# A Short History of the Data Network
# in the Department of Computer Science
# and the School of Informatics
# at the University of Edinburgh
# from the 1970s to 2022

George D M Ross
`gdmr@inf.ed.ac.uk`, `gdmr@dcs.ed.ac.uk`

October 11, 2022

This document relates the evolution of the data networks in the Department of Computer Science and the School of Informatics at the University of Edinburgh. It attempts to say why things ended up the way they did, as well as looking at what emerged. Note that diagrams may be simplfied slightly for clarity, though significant details are included. Related documents can be found on the Edinburgh Computer History site: `<https://history.dcs.ed.ac.uk/>`.

The Author was a member of the Department/School from 1974–1976 and from 1979–2022.

## 1 Departmental Link Boards

A significant step towards the development of a high-speed data network in the Department of Computer Science was the design in the early 1970s of the "Link" board. This provided independent RX and TX serial connections, often but not always used as a pair. The Links ran at up to 1Mbps over standard $75\Omega$ coax cable. Each connection was 8-bit clean and flow-controlled, with a positive ACK signalled to the sender for each octet transmitted. Implementations existed for the Interdata backplane, the DEC UNIBUS (PDP-11 and VAX-11) and the DEC PDP-9 and PDP-15 machines.

By 1975 the Department had two Interdata clusters interconnected by Links, as well as a PDP-8, PDP-9 and a PDP-15/20. Machines were gradually added, so that by about 1980 the following were interconnected by Links (see figure 1):

- the "Filestore," an Interdata processor with two 70MB disc drives and a number of Link boards, the purpose of which was to act as a central file storage repository for all the Department's other small machines

- the Interdata clusters

- the PDP-9 and PDP-15/20, now using the Filestore as virtual DECtape drives, as this turned out to be faster than using real physical hardware

- a PDP-15/40

- an ICL 7502 discless editing station, which ran VECCE and very little more

- the "LEGOS" systems: standalone Interdata machines with rather more memory than the small cluster machines
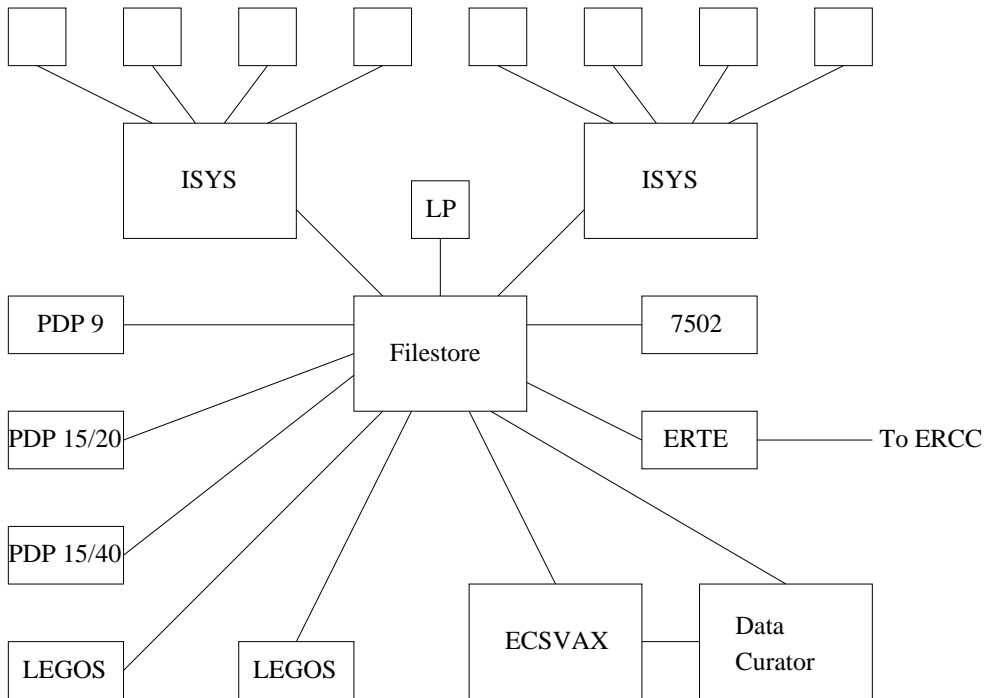
Figure 1: Network diagram from around 1980

- the Edinburgh Remote Terminal Emulator ("ERTE"), a research project into performance testing

- a DEC VAX-11/780 "ECSVAX" running VMS, which was the Department's main multi-user machine at the time

- the "Data Curator": a Perkin-Elmer 3220 used as part of another research project

- a lineprinter.

The Filestore was the hub of the network, and almost all traffic over the Links was filestore protocol (described elsewhere[1]), either directly manipulating the files in the case of discless machines, or to transfer them back and forth in the case of ECSVAX and the Data Curator. There were only two real exceptions to this.

One of these was the lineprinter. This was provided with a Link receiver (no need for a transmitter) and connected to its host by this means. Initially this was to one of the Interdata clusters, and latterly to the Filestore. The Link's inherent flow control meant that the despooler could simply throw characters down the line as space in the printer's buffers permitted.

The other exception was that there was a direct connection between the Data Curator machine and ECSVAX. This was originally intended to allow direct data transfer between the two machines. In fact, it was not used for this purpose, with file transfer being staged via the Filestore instead, but it did run a home-brew bidirectional remote login protocol for a while.

## 2   2Mbps Departmental "Ethernet"

There were a couple of drivers for the development around 1980 of the Computer Science Department's 2Mbps ethernet-like network. The Metcalfe & Boggs paper[2] had come out, showing what

---

[1] For example, `https://history.dcs.ed.ac.uk/archive/os/APM-filestore/FS.1976/FSDOC.LAY.html`

[2] Linked from `https://en.wikipedia.org/wiki/Ethernet`

could be done; and the sheer number of connections back to the Filestore were becoming unmanageable — indeed, latterly there was a multiplexor machine attached to the Filestore, mainly to provide additional connectivity.

The Departmental ether was different in a few ways from standard ethernet: it ran at 2Mbps rather than 10Mbps; it had a slightly different addressing structure; and by default it presented as a *reliable* datagram service between configurable end-station ports, in contrast to standard ethernet's unreliable datagrams between stations, which allowed filestore protocol to be layered directly on top. Ether stations were connected to their hosts (Filestore, LEGOS, ECSVAX) over serial Departmental Links, now 9-bit-clean and running at 2Mbps, and were designed to offload the ether protocol processing from the hosts. Like standard ethernet, the Departmental ether ran over 50Ω yellow "thick ether" cable, and used the same "vampire" clamps, though interfaced through purpose-built transceivers.

The arrival of 32-bit microprocessors, in particular the Motorola 68000 series, prompted several members of the Department to construct an experimental testbed, known as the "Fred Machine." This was a modular system, with processor cards, memory cards and graphics cards all plugging into a common backplane. In particular, there was an ether card, with a revised version of the Departmental ether station. These machines proved so successful that they were relabelled the "Advanced Personal Machine" ("APM") and gradually rolled out as the Department's main workhorse for staff and students.[3]

Unfortunately, the original version of the ether station as used by the Filestore turned out to be under-buffered, and unable to cope with the increasing load. Packets would be dropped quite frequently, resulting in broken connections, rebooted machines and lost work. It was decided to reimplement the Filestore using a 160MB Winchester disc attached to a Fred machine. This took place from 1983 onwards, in two phases:

1. as the Interdata Filestore's ether connection was too unreliable for the intensive traffic involved alongside its normal load, the filestore protocol interpreter was initially implemented on ECSVAX, calling out to VMS's native filesystem; then

2. having uploaded an image of the necessary file structure to the ECSVAX "filestore," the APM-based Filestore was developed using the relatively more reliable facilities it provided,[4] porting as much code from the ECSVAX version as possible.

Once the APM-based Filestore was brought into service, the original Interdata Filestore was retired, and run as a printer-despooler driving the original lineprinter over its Link interface. In addition to the 160MB Filestore, two 300MB Filestores were deployed, one mainly for staff use and the other mainly for student use, along with four based on 70MB drives for development and small projects.

# 3 10Mbps Ethernet

Economies of scale kicked in and researchers started to buy machines, mostly from Sun Microsystems, rather than designing and building them in-house. These machines came with standard ethernet, as had lfcs's VAX-11/750 and Pyramid machines ("cstvax" and "cheops"). A LANCE-based[5] ethernet card was built for the APMs and installed in a couple of the Filestores. Initially 10base5 ("thick ether") was used, but as clusters of Suns appeared in an *ad hoc* manner they were linked with segments of 10base2 ("thinnnet") and routed by adding second interfaces to one or more machines in each cluster.

The Department's machine halls at the time also housed a Gould PowerNode ("itspna"), shared with the Artificial Intelligence ("AI") and Electrical Engineering ("EE") Departments,

---

[3]https://history.dcs.ed.ac.uk/archive/scans/gordon/apm/

[4]The initial implementation used what was essentially the same filesystem structure as for the Interdata Filestore. The ECSVAX version showed, however, that the protocol could cope with a more complex filesystem, and a hierarchical version was developed later.

[5]https://en.wikipedia.org/wiki/AMD_Lance_Am7990

Filestore    Filestore    APM    - - - - - - - - - - - - - - -

"2base5"

Filestore    ECSVAX    Gould    GEC    **To EE**

10base5

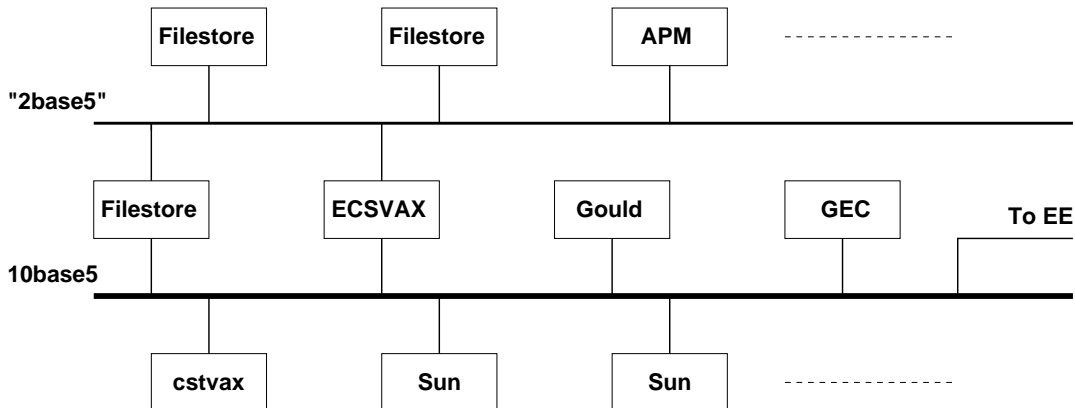cstvax    Sun    Sun    - - - - - - - - - - - - - -

Figure 2: Network diagram from around 1986

and a couple of GEC machines for ERCC's commercial arm.[6] AI accessed the Gould by serial lines connected to their South Bridge site using a statmux pair and a BT kilostream link, while EE had their own ethernet bridged to ours over a fibre pair run through one of the KB site's heating ducts (figure 2). This link in fact remained in service until building work in 2000 meant the destruction of the fibre.

Although there was this common infrastructure, there were actually at first a few separate islands of connectivity: all the BSD-based machines spoke TCP/IP; the GECs spoke Coloured Book protocols, for political reasons; and the Filestores and ECSVAX initially spoke neither. A TCP/IP stack was therefore written for these latter machines, as this suited our requirements better than Coloured Book would have. It included RIP-based routing, rlogin and rcp clients and servers, and bidirectional lpr-based printing.

Though perhaps not strictly part of the data network, it's worth mentioning that at this time the Department also ran serial lines to offices and a terminal room in JCMB-2501 from ECSVAX, cstvax, cheops and itspna using a Gandalf multiplexor.

Gradually as the APMs reached the end of their service life they too were replaced by Sun workstations, and the thinnet structure was necessarily rationalised. Discless Suns were heavy network loads, and so the network was explicitly subnetted into several staff segments and several student segments, with the Sun fileservers doubling up as routers. The "CS1 lab" in Appleton Tower was initially separate, before it too was linked to the rest of the Department at JCMB, first using UUCP and then SLIP over a 4800bps EPS8 BT line (fortunately both sites were within the same 031-667 exchange area). All subnets were /24, and routing information was distributed using Routing Information Protocol ("RIP").

## 4 Wider: RCOnet and EdNet

The wider University network is described elsewhere,[7] so only those aspects which were CS-related are mentioned here.

For the first period of time covered by this note (mid-1970s to roughly 1980) the Department used serial terminals attached to the University's multi-access EMAS systems for its teaching lab in Appleton Tower, until that too was replaced by Sun workstations. At JCMB there were also EMAS terminals in staff offices and Block O/P, but the Departmental and ERCC networks were initially almost completely separate. The only connection, such as it was, was the ERTE system, which could be used for small-scale file transfer between EMAS and the Filestore when it wasn't required for its research project work.

---

[6]One of these eventually became the original "tardis" machine. See http://www.tardis.ed.ac.uk
[7]https://www.wiki.ed.ac.uk/display/iticis/University+of+Edinburgh+Networking+Overview

When the ERTE project finished the hardware was mostly disposed of. Its HDLC interface (a DQS11) was, however, redeployed in `ECSVAX`, with a locally-written driver and protocol modules (written by Gordon Brebner and Paul McLellan[8]). This provided a full set of remote login, file transfer, printing and email applications.

`cstvax` also had a connection to EdNet, using an LSI-11 as a front-end (a "York Box"), used mainly for FTP and mail once the `ECSVAX` TCP/IP stack had been developed.

# 5  TP and the First Great Rewiring

As the number of Suns in the Department increased, it was necessary to control the allocation of machines to subnets ever more carefully. Discless Suns paged heavily, and the applications used by some groups of users were quite network-intensive. As a result, we had numerous 10base2 segments strung around the various parts of the Department's accommodation to spread the load.

Physically, 10base2 wasn't very robust. It was disruptive to add new connections, as the cable had to be cut and connectors carefully fitted. Connections were then subject to strains as machines were moved and feet tangled in the cables. Corrosion set in (not helped by drips from wet bicycles parked in offices!). Hum-loops could be created as connectors came in contact with building metalwork. Faults might affect everyone on the wire, or just the odd machine. Debugging was tedious and time-consuming, and could often "fix" problems for no obvious reason. Finally, we had too many subnets in use to be able to run them physically through all the offices, simply because of the size of the cable bundle, which put a severe constraint on the allocation of people to offices!

By 1991 we had concluded that we couldn't carry on with 10base2. We decided to have the First Great Rewiring, introducing 10baseT on cat-3 cables. Physically this is much more robust than 10base2, and gives better, but not perfect, fault isolation. Crucially it made it possible to patch any office to any subnet (figure 3), considerably easing the room-allocation problems of the admin staff. It was also now even easier to split subnets to add bandwidth, and eventually we had around 25 subnets in use for various purposes. We were back to point-to-point connections, which had proved restrictive with the Links and the Filestore, but we weren't now limited by backplane space as it was easy to add new hubs as required. We had active electronics instead of a passive piece of wire, but the overall robustness more than made up for the added complexity. Wiring was installed by the Department's technicians, with one hub site in an alcove in room JCMB-2501, covering the offices, and one in the South Machine Hall covering the servers. A few hubs were also spread around under the Machine Halls' floor for the students' client machines. At its peak this network had over 30 hubs in 2501 and another 15 or so in the Machine Halls.

This was a considerable improvement over what we had had before, but there were still issues. Although we had individual connections from machines and offices to the hubs, those hubs were interlinked using 10base2 (figure 4). We still had a shared medium — hubs are simply repeaters, and copy everything they receive out onto all their ports — and bandwidth was still under pressure. The hubs were connected together using the 10base2 cabling, which was rather old and fragile. And although faults were generally much more contained, hub faults could affect everyone on an entire subnet, often in strange ways, and were hard to diagnose.

And we now had the issue of how to manage the large number of connections at the 2501 hub site, particularly as people moved offices or had new machines to connect. Churn resulted in the 10baseT patch areas getting completely out of hand over time, so that on a couple of occasions we just ripped the whole lot out and repatched again from scratch.

The design of the network at this point did still cause some practical problems. We had partitioned functionally into subnets ("backbone," staff, students, various research groups, etc), each with several fileservers also acting as routers. Although staff would generally use their office workstations, and students the lab workstations, there was a requirement that anyone should in principle be able to use any workstation. However, because we had bandwidth constraints we

---

[8]Initially the driver kept crashing the machine for no obvious reason. Eventually after rummaging through the fiche Paul realised that the IPL being used by the DQS11 didn't actually have any handler assigned to it by VMS.
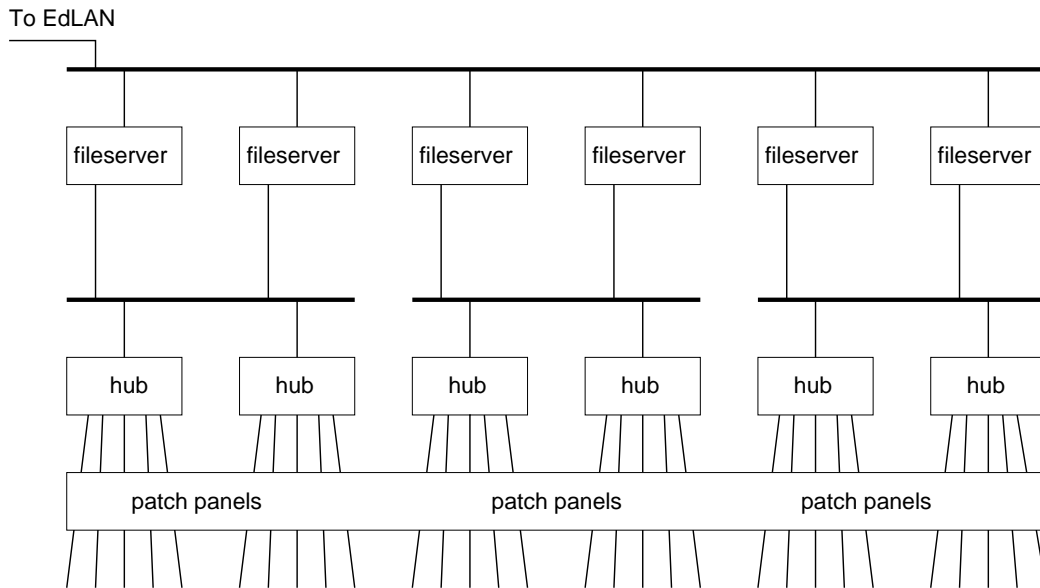
To EdLAN

fileserver fileserver fileserver fileserver fileserver fileserver

hub hub hub hub hub hub

patch panels patch panels patch panels

Figure 3: Outline diagram after the First Great Rewiring

10base2

Hub Hub Hub Hub Hub

10base2

Hub Hub Hub Hub Hub

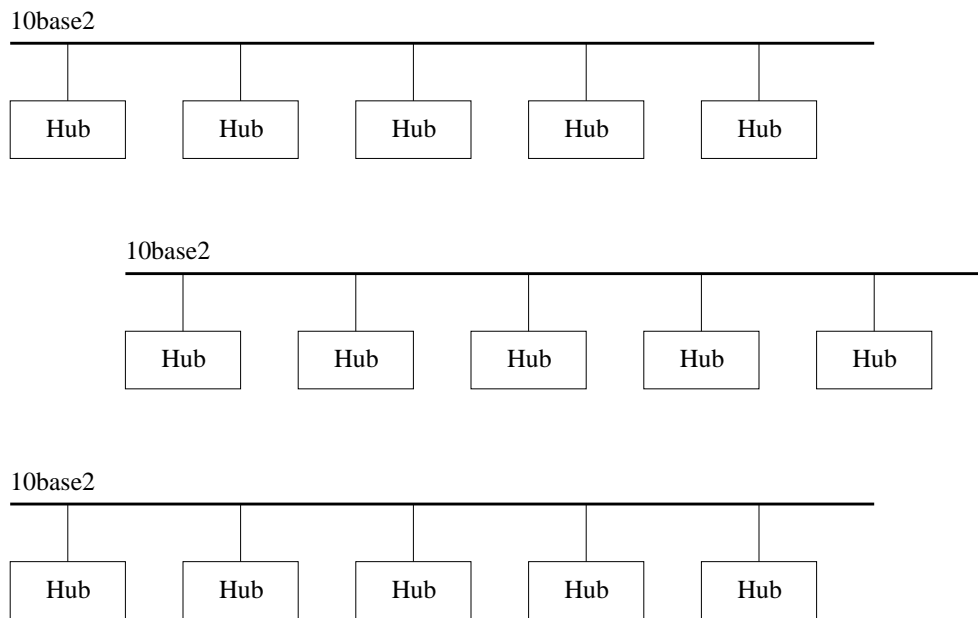10base2

Hub Hub Hub Hub Hub

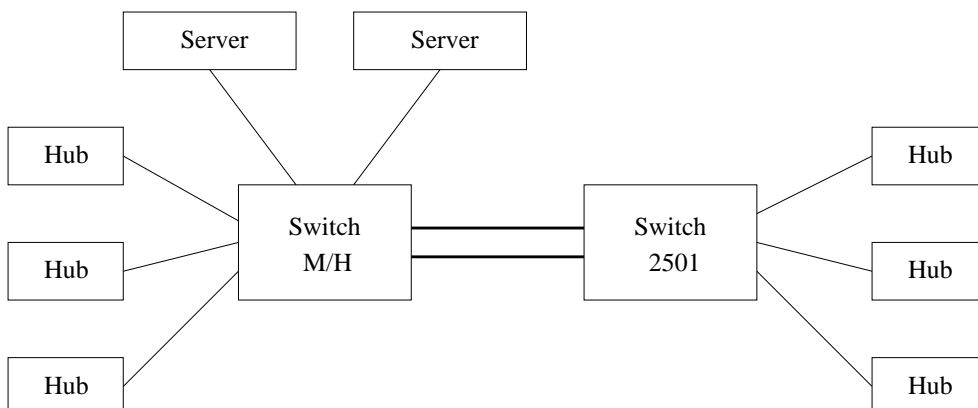Figure 4: Hubs connected by 10base2

Figure 5: Hubs interconnected using switches

had to make sure that packets were routed as efficiently as possible. In particular, we couldn't afford to have the `amd` automounter mount a filesystem from "the wrong side" of a fileserver. The automounter maps were therefore constructed in such a way that each specific mount was made conditional on the source address of the machine making it. This made them much more complicated to construct and maintain.

# 6   EdLAN

When EdNet was replaced by EdLAN in 1990 the Department connected both its JCMB and Appleton Tower sites, and retired the SLIP line as soon as we could. Up until then we had been using IP addresses chosen somewhat arbitrarily,[9] we but renumbered both sites onto the officially allocated addresses as they were connected to EdLAN.

We had been using DNS for our own purposes before this, but now set up a nameserver to provide part of the University's external provision.[10] We also renamed ourselves from "`cs`" to "`dcs`" to avoid the ambiguity caused by there now being two different name orderings in use (`cs`.ed.ac.uk *vs* uk.ac.ed.`cs`) and .`cs` being the country-code for Czechoslovakia.

# 7   Switches and the Second Great Rewiring

The advent of switches (marketing-speak for multi-port bridges) opened up a couple of avenues for improving the network. These were intelligent devices (well, slightly...) with much better fault diagnosis and reporting than was available from the hubs. They also implemented VLANs, allowing for several logically-separated subnets to be carried over the same physical wiring, and could group several links together into one "fatter pipe" to increase aggregate bandwidth.

We introduced a stacked pair of 3Com switches in the 2501 hub area and another pair in the machine halls server area. Hubs were now connected as individual leaf nodes off the switches, immediately reducing contention for bandwidth, and inter-switch traffic was carried over dual 100baseTX links (figure 5). Faults now had a much more localised effect, and the switches' diagnostic tools allowed for faster resolution and better forward planning. Once again we had complex electronics replacing simple wires, but the overall result was generally much more reliable. However, the patch areas were still completely out of hand!

Although the 3Com switches improved things, there were still a few rough edges. The firmware didn't initially support VLANs; and the switches would sometimes crash and lose their configu-

---

[9]Actually, based on the ASCII codes for 'E' and 'D' with 128 added, to give a range of class-C addresses: 197.196.N.N, now assigned to an Egyptian mobile phone network.

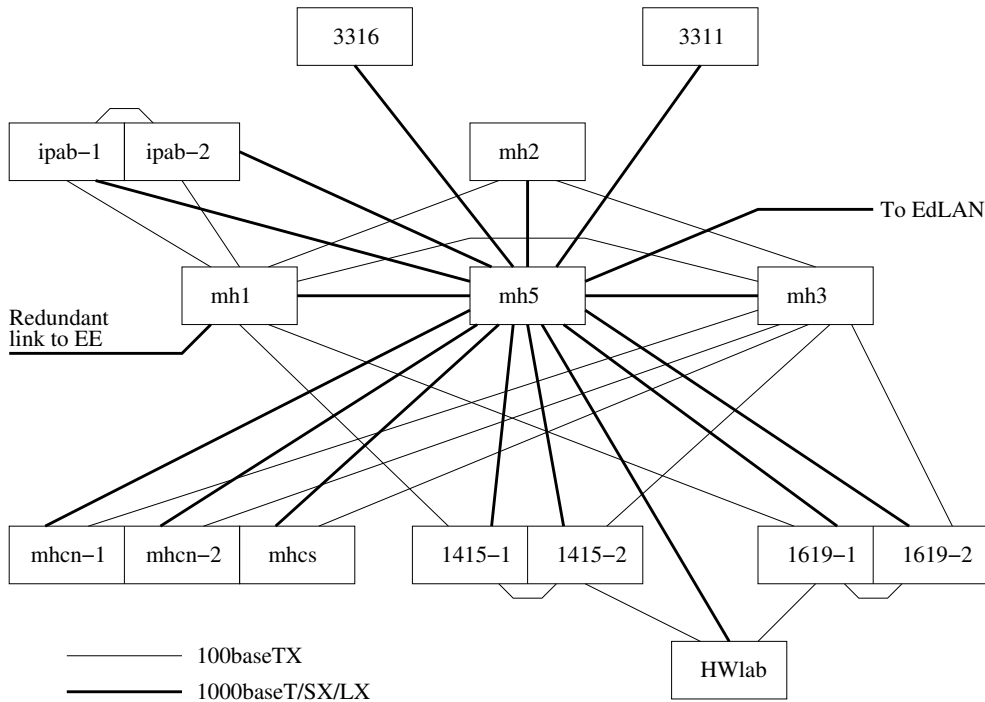[10]We still operate three of the University's external nameservers: `dns[012].inf.ed.ac.uk`

Figure 6: JCMB network using switches throughout

rations, though unstacking them did help with this. The biggest issue, though, was that we still had servers connected to shared segments at 10Mbps, albeit now at least single-hub segments.

The introduction by HP of their Procurve 4000M modular switches, at an affordable price and with a lifetime warranty, allowed us to rethink our network connectivity. Each machine could be given its own private 100Mbps link to one of the switches, and the switches themselves could be connected together using 1Gbps fibre and copper links. Ports could be assigned to VLANs based on configuration files, so there would be no more need ever again for physical repatching and the mess that resulted. Redundant links could be added, under Spanning Tree Protocol control.

Unfortunately, in order to make use of these opportunities we had to undertake a Second Great Rewiring (in 1998). Our initial TP cable installation had too many over-length lines, which passed the 10baseT electrical requirements but not the 100baseTX timing requirements. We had also doubled up many connections using the "spare" cable pairs, which 10baseT could tolerate but which caused unacceptable crosstalk for 100baseTX. Finally, most of our initial cable installation was cat-3. Once again the Department's technicians installed the new cat-5 cable, now terminated in two sites (rooms 1415 and 1619), and a pair of 4000M switches was installed at each. Further switches covered the teaching machines in the Machine Halls, and three more switches provided "core" connectivity as well as links to the servers. Most users had two outlets installed for them, with mini-hubs being used where more machines were needed, the realisation that the panels above the JCMB office doors were asbestos-based having meant that only existing holes could be reused for wiring! Though expanded later by the addition of the IPAB robot labs and some student workrooms, this design remained essentially unchanged until we moved from JCMB to the Informatics Forum in 2008 (figure 6).

Initially we didn't entirely trust this new hardware, so we ran one of the server-area switches as a hot-spare, with a shadow configuration. Later we concluded that this wasn't really necessary, and some of its ports were put to "real" use. We did, however, still try to make sure that we had enough spare ports overall to carry all the most important servers if necessary.[11]

Routing was still initially performed by the fileservers, and then by dedicated Suns with multi-

---

[11]We use bonding to achieve resilience these days.

port 100baseTX cards. For some reason these believed by default that all interfaces on a machine should use a common MAC address, which did not interact well with switches which used a common forwarding table for all their VLANs and believed that a MAC address should only be in use on one port. This was solved by running a script early on in the boot sequence to reconfigure the Suns to use separate MAC addresses for their interfaces.[12] (The Linux routers we run now do not have this quirk. Instead they believe by default that any IPv4 address configured on a machine can be used on any of its interfaces, which strangeness we also have to reconfigure away early in the boot sequence.)

Eventually we added a ProCurve 5308xl at the core, able to route at wirespeed, in order to handle increasing traffic levels. As a beneficial side effect of this, the fact that we didn't have to bother about which "side" of a fileserver we were speaking to from any given machine meant we could considerably simplify our automount maps. This marks the point at which we could really start to decouple host configuration from the ever increasing complexity in our network topology, and is definitely one of the more significant moments in the development of the Department's network.

## 8  Configuration and Monitoring

While we still had only a small handful of switches connected as a backbone for the hubs, it was reasonable to configure them using their CLI or web interfaces. Once we had decided to roll out "switch ports for all," however, it was immediately apparent that some kind of configuration tool would be needed to allow us to devolve sufficient network management features in a simple way. The requirements were that it should support both our 3Com and Procurve switches with a common user interface; and that interface should expose only those items which were necessary, and in such a way that a simple subset could be made available for the technicians and support staff to allow them to make routine changes. The tool should run on a commodity machine, with the configuration stored in its usual filesystem and backed up in the usual way.

At the time there wasn't a suitable tool commercially available, so we developed our own,[13] based around Tcl, `expect`, and the Tnm SNMP package. A Procurve or 3Com switch would be queried using SNMP for its current state, which would then be compared against the switch's configuration file and any updates applied as required. For the Procurve switches those updates could also be done using SNMP. The 3Com switches, however, had such a baroque VLAN implementation that it was simply not safe to try to use SNMP there too, and so updates for those were applied by sending commands to the CLI using an `expect` script. The tool was deployed using the "`rfe`" remote file editing tool and `make`. CogSci also used the tool, with a front-end to generate the configuration files. (A 2009 Development Project took another look and still couldn't find anything more suitable for our use than this in-house tool. It is, however, expected that things will change, for the edge at least, once the new EdLAN is finally rolled out.)

We also needed better routine monitoring, for error and basic statistical trend reporting. Again, this was written in-house, using Tcl and Tnm. All the switches are polled every 5 or 6 minutes, and their traffic and error counters extracted for every port, whether configured or not. These are stored using `rrdtool` for future use. Index files are generated for configured ports several times per day, as required, and the graphs are available as necessary using cgi-bin script wrappers around `rrdtool`'s facilities. `rrdtool` is also used to produce temperature and power-utilisation graphs, as well as composite network graphs based on several ports' data.

## 9  Traffic filtering

While we had an isolated data network, the issue of allowing or blocking traffic from outside didn't, of course, arise, though not all our internal users were as benign as might have been desired. When

---

[12]https://history.dcs.ed.ac.uk/archive/os/lcfg/ether-object.txt

[13]https://svn.lcfg.org/viewvc/source/trunk/ProCurve/ is the HP/ProCurve/Aruba version. This was originally checked into CVS in September 2002. The 3Com version is no longer maintained, and not on-line.

JANET's "shoestring" project developed into a service in 1993[14] and we connected to the wider Internet, we initially relied on the underlying strength of the systems we were running to keep outsiders out. As the Internet grew and connectivity improved, this position became clearly unsustainable.

We were using Suns as routers at the time, so the natural tool to use was `ipfilter`.[15] We reorganised our subnets slightly so as to have one "outside" subnet, linked onwards through EdLAN, with all the rest being "inside." We did consider setting up a DMZ in the usual way, but concluded that it didn't really fit into the way our network was used. We configured `ipfilter` initially by hand, and later under `lcfg` control, so as to block inbound traffic by default but permit outbound traffic by default. The more restrictive position of blocking everything by default would have resulted in an excessive management load then, as it would now. Later as we moved to using Linux perimeter routers we ported our rules to `ipchains` and then `iptables`, as this was the native filtering system available.[16] We did consider using the ProCurves' ACLs, but these were, and remain, less flexible.[17]

Blocking inbound and permitting outbound by default has remained the configuration to date. It's not ideal, but it is still the best fit for us from the well-known paradigm of "choose any two from 'secure,' 'cheap' and 'easy to use'." How this might change under new-EdLAN, with its much more powerful firewalling capabilities remains to be seen.

## 10    Formation of Informatics

With the formation of the Division (later School) of Informatics in 1998 came the requirement to integrate the constituent Departments' networks.

- The situation in CS was as described above: we had our main site at JCMB with the "CS1 lab" in Appleton Tower as a satellite operation. All machines were connected to a 100Mbps switch port, either directly or through a local mini-hub. Hard patching had been replaced by soft configuration of the switch ports under control of the configuration tools.

- Artificial Intelligence ("AI") was split across two sites: Forrest Hill and South Bridge. The Forrest Hill site had been rewired with cat-5. South Bridge was still mostly 10base2.

- Cognitive Science ("CogSci") was sited at the East end of Buccleuch Place. The buildings had been rewired with cat-5 and a switch-based installation was in use, very similar to CS's.

The requirements were as follows: Informatics machines should be treated as part of one unified network which happened to be spread across the five (initially) sites; and the network should be robust, there should be transparent fast paths between sites, and all external traffic should be filtered. We decided that it would be best to keep subnets and their associated VLANs within sites as much as possible, to minimise switch table use, and inter-site and EdLAN dependencies, a position we have maintained to date; we introduced a "transit" subnet for inter-site traffic, with only routers and a small number of network infrastructure machines on it; and we gradually introduced switch/routers throughout, to improve intra- and inter-site routing throughput. Inter-site traffic was carried by tagged VLANs across EdLAN, for the most part. We did have some dark-fibre links between AT and FH and between AT and BP for a while, but instabilities introduced by being part of the EdLAN spanning tree meant that we had to keep them mostly disabled.

Routing was still controlled using RIP, initially using `gated` and latterly `quagga`. Traffic was directed through preferred primary and secondary routers by adjusting the metrics advertised. We realised quite soon that the use of RIP at our privately-interconnected sites did not interact well with the OSPF that EdLAN was running — in particular, metrics were lost in the conversion,

---

[14]https://www.jisc.ac.uk/janet/history
[15]https://en.wikipedia.org/wiki/IPFilter
[16]This was very much a case of "build something three times" to get it right.
[17]We're taking a closer look at them at the time of writing (late 2022), with a view to making some use of them to segment our network a little more, but they aren't an ideal fit for the problem.

and so taking routes from EdLAN at one site, propagating them internally to another site, and then forwarding them from that site to a different EdLAN router was not safe — and so we had to introduce heavy filtering of the RIP routes we sent and received. The realisation that a large amount of Physics's traffic was transiting across our networks as a result of RIP/OSPF interaction helped to increase the priority of this work.

## 11   South Bridge fire and relocation to Appleton Tower

The South Bridge fire in 2002 did not actually have much of an immediate impact on the Informatics network, other than getting rid of all remaining 10base2. Our Appleton Tower core was upgraded by the addition of a switch/router, and temporary edge switches in level-3 and level-5 were connected to it. The subsequent relocation of the core from level-5 to the basement server room was done essentially transparently to the users by relying on the redundant inter-switch links to keep the traffic flowing while relocating the main links to new fibre. As floors were refurbished, a pair of edge switches was installed in each and linked back to the core. (When floors were later upgraded we moved to the "Forum" model of a hierarchy of switches as a more cost-effective solution.)

The Appleton Tower network was designed in the light of experience from our other sites (figure 7). Core and edge switches were clearly distinguished by role, and models were chosen for their particular role, not as a compromise for mixed use. The initial design had one switch/router at the core (`atba`), alongside one non-routing switch (`atbb`), with Linux routers to provide redundancy. The edge switches, two per floor, each had a "main" 1000baseSX link to `atba` and a backup 100baseFX link to `atbb`. A PoE switch per floor was added, with one link to each of the floor's edge switches. Those original core switches have since been replaced by a pair of more capable switch/routers, with one 40Gbps link between them and another to the Forum core, and with one edge switch from each floor connected to each AT core switch at 10Gbps.

## 12   Informatics Forum

The design of the network in the Informatics Forum once again drew on experience at our other sites. As the building was being constructed from scratch we were able to get a network design which was reasonably close to the way we wanted it, rather than having to be retro-fitted to an existing building. The funding included provision of new switches, as the ones we had installed in our previous sites would have been due for replacement anyway.

At the core of the building network was a cluster of three ProCurve 5412zl switch/routers (since upgraded), each connected to the other two by 20Gbps trunked links (now 40Gbps links). Each of these switches also had a link to one of the Appleton Tower core switches, the whole forming one tree under Rapid Spanning Tree Protocol control,[18] with priorities and link costs chosen so that the tree is stable and traffic is contained within the two sites and passed preferentially over the desired links. These switches have four power supplies, connected to three different power circuits and two bottom-of-rack UPSes, for robustness, but the intention of the design is such that one or even two of the core switches can fail without the service for the rest of the building being much affected. (This also proved very useful during the upgrade process, as we could take a switch down for replacement without the users noticing much effect, other than briefly while the spanning-tree rebalanced.)

The physical cable installation was designed in detail by the building's engineers, more or less to our specifications. Each floor has from one to three "IT closets," located to fit in with the building's other services. In the upper floors the closets are sited in three columns, vertically one above the other, with a direct cable path between. Unfortunately the ground floor and basement closets are offset slightly from this arrangement. Fixed copper wiring radiates out from each closet

---

[18]We filter BPDUs where we connect to EdLAN, so that neither side is affected by the stability or otherwise of the other.
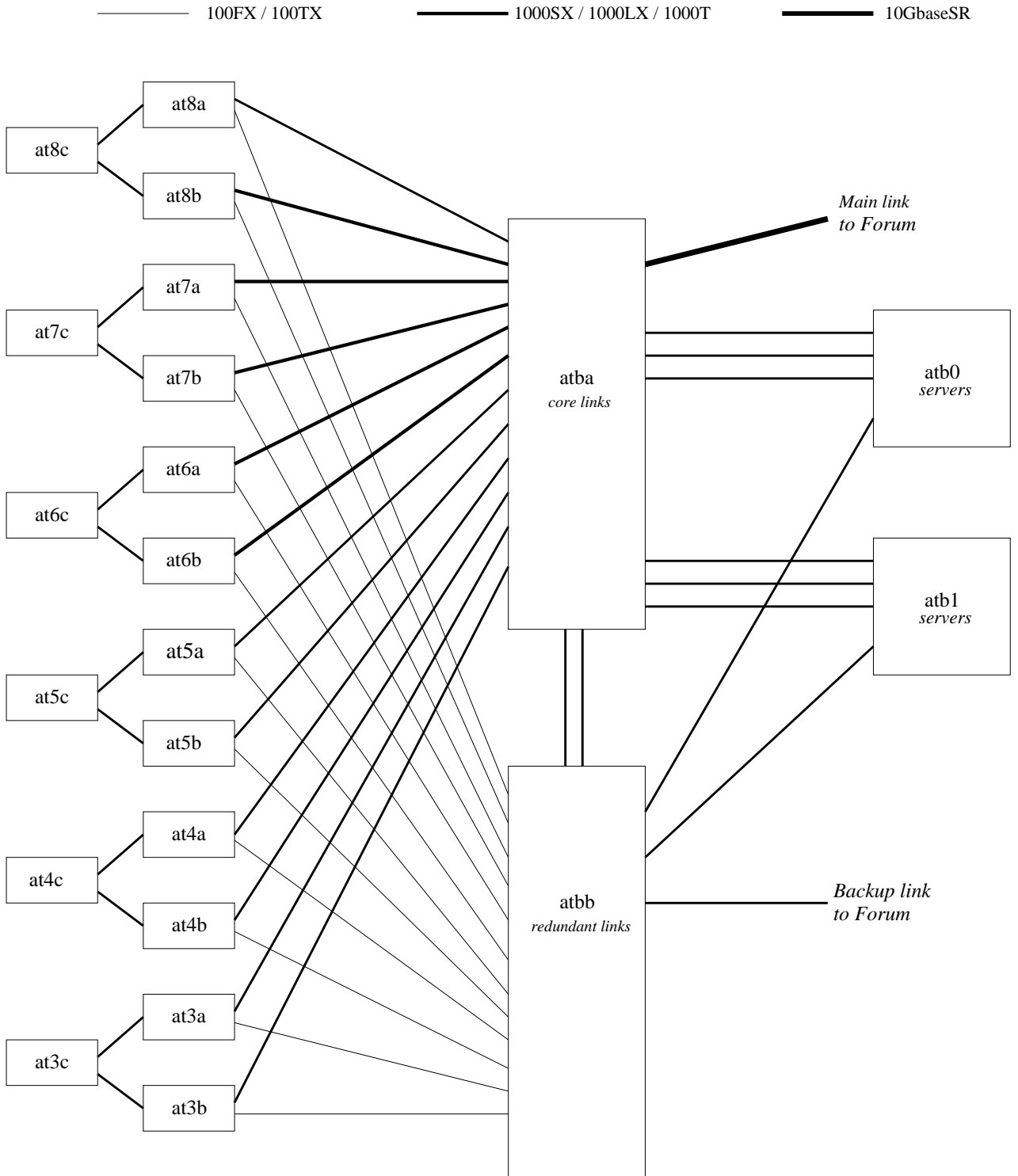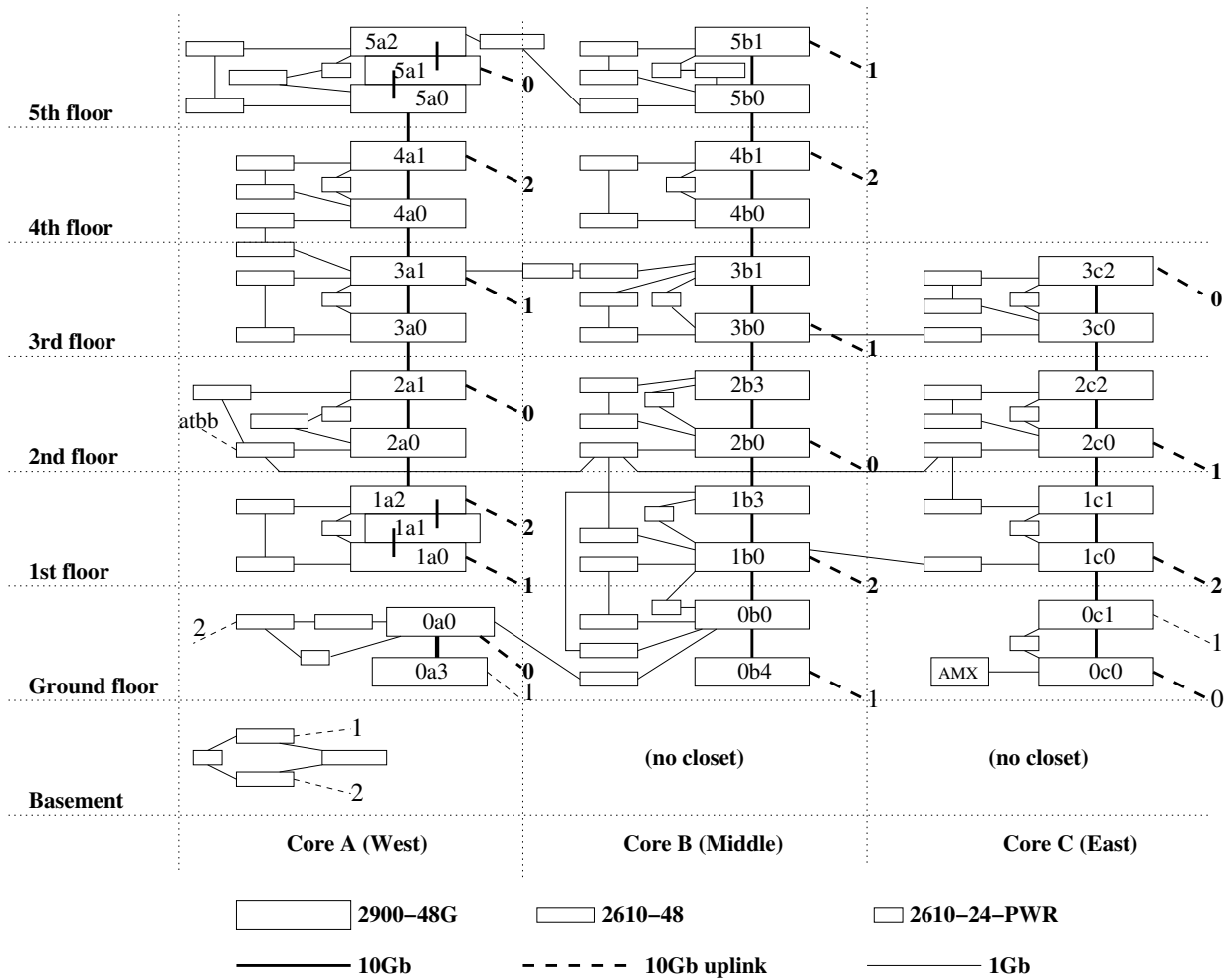
Figure 7: Appleton Tower network

Figure 8: Forum edge date network (as initially installed)

to a number of distribution boxes under the raised floors. Floorboxes, with six data and four power sockets, are then connected to these under-floor distribution points. The intention was to allow flexibility for floorboxes to be added or moved later without the disruption of having to rewire them all the way back to the closets. This has proved to have been very useful over the years, as more and more occupants have been squeezed into the building and rooms reconfigured to suit. There are also a number of fixed data outlets spread around the walls of the building, for printers, wireless access points and the like.

Of the six data points in each floorbox, four were initially patched in. Each user was provided with one 1000baseT link, two 100baseTX links and one VoIP phone link. As with our other sites, these could be assigned to VLANs under the control of the configuration system. This arrangement was chosen as being a reasonable compromise between cost and speed. Each closet then had fixed-format ProCurve switches installed, sufficient to service the required number of each type of port. At the time of writing there were more than 150 switches in the Forum, providing for almost 4000 data connections and over 500 phone connections, patched in 17 closets and three server rooms. The 100Mbps switches were later upgraded, so that all data outlets are now patched to 1Gbps ports.

Unlike Appleton Tower, where each main edge switch had its own direct link back to the core, not all the Forum switches were directly connected, due to the cost of doing so being prohibitive. Instead (see figure 8) all the Gbps switches in each column of closets were interconnected vertically

using 10GbaseCX4 links (now replaced by 10GbaseT). Roughly every third switch in each column then had its own direct 10GbaseSR link back to the core, with the remainder being indirectly connected. 100Mbps switches were then connected indirectly via one or other of the Gbps switches. In all cases, connections were arranged so that each switch had at least two independent paths back to the core, all still under RSTP control, so that the failure of one switch would not result in a loss of service for any other. There were also some cross-links between the columns, partly for robustness and partly because it turned out in some cases that this was the best way to provide dual paths back to the core. Each closet also had one PoE switch, to power wireless access points. The phone network effectively paralleled the data network, to minimise the chances of either impacting on the service of the other. In all cases "close" switches were linked back to different core switches for added robustness. Although most of the switches have since been replaced and upgraded, the basic structure remains.

We have not stacked switches since our 3Com days, for a couple of reasons: firstly, it means that we do not have to have one single model in a closet, which makes upgrades and end-of-life replacements easier and with the cost spread to suit the available budget; and secondly, our configuration management tools mean that there is no particular advantage in stacking. And in a few closets we have too many switches to be able to combine them all in one stack anyway, so the simplicity case is lost there. However, the new-EdLAN edge is based around switch stacks, so there is expected to be some change as a result of its roll-out.

The main server room in the Forum is fairly straightforward. Each group of three racks has a pair of switches installed at the top along with a redundant power supply. The pair is connected together using 10GbaseT, and each switch is connected back to a different core switch, again using 10GbaseT. Finally, each switch has a 1000baseT link to a switch somewhere else in the room, chosen so that it links back to the third core switch. Servers are then normally connected to both of the switches in their "group" of racks, using bonded interfaces for resilience.

The self-managed server rooms are similarly arranged, though we do not normally provide bonded interfaces for self-managed machines.

# 13   Routing using OSPF

We had been using RIP as our routing protocol ever since we had started running TCP/IP, in spite of its awkward interface with EdLAN. However the move to the Forum prompted us to re-evaluate. A larger building suggested that sticking with /24-sized subnets throughout would make client management rather more complicated than it really needed to be, as well as preventing fully efficient use of the (scarce) address space. We did consider RIPv2, but as EdLAN was using OSPF as its native routing protocol at the time it made much more sense for us to adopt that too.

In conjunction with Information Services ("IS," formerly UCS, formerly ERCC), it was decided that all of Informatics should be in one OSPF area (area 64). The Area Boundary Routers would be the EdLAN routers at JCMB (`kb6` at the time) and Appleton Tower (`at5` at the time; both have since been upgraded, and are now being replaced by the new-EdLAN distribution routers). The two "external" subnets (42 and 160) would each have one EdLAN router and all our Linux perimeter routers on them, to provide for better control of external traffic as well as robustness against router failures. The transit subnet would remain, but now with only switch/routers (and a handful of infrastructure machines) but no Linux routers. Each site would have one subnet where Linux routers and switch/routers speak OSPF, and routes for almost all other subnets would be injected by the switch/routers acting as ASBRs. Traffic would be directed preferentially by setting the routers' path costs — this is necessary to ensure that external traffic is routed symmetrically, so that `iptables`'s connection-tracking works properly. Figure 9 shows this in outline form: indicative switch/routers are labelled "Sxx" and Linux routers "Lxx." As well as allowing for non-/24 subnets (we now have a variety of /22–/27-sized subnets in use) this arrangement provides improved robustness, by allowing any site to act as a standby perimeter router for any other site. New-EdLAN has prompted a revision of this arrangement; see section 19 below.
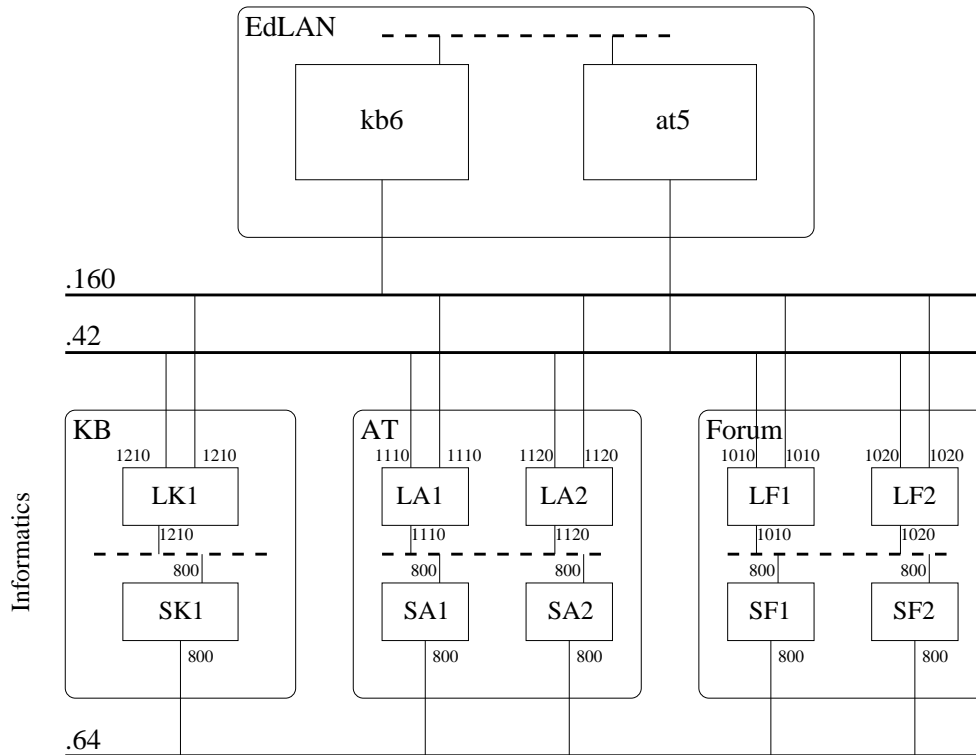
Figure 9: Outline OSPF routing arrangements (pre-BGP)

The transition from RIP to OSPF was handled in several stages. The first step was to enable OSPF on all the core switch/routers, but with the administrative distance set so that RIP continued to be preferred. Once that was in place, OSPF was introduced on all the Linux routers. Finally, RIP was disabled everywhere, and we had a pure OSPF site.

Moving to OSPF also meant that we had to change the way routes are distributed to non-routers. While we had RIP this was done by those non-routers snooping on the RIP traffic. We didn't want non-routers participating in the OSPF exchanges, as this would have loaded the DRs and BDRs unnecessarily as well as weakening security. We considered RIPv2, but concluded that as we only really needed to distribute default routes it would have no advantage over router-discovery.[19] We have, however, tuned the router-discovery parameters to provide a faster failover time. Those machines which are not able to listen to router-discovery are expected to honour the values sent to them by DHCP. (Similar arrangements are now in place for IPv6 using RA.)

## 14  AT Basement Refurbishment

The Appleton Tower basement server rooms were refurbished in 2009, and we had to decant all our kit out, mostly to Forrest Hill, for the duration. The comms rack was able to stay *in situ*, but behind a temporary wall and with no power or ventilation. Although spread across three locations (Appleton Tower upper floors, Forrest Hill server room and the Forum server room), we decided to maintain the "AT network" identity throughout, to make management of clients, servers and self-managed systems easier.

We installed a pair of temporary switch/routers in the Forum server room, initially running in parallel with the AT core. They were connected to the Forum core switches, which had had

---

[19]This is apparently now being deprecated by our upstream Linux distributions, so we may yet have to revert to some form of RIP to distribute (only) default routes.
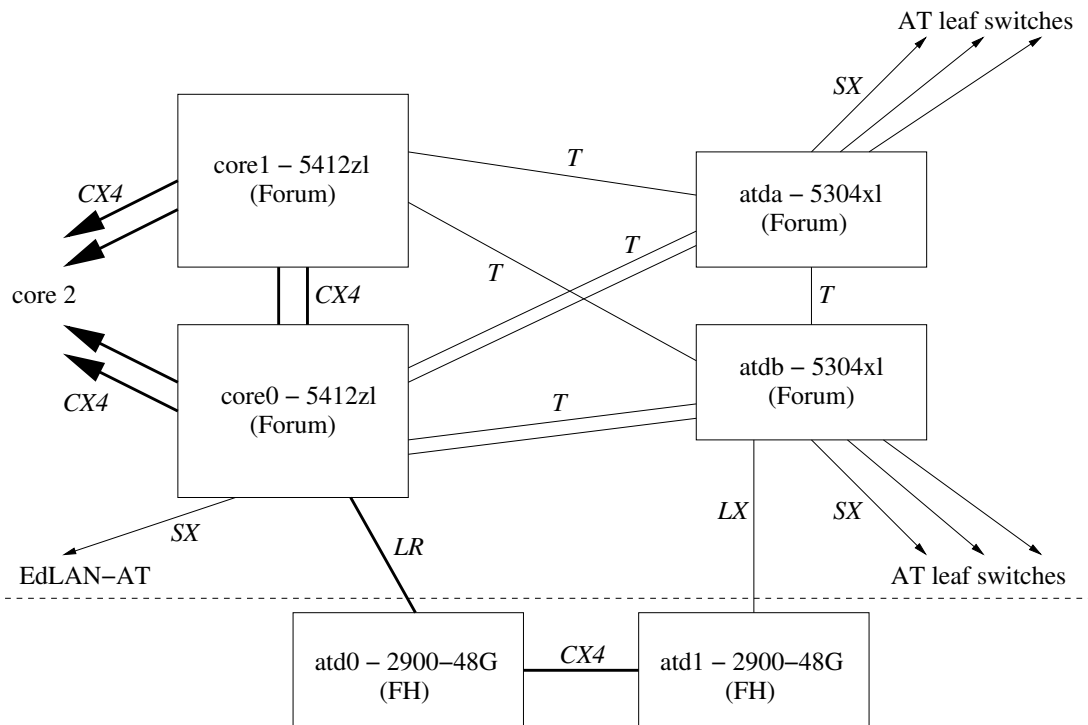
Figure 10: Appleton Tower decant arrangements

the AT VLANs added. Once those switches were up and running we were able to re-patch the upstairs AT switches to them using the fibre we had had installed between the buildings when the Forum was built. The decanted servers in Forrest Hill were catered for by installing a pair of switches, connected back to the Forum using 10GbaseLR and 1000baseLX over single-mode dark-fibre. Figure 10 shows the arrangement, with the four switches above the dotted line being in the Forum and the two below in Forrest Hill.

Once the refurbishment was complete, the AT core switches were reinstalled and the upstairs floors re-patched. A pair of new switches was installed in the server racks, connected back to the core using three 1000baseT lines trunked together as the distance was too far for CX4 and the expense of fibre wasn't really justified for the switches involved.

When these server room switches were later upgraded we did also upgrade their uplinks to use 10GbaseSR. We decided on fibre rather than using 10GbaseT for a couple of reasons:

1. the links were run on cable baskets shared with power and other unidentified data lines, and we weren't confident about 10GbaseT's noise immunity under this arrangement; and

2. we needed SFP+ cards in the core switches anyway, for links to the upstairs floors. to the Forum and to EdLAN, and this allowed us to share resources.

## 15    Moving out of Buccleuch Place and Forrest Hill

After we had moved out of the Buccleuch Place and Forrest Hill offices into the Forum we still had machines in the server rooms there for quite a while afterwards. In conjunction with IS, therefore, we decided to continue to manage the edge networks for the new occupants of the offices in the buildings using the existing switch infrastructure. This was reasonably straightforward on the whole, as there was not nearly as much VLAN churn as would have been produced by our own users, as the rest of the University was generally using "building" subnets rather than per-function subnets. Once all our machines were out of Buccleuch Place we reset all the switch passwords and

SNMP communities to temporary transitional values and handed the entire network installation over to IS to manage from then on. The same procedure was followed when we moved out of Forrest Hill, though this was delayed by the Appleton Tower basement server room decant.

# 16  Decanting out of Appleton Tower again!

In 2015 the Appleton Tower recladding project[20] required us to decant out of the upper floors of the building while the work was in progress. Teaching went back to Forrest Hill, which had been extensively refurbished to suit our needs as much as possible, while non-teaching moved to the Wilkie Building. We moved out in Summer 2015, with the intention of moving back in Summer 2016. Additional building work, however, resulted in this being delayed by a year.

For simplicity it was decided to run both Forrest Hill and Wilkie as "virtual floors" of Appleton Tower, connecting back to the core using 10GbaseLR links over dark fibre. There were a couple of advantages to this:

1. machines could be taken off their AT desks and connected to identically-configured ports at their new locations, and would then continue to work with no further reconfiguration required; and

2. there was a cost saving, as core infrastructure was not required at either site.

The disadvantage, of course, is that it introduced outside dependencies, but given the expected duration of the decant this was felt to be an acceptable compromise. We did finally return to Appleton Tower in 2017.

# 17  IPv6

Recognising that IPv6 support was becoming more and more necessary, a Development Project[21] was started in 2015 to investigate how this might be provided on the Informatics network, and to set up an initial service for managed machines. We joined the IPv6 Internet in February 2016, and by September 2016 all of our managed Linux subnets were IPv6-enabled. As of August 2019 there were over 3500 `AAAA` entries in the `inf.ed.ac.uk` zone, most of them automatically generated as "SLAAC-style"[22] addresses, derived from the MAC addresses in our `lcfg` host configuration database. A number of enhancements were also required to our switch configuration tools, some of which had to be made conditional on the particular switch model as some older ones had incomplete IPv6 implementations. Our `iptables` support was mostly already in place, though some scripts to generate ruleset fragments did require modification to take into account IPv6 addresses.

We decided to use BIRD[23] as our IPv6 routing mechanism, and having found it easy to use and well supported we moved all of our IPv4 routing to use it too, ahead of the expected new-EdLAN changes.

# 18  Bayes Centre

The Bayes Centre occupies the remaining North-East corner of what was the "Potterrow car park" site.[24] The building was completed in September 2018, with Informatics occupying three of the floors. For the most part, we manage the network on "our" floors, and IS manage the network on the rest.

---

[20]http://www.ed.ac.uk/projects-estates/current-projects/appleton-tower-recladding
[21]https://history.dcs.ed.ac.uk/archive/docs/Informatics-IPv6.pdf
[22]https://www.rfc-editor.org/rfc/rfc4862
[23]https://bird.network.cz/
[24]https://dtisite.inf.ed.ac.uk/

The initial question was to what extent we should manage the building as a separate entity, or integrated with the Informatics Forum. The decision was the latter, for a couple of reasons:

1. Cost: we wouldn't need core switches and the three Linux servers that a full "site" of this size would require; and

2. Shared VLANs: there would be many VLANs which would naturally be common across the two buildings, and quite a number which strictly needn't be but would be simpler to implement if they were regarded as Forum offshoots.

The original intention had been to configure the network hierarchically, as we do in the Forum and Appleton Tower, but unfortunately some of the expected inter-closet copper wiring was not installed. Instead we operate each floor independently, with the switches connected in a 10Gbps ring, with connections back to the (Forum) core every second or third switch. The switches are not stacked, but instead connections are under spanning-tree control. Again, this arrangement may change with the new-EdLAN edge roll-out.

We're not quite completely independent of IS, however. Because high-level and soffit-mounted data outlets are connected to the switches on the floor above, we do share one closet with IS. Fortunately the way the wiring is installed means we don't really intermingle with their patch leads and switches.

There are also services in the building (such as BEMS, locks, café tills, CCTV and a bus-timetable display) which are attached to our part of the network but naturally connect back to IS's part. As a result there are several VLANs which span across between the two. It's not ideal, and does take a bit of coordination, but the alternatives (all managed by us or all managed by IS) were less appealing.

## 19   New EdLAN

After what many would see as years of chronic under-investment by the University in EdLAN, and the edge in particular, the decision was eventually made that the active kit should all be replaced. By August 2019 the procurement phase had been completed and the project had moved on to the detailed design phase. Although input from Informatics and other Schools was provided into the process, it remains to be seen to what extent the features (mainly automation, devolved management and network login) will meet our needs.

Unlike the "old EdLAN," which was mostly a bridged network, the new EdLAN was designed to be a fully-routed network, divided up into separate "VRFs"[25] for overall robustness. VLANs can still be stretched across the network where necessary, but are carried in IP tunnels rather than as native bridged traffic. The core comprises two routers in AT and two at KB, cross-connected in a mesh, with distribution routers ("DRTs") attached to the core; and edge switch stacks are then connected to the distribution routers. There are separate appliances for network management, wireless control and firewalling. BGP[26] is used as the underlying routing protocol. Note that even for those sites where there is a core-router presence there are still separate distribution routers, so as to keep the functionality clearly distinct as well as to allow for different connectivity options.

The new EdLAN edge will look considerably different from now. The current (non-Informatics) edge is essentially statically configured, with subnets and VLANs assigned more on a per-building basis than for functionality (in contrast with the Informatics approach of having separate subnets and VLANs for specific purposes, along with an easy way to configure switch ports to use them). The new edge will instead be dynamically configured where possible, ideally using 802.1X network logins, failing which based on MAC addresses, with hosts being assigned to functional subnets based on a configuration database.

Because EdLAN is transitioning away from OSPF to BGP, it was necessary for us to do likewise, at least so far as our interaction with it goes. Unlike OSPF, where we can basically set a

---

[25]https://en.wikipedia.org/wiki/Virtual_routing_and_forwarding
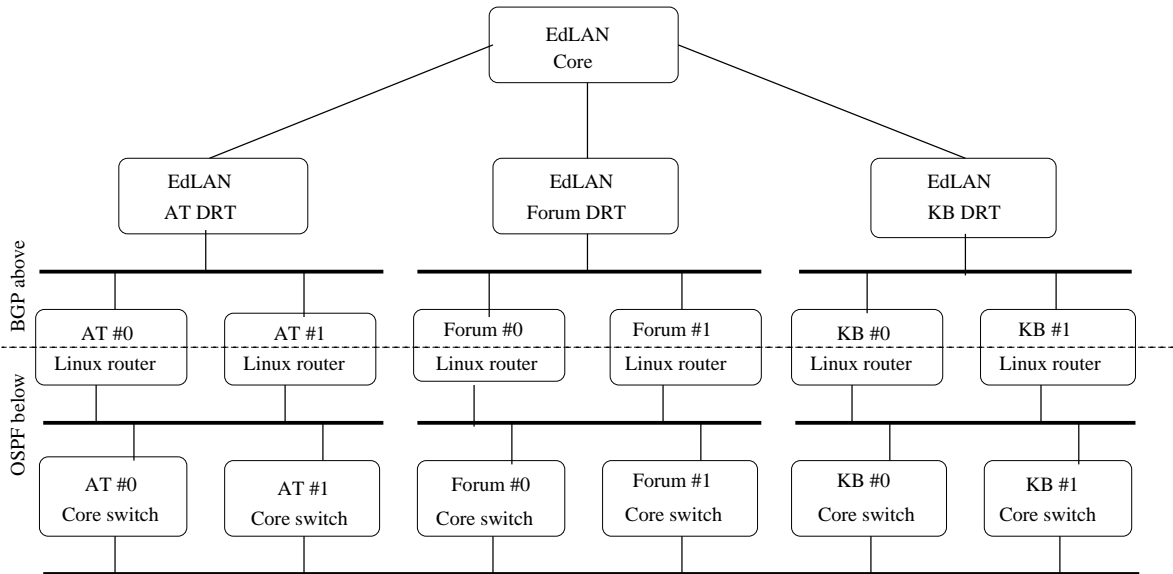[26]https://en.wikipedia.org/wiki/Bgp

Figure 11: OSPF and BGP

few parameters and leave the routers to get on with it themselves, BGP requires a bit more in the way of explicit configuration. On the plus side, though, it does give us rather more control over which routes we send to and receive from EdLAN. Figure 11 indicates how we now split between OSPF and BGP. Our three sites (Appleton Tower, including Wilkie Building, Informatics Forum, including Bayes Centre, and the College Server room in the JCMB) each have one private subnet shared with their local distribution router, along with a per-site OSPF subnet and the inter-site "transit" subnet. The distribution routers are configured to listen for BGP connections on their local subnet, and our edge routers at each site are then set up to speak eBGP to their local distribution router, iBGP to each other, and OSPF to each other and the local core switches. OSPF costs and BGP MED values are then set so that we have a well-defined symmetric path through each site's perimeter, so that `iptables`'s connection-tracking works as desired.[27]

At the time of writing (October 2022), we were still using EdLAN essentially as an ISP, while managing our own edge. We have direct 20Gbps connections from our site-local core switches to the Informatics Forum and Appleton Tower distribution routers for our routed traffic, and separate 20Gbps connections for bridged traffic (mostly for VLANs relating to services managed by other University bodies, such as wireless, phones and building-management); and an indirect connection to the Kings Buildings distribution router, carrying both routed and bridged traffic. We take only the IPv4 and IPv6 default routes, and by default we announce only our Forum routes to the Forum distribution router, our AT routes to the AT distribution router and our KB routes to the KB distribution router. As a transition arrangement, we do still (October 2022) speak OSPF to the old-EdLAN routers. We anticipate that this will no longer be necessary once the migration to the new distribution is complete, and some transitional configuration is removed.

Having found BIRD easy to set up and configure, we use that for both IPv4 and IPv6 routing, speaking OSPF internally as well as BGP to EdLAN. We use it on some of our subnets for IPv6 Router Advertisement, and may yet use it to advertise IPv4 default routes using some version of RIP.

---

[27]The `lcfg` host-configuration system makes this all very straightforward to automate.

# 20   Looking ahead

As noted already, this is a time of change for EdLAN, and hence inevitably for the Informatics network too. For various reasons, not least Covid-19, the original EdLAN replacement schedule had to be changed, so that at the time of writing (October 2022) the following were in place:

- the core and distribution layers

- the firewalls and external connectivity

- the network mangagement appliances and wireless controllers

- some of the wireless APs

- some of the edge switch stacks

In particular, the edge switch stacks covering the Informatics areas are not expected to be installed now until next year, so we are essentially still running a more-or-less independent network with EdLAN as our ISP, as we have done since its transition from EdNet in 1990.

This is therefore a nice clean point at which to end this history, with the Informatics network in a stable configuration; and indeed the Author will also be retiring from the University shortly.

On the horizon is the edge replacement, mentioned above, which is likely to require many of our current management practices to be adapted. Further ahead, there will be questions over whether a separate Informatics network core and firewall configuration remains necessary, or whether some or all of the functionality could be folded into EdLAN. And no doubt, in a few more years yet another new piece of technology will render all of these current developments obsolete in their turn.

# Epilogue

It would be wrong to say that the current Informatics network design was inevitable, or indeed that it's the best that could have been achieved. It was the result of many developments, some incremental and some less so, and with many internal and external influences.[28] Much of what we were doing might reasonably be described as "network engineering" rather than design as such — essentially, taking the kit we had, or could afford, and adapting it to suit the requirements at the time as best we could, or indeed taking it apart and restructuring it to reduce the overall cost.

With the benefit of hindsight, however, it's perhaps possible to draw out a few threads:

1. More bandwidth is better. The Link boards allowed us to develop the discless Interdata clusters, and the virtual DECtapes provided by the Filestore were faster than the real hardware. On the other hand, being constrained by our thinnet segments definitely affected what we could do, and made management considerably harder.

2. Economies of scale matter. However, you do also have to bear in mind that it might still be better to build parts of a system yourself, rather than buying it all in.

3. Layering is good. Being able to decouple our physical layout from its logical arrangement was a major step forward, and meant that both could develop independently, each unconstrained by the other. And it can help hide "irrelevant" details from the users too.

4. Adding complexity can actually make things more reliable overall. For example, hubs were better than thinnet, because faults were more contained; and switches were better than hubs, because of the better segregation, VLANs and traffic monitoring they provided.

---

[28]Including manufacturers' decisions regarding their products, those products' facilities, and their pricing strategy.

5. Design your management tools and interfaces appropriately for their intended users. Keep them simple enough that they're easy and clear to use, and don't require too much additional knowledge. However, make sure that the necessary features are there, so that those users can do all the tasks they need to. As a result you'll be able to delegate management, which will save time and effort all round.

6. While you're at it, incorporate existing sources of data, rather than requiring things to be copied and pasted, as there'll be fewer mistakes that way.

## People

The following were involved in the events described above:

- Peter Lindsay designed the Link boards

- Rainer Thönnes and Paul McLellan wrote the original Interdata Filestore system, with input from Hamish Dewar

- Gordon Brebner wrote the WAN software for `ECSVAX`; Paul McLellan wrote the DQS11 driver

- Fred King designed the Fred machines (APMs)

- Rainer Thönnes designed the 2Mbps ether stations and wrote the Fred machine system

- the Author wrote the APM-Filestore system, based on the "1976" Interdata version and with contributions by Rainer; the TCP/IP stack for APM and VMS; and the switch configuration and monitoring tools; and designed and implemented the Department's and School's data networks from the First Great Rewiring to date.

Many other members of the Computing, Technical and Academic staff of the Department and the School have contributed to the network's development over the years, not least with the `lcfg` system-configuration tools which allow much of the host provisioning to be automated..

The following made some very useful comments on previous versions of this document: John Davis, Hamish Dewar, Gordon Hughes, Mike Hurst, Nick Rothwell, Frank Stacey, Rainer Thönnes, Graham Toal.

Thanks to Alan Boyd for the EdLAN documents.