From Thinnet to 10base-T,
From Sys Admin to Network Manager

Arnold de Leon
Synopsys, Inc.

From Thinnet to 10base-T, From Sys Admin to Network Manager

Arnold de Leon – Synopsys, Inc.

ABSTRACT

Once you have more than one computer at a site, a network is usually required. Often the system administrator also becomes responsible for the network. This becomes yet another task on an already full schedule. To be a successful system administrator, you need a working network; therefore, you now have to become the network manager with a limited amount of time. I will go through the techniques and strategies, including examples from the implementation used at Synopsys, for supporting a network in your spare time.

Introduction

In an ideal world, every organization would have a network manager who has enough time, staff and information to plan network growth and changes. In reality, system administrators are often called upon to act as the network managers for their sites. This paper is about techniques for making that transition.

There are no magic bullets in this paper. It is not about a single tool that will solve your problems. Instead, I hope to present ideas that can help you survive your tenure as the keeper of your network. Here I present the approach that I took and the lessons I learned in designing, growing and supporting the Synopsys network.

The examples herein are from the Synopsys campus in Mountain View, during the deployment of our 10base-T and 10base-FL network and later FDDI/CDDI. I will go through the architecture of the network, including decisions on avoiding the bleeding edge of technology while providing an affordable high-speed network.

I have also included a glossary of working definitions. The definitions are not complete or necessarily even fully accurate. Hopefully they are sufficient to get through this paper.

Background

Synopsys manufactures software and hardware tools for the EDA (Electronic Design Automation) industry. Our products are used by engineers in the design and implementation of chips and systems.

The majority of our computing activity is split among software engineers developing code and users of business applications (management, finance, marketing, etc.). In 1990, when I joined the company, the network was primarily used by developers. Today, all aspects of the business depend on the reliable functioning of the network. NCS, the Network and Computing Services department, is responsible for the network as well as the rest of the computing equipment deployed throughout Synopsys.

In this paper, when I say "we" I am referring to NCS, unless otherwise noted.

Five years ago, the primary Synopsys network was an Ethernet network composed of several pieces of thinnet coax connected by a pair of repeaters (see Figure 1). There were 60 hosts all located on a single floor of a building. The nodes were Unix workstations used in the development of product. There were no routers or bridges in this network. A separate network of 50 Macintoshes also existed. The IP addresses were from the infamous network 192.9.200 (sun-ether). NCS had two Unix system administrators, 1.5 Macintosh administrators and a manager.

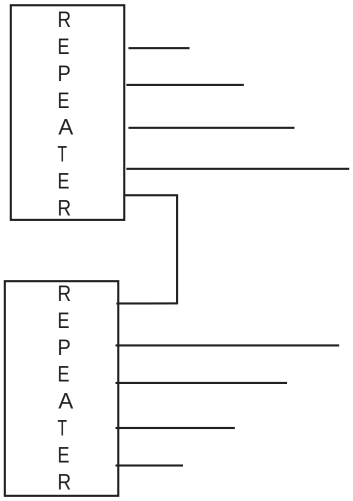

Figure 1: The Synopsys network in 1990: two repeaters and a bunch of thinnet

Today, there are more than 2,500 hosts in a network that reaches from Asia to Europe. The Mountain View campus includes three buildings that share a computer room. There are at least seven FDDI networks, many routers, more than a dozen Ethernet switches and at least 70 10base-T hubs. We have a registered class B network with more than 100 active subnets (more than 30 in the Mountain View campus). We have a 10 Mbit/s connection to the Internet.

The focus of this paper is our campus LAN. I will begin with the general strategy that we used in approaching the problems. Then I will start from the wires and lead into the design of our workgroup and backbone networks.

The challenge

Everyone wants a network that is:

- reliable
- affordable
- high-performance

Synopsys is no exception. We also need a network that can keep up with the rapid company growth, both in terms of the number of users and the capacity per user.

Even though Synopsys was already a successful company, it had a limited staff to deal with the network. As the second Unix system administrator I also became the network manager. I suspect this case is not unusual.

System administrators face enough tasks to keep them very busy. NCS has the added problem of a growing company: we are adding new machines on top of maintaining what we have. This leaves very little time for the network. As a result, planning often becomes an early casualty. Formerly, I expected that once I had a working network I would be able to measure and fine tune it to keep it running in top form; now I know that's not likely to happen in the near future.

Instead, there will be the need to deal with emergencies like having to move 60 people with minimal downtime while adding 10 new users. Reorganizations and moves will also wreak havoc on a network.

Documentation can easily get neglected, posing a potential problem. You may end up following what the documentation says only to discover later that it was incorrect, wasting time and effort. Knowing that the documentation does not exist will at least save the mistakes due to following the incorrect documentation.

Strategy

We needed to have a strategy to approach these problems.

Make Assumptions

First, several assumptions went into the planning process:

- The network is going to grow.
- The network is going to get shuffled.
- Technology is going to change.
- The staffing will eventually catch up.

I discovered that:

- The network grew and is still growing.
- The network got shuffled and is still getting shuffled.
- Technology changed and is still changing.
- The staffing may eventually catch up.

For Synopsys, "the network is going to grow" means more users and more network bandwidth per user. Estimating how much the network is going to grow is important. As a system administrator you should understand how your company is doing. If you can get projections on the growth then you're lucky; otherwise, you will have to guess.

Once we had our assumptions and the problem defined we wanted the solutions to last as long as possible.

Use Standards and Models

Synopsys adopted a strategy of using standards. This idea is not exactly revolutionary. Using industry standards gives us flexibility, which in the short term sometimes appears more costly, but in the long term is generally a better value. Technology moves rapidly, so it is easy to get trapped. Standards allow us to migrate instead of having to do cutover changes.

NCS created standard models for the computing equipment, which have proven to be even more valuable. These models incorporate industry standards as appropriate, but use our own if needed. This means we have specified what kind of equipment is required for most situations. Once a model is in place the effort that went into specifying it can be recouped over and over again with little demand on the network manager's scarce time.

The resulting consistency also allows us to reduce the complexity of maintaining or expanding current configurations. The network management staff can apply knowledge from one installation to make subsequent ones easier. Later in this paper, you will find examples of the standards and models we used as we built our network.

Design Conservatively

Designing conservatively means avoiding being on the edge. If you are not close to capacity or other design limitations you do not have to spend as much time determining if you are about to go over the edge. For example, we try to size things large enough so that we do not have to spend a lot of time measuring and tuning. This may actually mean spending more money on equipment in order to reduce administrative complexity.

Make It Easy

Even the best network will require some work. We set out to make things as easy to support as possible.

Make things easy on yourself and others. To paraphrase [Wall 90], one of the great virtues of a system administrator is laziness. This meant trying to make things easy to use and maintain. If something was too complicated it would either be bypassed or be done wrong. Being a member of the lazy club, I also sought ways to make it possible for other folks to do the work. This usually meant, once again, making things easy to support.

Learn To Do Nothing

Another important survival technique is doing nothing.

Some of the glitziest items on your network wish list may be difficult or time-consuming to implement. Sometimes the best thing to do is to ignore them. In our case this meant for a long time we did not have a functional network management station (NMS); our staffing was just not sufficient to support one. We have no fancy network maps to this date; our network administrators can draw one from scratch as needed because of the simple architecture of our network.

Look For Value

An underlying principle is the search for value. Implementing the preceding principles often requires spending additional dollars. While we often made the tradeoff to spend money now to save money later, we did not ignore opportunities also save money during initial installation or procurement.

We learned that equipment can be cheap relative to other things. We could have spent money on engineers who would have had to wait for network connections. The “extra” equipment allowed us to accommodate growth and reorganizations with minimal down time. During moves we have been able to bring up the network at the new location before taking down the old network.

Principles in practice

Now we can see how it all went into practice as Synopsys left its old thinnet network for an all new 10base-T one.

Wires

First, we need wires for a network. The opportunity to the use the principles on a new wiring scheme came when we moved to a new building.

General Wiring Design

Our wiring is hierarchical. Nodes are leaves to the workgroup network, a workgroup network is a leaf on our backbone network. This structure makes it easier to debug. When debugging a problem you can just work your way up the tree. Having a simple hierarchy is valuable. At one point, I started creating configurations that had webs of dependencies instead of a simple hierarchy. The result was not a fun time; debugging sometimes became an involved process. A problem in a given area could be caused by what seemed to be an unrelated area of the building or even campus. We have since switched back to simple hierarchies.

The resulting wiring should be easy to trace. I had a limited time to spend on the network. I thought about documenting every cable in network, when I naively thought that documentation was the answer to all problems. I actually tried going through the steps it would take to create and maintain such documentation. I quickly realized that this was not going to be practical with the amount of time and resources available. What was important was to make the system require a minimal amount of documentation. It should be possible to work most problems without having to track down a lot of documentation.

Users being moved cause premature graying in our network managers. We planned for these changes by supporting “any net anywhere.” This design meant that any port on the network can be connected to any of the campus LANs. This strategy is necessary to make moves easier to support.

Wiring Our New Building

We defined what we wanted at every station drop. Synopsys had many Macintoshes on PhoneNET so we decided that we needed a PhoneNET port. 10base-T had just become a standard, and we decided that it was going to be the wave of the future. Since 10base-T is a point-to-point system, we provided two ports at every station drop to give us flexibility. The facilities department wanted two phone jacks in each office.

Version 1 of our network jack contained three cables, each with four twisted pairs (see Figure 2).

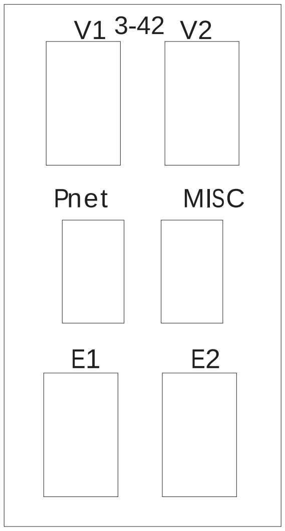

Figure 2: Original ports per station drop, ports in today’s station drop

Cable 1 was dedicated to voice. The pairs were split between two jacks – three pairs in one and one pair in the other. The three pairs in one port allowed it to be used for digital phones. Only one pair is required for an analog phone.

Cable 2 was assigned for PhoneNET. Since PhoneNET needed only one pair in an RJ11 jack, we decided to connect the remaining pairs to a second RJ11 for "misc" use, such as a serial device.

Cable 3 was for Ethernet. On our vendor's advice we decided to use a new high-grade cable for this one. This cable would later be classified as Category 5. The wires were split evenly between two RJ45 jacks to support two 10base-T ports.

We chose configurations that we believed to be flexible. We had some pretty good success. The wiring based on those standards, over four years ago, is still in use today. I learned a great deal from the experience.

Those insights have been applied to what we do today. Our new wiring is much simpler (see Figure 3). The voice part is unchanged. Instead of having four different data jacks with different wiring in them we now use two identically wired RJ45 jacks for every drop. A good specification to follow today is the EIA/TIA-568 standard.

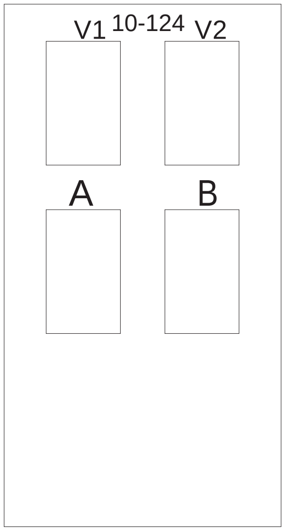

Figure 3: The current version of our network drop.

This configuration is quite flexible. It can be used without changes for:

- RS-232
- PhoneNET
- 10base-T
- telephones (analog and digital)

Many of the new technologies require Category 5-rated systems so most new wiring installations being installed today are being built to Category 5 specifications (see the EIA/TIA-568 standards). The resulting increase in demand has driven the cost of a Category 5 system lower and lower. It just does not make sense to do any less today. And if you do get a Category 5 installation you will be able to support any of the following newer technologies:

- CDDI
- ATM
- 100base-T
- 100base-VG

We learned to avoid overly specialized wiring. Our initial install had a jack dedicated to PhoneNET. The value of those ports is diminishing rapidly as we move our Macintoshes to Ethernet.

Another thing we learned was not to split wires between two jacks. 10base-T only requires two pairs. In our first installation we had a four-pair cable split into two 10base-T jacks. It seemed clever at the time, but we failed to consider what future networking technologies might require. Now, we are regretting that decision as we face the prospect of reterminating those jacks to support emerging technologies. If you must split wires, make sure you leave enough slack in the cable to allow you to fix it in the future.

It is important to provide enough ports per station drop. The two RJ45 jacks allow us to connect two different devices to two different networks without any trouble.

In case two Ethernet ports are not enough we can "split" a fully populated RJ45 into two 10base-T ports using an adapter. And if that is not enough we can install a hub or a fanout in the office.

The station jack is connected to a wiring closet (sometimes referred to as a phone closet, IDF or SDF).

We specify that the cable runs be 90 meters or less. The commonly accepted limit for 10base-T runs is 100 meters. We want to leave room for the patch cables. Also note that a good Category 5 installation can support runs significantly longer that 100 meters for 10base-T. A conservative installation creates a safety margin that can save work down the line.

Our initial install had the wires going to punchdown blocks. The wires were then cross-connected to patch panels. To save money, we only cross-connected one of the Ethernet ports and the PhoneNET port. The "misc" port and second 10base-T port were unconnected. We thought these ports would be rarely used and we could connect them on an as-needed basis. This second Ethernet port is now used far more than we had anticipated, and any money saved initially has long since been spent in the trouble that we go through when enabling the second Ethernet port. We failed to apply the principle of making things all alike to simplify support.

The punchdown blocks were also supposed to provide the flexibility of changing the pair allocations. By changing the cross-connects we would support future technologies. This has not turned out to be useful. We have not needed or wanted to fix the ports this way. Sometimes, predictions go wrong and we just take the lesson and move on.

We now have the station cables terminated onto panels directly. This reduces the number of connections that need to be made. This saves on installation costs and the higher-speed technologies require (or at least prefer) fewer interconnects. Instead of worrying about changing how the pairs are split and what pairs to cross-connect, we have all the wires connected.

Make sure that slack is left on the cable. Sometimes even the best designs will need revision. We have a floor that was wired before the Category 5 standards were defined completely. The patch panels may need to be reterminated to be fully compliant with Category 5 at a future time.

You may also need the slack to move things around later. Our first building did not have a wire management system at all. Imagine a wiring closet in use: there will be many patch cords. It will need wire management. We have spent a of lot time and money cleaning up after this oversight. We still have a lot of time and money to spend before we recover completely.

Backbone Cabling

If a building has multiple wiring closets, install "building backbone" wire. You may need fiber to do this. We installed 12 pairs of 62.5 µ multimode fiberoptic cable between the main wiring closet in the building and each of the other wiring closets. When we chose 62.5 µ multimode fiber, we knew we could use it for FDDI and 10base-FL, and today ATM, 100base-FL, etc., also can run on it.

Our wiring closets were close enough so we could also run a copper backbone network from the building closet to each of the secondary closet. We ran 24 jacks. Since fiberoptic cable is significantly more expensive we were able to save money with this installation.

If you cannot run a copper backbone you should consider running more fiber. Look at the size of the area the closet is serving. Keep in mind that we had a design goal of supporting any subnet anywhere so it was possible to have large number of networks in an area. This required that we be able to bring a large number of networks from elsewhere to any wiring closet.

We also took advantage of the copper backbone between the closets to share a terminal server between the closets. We connected the serial consoles of our network devices to the terminal server. The serial consoles on the terminal server allowed us to access the consoles of the devices even when they were off the net. Console servers are discussed in a paper from LISA IV [Fine and Romig 90].

Campus Cabling

Our campus backbone resembles the building backbone, except it was all fiber since we were going between buildings. We have a strong emphasis on star topologies, so we ran enough fiber between the buildings to take all the fibers in the closets back to our computer room. For example, in each building with four closets, we ran 48 pairs of fiber from the building main wiring closet to the campus computer room.

We have not had a need for single-mode fiber yet. I suspect this will be more of an issue if our campus gets larger and we need the longer distances that single mode supports.

Labeling

Labels are critical to make something self-documenting. You can skimp on other documentation but it is painful when you do not label well.

As you go through the trouble to label, make sure the labels are correct. It is better to have something unlabeled rather than incorrectly labeled. Whenever possible use labeling systems that are permanent. Examples of this include station numbers and port numbers. Those kind of labels should be well-attached (permanent). You do not want it to change. Other labels, like the name of the network or machine, should be removable or erasable since they can change. They should be easy enough to remove so you do not get stale labels, but attached securely enough so they do not fall off accidentally.

In our network, each station drop has a number in the form *nn-mmm*, (e.g., 10-102). The first number signifies which wiring closet it is in. We decided to assign a unique number to all our wiring closets in the campus. This means we do not need to know what building a jack is in. We can infer it from the station number. We also make sure that the there is only one number for the jack. The entire set of ports on a given plate are 3-42. We had sub-identifiers, like V1, V2, P-net, E1, A and B, for each port.

Do not take for granted that your vendor will its works correctly. Make sure that you discuss this issue with your wiring vendor. We had our vendor redo labels in a section when they failed to communicate our requirements to their technician. Consistent labels are important in troubleshooting. When there is a problem, we can get the user to read the number and we know which wiring closet to go to.

In the closets we have patch panels and equipment labeled with the subnet/network number that they are on. This makes checking a connection quick and easy. All the information is out where you need it, not hidden away in some book or file.

An important result of good labeling is that it is system administrator friendly. Good design allows our system administrators to do a lot of the first-level debugging. They can quickly check a given station to see if it is connected to the right place. If a user moves to a new office the system administrator can easily move the connection by following the labels and cables that are in place. They can also just as easily add another connection for a new system.

Patch Cords

You are going to need patch cords. Lots of them, actually. Here are some rules of thumb to make them easier to manage and take up less support time.

1. Avoid making your own. You are probably not good enough! Making good cables takes practice and patience. It is not cost-effective for you to make them. You will probably pay for the cables twice, first in the time it takes to make it and second when you debug a problem caused by your cable. Making good cables requires good equipment. I have wasted a lot of time fighting with a cheap crimping tool.
2. Make sure your patch cords are correct. Qualify your vendor. You will be surprised at the number of vendors that produce incorrect wiring. A well-known workstation manufacturer used to ship an incorrectly wired 10base-T cable. Also, make sure your purchasing department does not switch sources on you. Qualify multiple sources to make the purchasing folks happy. Spot check your cables as they come in. You need to learn enough about cables to know if you have good ones or not. A cable tester, one that can check the quality of the cable, can help provide the needed expertise in checking patch cords and wiring.
3. Get your vendor to serialize your patch cables. It's cheap. Make sure the serial number is on both ends. This will make it easier to identify ends of cables. I initially tried serializing them on my own. Don't even bother. It was time-consuming, boring and the labels fell off. Get your vendor to do it. Our serial numbers are of the form *ttt-lll-nnnn*, where *ttt* is the type, *lll* is the length and *nnnn* is a sequence number. This is particularly handy for long cables since you don't have to follow the cable all the way from end to end to know that you have the right one.
4. Consider color coding patch cords. We use the colors to identify groups (subnets) within a wiring closet. The downside is you will have to stock more cables. If you don't keep the right colors handy, your color coding will become almost useless as the wrong color will be used if it was the only one available.
5. Make sure you keep a good stock of the cables that you need on hand. Keep multiple lengths, store some in the wiring closets. Again they are cheap. An inadequate stock of the right cables will result in all kinds of cables appearing in the wiring. There are many RJ45 cables out there, including ones that use flat untwisted wires. These cables are not suitable for 10base-T. What is even worse is they will often work, but only marginally. You do not want to have to debug that problem.
6. Make sure that it is possible to tell the cables apart at a glance. A crossover replaced with a normal patch cable will not work. If you have CDDI, a CDDI crossover cable will not work in place of a 10base-T cable or a 10base-T crossover. We now get colored boots on the connectors of our cables. These boots serve two functions. First they protect the tabs on the cables so that it is possible to pull a patch cable through a group of wires without getting it caught. And second, they identify the cable type. (Since we used the color of the cable already for a different purpose, we could not use that to identify the type of cable.) Boots of different colors help make it easy to quickly identify the type of cable.

Plan For the Space

When sizing and building a wiring closet make sure that the closet has enough room for the space that it is supporting. It is unlikely that you will get a chance to rebuild your wiring closet and get more space in the future. For example, an area that is sparsely populated today because it is a warehouse may hold a dense population of engineers in the future. The Synopsys warehouse has moved three times over the last four years. The previous locations now have cubicles in them. Make sure that the wiring closet can handle those kinds of changes. You can approximate a maximum port count by taking the square footage and dividing by 100 (a typical office in Synopsys is between 80 and 140 square feet). You will also have to take into account exceptions. For instance, lab space may require a significantly higher density of of network connections.

Planning for space also means making sure that a large conference room has enough drops. Yes, it may seem like a waste but it is just about impossible to know how the room is going to be used in the future. The extra drops will make it easier to convert the room into temporary office space. It also means that it is more likely that the drop will be on the correct wall when you want to set up a demo in that room.

Large executive offices also need extra drops. I will guarantee that if you place only one drop in those offices, the vice president will place their desk at the farthest possible point from the network drop.

Plan For Change

Plan enough capacity to minimize, if not eliminate, the need to make changes.

If you must make changes then make additions instead of modifications to the existing wiring. For example, if you discover that a drop is not where you need one, don't move a nearby one, add another one instead.

The vast majority of our wire problems are caused by additions and changes in the physical plant. It is generally better to have planned to have enough drops to begin with.

If you are using a contractor for wiring, the additions or changes will typically be done by the "most available" technician. Often "most available" is the "least able." This is usually not the same person that did the initial install. So you end up supervising more or correcting errors.

Plan for Growth and Spares

We have two RJ45s per drop. This allows us to support the growing number of machines that users ask us to connect to the network. The extra port is also useful when a problem occurs in the wiring. We can "borrow" the spare port from the cubicle or office next door. This is an important trick to be able to pull; we want to avoid absolute emergencies. We would rather schedule the work at our convenience.

Plan for Future Technology

Try to look at what new technology is being developed. We now use Category 5 wire; 10base-T does not need it, but higher-speed stuff almost certainly will. On the other hand, accept that you can not predict the future perfectly; we already know parts that we got wrong. There are parts of the network that will not support the new stuff without some work. The ports with split 10base-T; for instance, ports will require retermination.

Putting the wires to work

Once we had the physical plant done we needed to actually start using it. We would need to define a backbone and the networks that would feed it.

Collapsed Backbone

Instead of a traditional backbone which might require high-speed media, we used the high-speed backplane of today's routers (and bridges, or "switches"). Our first implementation used a multiport router with all Ethernet ports. An Ethernet backbone could become a bottleneck. It could be FDDI to make it faster, but that would have been expensive. By using a collapsed backbone, we are able to avoid investing early in an FDDI backbone.

Hosts Do Not Route

Hosts generally are lousy routers and routing usually detracts from their performance. Using hosts as routers complicates configurations and the network becomes harder to debug. Simplification is an important win when you have limited time and staff to deal with network problems. Make sure that the machines have the *ip_forward* (or equivalent) kernel variable turned off. Tracking down packet loops in networks is not fun.

Services Should Be Local

Services for a group should be as network-local as possible. We do not want to depend on the performance of routers to provide reasonable access to most services. It is too easy to get caught in the trap of trying to figure out how to deliver "wire speed" from one place to another when the correct answer is to locate the service closer. Some servers may require multiple interfaces to get them "local." You will need to pay attention to how packets get to these hosts [Swartz 94]. This may be cheaper, and may perform better, than having a high-speed device in-between.

Services Network

A group may need to "publish" or provide a service to the rest of the network from one of its servers. In order to minimize the impact of these shared services, I created a "services" network. A group's server that is providing the service will have an interface on this network. Access by "remote" users go through this network and interface. This has an important benefit: non-workgroup-initiated traffic is kept to a minimum on the workgroup backbone interface.

Workgroup Model

First, we defined a workgroup as group of systems and their users that shared resources (typically computer systems). Our workgroup model resulted in the workgroup sharing a network.

Our initial workgroup network was a single Ethernet segment (see Figure 4). It would have 12-36 standard workstations (per our computing model) with a target of 24. As a workgroup grew it would be split into two or more groups as appropriate. This model gave us the ability to project our networking equipment needs.

When we were implementing this network, FDDI was the showcase technology at Interop. Showcase meant that the technology was still expensive and too new. Instead of going for the latest technology, we gave our large shared servers multiple Ethernet interfaces. This approach allowed us to keep servers and clients topologically near each other. The performance of the backbone router was not as critical since it was not a part of the routine traffic between workgroups clients and servers.

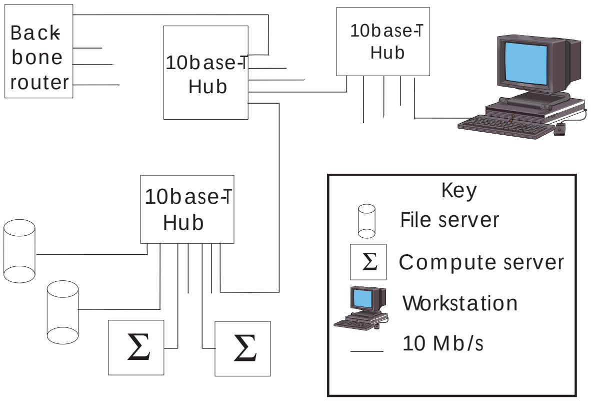

Figure 4: The original workgroup network: One Ethernet connected to the backbone router

As workstations became more powerful it became clear that a single Ethernet segment was not going to be sufficient for a workgroup. It was not practical to split a workgroup network into separate sub-networks that were each smaller than the actual workgroup. Part of the solution is to split the workgroup network into multiple Ethernet segments. Our workgroups today look very similar to the ones we had four years ago. All that we did was modify the root or base of the group. Instead of a hub at the root we now use a high-performance bridge. This modification allowed us to give each group more than a single 10 Mbit/s Ethernet.

The new lower-cost, high-performance bridge/routers (switches) also allowed us to redefine the size of a workgroup. Previously we had to make a workgroup fairly small to fit within a single Ethernet. We are now able to allow a workgroup to be bigger since we can have a fairly large number of Ethernet segments available. This simplifies the administration of the groups. Since there are fewer logical workgroups, there are fewer moves that actually crossed workgroups. This also simplifies some of the resource sharing problems, the larger workgroups can share a server more readily.

A mismatch between the client and server network pipes can occur when switches are deployed. In our computing environment the traffic predominantly is between client and server instead from peer to peer. We have many devices talking to a few servers. It was possible for the clients in one group to generate combined traffic in exceeding the bandwidth of a of single Ethernet. Our solution was to give the servers FDDI interfaces. Note that we were able to delay this decision until FDDI interfaces and hubs had dropped significantly in cost. The introduction of FDDI over twisted pair wires (TP-PMD or CDDI) further reduced the expense required. This led us to implement using workgroup bridges with FDDI interfaces.

There are also additional factors that influenced our workgroup model.

Keep High Speeds in the Computer Room

All along we tried to keep the need for high-speed networking away from the desktop machines because it is much easier to retrofit a computer room for new technology than a campus. We have been able to restrict high speeds to the computer room by keeping the servers in the computer room.

We have taken advantage of our current configuration by providing our compute servers and file servers, located in the computer room, high-speed (CDDI) connections. The higher-speed connections give the compute servers improved access to the file servers without having to rewire the buildings.

At Synopsys, the average desktop machine has not required more than Ethernet. When we started four years ago a workstation shared an Ethernet with about 30 other workstations, and today it is closer to 10. We will eventually be able drop it down to a dedicated Ethernet.

By restricting higher speeds to the computer room we have been able to continue to use the same wiring we had four years ago. In that time, we have significantly increased the available bandwidth between servers, as well as between the servers and the desktop machines, without having to redo our station wiring.

Segregate Traffic

Users' perceptions of network latency are important. The workgroup Ethernet switch architecture also allows us to dedicate a segment to X terminals. The logic behind this approach is that NFS uses large packets and can delay the small packets X usually generates. The effectiveness of this change has not yet been measured.

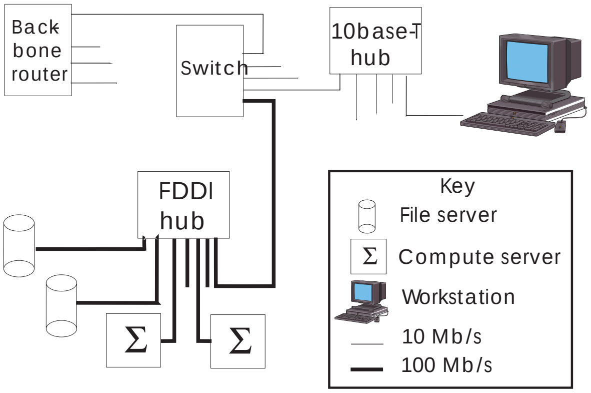

Figure 5: The current workgroup network, built around a local high-speed network and switches

The result is our standard workgroup now consists of a FDDI/CDDI network connected to some number of Ethernets by high-speed bridges/switches (see Figure 5). Compute servers are attached to the the high-speed network (FDDI/CDDI) along with the file servers. An Ethernet segment is intended to be used by 6-12 standard workstations. An Ethernet segment is reserved for X terminals.

An interesting problem arose in the specification of the equipment for this configuration. Choosing to bridge between Ethernet and FDDI led to an interesting trade-off. I could have chosen a simple, less-expensive device to link the Ethernet to FDDI, with the caveat that it would not handle the different MTUs between Ethernet and FDDI. We would have had to ensure that the different host configurations deal with the different MTUs correctly. Instead we chose to specify the more expensive box that could deal with the problem correctly. This choice allowed us to not to have worry about the MTU problem.

Note that in four years a workgroup has gone from a single 10 Mbit Ethernet segment to several (8 to 24 today) and a local 100 Mbit (FDDI) network. The hub and wiring from four years ago are still in service. Even though the available bandwidth for a workgroup has changed by at least an order of magnitude, we have been able to keep the backbone interface at 10 Mbit. We have not added a higher-speed backbone connection because we do not require one yet. I expect that by the time we need one, the prices will be lower. Continuing to use existing equipment helps keep our costs down.

Manage the technology

Managing the technology is an important component of supporting a network.

Simple Routing

Because of the structure of the network, our routing configuration is fairly simple. This simplicity helps cut down the debugging time. It is extremely rare that we have to debug the network using *traceroute* (8). Our system administrators also know the standard IP configuration for an end node. We have a convention for the IP address of the default router on a subnet. A system administrator can configure a node on almost any of our subnets by just understanding this convention.

Ignore the Hype

We were conservative in that we ignored a lot of the hype about new technology. The leading edge can often be the bleeding edge and new technology usually has higher costs. The vendors and the market have a learning period to go through. First-generation products always have bugs. Performance may not yet meet expectations. Unless you have time to spend helping the vendors make the products into the solutions they are hyped to be, it is best to avoid this time sink.

Stay Tuned In

On the other hand, you will have to make a lot of assumptions. To be effective you do need to listen. You need to guess when it's safe to take on board a new technology. If you are unsure about a technology then you may need to test and prototype the new technology. It is a tough balancing act, avoiding products that are not ready yet versus deploying a new technology that is valuable. We ran FDDI in a test configuration for a long time before adopting it as standard. I kept an eye on trade shows. If a technology is taken for granted and is just used without a lot of hype then it is probably mature technology.

Limit What You do

Just because something is possible does not always make it a good idea. Some well-known routers can route just about anything, but being able to route something doesn't mean that you should. I tried to make it easy on myself and the other system administrators. For example, we defined our backbone protocols. We route only TCP/IP and AppleTalk. This simplifies router configurations. Remember: limited staff, limited time.

Augmenting your staff

Synopsys went from about 100 to 1000 people with only a part-time network manager. I did not have the time to do all the work, so I looked for ways to have other people do the work.

Vendors

Vendors can help you track the market. They may present solutions to problems. By asking them how to solve a problem and having them explain their solution, I was able to learn something new. I did have to remember to be skeptical. I also sometimes just had to take (believe?) what they offered, either because I did not have time to research the topic further or because it was an area where I did not have enough expertise.

You should make sure that vendors document the work they do. Make sure you agree up front what the requirements are. This includes having them test the installation and document the test results. Make sure the agreement is in writing. I have made the mistake of not getting the details documented earlier on. In these cases I had to live with what I got since I did not have time to track it down later.

The Other System Administrators

System administrators can be network administrators. We try to make it possible for the system administrators to do some of the work. We choose to give system administrators access to wiring closets. This means they can do a lot of the work normally handled by the networking staff. Because the scheme is simple enough, they can handle adds and most changes. This also requires some tolerance; errors and mistakes will happen. You can decide for your site if it's more trouble that it's worth.

Tuning, growing, maintenance, etc.

Tuning is a tough problem, sometimes you have to guess what you need. You can try to increase the speed of your nodes and your users will tell you the network became worse [Seifert 95]. You may discover peculiar performance characteristics, some machines performing well while others suffer [Molle 94]. Ethernet has been around long enough to be a mature technology, yet unexpected problems still crop up. Fortunately, there are people looking at the problem. You may have to wait for someone else to solve the problem for you.

One of the reasons for designing the network conservatively is to avoid these problems. By having enough capacity you can ignore tuning for long stretches of time. You will also hopefully avoid triggering some of the bugs that appear at the edge of the performance curves by having this extra capacity.

You should also design conservatively so that you do not have to track down all the details. We avoid pushing things to the edge so we do not have to sweat the details. Examples include keeping wiring short so we do not have worry about a station being outside the design limit if we use a slightly long cord. Another example is our Ethernet repeater configuration. We designed it to be shallow. Ethernet has a limitation on how many repeaters can be between two stations. By staying away from margin we know that if we install an extra hub (repeater) at the edge that it will not push us over the limit.

Metrics are hard. Our management wanted a single number to reflect the health of the network. I deferred this problem. See [Hogan 95] for an approach NCS took at trying to address this problem.

We want minimal maintenance. Equipment is cheap, but cheap equipment is expensive. Reliability should be one of your purchasing criteria. Keep spares on site for anything that you can't live without. This seems obvious but it is easy to forget.

Do not forget that standardized parts yield in spares for "critical" parts by stealing them from "less critical" areas. Standardization also helps in stocking spares, since you have fewer spares to stock. Finally, standardization lowers your support costs: once you know one, you know the rest.

Another use for a spares inventory is to deal with unexpected growth. By keeping a stock of parts we can install in a timely manner and deal with a surprise move or acquisition.

Make sure network parts are reliable. Where appropriate, use equipment designed for the job. This may mean using things like redundant power supplies. You do not want to be paged because of simple failure. Save the heroics for the real hard stuff.

Miscellaneous

Don't be afraid to guess. We had to guess. There is just not enough time to get all the details. Chances are you are the expert for your site. You know how systems are used and can design the network to support it.

There are advantages to being both the system administrator and the network administrator. Instead of choosing a path that was just optimal for network administration we chose one that was good for the group. When one our of workgroups moves, the machines do not have to be reconfigured. We make it easier on the system administrator by making the same network available at the old and new locations. This also reduces the downtime for the users involved. It requires a little more preparation and equipment on the networking side but it saves work for the system administrators.

Radios are step savers. Consider having at least a couple of walkie-talkies available. Two people can often debug a network problem a lot quicker than one. We have vendors who do work for us who end up borrowing our radios.

I am a big fan of blinking lights. I think vendors do not provide enough of them. Our standard transceiver has power, link, transmit, and receive lights. The hubs have link lights. Our system administrators can easily determine if a station is on the network or not. Our help desk can also quickly determine if a user has basic connectivity without having to walk to the user's office. We are lazy, after all.

Make things easy on yourself. We made our AppleTalk network number be the same as the IP subnet number. This makes debugging easier for us.

Network management

Network management stations (NMS) are thought to help by giving advance warning of when upgrades or expansion or repair is needed but they are are not panaceas. Deploying and supporting an NMS is not a simple undertaking; additional resources are usually required to configure and maintain it. Even if one managed to get one deployed and supported, there may still be no one to dispatch to solve the problems that it, the NMS, reports.

I tried deploying an NMS three years ago, and it didn't work. The problem is that an NMS requires time to support and maintain and I did not have the time. We did, however, continue to plan for one by specifying and purchasing equipment that supports SNMP.

We are deploying a new one now. Management decided that this is important. We have more staff now so maybe it will work and our networking hardware was ready to support it. We will see how it goes.

Conclusion

The network is alive. We have somehow managed with a limited staff (mostly me) taking care of it. Our design has been surprisingly resilient in the face of all the demands and changes. Most importantly, my successor has not killed me yet.

We now have two full time folks supporting our network. The problems strangely haven't changed too much. Synopsys continues to grow so our challenges continue.

Acknowledgments

I would like to thank my merry band of proofreaders: my wife Laura, Dave Stuit (who can tell the difference between an em-dash and an en-dash), Jeff Jensen (who took over running the Synopsys network), Becky (who let her husband, Jeff, take over running the Synopsys network), Ted Lilley (who recently joined Jeff in running the network), Dave Clark and Christine Hogan. Additional thanks to Erin Elder and Michael Jensen, who had the misfortune of being at the wrong place at the wrong time. They all tried valiantly to find all of my typos, missing words and poor writing; any remaining errors are my fault.

Geraldine de Leon created the figures based on my crude sketches.

Also thanks to Paul Evans for encouraging me in this process, Eric Berglund for making this paper my project, and finally, to Randy Collins for letting me build the Synopsys network.

Author Information

Arnold de Leon is a senior system administrator at Synopsys, Inc. Arnold is the current president of BayLISA. He was also a founding board member of SAGE. He has a bachelor's degree in mathematics from Harvey Mudd College. Reach him via U.S. Mail at Synopsys, Inc.; Building C; 700 East Middlefield Road; Mountain View, CA 94043-4033. Reach him electronically at <arnold@synopsys.com>.

References

[Wall and Schwartz 90] Wall, Larry and Randal L. Schwartz, *Programming Perl*, Sebastopol, CA: O'Reilly & Associates, Inc., 1990, p. xviii.

[Fine and Romig 90] Fine, Thomas and Steve Romig, "A Console Server," *Proceedings of the Fourth Large Installation Systems Administration Conference*, The Usenix Association, 1990.

[Swartz 92] Swartz, Karl L., "Optimal Routing of IP Packets to Multi-Homed Servers," *Proceedings of the Sixth Large Installation Systems Administration Conference*, The Usenix Association, 1992.

[Molle 94] Mole, Mart L., *A New Binary Logarithmic Arbitration Method for Ethernet*, Computer Systems Research Institute, University of Toronto, Toronto, Canada, April 1994, available for anonymous ftp from ftp.utexas.edu as /pub/netinfo/ethernet/ethernet-capture/report.ps.

[Seifert 95] Seifert, Rich, *The Effect of Ethernet Behavior on Networks using High-Performance Workstations and Servers*, March 1995, available for anonymous ftp from ftp.utexas.edu as /pub/netinfo/ethernet/techrept13.ps.

[Hogan 95] Hogan, Christine, "Metrics for Managers," *Proceedings of the Ninth Systems Administration Conference (LISA IX)*, The Usenix Association, 1995.

Working Definitions

This is not intended to be a definitive glossary. The definitions will not be perfectly accurate. They are what could be called working definitions, sufficient to get through this paper.

100base-T, roughly, Ethernet sped up 10 times.

100base-VG, a technology for 100 Mbit networking.

10base-FL, a standard of running Ethernet on one pair of fiberoptic cables.

10base-T, a standard for running Ethernet on two pairs of twisted wires.

62.5 μ is a common size for fiberoptic cables. It refers to to the inside core diameter; the cable itself is much bigger.

AppleTalk is a network protocol designed by Apple, commonly used by Macintoshes.

ATM (Asynchronous Transfer Mode), a faster networking technology that is expected to become the standard of the future.

bridge is a device that connects networks together and makes them look like one network. It does not understand the protocols of the packets it forwards. Forwards broadcasts to all ports.

Category 3 is a standard for cables. See **Category 4** and **Category 5**.

Category 4 is better than **Category 3** but not **Category 5**.

Category 5 is a standard for cables; supports "high-speed" networking. The most common cable used for data applications today. Also used to described an entire wiring installation. A Category 5 installation requires Category 5 wire and Category 5 rated equipment. See **TIA-568**.

CDDI, see **TP-DDI**.

compute server, a computer system assigned to run compute-intensive applications. See **file server**.

concentrator, a box that has ports and looks logically like a wire; a multi-port repeater. For example, a 10base-T concentrator connects several 10base-T ports together into one network. Also known as a **hub**.

crossover cable, a cable that changes the signals so that transmit is now receive and and receive is now transmit. A 10base-T crossover is not the same as CDDI crossover. See **straight-through cable**.

drop is short for **station drop**.

EIA-568, see **TIA-568**.

Ethernet is networking technology for connecting devices at 10 Mbit/second.

fanout is a multiport transceiver.

FDDI is a networking standard that runs at 100 Mbps.

file server, network device that offers file service.

FOIRL is the precursor to **10base-FL**.

hub, see **concentrator**.

IDF (Intermediate Distribution Frame), telecommunications speak for a wiring closet.

MTU (maximum transmission unit), the largest packet that can be carried on a network.

multimode fiber, a kind of fiberoptic cable. Usually used for FDDI and 10base-FL.

multiport transceiver allows multiple devices to share a network port.

network administrator, the networking analogue of system administrator.

network manager, another name for network administrator.

NMS (network management station), a big expensive collection of hardware and software that is supposed to make the life of the network administrator easier. Usually requires a lot of work to maintain.

patch cables, a cable used to connect two ports on patch panels.

phone closet, see **wiring closet**.

PhoneNET, an AppleTalk network that was designed to use one pair of wires in a two-pair phone cable.

repeater takes a signal into a port and regenerates (repeats) it to its output ports.

router, a device that connects networks together. It actually understands the part of protocols that it forwards (see bridge).

RJ11, the connector used by phone jacks; it can contain up to to six conductors.

RJ45, a bigger version of the RJ11; it can contain up to eight conductors.

RS-232, a standard for serial communications.

SDF (secondary distribution frame), see **IDF**.

single-mode fiber kind of fiberoptic cable. Allows the use of lasers to support longer distances.

SNMP (simple network management protocol), a standard for obtaining information and managing devices on a network.

star topology, wiring goes to a central point.

station drop, a box that contain the network jacks.

straight-through cable, a cable that just passes the signals through pin to pin (e.g. pin 1 to pin 1, pin 2 to pin 2, etc.).

sun-ether, 192.9.200, the network number that Sun used as an example in their documentation. Many sites used it as their network number.

switch, the marketing speak for multiport bridge. Usually implies high performance.

TIA-568, Telecommunication Industry Association standard for structured wiring. Describes practices needed to meet various wiring standards.

TP-DDI, FDDI over twisted pair wires.

TP-PMD, describes the part of FDDI that was changed to support twisted pairs.

traceroute, a tool for identifying the TCP/IP route from one node to another.

transceiver, connects a device to the network.

wire management, pieces of metal or plastic to guide cables around.

wiring closet, the room or space where the wiring for a floor or area goes.