

# THE MARRIAGE OF CLOUD, HPC AND CONTAINERS

...AND SERVERLESS?

# ADAM HUFFMAN

Senior HPC and Cloud Systems Engineer

The Francis Crick Institute

[@adamhuffman](#)

[adam.huffman@crick.ac.uk](mailto:adam.huffman@crick.ac.uk)

# THE FRANCIS CRICK INSTITUTE

Biomedical discovery institute dedicated to understanding the fundamental biology underlying health and disease

Biggest biomedical research facility under a single roof in Europe

1,500 staff, £650m investment

# HPC STATUS QUO IN PERIL

Pressure from funding bodies

Pressure from procurement

Deadline-driven bursting

# EMEDLAB

<http://www.emedlab.ac.uk>

Biomedical research cloud, with 6,048 cores and 5.5PB  
storage

Share data, methods and expertise

OpenStack, GPFS

# GPFS

Benign environment assumption

Licensing complications

NFS export

# IMPORTANCE OF SHARED FILESYSTEM ACCESS

Scientific reference datasets

Object storage?

Application support for objects?

# DATA ACCESS EXAMPLE

800TB dataset from TCGC

Restricted access

Special client

Multiple instances



# VIRTUAL CLUSTERS

Elasticcluster

Heat

Ansible/Salt etc.

Back to importance of shared filesystem access

Are they actually the best fit for clouds?

# THE COMFORT OF CLUSTERS

How well are the characteristics of pipelines known?

Only start VMs as needed

Reflects focus on high utilisation, rather than "time to science"

# PERFORMANCE/FLEXIBILITY TRADEOFFS

Performance tuning of clouds

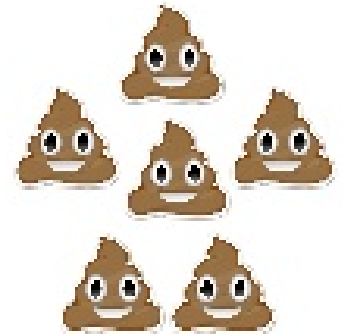
Flexibility cost

Rising software complexity

# Monolithic vs Microservices



Monolithic



Microservices

# CONTAINERS THE UNSTOPPABLE HYPE MACHINE

Hyperscale mindset and capabilities

Maturity of container support for scientific  
applications

Singularity, Shifter "fixing" containers for scientific use

Are we just reinventing packaging, badly?

# SERVERLESS

Lambda

Google Cloud Functions

OpenWhisk

Fission

Will science have a say?

# ABSTRACTION LAYERS

Users focus on images and snapshots

Is this the right level of abstraction

Use repeatable configuration  
management/provisioning instead?

Criticism of TripleO for being image-centric

How to update them

Image catalogues

# CMS/CERN HIGH ENERGY PHYSICS

Simply treat clouds as nice new resources

Already have tools to handle remote data  
transparently

Already have accounting tools

Advantages of grid experience

Concerns about security



# THANK YOU

Image Credits:

The Guardian

Alvaro Sanchez