Interviewee: **Emhof, Kalen**
Interviewer: **Valerie Modoux**
Date: **February 21, 2008**
Category: **Florida history; Florida Archives digitization**
Status: **Open**
Tape location: **CD** (Box #57)

Modoux: To start with, can you briefly describe your position in the firm you work for?

Kalen Emhof: My position is part customers' support, part technical support. I help the customers with requests they have regarding services we provide for them. Our services being web-hosting company. And I also provide technical support for the company, by designing servers, and web applications; I build work stations for my co-workers. So, those are my two primary tasks: IT support for the customers and technical support for the firm.

Modoux: For the purpose of this project, we are going to focus on storage mediums now. So, let's start with magnetic medium storage. Can you explain to me what they are, and how they function?

Kalen Emhof: Magnetic medium is commonly known in the form of hard disks and floppy disks. Those are the two most common forms of magnetic mediums in use today. Well, the floppy disk is actually dying out. They store data by manipulating the magnetic fields, and the magnetic flux .Magnets have a North and a South polarization. And it flows from the North pole to the South pole. So, basically, magnetic media uses the binary format, which consists of 0 1, and maps them to which direction the magnetic flux is pointing, on whatever magnetic medium you have, be it a hard disk or a floppy disk. The flux can be pointed in one way, North, or the other way, South. And the computer will interpret it in binary format, 0 or 1. And how that works is, you have a platter of magnetic medium, it can either be a floppy disk...what they have imbedded on them is what they call sectors. Those sectors are just a very small isolated grid of magnetic ions and the computer can read those or can write those by sending an electrical charge and reversing the magnetic direction. Either changing it from pointing to South, or changing it to pointing to North. So it will store the digital format, being binary. Computer language is all.. . in its most basic format, it is zero or one; yes or no. Computers use a binary format. So, that is how the magnetic media of today works for storage.

Modoux: What are the advantages of such mediums?

Kalen Emhof: It is currently the cheapest format to produce it. That is really the huge advantage. It is the industry's standard; it has been around; it has developed, as you can see now —. You know, computers used to take up an entire room, for storage and everything. And now, we can fit a thousand gigabit in a standard hard drive. Hard drives are getting more complex, built in better and better quality.. . Basically, it is the most matured media device, storage device

we have on the market today.

Modoux: Just one clarification. When you talk about hard drives or hard disks, you are including both external and internal hard drives?

Kalen Emhof: Yes. They use the same technology. It does not matter where they are in the physical location of the computer -inside or outside the computer, what interface they use (be it parallel ATA, Serial ATA, USB, Firewire, external) -it still uses a magnetic media. Basically, your external hard drives are just hard drives in their own little box, with their own power supply. And that's it. And you can basically build your own portable hard drive, by going to a computer store, buying an external enclosure purchasing the hard drive, installing it in the enclosure, and then you are good to go!

Modoux: Now, are there any risks coming from using external drives like that?

Kalen Emhof: Well.. . Now, I know you are coming from a historical vintage point, so you are going to look at long term storage. I would say "yes", there can be some long term storage risks with magnetic media. If what you are doing is taking whatever it is you are preserving, whatever it is you want to store, sticking it on a hard drive, taking the hard drive and then putting it up on a shelf. I do not know if you have ever done this before, but you can take a screwdriver, or a paper clip, or anything metal, take a magnet, and stroke that magnet in one direction along whatever piece of metal you have. What you are doing is that you are aligning the metallic ions in there and you are basically turning it into a magnet. And you can take it and pick up several things. For example, if I had a paper clip in here and a fairly strong kitchen magnet -the type you put on your refrigerator -run it along that; I could collect staples at the end of the paper clip. Now, if you were to come back here maybe three or four weeks later, and I had not done anything with that paper clip, I would not be able to pick up as many staples as I could. Because it is not magnetic in itself -it is not magnetic as in what we would mine out of the ground -it will lose its magnetic properties. And that is what will happen. And that is actually a danger I foresee as hard drives are growing larger in the capacities storage, we have not actually physically grown the size of the hard disk itself. I am not talking in size as in storage, but physical properties. They are still 3.5 inches wide by 5 inches long and one inch high. And basically, you are taking the same size drawer you had, and you are fitting more and more stuff in it. Now, what can happen is the same thing that will happen with the paper clip. It will lose its magnetic capabilities if just left alone. And the sectors that I spoke of earlier, in order to store more and more data, we have to be able to fit more and more sectors in the drawer; and the drawer is not growing bigger. In fact, it is actually growing smaller. So, we have to create the sectors even smaller. Remember, all the sectors are storing is the direction of the magnetic ions. So they can be really small. But what happens is, when they are smaller, it is like a radiation half-life. Radioactive material has a half life of two days, meaning after two days you have half of the radioactive material that you had to start with. If you have ten pounds of this radioactive material, and one pound of radioactive material, which one is going to last longer? The ten pounds, because it is going to take longer for it to degrade to the extend of the one pound. Now,

the half-life works exponentially, so the math will not work out exactly. But the point is, it will take longer for the ten pounds to degrade than it will take for the one pound. So, with magnetic sectors shrinking smaller and smaller, it will not take as long for them to lose their magnetic reference, as it will for the same size disk that only carries 100 megabytes, versus 320 megabytes. The sectors are larger, and it will take longer for those magnetic properties to degrade. I have 5.5" floppy disks .that only store 500 kilobytes, compared to my floppy disks which carry 1.44 megabytes. I have had floppy disks die on me; information I had put on them, drivers, bios updates, I have put them up on my shelves, five-six years later I come back to use them, and the data is not there. In fact, I have to throw out the floppy disk, buy a new one and reacquire my data. Because it loses its magnetic properties. Now, hard drives in a computer are not so susceptible to that, because they are constantly spinning; the computer is accessing, reading them, and ensuring that the little magnetic ions point where they are supposed to. So that is a potential danger I see in magnetic mediums. And, like I said, there is no way you can test it unless you stick it in a time machine so you can speed up the amount of time you have to wait! But it is a real physical problem. It is scientific: they lose their magnetic capabilities.

Modoux: I know there is another type of medium storage that you can use, which are the optical ones. So, again, can you explain what they are, and the ways in which you can use them?

Kalen Emhof: Optical mediums manifest themselves in the market as the common CD and DVD. Those are read by a laser. And again, how it takes that optical medium and transforms it into the binary system of 0 and 1, yes or no, true or false, is they take a very thin metal disk -very very thin, it is almost like tin foil..If you have ever taken a CD disk apart, which I have [laughs], it is very much like tin foil and it is very reflective. And this is important, because optical disk use a laser to read the data on the disk. What happens is: it has very minor imperfections, designed imperfections, that are stamped into it, and then protected by a coat of plastic beneath it and then on top of it. What happens is, those imperfections, when the laser light shines on the compact disk, it either reflects right back down, or the imperfection throws the laser off. So if the laser hits the receptacle back in your CDRom, or your DVD player in the computer, let us theoretically say that it hits it at zero, and if it misses the receptacle, it reads it as one. So there you have the transformation into binary format. Now, physical imperfections in the metal -which they are not physical imperfections, they are designed that way. So the deformities in the design tell the CD to reflect the light differently are not susceptible to the same degradation as magnetic media is. If you take a CD from retail -and I will differentiate between burnt CDs and stamped CDs -wrap it up, protect it from the light, protect it from physical damage, pretty much like anything else you are going to store, put it on the shelf, walk away.. . In ten years, you will come back and those physical bumps, the designed deformities in the DVD, are not going to magically disappear. They are not going to lose their shape, like magnetic media does degrade. Now, because your reading surface is exposed, unlike magnetic medium in a harddisk -you know, it is enclosed in metal -yes, it is susceptible to another level, another type of damage. And that is why burning CDs, copying DVD, is so popular: because just due to regular use, they suffer from wear and tear. You know, something in the disk, in your CDRom, scratches it, and the scratches in the plastic throw off the light. So, that is why people burn

copies all the time. But now, let me get to burnt CDs. Burnt CDs, when they first came out -and I have not checked up on the latest ones (the thing about the computer industry is that it is always changing) -but I know it is the same with DVDs what I am going to reach to: burnt CDs use a dye, not an actual physical deformity. And a dye, when exposed to light, or over time, can degrade just like magnetic media. So for an archive place to take their items, burn them on a DVD and then stick that DVD up on a shelf.. . Again, DVDs have not been around as long as hard drives so I cannot say the shelf-span like we have the experience with hard drives, but just according to its physical properties: yes, it will suffer degradation from sitting there. All things tend to disorder, it just depends where the source of the disorder comes from. When I say stamped CDs, yes they tend to disorder and degradation just from regular use. From non-use, they are not going to suffer so much, unlike a hard drive. So, stamped CDs, those to not suffer from the same type of degradation as burnt CDs.

Modoux: Can you briefly explain what stamps CDs are as opposed to burnt ones?

Kalen Emhof: Stamped CDs are basically what you buy from the store. They have huge CD manufacturing implants, they take the data, they create a cast, and then they stamp the thin sheet of metal that is coated by those plastics on either side, verify the integrity of the data, and then ship it out. Whereas on your computer, they build the CDs with little dyes in them, so that when the laser hits it at a certain wavelength frequency, it changes its reflective properties -that is why you need to have a special CD burner; the laser is modulated, you can change the energy of the laser -at a certain energy, it will change the dye chemical composition, and therefore the light of the laser will be reflected differently. Again, equaling your 0 and 1 binary format, which the computer understands. So, that is the difference between what I refer to as stamped CDs or burnt CDs. Stamped CDs have a much longer shelf life than burnt CDs, because of the way they are produced and designed. Metal does not degrade like a chemical dye would.

Modoux: So, they might suffer from time too, but in a different way? Or it will take them longer?

Kalen Emhof: It will take them significantly longer. But CDs are still young; I think they made their huge debut back in 1991 -1992 -this is when they first started to be widely accepted. So we only have about sixteen, eighteen, twenty years to go. Whereas hard drives have been around since the advent of magnetic media, so we have more historical data on them, than we have for CDs and DVDs.

Modoux: Just to make sure that I grasped all the information: How long would you evaluate the life of a hard drive that you stick on a shelf?

Kalen Emhof: I really have any hard data on that. [ponders] Well, I have had that hard drive sitting on my shelf for quite a while now, and when I took it the other day, it would not work anymore. [goes look for the hard drive, and eventually finds it]. This is a Fujitsu brand IDE hard drive. And, let's see here [looks at the hard drive]...This is an eight gigabyte hard drive,

manufactured in...I am not seeing a date on here, but I would say it is at least eight to ten years old, considering it only stores eight gigabytes, and now we have hard drives of the same shape and size, that will store one thousand gigabytes. I have hard drives in my computers that will store two hundred and fifty gigabytes, and that is actually pretty low for right now -for desktop storage. Right now, it is becoming hard to find a regular, main stream desktop being sold for retail that carries less than five hundred gigabytes of storage. This one drive has only eight gigabytes...But before you think: "oh, you know, maybe he does not know what he is doing, what he is saying." Like I said, my job is, building servers, building desktop computers. I had a server that I just wanted to use this hard disk to boot it up, just to run to properties test, so I would not have to stick a huge hard disk in, reformat it, install the operating system and all. It would have gone faster with the eight gigabytes, but it would not read the disk, so...It could be due to multiple problems, but you know...

Modoux: They are not really stable?

Kalen Emhof: Well, it has been sitting on the shelf for who knows how long. I honestly believe it has lost its magnetic property. And it is very physically possible and real. I do not have any lab to run that sort of tests, but from my knowledge of computers, and how they work, this is a strong possibility. So, I would highly recommend against taking a five hundred gigabyte drive, throwing a massive amount of data on it, taking it out of the computer, wrapping it up in an anti-static bag, sealing it, and putting it on the shelf somewhere safe...No, I would not recommend that. Not if you expect to come back five-ten years later and still recover everything. No, I really do not see that happening.

Modoux: Now, what would you think about storing data on a server?

Kalen Emhof: That is an industry standard. It is what all the companies do. They are called file server. Basically, all they are is a computer specifically built.. . I actually just built one the other day for our firm.. . specifically built to handle massive amount of storage, and to maintain that storage, i.e. the computer is constantly on; the hard drives are constantly on, they are being fed electricity to maintain those magnetic sectors and make sure each stays align as it is supposed to be properly aligned. And, there are techniques -I really cannot say it is software because it can be done both ways -there is a technology that has been around for over twenty years, called RAID and it stands for Redundant Array of Inexpensive disks. Basically, what this does is that it takes multiple hard drives -and what it does is very basic in essence -combines them together so that the computer reads them as one single disk. And there are very different ways of implementing RAID. There are several different RAID levels; Raid 0, Raid 1,2, 3, 4, 5, 6.. . and they are combining more and more. Each one serves its own purpose. For example, a RAID 0 takes two hard drives -you have to have a minimum of two hard drives -and it stripes the data across the hard drives. Say you have a file, it writes half of the file on one disk, and the other half of the file on the other disk. In effect, you double your storage. So if you have two one hundred gigabyte and you stick them in a RAID 0, you have two hundred gigabytes of storage. Now the problem is this: if one hard drive fails, you are left with a hard drive that only has half

the data of the files in the computer. So, in essence, it is a catastrophic failure should one drive fail. You lose everything. Now, RAID 1 mirrors the data. So you take two one hundred gigabyte drives, and combine them together to form one 100-gigabyte drive. So you take a file, and it writes it on drive A and it writes it on drive B. So you have complete data redundancy. You do not lose any data, should one hard drive fail. You just stick a new hard drive in, and depending upon which raid array you built, and if you built it with a fully capable controller, it will just take the data from disk A and write it on disk B. And you have preserved all your data. And as you get further on and on, then it becomes more complex and you can institute data checks, to check for data integrity, to make sure that what is actually written is what you really wanted and it says that way. For me to go into the technical details of that would be beyond the scope of this interview, I guess. The most popular one is RAID 5, when it comes to massive amount of storage and what that does is that it takes multiple disks..So let us say four disks of a hundred gigabyte, you do not end up with four hundred gigabytes storage; you end up with about three hundreds and twenty-five, or three hundreds and fifty gigabytes.
It depends on what RAID controller you use. There are so many variables. [someone knocks on office door. Short interruption]

Modoux: So, can we start again from RAID 5?

Kalen Emhof: Yes. What RAID 5...For example, let us take four one hundred gigabytes drives and instead of having just one hundred gigabytes with RAID I, or four hundred like you would have with RAID 0, what you have is almost a combination of the two. So you probably end up with about three hundred and fifty gigabytes; it all depends how your RAID controller implements it. What it does, is that it does some striping, mirroring and parity checks. And it spans those not only on one individual drive -like RAID 3 which does all of that, but only puts it on one drive. So if that one drive fails, you lost everything. But RAID 5 stores this information across all the disks in the raid array. So that if one drive fails, you do not lose any data. You just stick it in, and it rebuilds the data from the other drives in the ray. And it can hold, theoretically, a unlimited amount of data. However, if you lose two drives -and you know, the possibility of losing two drives at once is kind of low, but it becomes more of a probability with the more drives you have -it cannot rebuild the data. But if you lose one drive, you can still recover the data and rebuild the array. So, if you have between twelve and twenty disks in a RAID, then you would probably be better off running a RAID 6. It is identical to a RAID 5, it just stores the array specification and the information better across the hard drives. So if you have two hard drives failure, you do not lose the data like you would in a RAID 5. But RAID 5 is pretty much an industry standard when it comes to the enterprise world. RAID 0 is also quite common in the desktop world -people wanting more storage and they are not handling sensitive data, so if they lose it, it is not that bad. They just re-install their operating system, put on their games or what not, and usually the desktop market has an external hard drive where they regularly back up things to, if they are wise when running a RAID 0.

Modoux: Now, my last question is: were you to be an IT for, let us say an Archives or a library, and the institution needs to digitize different items (audio files, pictures, documents).

from your experience and knowledge, how would you tell them to backup all those items?

Kalen Emhof: One of my first questions when anybody asks me this in my industry is: "what is the budget you are working with?" Because in computers there are so many ways to accomplish what they want and usually the biggest determining factor for what path you take to achieve your goal is your budget. And also, what are your intend and purpose? Probably the most feasible solution would be putting together a file server, that way you can add more hard drives to your server. You cannot add more data to a CD once it is full. With a raid array, once you start running out of space, you add more disks.

Modoux: Would you think that combining different types of storage would be a good thing too?

Kalen Emhof: Yes.

Modoux: And what would you mostly likely combine?

Emhof: [thinks about it a little]. I am not an archivist, so it would be hard to say [laughs]. But for me personally.. . I do not know if my personal example is applicable in the sense, I do not know if it is...But, I have a lot of clientele information in my computer, so I actually run a RAID 1, that means I mirror. If one hard drive fails on my work computer. I just replace it and I do not lose any data. I also back those up to our work servers, and we have a huge external drive, which we back all that data back on. So basically, we have three working copies of all the information of our company.

Modoux: So that is RAID, external hard drive and the server?

Emhof: Yes. Well RAID and the server are the same...The server has a raid array also, but it
runs a RAID 5. My computer is not capable of a RAID 5 because it is a desktop. I cannot build a RAID 5, because it costs money I do not have [laughs].

Modoux: So, if the institution does not have the money to get CDs or DVDs actually stamped,
they should try to combine different types of storage?

Emhof: Yes, that might be a solution. Honestly, I really do not have any prices on what it takes for CDs to get stamped. I have never had to get CDs get stamped. I have always burnt my CDs because they have proven themselves as reliable format for home use. But I do not have to store my data for the next fifty years. Chances are, I will not even care about the data fifty years from now [laughs]...which is not the same for archives!

[End of transcription]