructures comparison
    - Performance overhead as indirect costs
    - **Inclusion of an additional weight factor in the cost model**

Scope

- What we are not doing?
  - Benchmarking individual components (e.g. memory, network or I/O bandwidth)
- Why?
  - Because resources cost money
  - Aim is to compare HPC, HTC and Cloud infrastructures to identify ranges for performance variations

Benchmarks

- LINPACK – Top 500
- SPEC06 – CPU intensive benchmark
    - HEP-SPEC06 (for HTC vs. Cloud instances)
- HPC Challenge (HPCC)
    - Linpack, DGEMM, STREAM, FFT…
- Graph 500
- MPPtest – MPI performance
- NAS Parallel Benchmark (NPB)
    - (for HPC vs. Cloud HPC instances)

e-FISCAL

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
- Performance metric – Kernel execution time (in sec)

e-FISCAL

NPB Kernels

Kernel	Description	Problem Size	Memory (Mw)
EP	Monte Carlo kernel to compute the solution of an integral – Embarrassingly parallel	2^{30} random-num pairs	18
MG	Multi-grid kernel to compute the solution of the 3D Poisson equation	256^3 grid size	59
CG	Kernel to compute the smallest eigenvalue of a symmetric positive definite matrix	75000 no. of rows	97
FT	Kernel to solve a 3D partial difference equation using an FFT based method	512x256x256 grid size	162
IS	Parallel sort kernel based on bucket sort	2^{25} no. of keys	114
LU	Computational Fluid Dynamics (CFD) application using symmetric successive over relaxation	102^3 grid size	122
SP	CFD application using the Beam-Warming approximate factorisation method	102^3 grid size	22
BT	CFD application using an implicit solution method	102^3 grid size	96

Cloud Cluster Setup

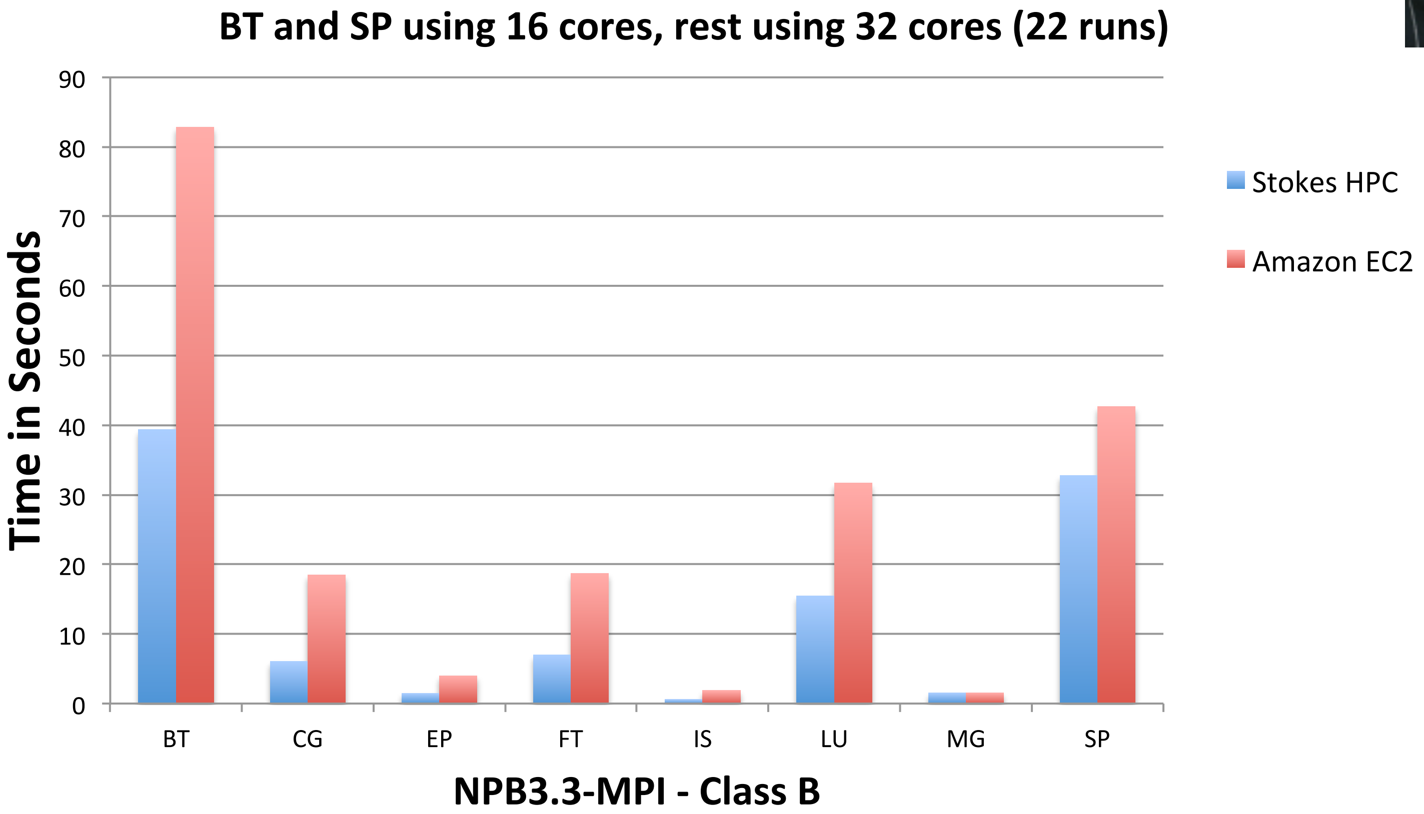
- EC2 instance management
  - StarCluster Toolkit
    - http://web.mit.edu/star/cluster/
    - StarCluster AMIs – Amazon Machine Image
      - Resource manager plugin
- Login vs. compute instances
  - EC2 small instance as login node
  - File system shared via NFS across nodes

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

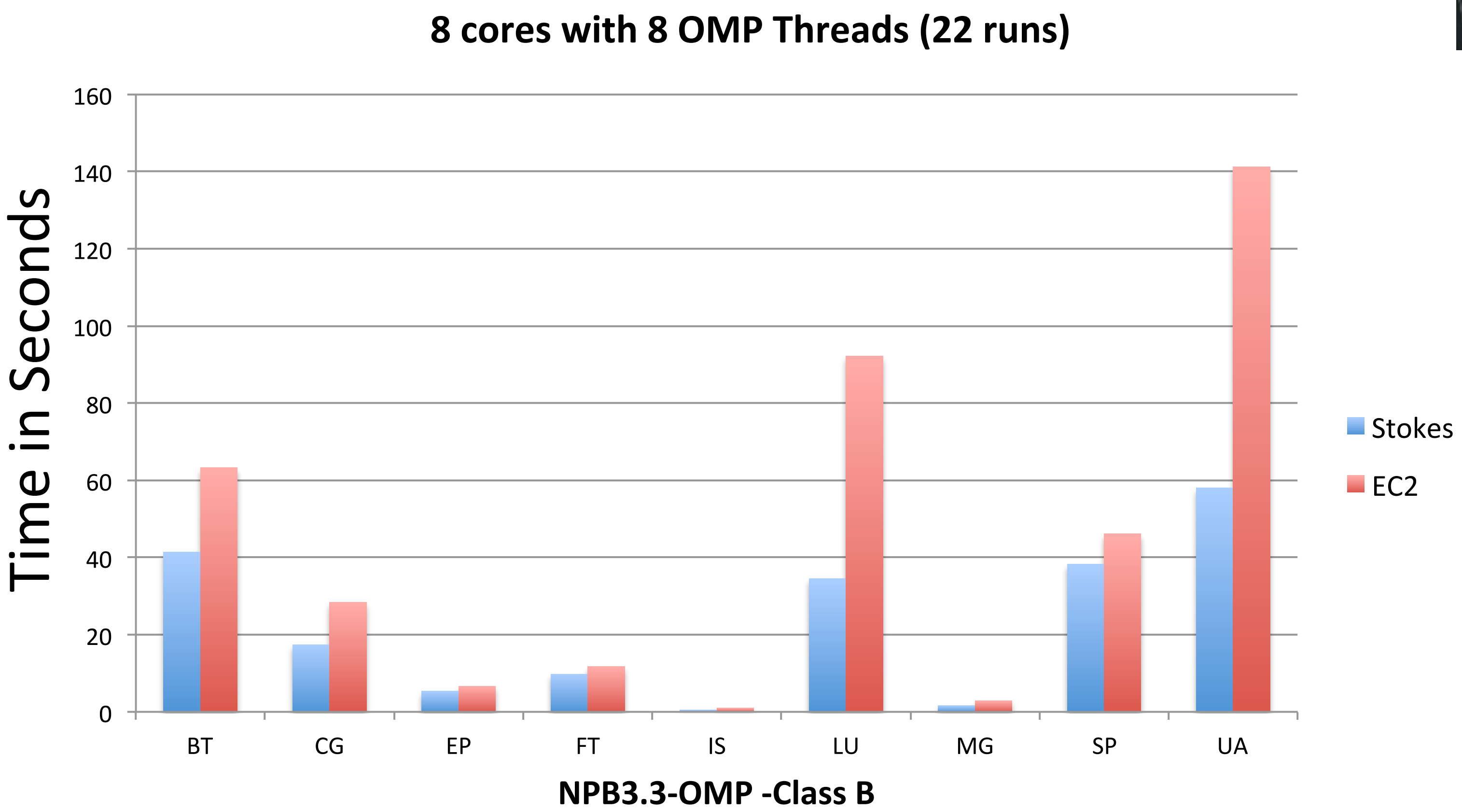
- Non-trivial to replicate runtime environments
- Large variations in performance possible
- Logical vs. Physical cores
  - HT/SMT – Hyper or Simultaneous Multi-Threading (i.e. 2 X Physical Cores)

NPB – MPI



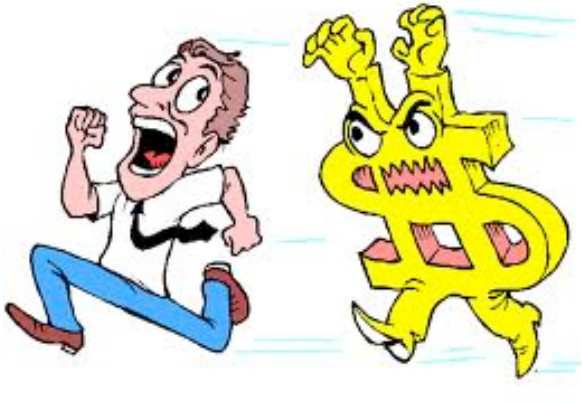
The average performance loss ~ **48.42%** (ranging from 1.02% to 67.76%).

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 facts:
  - On-Demand 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
  - – Performance Metric: HEPSPEC Score

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

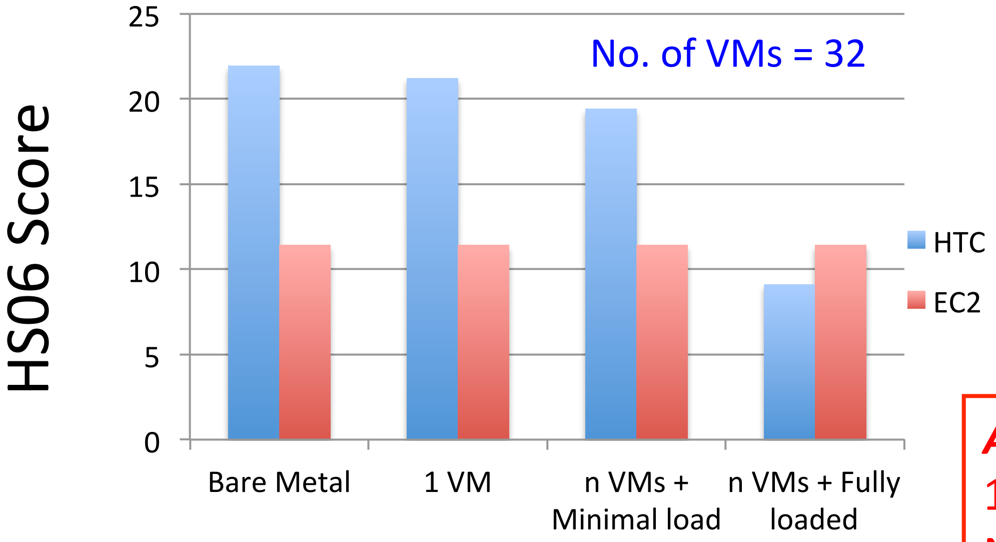
Benchmark Environment

	Amazon EC2	HTC resource at INFN
Compute Nodes	Medium: 2 ECU Large: 4 ECU XL: 8 ECU M3 XL: 13 ECU M3 2 XL: 26 ECU **1 ECU = 1.0-1.2 GHz**	**Intel(R) Xeon(R)** CPU E5-2660 @ 2.2 GHz, 2 X 8 cores, 64 GB memory **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** **M3 XL Instance Quad-core VM** **M3 Double XL Instance Eight 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

Benchmark Variations

- Bare-metal vs. EC2 imply:
  - *Cloud migration performance implications*
- 1 VM vs. Bare-metal:
  - *Virtualisation effect*
- n VMs + 1 loaded (HS06) vs. Bare-metal:
  - *Virtualisation + multi-tenancy effect with minimal workload* - Best Case Scenario!
- n VMs + n loaded
  - *Virtualisation + multi-tenancy effect with full workload* - Worst Case Scenario!

HS06 for Medium



NGIs – Possible variations

Bare Metal – NGIs current

1 VM + idle
N VMs + minimal load
N VMs + Fully loaded

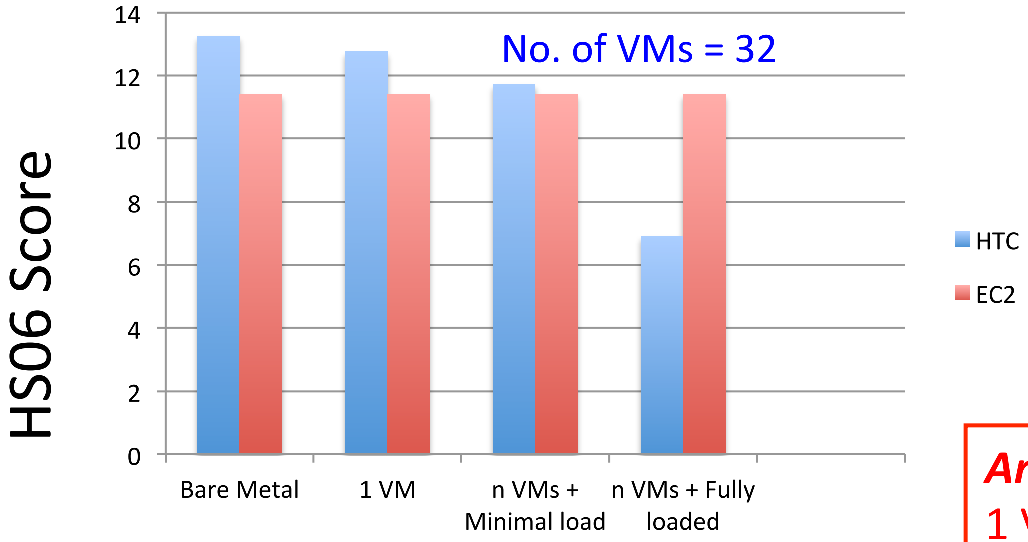
} Future

Amazon EC2 - Possible variations

1 VM + idle – Unlikely in EC2
N VMs + minimal load – Best-case!
N VMs + Fully loaded – Worst-case!

- SPEC score < with the > no. of VMs
- Virtualisation + Multi-Tenancy (MT) effect on performance ~ 3.28% to 58.48%
- 47.95% Performance loss for Cloud migration

HS06 for Medium



NGIs – Possible variations

Bare Metal – NGIs current

1 VM + idle

N VMs + minimal load

N VMs + Fully loaded

} Future

Amazon EC2 - Possible variations

1 VM + idle – Unlikely in EC2

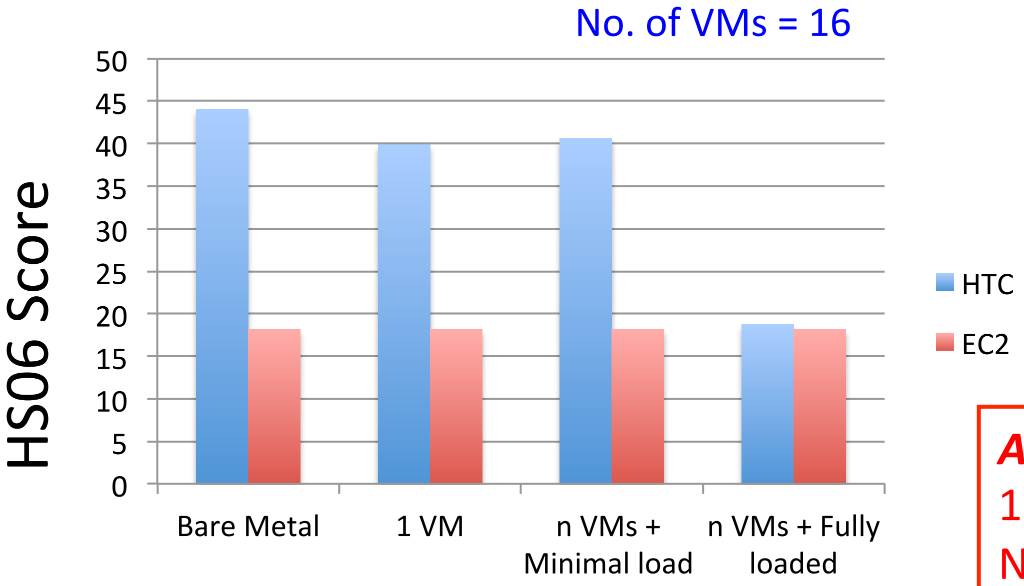
N VMs + minimal load – Best-case!

N VMs + Fully loaded – Worst-case!

- Virtualisation + MT effect on performance ~ 3.77% to 47.89%
- 18.19 Performance loss for Cloud migration

e-FISCAL

HS06 for Large



NGIs – Possible variations

Bare Metal – NGIs current

1 VM + idle

N VMs + minimal load

N VMs + Fully loaded

} Future

Amazon EC2 - Possible variations

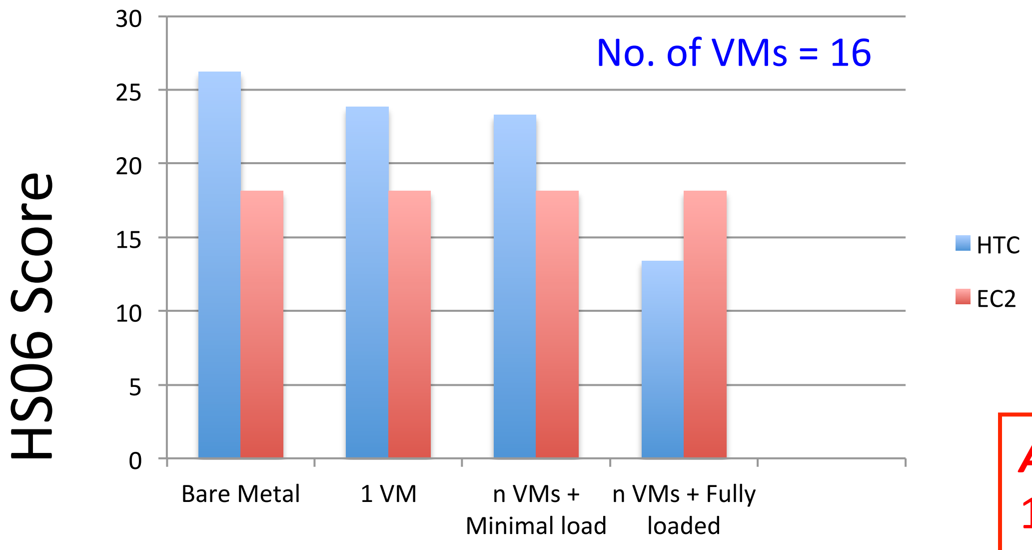
1 VM + idle – Unlikely in EC2

N VMs + minimal load – Best-case!

N VMs + Fully loaded – Worst-case!

- Virtualisation + MT effect on performance ~ 9.49% to 57.47%
- Note the minimal effect on performance with > no. of VMs
- 58.79% Performance loss for Cloud migration

HS06 for Large



NGIs – Possible variations

Bare Metal – NGIs current

1 VM + idle

N VMs + minimal load

N VMs + Fully loaded

} Future

Amazon EC2 - Possible variations

1 VM + idle – Unlikely in EC2

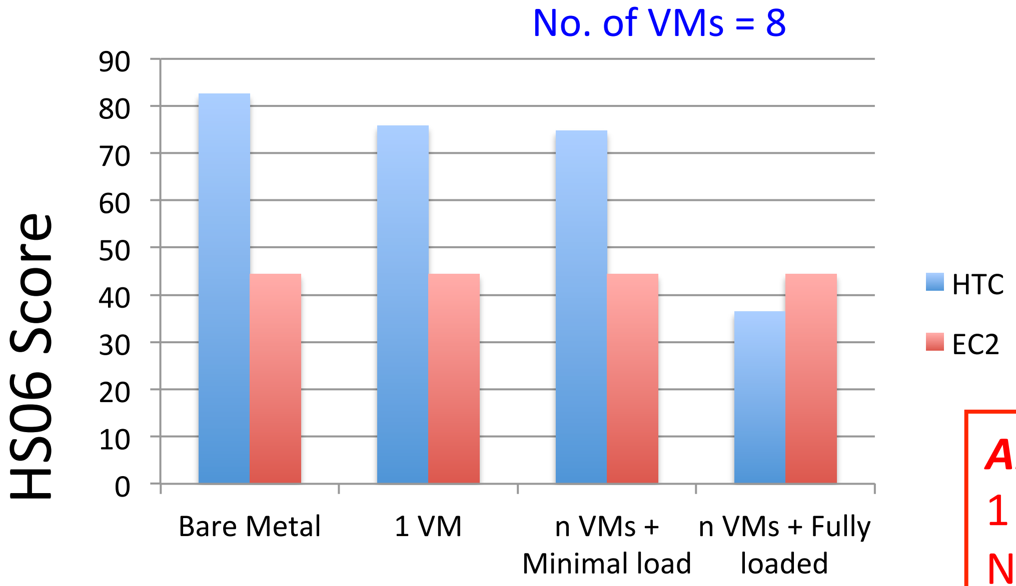
N VMs + minimal load – Best-case!

N VMs + Fully loaded – Worst-case!

- Virtualisation + MT effect on performance ~ 9.04% to 48.88%
- 30.75% Performance loss for Cloud migration

e-FISCAL

HS06 for Xlarge



NGIs – Possible variations

Bare Metal – NGIs current

1 VM + idle
N VMs + minimal load
N VMs + Fully loaded
} Future

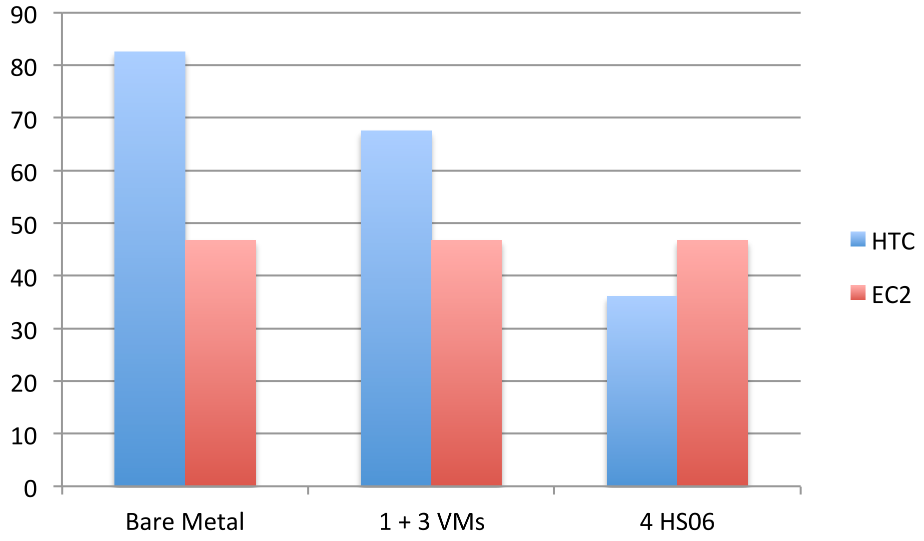
Amazon EC2 - Possible variations

1 VM + idle – Unlikely in EC2
N VMs + minimal load – Best-case!
N VMs + Fully loaded – Worst-case!

- Virtualisation + MT effect on performance ~ 8.14% to 55.84%
- Note the minimal effect of > no. of VMs
- 46.21% Performance loss for Cloud migration

e-FISCAL

M3 X-Large


No. of VMs = 8

NGIs – Possible variations

Bare Metal – NGIs current

N VMs + minimal load

N VMs + Fully loaded

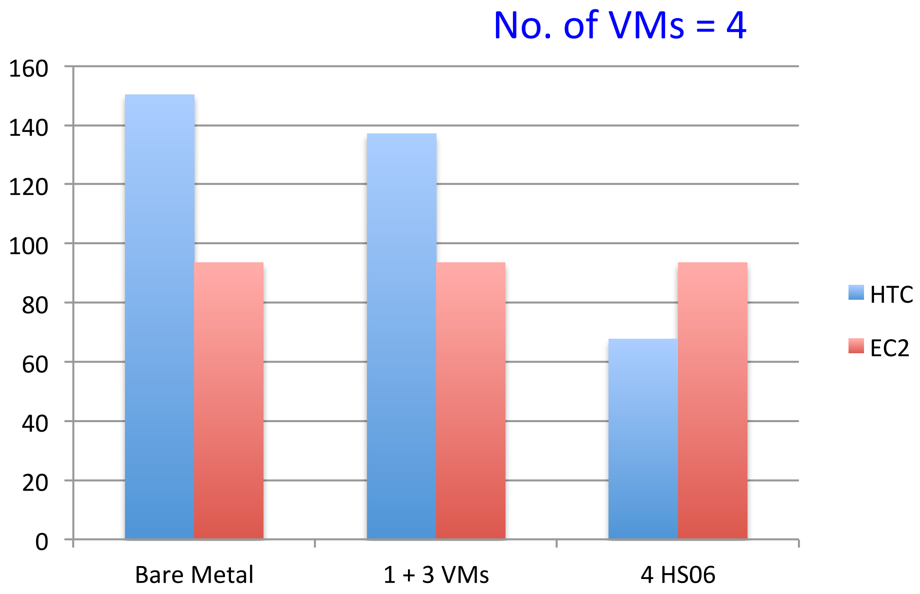
Amazon EC2 - Possible variations

N VMs + minimal load – Best-case!

N VMs + Fully loaded – Worst-case!

- Virtualisation + MT effect on performance ~ 18.24 – 55.26
- 43.37% Performance loss for Cloud migration

M3 Double X-Large

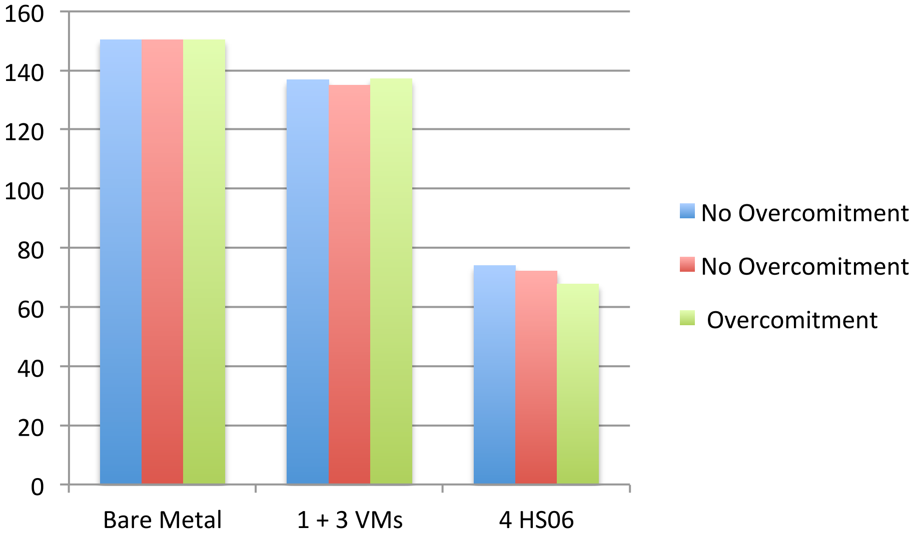


NGIs – Possible variations
Bare Metal – NGIs current
N VMs + minimal load
N VMs + Fully loaded

Amazon EC2 - Possible variations
N VMs + minimal load – Best-case!
N VMs + Fully loaded – Worst-case!

- Virtualisation + MT effect on performance ~ 8.81 – 54.98%
- 37.79% Performance loss for Cloud migration

Interesting Finding



Memory: 28GB of 64GB
Memory: 48GB of 64GB

Memory: 120GB of 64 GB

- Over-commitment of resources vs. Cost effectiveness
  - That is more reasonably why 1 ECU = 1.0-1.2 GHz 2007 Xeon or Opteron!!

Conclusions – HPC vs. EC2

- 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.

Conclusions – HTC vs. EC2

- 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!
  - Results with M3 L instance: 46.92, 46.61, 47.79
- HS06 Scores variations in the order of 40-48%

Thank you for your attention!

Questions??

kashif.iqbal@ichec.ie