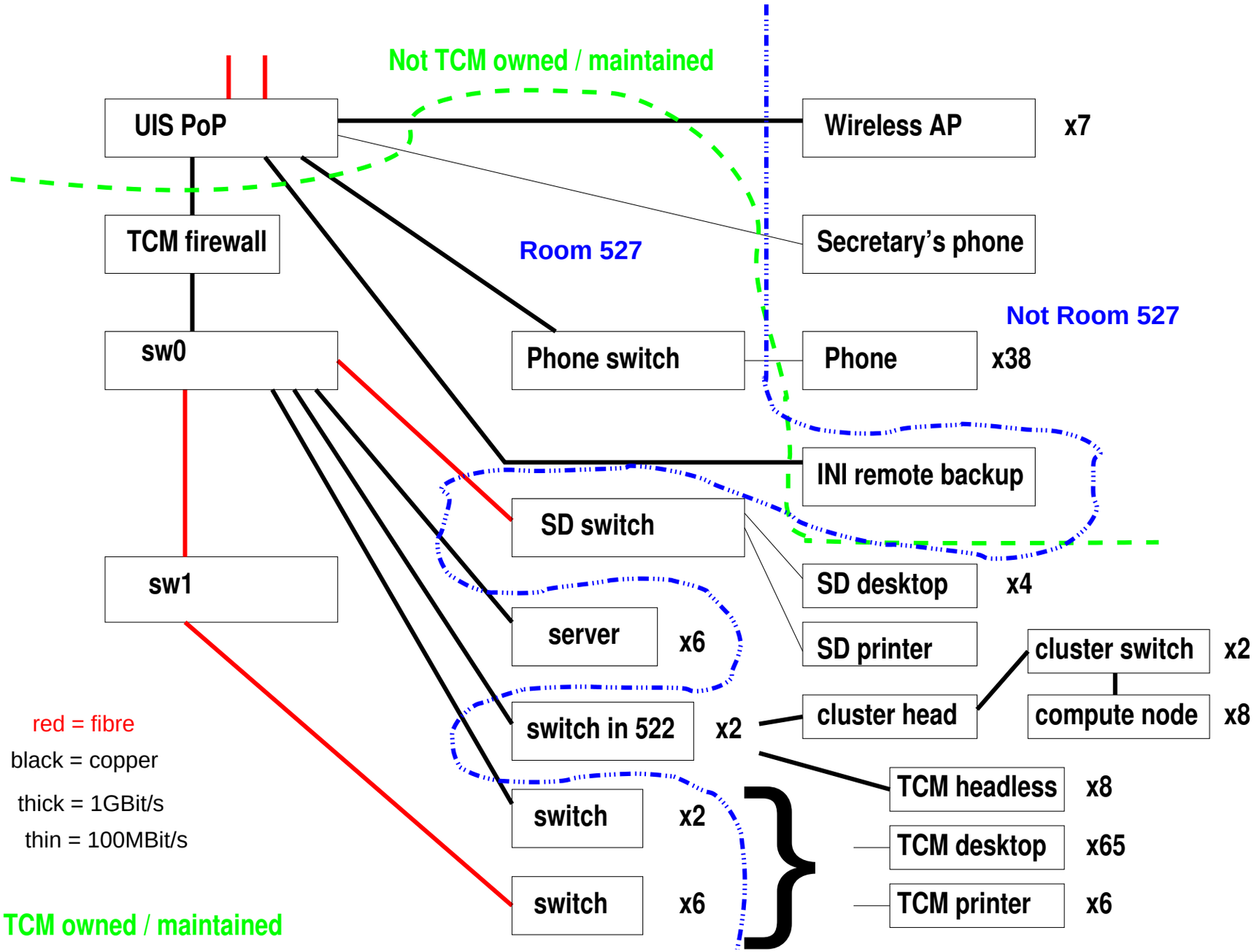
IT Services in TCM

23rd February 2015


Not TCM owned / maintained
UIS PoP
Wireless AP
x7
TCM firewall
Secretary's phone
Room 527
Not Room 527
sw0
Phone switch
Phone
x38
INI remote backup
SD switch
sw1
SD desktop
x4
server
x6
SD printer
cluster switch
x2
switch in 522
x2
cluster head
compute node
x8
TCM headless
x8
switch
x2
TCM desktop
x65
switch
x6
TCM printer
x6
red = fibre
black = copper
thick = 1GBit/s
thin = 100MBit/s
TCM owned / maintained

The PoP

The UIS owns the ‘Point of Presence’ which brings the internet to TCM. We pay around £2,600 a year for this service, plus a few pence per GB of network traffic.

The PoP has redundant links back to the UIS’s network, and the UIS provides its UPS which gives over an hour’s runtime in the event of power failure. Its link back to the UIS’s network runs at 1GBit/s.

(Our traffic charges are generally modest. The bill appears quarterly.)

Lapwing

The wireless access points which carry the Eduroam and Uni of Cam wireless services are owned by the UIS, and directly connected to its PoP. TCM is responsible merely for the cabling between the two.

They are powered by the PoP, so continue working for a while in the event of power failures, and survive power glitches happily.

(Although ethernet generally does not provide power, an optional extension allows for up to 13W to be provided. Cf. USB 2 at 2.5W, or USB 3 at 4.5W.)

Currently there are seven APs of three sorts in TCM, all supporting 802.11n, and the newest, of which we have four, also supporting .11ac. Their lights can be admired in order to see if they are happy.

Traffic over TCM’s access points is not charged to TCM.

Lapwing Lights

Green lights showing the presence of power (supplied over ethernet via a network switch capable of supplying 13W at a nominal 50V), wired ethernet activity (flickers), and that the two radios are working (b & g are 2.4GHz, a is 5GHz, n is both).

Lapwing faults are rarely TCM’s fault, and many cause the pattern of lights to change.

Penetration of our weak breezeblock walls is surprisingly poor. The signals travel in remarkably straight lines, and survive about three walls if penetrated in a perpendicular fashion.

Telephones: VoIP

One telephone, found in the secretaries' office, is connected directly to the UIS PoP.

The others are all connected back to a 48-port HP switch, which in turn connects directly to the UIS PoP.

This switch provides power over ethernet (PoE), and is itself connected to a UPS, so short power failures should have no impact. A failure of the switch itself would be embarrassing, as PoE switches are rare-ish – we have no spares, and it might take a couple of days to locate / procure a substitute. A couple of days without phones? Heaven! (Two spare power bricks for powering phones from a non-PoE switch exist in my office.)

VoIP cables in our patch panel are pink, and all pink cables are VoIP.

We have to pay the UIS £5 per month per phone, plus call charges. So the fixed charge for TCM is about £2,300 p.a.

Isaac Newton Backup

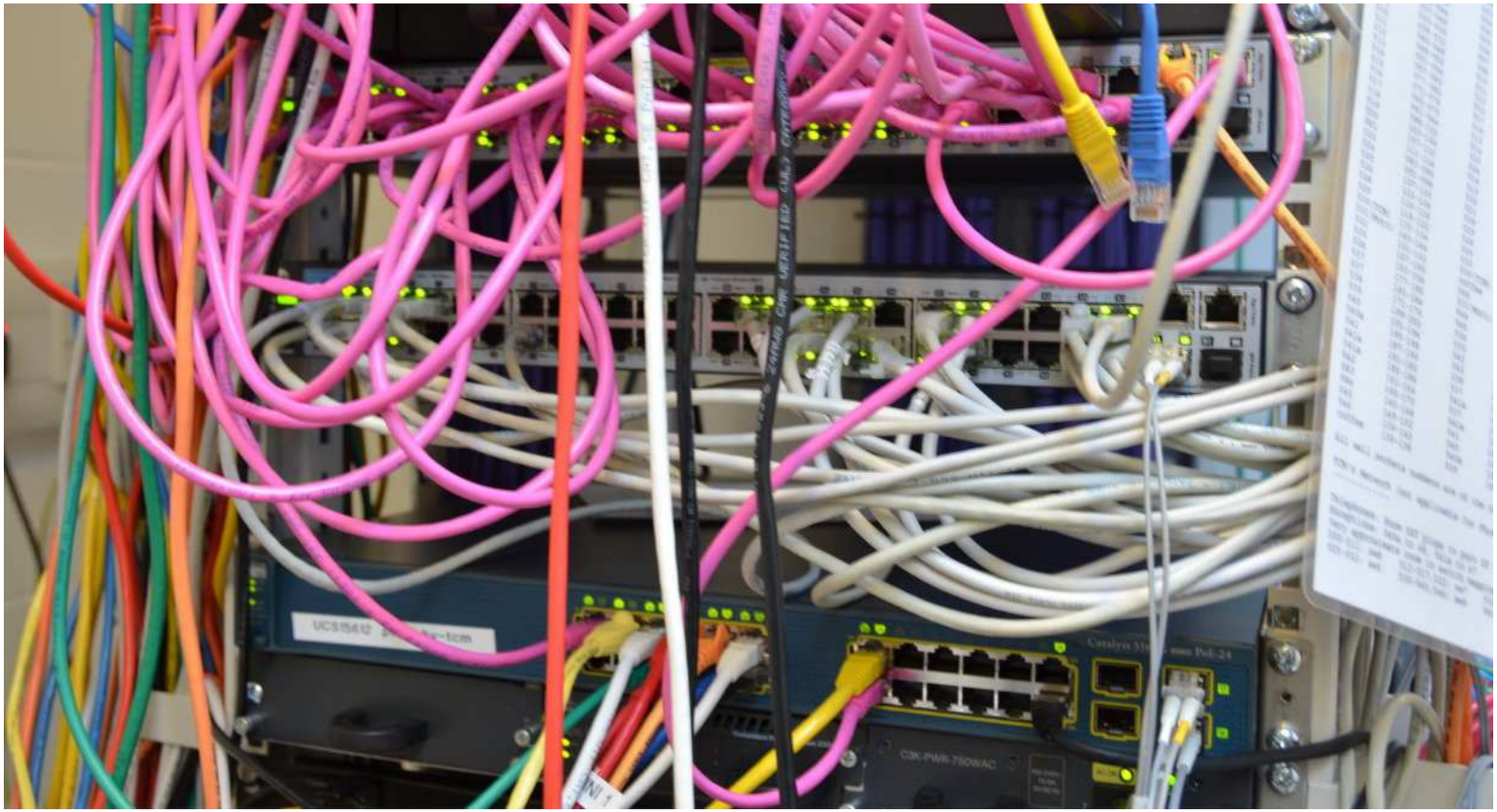
We host a backup system for the Isaac Newton Institute. This machine is physically quite large and heavy, and is directly connected to the UIS PoP. It has two network interfaces, one for the machine itself, and one for its ‘integrated lights out management’ device.

It uses the addresses `newton.tcm.phy.private.cam.ac.uk` and `newton-ilom.tcm.phy.private.cam.ac.uk`.

In return, TCM is able to put an offsite backup server in the INI machine room.

The INI’s machine is bigger than we were hoping / expecting. Its 75kg weight is a significant fraction of our 300kg maximum gross rack weight for floor loading considerations, and its 1.1kW rating, though no doubt containing a considerable margin, is also a significant fraction of the room’s capacity. It does contain 48 disk drives, which probably take over 10W DC each when writing, with the conversion from mains being less than 90% efficient. Add a couple of CPUs and some memory, and the total is quite close to 1kW.

The PoP



From the top, TCM’s VoIP switch with its pretty pink cables. The Mott network’s data switch with dull grey cables. The UIS’s PoP for TCM. On the far right a redundant pair of fibres connecting to the outside world. On the left direct connections for Lapwing, one telephone, the INI’s remote backup server, and to our firewall. Below the PoP’s redundant PSU.

To TCM

The link from the PoP to the rest of TCM first goes through a firewall. This firewall currently passes all traffic, but its presence means that we can easily block specific traffic should the need arise. It is merely a Linux desktop, the second one to perform this service, and it is rather elderly – a dual core 2.4GHz Core 2 with 4GB of memory. Its network interfaces are 1GBit/s.

There are many reasons why a more active firewall could be useful. In the first 12 days of January, the subnet 103.41.124.0/24 was responsible for over 300,000 failed password login attempts on each of our desktops…

Web Server and Firewall

TCM’s Core Switches

TCM’s network moved in 1999 to a collection of edge switches offering 24 ports of 100MBit/s each, connected to a central switch over 1GBit/s fibre links to form a ‘star’ configuration. In those days the 1GBit/s links had to be fibre.

That the links between the switches, which carry the aggregated traffic of many desktops, are faster than the links to the desktops, is important. If this is not the case, these uplinks become saturated, and the only response is to discard packets of data. Whilst internet protocols cope with packet loss, the coping can be slow and inefficient. For most situations, a 100MBit/s link with zero loss is much better than a 1GBit/s link with 50% loss.

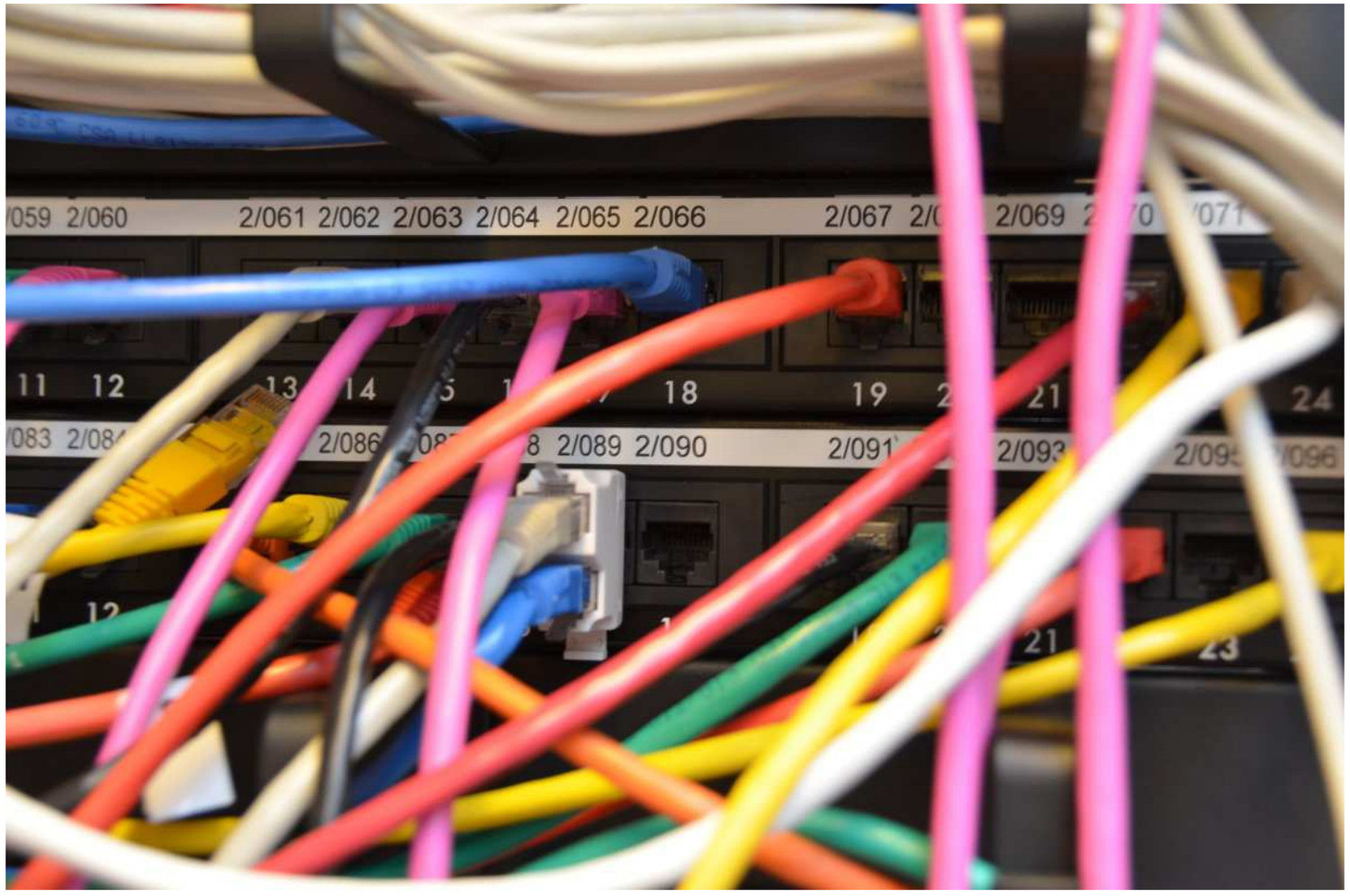
Further expansion has left the configuration in a strange double-centred star formation.

Patch Panel

Each office socket is joined to the socket in the patch panel with the same number. Depending on where the cable plugged into the patch panel socket goes, it could have VoIP, Lapwing, TCM's network (at 100MBit/s or 1GBit/s), the Mott network (i.e. QM), the CTA network (i.e. the Mott Seminar Room), or nothing connected to it.

In some cases nasty socket doublers are used. CAT5 cable has four twisted pairs of wires, and 100MBit/s ethernet uses just two. So one can run two separate 100MBit/s connections down a single cable. Unfortunately GBit/s uses all four pairs.

Patch Panel Closeup



Pink is telephone, so socket 65 is telephone, 66 & 67 TCM, 68 & 69 empty. Socket 89 shows a doubler in use.

Disk pack for old Sun server
Old Sun server (spare) `tcmsf2`
Old Sun server `tcmsf1`

New fileserver (`fs1`)
New fileserver (spare)(`fs2`)

INI backup server

TCM backup server (`backup`)

`ladon`
`henri`

`fs1` and `fs2`

Our main fileserver is `fs1`. This has two quad core 2.4GHz Sandy Bridge processors and 32GB of DDR3/1066 memory along with space for twelve hot-swap disks. The home directories are on a three-way mirror of three 3TB disks, and a two way mirror of 3TB disks is used for
`/usr/local/shared`. Software RAID is used for both mirrors.

The machine is also the mailserver for the `@tcm.phy.cam.ac.uk` domain, and serve TCM's password and group files.

It friend, `fs2`, is mostly identical and acts as a spare. Should `fs1` fail, there would be no automatic failover, by `fs2` has suitable hardware, and suitable data, to be configured to take over relatively rapidly.

scratch

This machine is responsible for `/rscratch`.

It is an 8GB quad core 2.33GHz Core 2 based system with a MegaRAID hardware RAID card. This has six 1.5TB disks in a 6TB RAID 6 array. It has a 1GBit/s link to the rest of TCM.

As there are no backups, space is cheap.

It accepts direct sftp connections, and one can log into it directly to perform a limited range of file manipulations. If copying data to it via one's TCM desktop, you will be limited to 100MBit/s, whereas direct connections (or via pc52) could be ten times faster.

For historical reasons (a full fileserver), it also hosts the majority of our centrally-installed software (about 160GB).

backup

Our local backup server is a 2.27GHz quad core Nehalem with 6GB of triple channel DDR3/1066 memory and six disk drives. Three 1TB drives form a 2TB software RAID volume and are used for the backup of home directories. These backups are exported back to the desktops so that anyone can browse them and recover files from them.

```
pc0:/backup/fs1$ ls
Dec  Jan  Nov  Sat  Sun      Sun_older   Thu  Wed
Fri  Mon  Oct  Sep  Sun_old  Sun_oldest  Tue
```

The precise date of each backup can readily be determined

```
pc0:/backup/fs1/Sun$ ls
2015-01-11_22:38:04
aa606
aam24
...
```

backup

The backup schedule, for those on the new fileserver, is that a script which runs nightly makes an incremental backup with `rsync`. It also keeps three extra weekly (Sunday) backups, and monthly backups for a year. Two offsite backups are maintained, one weekly, one monthly, on the same server and disks…

The nightly backups are also pushed to the local second fileserver, `fs2`.

A separate script makes a weekly backup of applications from `fs1`, and copies this to `fs2`, and once a month to the remote backup server too.

ladon

Ladon exists merely to host four virtual machines.

It is a quad core 2.67GHz Nehalem with 8GB of dual channel DDR3/1333 memory.

It uses the ‘kernel virtual machine’ VM system which has been part of the Linux kernel for several years.

`plover`: the Print Server

A virtual machine running on `ladon`.

Serves two laser printers in coffee area, two in room 523, one in secretaries' office, one in SD, and the A0 inkjet printer in 523.

Why have a print server? Mostly so that if a client repeatedly sends a job which crashes a printer, or causes it to spew rubbish, there is one place to go and look for the job and stop it.

Older printers also dislike dealing with large numbers of simultaneous connections. The print server won't care.

`lic`: the licence server

Some machine has to do the job of serving ‘floating’ licences. In TCM there are three major products which are licenced in this fashion: Matlab, Mathematica, and Intel’s compilers. This VM runs their licence servers.

(Some licences are fixed to a particular machine. Others use a network server which is entitled to issue licences to clients up to some fixed number. For software which could be run on any of our seventy computers, but is unlikely to be run on more than half a dozen at once, floating licences are usually much cheaper.)

cvs and svn

A pair of virtual machines running on ladon which provide an SVN-based repository service, complete with web interface. This is configured so that it is easy to add external users and to form collaborative projects.

cvs and svn

[check2xsf] Diff of /cell_read.c - Mozilla Firefox

[check2xsf] Diff of /cell_...

src.tcm.phy.cam.ac.uk/viewvc/mjr/check2xsf/cell_read.c?r1=30&r2=32

bmp2eps

/ [check2xsf] / cell_read.c — Logged in as: **mjr19**

Diff of /cell_read.c

viewvc

Parent Directory | Revision Log | Patch

#	revision 30 by *mjr*, Thu Jul 17 20:41:18 2014 UTC	revision 32 by *mjr*, Sun Jul 20 19:57:48 2014 UTC
	Line 214 void cell_read(FILE* infile, struct unit	**Line 214** void cell_read(FILE* infile, struct unit
214	}	}
215	kp->n=i;	kp->n=i;
216	}else if(!strcasecmp(ptr,"symmetry_ops")){	}else if(!strcasecmp(ptr,"symmetry_ops")){
217	sym_mat=malloc(48*9*sizeof(double));	sym_mat=NULL;
218	sym_disp=malloc(48*3*sizeof(double));	sym_disp=NULL;
219	if((!sym_mat)||(!sym_disp))	for(nsym=0;;nsym++){
	error_exit("Malloc error for sym ops");	
	for(nsym=0;nsym<48;nsym++){	
220	cellreadline(buffer,LINE_SIZE,infile);	cellreadline(buffer,LINE_SIZE,infile);
221	if (!strncasecmp(buffer,"%endblock",9)) break;	if (!strncasecmp(buffer,"%endblock",9)) break;
222		sym_mat=realloc(sym_mat,(nsym+1)*9*sizeof(double));
223		sym_disp=realloc(sym_disp,(nsym+1)*3*sizeof(double));
224		if ((!sym_mat)||(!sym_disp)) error_exit("realloc error in cell_read");
225	for(i=0;i<3;i++){	for(i=0;i<3;i++){
226	if(sscanf(buffer,"%lf %lf %lf",sym_mat+9*nsym+3*i,	if(sscanf(buffer,"%lf %lf %lf",sym_mat+9*nsym+3*i,
227	sym_mat+9*nsym+3*i+1,sym_mat+9*nsym+3*i+2)!=3){	sym_mat+9*nsym+3*i+1,sym_mat+9*nsym+3*i+2)!=3){

Colored Diff | Show

Legend:
Removed from v.30
changed lines
Added in v.32

ViewVC Help

Powered by ViewVC 1.1.20

web

This is a physical machine, and serves multiple site using HTTP's virtual hosts. (HTTP requests contain the host name of the intended target, and the web server can modify its output accordingly.)

It is old, and runs on a 100MBit/s link, as we do not want the web server taking the full bandwidth of our network connection.

It imports directories read-only from our main fileserver. If it were compromised, it is intended that the result is merely embarrassing, rather than catastrophic.

Though it does not support CGI, PHP or https, it does support SSI and also Raven protected pages.

henri

Henri is a spare server. This is actually very valuable – it ends up doing all sorts of things which do not really have another home to go to. It was purchased for a project for which it proved unsuitable (mea culpa), but has become quite useful.

It acts as a DHCP server for those using wired notebook connections, and is the SuSE network install server which has installed every machine in TCM. It performs monitoring of network traffic volumes, and runs the backend of the Onetep wiki. It has two 2.13GHz quad core Nehalems, with 24GB of memory. Overkill for its current usage, but still less GB and GFLOPS than our best desktops.

UPSes

UPSes

As heavy as lead. On the left is an extra battery for extended run-time, and on the right the UPS itself, the top part being the electronics, the bottom part being battery.

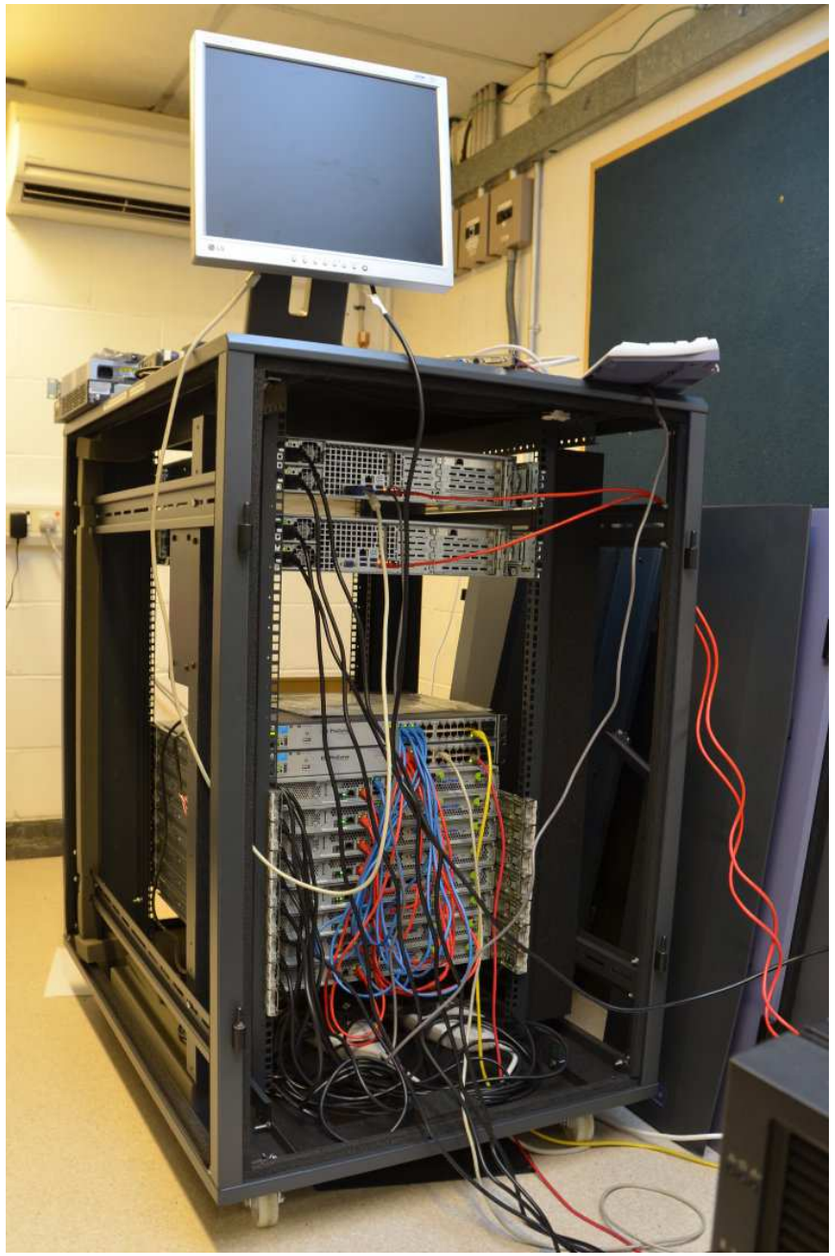
The battery is an APC RBC11, four 17Ah 12V sealed lead acid batteries, weighing about 25kg between them. The UPS thus runs on 48V DC, and is rated for up to 2.2kW, although we load them less and therefore enjoy longer run-times. We hope for at least 30 minutes, and should get well over an hour if the batteries are newish.

Cluster Room

This room contains both TCM's small cluster (eight dual socket Nehalem nodes in a rack, plus a floor standing head node), and also various dual socket headless machines purchased on various specific grants. It often has the odd spare PC in case one is needed in a hurry.

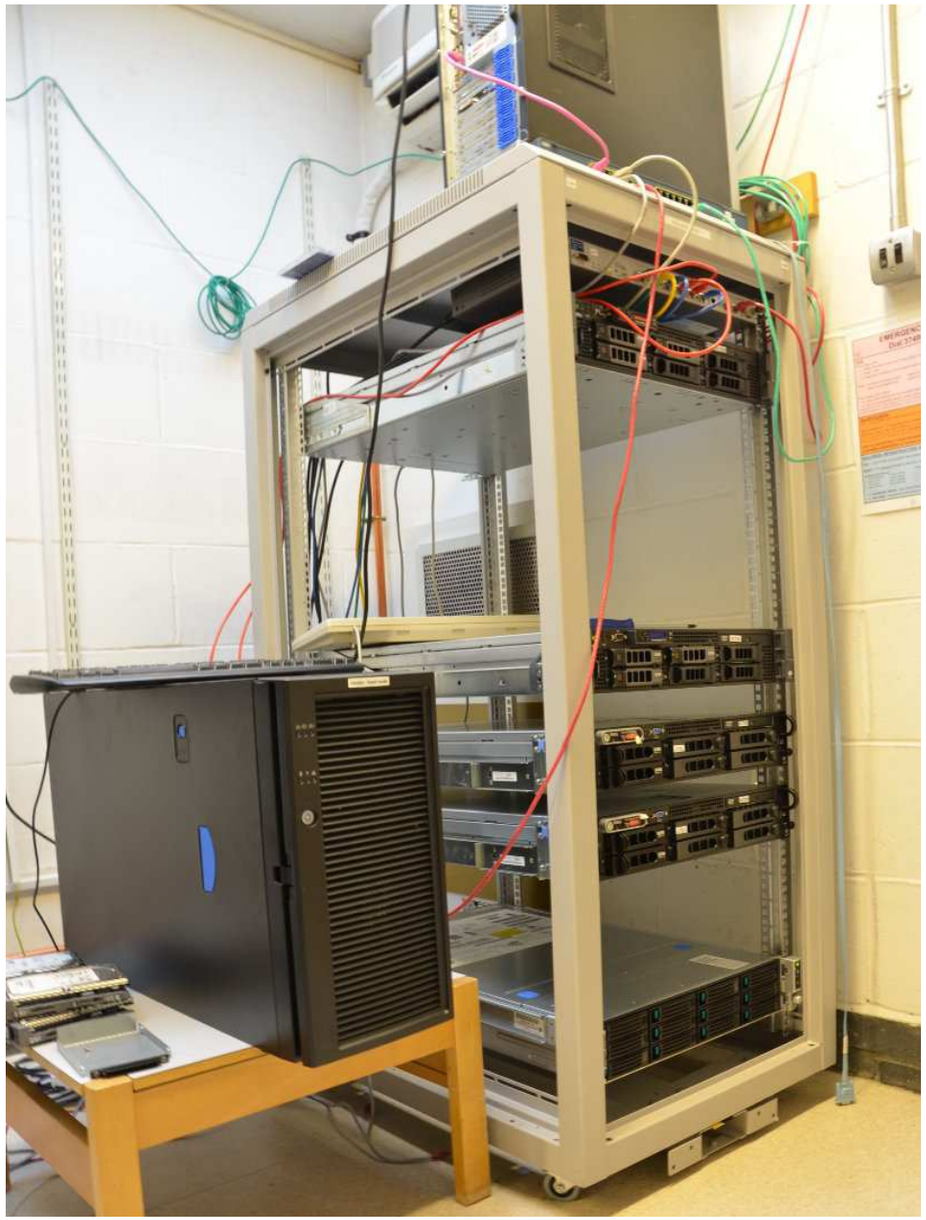
It is connected back to the server room by CAT5e cables training through the ceiling and running at 1GBit/s. It contains two 1GBit/s switches, both repurposed from elsewhere.

The Cluster



From the bottom, eight compute nodes, two switches (one NFS, one MPI), a large gap, then `s4` and `s5`. The air conditioner on the far wall blows cold air down onto the fronts of these machines, and the picture shows the backs. The corner of the cluster's head node / fileserver is visible in the bottom right.

Not the Cluster



On the low table to the left, the head node of TCM's cluster.

In the rack, `s1` to `s8`, save `s4` and `s5` which escaped to the cluster rack. (Photo taken before `s2` arrived.) More machines expected…

At the top of the rack, an eight port GbE switch for the 's' machines.

On top of the rack, a second GbE switch for the cluster's head node and any odds and sods. Also a spare desktop PC (`pc57`) as an example of an odd or sod.

SD

Downstairs in SD we have an old 3COM switch with a fibre link back to TCM, an office of PhD students containing three Linux PCs and one old B&W laser printer, and a further PC is in the Head of Group’s office.

Regrettably one of the PhD students’ PCs has a Windows VM.

Your Desktop

Connected with a 100MBit/s link, your desktop was purchased as a relatively high end machine suited for running calculations. It may have aged ungracefully since. It will have error detecting memory, and a reasonable amount of local scratch disk. It will have a reliable, silent, passively-cooled GPU.

It will be running OpenSuSE 13.1 with a lot of additional software, some commercial (Matlab, Mathematica, Intel compilers, . . .), some free but not distributed by OpenSuSE (Open64 compilers, VMD, Jmol, Skype, . . .), some updated versions of things SuSE does distribute (Libreoffice, valgrind, numpy, gnuplot. . .).

It includes most of both the Gnome and KDE application suites, and things get added when their absence is pointed out.

It can mount USB memory sticks. It might, or might not, be able to play audio or read/write DVDs. It will work with USB webcams.

and its c. 65 friends

There are probably bigger computers in TCM than your desktop. Certainly the combined power of the others outclasses your desktop.

That is why they are set up to make it easy to check what jobs are running on the others (with commands such as `rbusy`), and to log in to other computers and run jobs on them. Property is theft, and (almost) all is held in common.

A pair of computers in 523 have a particularly interesting selection of bells and whistles – both have scanners, one has an HD webcam with microphone (removeable), both have working audio, one (`pc52`) has a 27″ monitor (2560x1440), updated GL libraries, a 1GBit/s ethernet connection, and a ban on long-running jobs to make it particularly suited to interactive work.

There is also a colour photocopier in 523, an A3 colour laserprinter, and an A0 inkjet printer.

Some Data Flows