ABRAHAM MARTIN AND JON WARBRICK

VIRTUAL MACHINES ALL THE WAY DOWN

THE UIS'S NEW MANAGED WEB SERVICE

FIRST, A LITTLE HISTORY

tagxedo.com

2000

- Sydney hosts the Olympic games
- Concord crashes in Paris
- Y2K problem
- UCS launches the Managed Web Server

MWS 1: Solaris 7, Apache 1.3, chroots & loopbacks, static content only, single server

2008

Beijing hosts the Olympic games
Almost nothing else happens
Except UCS launches revised/improved MWS

MWS 2: Solaris, Apache 2, PHP, MySQL, Solaris Zones and ZFS, manual dual-site resilience
~230 sites, ~320 hostnames

WHY A MANAGED WEB SERVICE?

- Eliminate badly managed web servers
- Reduce duplicated staff effort
- VMs <—> MWS <—> [Falcon, Drupal]

NEW MWS DESIGN PRINCIPLES

- Preserve most of the functionality of the current MWS
- Provide an environment similar to that provided by commercial web hosting providers
- Reduce the amount of Unix knowledge required to successfully manage a MWS site
- Delegate as much configuration control as possible directly to MWS site administrators

"MANAGED"

We provide
- Hardware
  - Dedicated virtual server (1 GB RAM, 20GB SSD disk (10GB usable))
- Physical hosting
  - Automatic geographically-diverse redundancy
  - WCDC and RNB
- Operating system (Debian Jessie)
- Software
  - Apache 2.4
  - MySQL
  - PHP
- Backups
- YOU provide and MANAGE everything else

SO, WHAT DO YOU GET?

Mention some or all of these
- Delegated management - Web-based control pannel
  - All users need UIS accounts
  - New sites on request - **demo**
- Public IPv4 and IPv6 addresses, unlimited name-based vHosts
- ssh-based shell access, scp, sFTP, etc.
  - Independent, non-shared accounts
  - No plain Telnet or FTP
  - ssh public key authn, SSHFP records
- Apache, PHP, Python, MySQL, file system permissions, Cron jobs, email
- IP-register integration
- UIS Passwords
- Snapshots and backups; rollback
- https support

https://panel.mws3.csx.cam.ac.uk/

http://mws-help.uis.cam.ac.uk/

Extensive documentation

Probably works with…

WORKS WITH

- Static sites
- PHP sites
- Python sites
- Wordpress
- Drupal
- …and many more
- …but you have to look after them all

- Automatic payment management
- Requires purchase order
- Runs from the day you create the server
- Initial month to sort out payment

- No root/sudo access
- £100/year

TRANSITION

- Old sites need to go somewhere
- Old hardware
  - Difficult to manage
  - Getting too old for comfort
  - Causing operational problems
    - Can't deploy IPv6 multicast on data centre network
- Payment discount
- Transition advice on the web site

TALK TO US BY THE END OF JUNE

We've been trying to talk to current users, but not everyone seems to want to talk to us.

If you don't, your sites will be switched off first week July.

IF YOU DO, YOU HAVE TO THE END
OF SEPTEMBER TO TRANSITION

https://mnementh.csi.cam.ac.uk/mws3-progress.html

Declared plans

- ?
- MWS3
- Cancel
- Thinking
- CANCELLED
- Falcon

32.8%
48.3%
12.7%

Current status in the DNS

- mws2
- Other cam
- MWS3
- External
- Falcon
- NXDOMAIN

85.1%
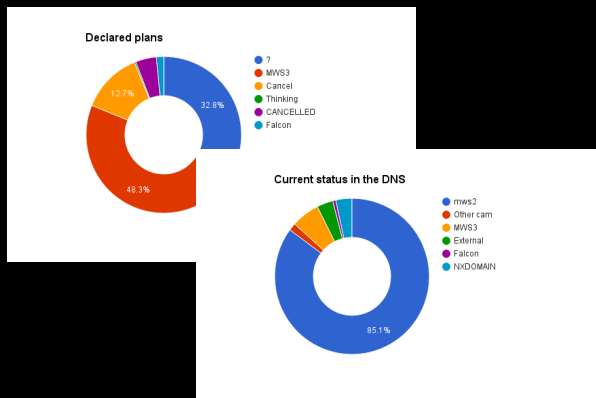
- Mismatch between plans and progress

These are the sites that we don't yet know about

intranet.ames.cam.ac.uk
www.unite.cam.ac.uk
www.archanth.cam.ac.uk
www.astronomicalimages.group.cam.ac.uk
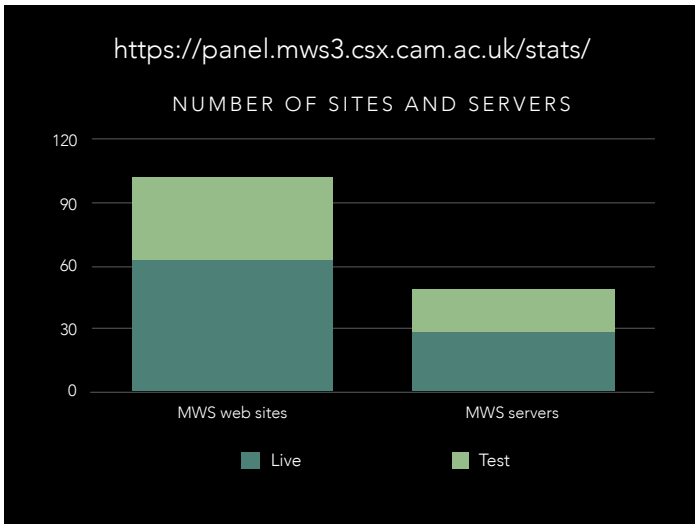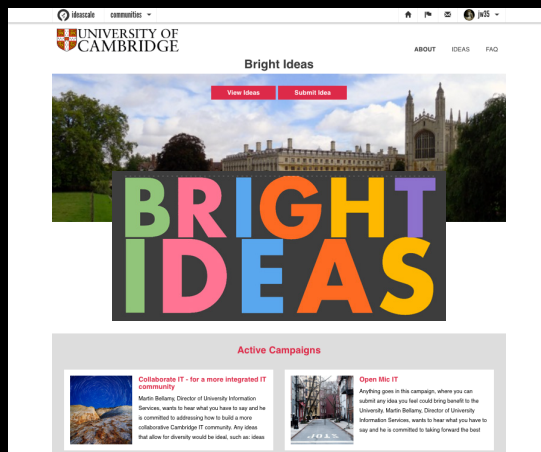**www.biotech.cam.ac.uk**
www.biot.cam.ac.uk
www.mcr.cai.cam.ac.uk
www.cambridgecancercentre.org.uk
www.cancer.cam.ac.uk
www.cag.group.cam.ac.uk
www.cilr.cam.ac.uk
mws.cimr.cam.ac.uk
financehub-old.cisl.cam.ac.uk
new-www.clarehall.cam.ac.uk
www.dev.clarehall.cam.ac.uk
www.test.clarehall.cam.ac.uk
www.demo.clarehall.cam.ac.uk
www.cis.cam.ac.uk
www.cmi.cam.ac.uk
cms.homerton.cam.ac.uk
www.cnig.group.cam.ac.uk
www.comodo.group.cam.ac.uk
www.conservation.cam.ac.uk
www.physics.corpus.cam.ac.uk
www.cria.polis.cam.ac.uk
cria.polis.cam.ac.uk
www.cric.arch.cam.ac.uk
www.ipscc.cruk.cam.ac.uk
www.lightmicroscopy.cruk.cam.ac.uk
www.ctc.cam.ac.uk
www.qi.damtp.cam.ac.uk
www.ditg.group.cam.ac.uk
www.ct2014.dpmms.cam.ac.uk

www.econsoc.hist.cam.ac.uk
www.edeg.org
www.edwinstowe.group.cam.ac.uk
www.eeci.cam.ac.uk
development.eeci.cam.ac.uk
cit.eng.cam.ac.uk
legacy.mrc-epid.cam.ac.uk
www.gates.scholarships.cam.ac.uk
www.genet.ac.uk
henge.bioanth.cam.ac.uk
christianization.hist.cam.ac.uk
www.econpublic.hps.cam.ac.uk
www.jobs.hps.cam.ac.uk
www.lara.psychol.cam.ac.uk
lipkow.sysbiol.cam.ac.uk
www.many-core.group.cam.ac.uk
www.genetics.mcdonald.cam.ac.uk
new-www.mediaplayer.group.cam.ac.uk
mega.bioanth.cam.ac.uk
www.bartlett.psychol.cam.ac.uk
www.ims.cam.ac.uk
www.aim.msm.cam.ac.uk
www.t1.aim.msm.cam.ac.uk
www.mv.crassh.cam.ac.uk
mvt.psychol.cam.ac.uk
mws.psychol.cam.ac.uk
old-www.mdrs2015.psychol.cam.ac.uk
www.camcompcog.psychol.cam.ac.uk
www.mzs.csx.cam.ac.uk
www.northasianborders.net
www.p-ace.group.cam.ac.uk
www.parenchyma.archanth.cam.ac.uk
www.pave.arch.cam.ac.uk

new-noggin.pdn.cam.ac.uk
www.peersupport.cam.ac.uk
www.phagedisplay.org
www.gscc.phy.cam.ac.uk
www.qm.phy.cam.ac.uk
www.sd.phy.cam.ac.uk
mws.polis.cam.ac.uk
old-www.cirmena.polis.cam.ac.uk
www.prime.bioanth.cam.ac.uk
www.purpleshin.phy.cam.ac.uk
www.pwf.csi.cam.ac.uk
www.rock-art.group.cam.ac.uk
manual.recoup.educ.cam.ac.uk
www.manual.recoup.educ.cam.ac.uk
www.wyse-applications.socanth.cam.ac.uk
old-www.socanth.cam.ac.uk
www.scopic.ac.uk
www.shef.group.cam.ac.uk
www.siphs.mml.cam.ac.uk
mws.sociology.cam.ac.uk
www2.synbio.cam.ac.uk
ad-consortium.sysbiol.cam.ac.uk
logic.sysbiol.cam.ac.uk
www-mws.sysbiol.cam.ac.uk
ralser.sysbiol.cam.ac.uk
trailblazing.group.cam.ac.uk
www.trailblazing.group.cam.ac.uk
www.techdem.crassh.cam.ac.uk
transgenerational.zoo.cam.ac.uk
www.viscog.psychol.cam.ac.uk
www.westcam-safety.group.cam.ac.uk
www2.zoo.cam.ac.uk

https://panel.mws3.csx.cam.ac.uk/stats/

NUMBER OF SITES AND SERVERS

120

90

60

30

0

MWS web sites   MWS servers

■ Live   ■ Test

THE FUTURE

- Current work coming to an end
- Backlog of 40-50 items that won't get done this time, for example
  - PostgreSQL
  - At-cost upgrades
  - Remote access to databases
  - GUI database admin tools pre-installed
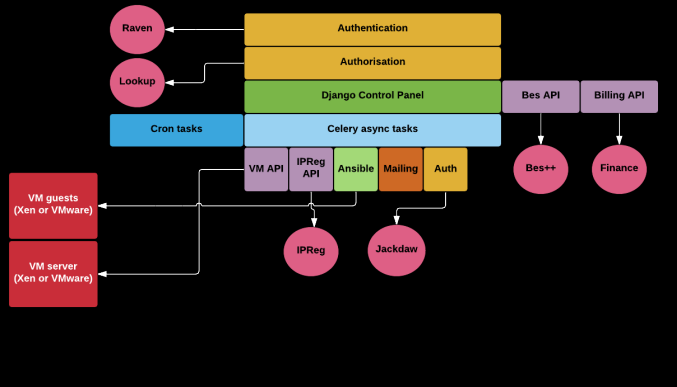 - Support for LetsEncrypt
  - Nginx/Tomcat/…
  - Varnish cache

---

http://cambridgeuniversity.ideascaleapp.eu/

Plan is to put them into a Bright Ideas campaign

…and ask you all for other ideas

…and ask you all to vote

ARCHITECTURE

Pink: External services
We manage all of it with ansible
Application Deployment + Configuration Management + Continuous Delivery
Dynamic Inventory
Raven Webauth
Lookup

AUTHORISATION (LDAP ISH BASED)

ARCHITECTURE

Jackdaw for UID of user

User is installed in the VM (Using Ansible)

UID (important for shared file storage) taken from Jackdaw (User central database)
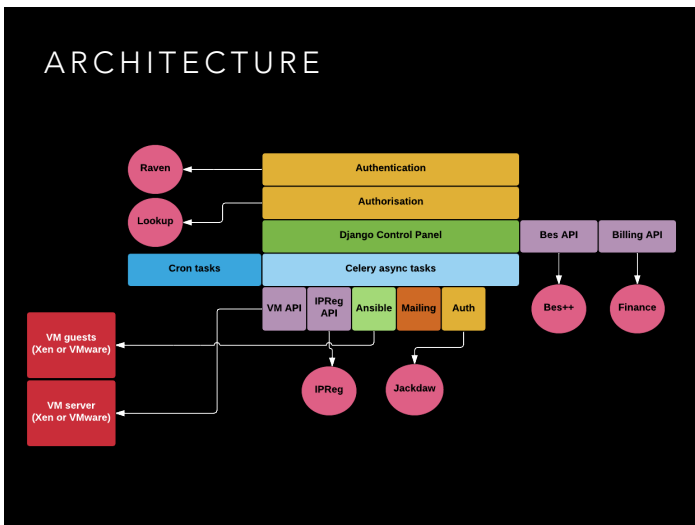
Periodic task to refresh installed users (in VMs) authorised via LDAP groups

SSH public key uploaded to the web panel

ARCHITECTURE

IP Register

Preallocated IP addresses
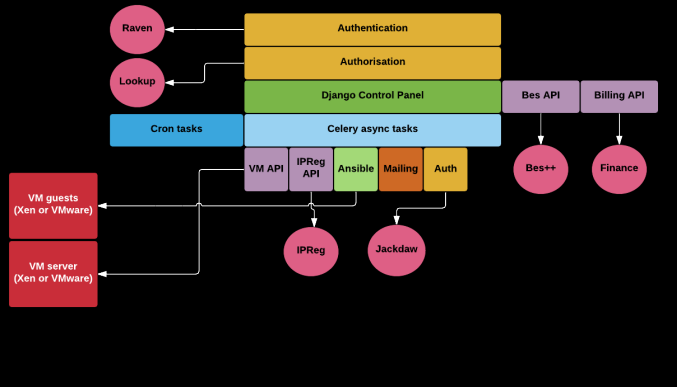
*.cam.ac.uk hostnames available for users (API)

Service/Host addresses

SSHFP records and DNSSEC

Preallocated: we can have instant access instead of waiting for DNS servers updates.
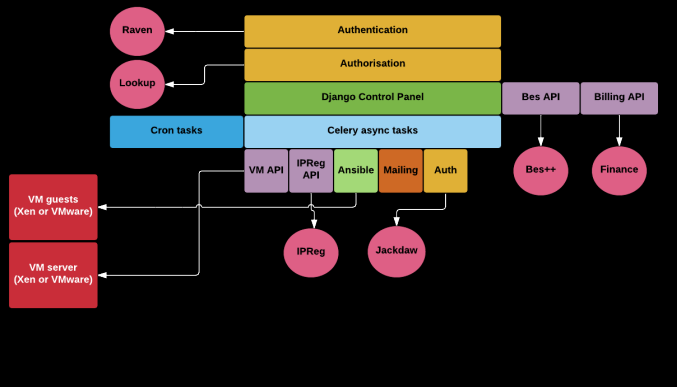
IP4&6 addresses and hostnames

Bes++ (central inventory)
JSON file with information about all hosts:
Location, IP, hostname, VM properties
Used for other purposes like monitoring

Celery + Redis
Some API calls
Background processes
Cron jobs (update machine with last changes in lookup / system packages / backups and snapshots)
Message queue storage
Mutex
Different VM APIs (compatible)

The service is designed to be scalable and has a small cluster of AMQP message brokers (which receive metrics) and a small cluster of carbon/graphite machines to store and graph data..
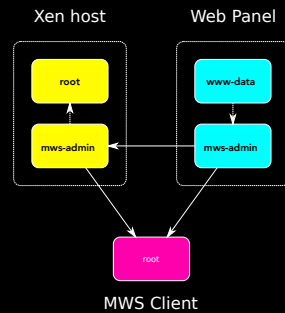
statsd & collectd

cluster AMQP message brokers
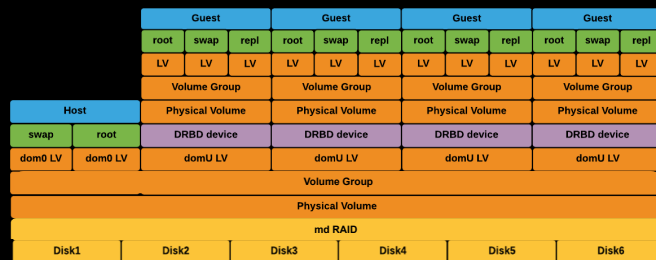
cluster carbon/graphite (storage)
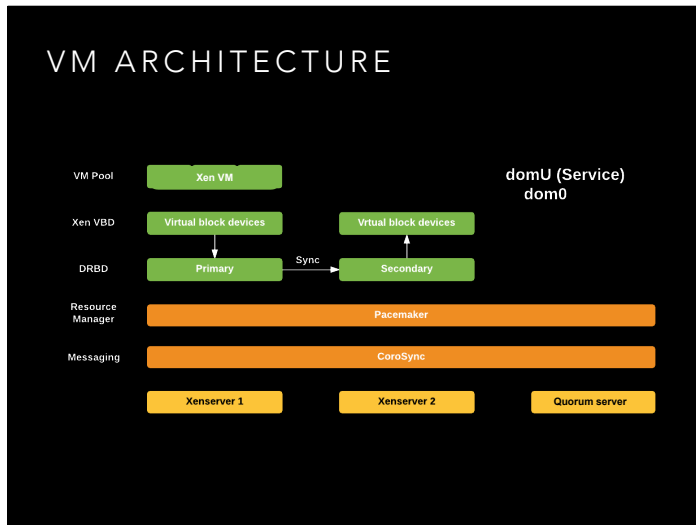
CLIENT STORAGE

- 1 LVM PV (25G)
  - root LV (4.5G)
  - swap LV (0.5G)
  - thin pool (20G)
    - user data LV (10G) /replicated
    - snapshots

UNDERLYING STORAGE

A DRBD device is generated for each one of the VMs at dom0 level. These devices have each one a domU logical volume as their storage. On top of the DRBD solution, a physical volume is created. On top of that physical volume three logical volumes are created, one for swap, one for root and another for replicated. Guests (domU) only see the DRBD as their storage, therefore only a physical volume with three logical volumes.

VM ARCHITECTURE

			domU (Service)
VM Pool	Xen VM		dom0
Xen VBD	Virtual block devices	Vrtual block devices	
DRBD	Primary	Secondary	
Resource Manager	Pacemaker		
Messaging	CoroSync		
	Xenserver 1	Xenserver 2	Quorum server

Storage replication is done at a higher level in dom0 insted of in domU.

Clusters are formed by two xen servers and a dummy VM as a 3-node cluster with each xen guest being a "resource" in the cluster.

Pacemaker tend to divide these "resources" up between the two Xen hosts (but they can be moved from one to the other as necessary)

domU VMs are treated as pacemaker resources and therefore are moved from one xen server to another in case of failure.