

It's all about the money, money, money... ISC 2017

21 June 2017





Who are we?



Why do we feel qualified to talk about this?

- Founded in 2004, Red Oak has a substantial customer base in government, industry, research and academia
 - Combined HPC industry experience in excess of 80 yrs
- Red Oak Consulting are uniquely placed in the sector as the only pure play, client side HPC consulting organisation in the UK
- Worked on 30+ HPC installations in last 10 years
 Ranging from \$100k to \$100m (incl. Top 500)





What do we do?



Advise on all aspects of the HPC lifecycle

- Some of the services we provide:
 - Strategic and business analysis
 - Change and programme management
 - Targeted technical advice and solutions
 - Benchmarking, system integration and tuning
 - Procurement support through to system acceptance
 - Highly experienced HPC project management
 - Business case creation and benefits realisation
 - Service transition and decommissioning
- HPC and its application are Red Oak's core expertise





Why are we here?



Not just AWS Certified Consulting Partners

- Red Oak is the *only* decision support HPC consultancy working with AWS in the UK
- A business built on knowledge, loyalty, and reputation with outstanding longevity in our customer relationships and partnerships
- Red Oak recognises that conventional onpremise HPC deployment are not necessarily a panacea





HPC: Why bother?



Only one answer

• Because it provides value

Alternative, slightly better answer

• Because it provides more value than it costs





Where is your HPC tomorrow?



The traditional answer ...

- The same place as it was yesterday
- "IT is a truth universally acknowledged ..."
 - That an organisation in possession of a large fortune must be in want of an on premise HPC system appologies to Austen
- Until recently it's certainly been difficult to challenge this accepted truth
 - Though certain classes of problem (esp capacity vs capability etc.) have mapped well onto public cloud
- However there are lots of confounding factors





Where is your HPC tomorrow?



The "heretical" answer ...

- In a public cloud environment such as AWS
- So can we can build a model that allows us to challenge the accepted orthodoxies of HPC and the Cloud Computing
 - Caveat: Assuming workflow which perform well on AWS
- Observation: Traditional TCO models do not account for the future
 - When this is factored into TCO models, the answers become ... interesting





Where is your HPC tomorrow?



The caveat

- Measuring 'value' is a complex topic with many different treatments available
- We'll show in the next few slides that the 'value for money' argument isn't always as clear cut as perhaps the traditional view suggests
- In particular, allowing for Moore's Law and using Internal Rate of Return (IRR) models, public cloud starts to look more competitive with traditional HPC on a cost basis





HPC exceptionalism



Where IT gets interesting

- HPC is special
- Pace of technological change is (very) high
- Performance of hardware increases rapidly and exponentially over time – "Moore's Law"
 - Not just gate density and historical Dennard scaling
 - Also improvements in the peripheral areas such as networking, storage, memory, manufacturing, ...
 - There are more innovative and disruptive technologies coming now than for a number of years ...





So about Moore's Law...



Is it dead?



Time

Moore's Law

Observations

- No it's not dead yet
 - Though it is becoming much harder to maintain
- Headline version focuses on the CPU, but equivalent curves also apply to networking, storage and others
 - Though exponential terms are different
- One of the interesting aspects of public cloud is that as it is exponentially growing the pricing more or less tracks Moore's Law

The 'value' of money

Internal Rate of Return (IRR) in TCO models

- IRR is a metric used in budgeting to calculate the return on investments into projects
 - You can operate on the assumption that any project will have at least an IRR of 20% (otherwise why are you spending the money?)
- What does this mean in practice?
- If you can delay your spend (on HPC) until next year then you can use the same spend for something else this year, make a profit of that, and still spend it (on HPC) next year

A simple TCO spreadsheet

Vastly over-simplified but illustrates the point

- Start with a traditional simplified TCO model and compare value for money with public cloud
- Spoiler alert:
 - Spend the whole day discussing TCO models at SC17
 - Discusses all the assumptions used here and many more besides
- The purpose of the following slides is to thought provoking, a conversational opener rather than to try and be a definitive answer

State the assumptions

Really simple TCO model for On Prem vs AWS

Purchase price per core (2017)	£185	Based on recent 5000 core
Size of system for examples	2000	Cores
Manufacturer support &		Come see our TCO Workshop at
maintenance	5.00%	SC17
Procurement overhead	10.00%	
Installation facilities costs	10%	
Electricity cost	9%	Based on 10p per kWhr
-		Not an unerasonable average for many older
PUE	2	DCs
Data centre costs	18%	
Internal support team	2.00	FTE
FTE cost	£100,000	per annum
Effective utilisation	70%	
Moore's Law per year	70%	
Discount rate (IPP)	20%	
	2076	
Dollar rate	\$ 1.28	USD/GBP
Hours per year	8760	

AWS Costs

AWS price per core (June 2017)

Using c4.8xlarge pricing

18 cores of c4.8xlarge per hour	£	1.48 \$	1.90	On demand
	£	1.01 \$	1.29	Spot reserved 6hr
	£	0.97 \$	1.24	Reserved
	£	0.82 \$	1.05	Spot reserved 1hr
	£	0.20 \$	0.25	Spot
Per core hour	£	0.06 \$	0.08	On demand
	£	0.04 \$	0.06	Spot reserved 6hr
	£	0.04 \$	0.05	Reserved
	£	0.04 \$	0.05	Spot reserved 1hr
	£	0.01 \$	0.01	Spot

- We have some controls to use in the calculations
- Including recurring costs such as support and maintenance, electricity, cooling, data centre costs, internal support, utilisation factors
- But also Moore's Law and Internal Rate of Return

The really simple treatment

Bit more realistic for on-premise

More realistic AWS pricing (res.)

What about utilisation?

But what about Moore's Law

And Internal Rate of Return

A diversion: Value

What is value?

The value of value

- Value and benefit poorly understood terms
 - Lots of confusion , misconception and misrepresentation
- Human nature to overvalue benefit, and downplay risk and cost
 - Lots of double counting
 - Confusion between necessary and sufficient
- Broadly two types of benefit/value
 - Direct organisational (or personal)
 - Socio-economic
- Implications for funding route

What is value?

Counting value

- Generally if it can't be measured it doesn't exist
 - Gold standard: £
 - Other metrics are available but need to be applied carefully
- Generally easier to metricate value than is assumed
 - Change in profit/loss
 - What would you pay
 - What would you pay to find out
- Detailed measurement can be hard though

What is value?

Show me the money

- Timeliness is often ignored in business cases
 - "Too hard" to account for
 - (In contrast to "time")
- Generally a mistake
 - Time is the most precious commodity we have
- Accepted accounting technique for understanding this: NPV
 - Then becomes an argument re discount rates
 - Not as straightforward as may be assumed
 - LIBOR: 0.5%, Wonga APR: 1,500%

A diversion: Cost

Cost

The price of everything ...

- Not as straightforward as assumed
 - Too often just focused on "cash out of the door"
- Example "cash costs" tracked carefully
 - Cost of the supercomputer (purchase and support)
 - Software costs (purchase and support)
 - Cost of installation and changes to the data centre
 - Electricity (and cooling)
 - Training, travel and subsistence!
 - Contractor costs

Cost

But that comes out of a different budget ...

- Other costs often ignored
 - Project staff costs
 - "Overhead" function costs
 - Data centre costs
 - Cost of money
 - Electricity (!)
 - Complexity of internal support (hardware and software)
 - Lost or delayed benefit
 - Not just NPV related
- Your project management may not "care" but your FD does

Conclusions

Don't loose sight of the real goals

- To deliver value (however you choose to measure it) at a reasonable cost
- More traditional HPC in the cloud is still viewed as hard but for some classes of HPC problem there are undoubtedly good reasons to look at the economics
- Accounting for effects such as Moore's Law and IRR in TCO calculations can significantly change the balance point for a move to a cloud strategy

