

Figures:

File for Subscriber Downloads:

Where is My Data?

Whil Hentzen

I have a couple of friends who have a computer in their family room – and the whole family uses it – both mom and dad, both teenage girls, and the two pre-teen boys. They use it primarily to surf the web and send email through AOL, but also do some family accounting, digital camera work, and writing of term papers and memos for volunteer groups.

One day I was talking to one of the parents, and we were trying to figure out how to send pictures through email to them – since I'd never used AOL, I didn't really understand how the mail client software worked, and I couldn't tell how they were receiving mail and the associated attachments.

In order to do this remotely, I started at the beginning – with how they logged onto AOL, with a 'screen name' and password, and then how their mail looked – in a list, one by one, or whatever. And, then, how they opened an email, how they read it, responded to it, and so on – just like you'd do when walking through a customer who was having a problem with software that wasn't acting the way they thought it should.

Once we had that down, we looked at how attachments were indicated, how they opened mail with attachments, how they determined what type of attachment it was – a DOC file, a ZIP file, a JPG, or whatever. And then, given the type, how they viewed, read, or otherwise handled the attachment. At one point, I asked them where they stored the attachment for later use, and the light bulb suddenly appeared – they had no idea.

In fact, they had no idea about where their mail was stored. They knew they had to be online to look at their mail, even if they'd already gotten it once and read it – in order to re-read a message, they had to be online. This meant to me that their mail was not stored on their own PC, but on the AOL servers. This stunned me. Actually – two things stunned me, in different ways.

The first was that they didn't "own" their own data. All of their mail, and, then, presumably, the files that went along with their mail, was stored elsewhere, and they were subject to the whims of a third party. Imagine if you had to go to the post office each time you wanted to read a letter someone sent you, or had to go to the library to read a magazine that you subscribed to! OK, those examples are a little far-fetched, but they were the first things that went through my mind.

Gradually, though, I became a bit more comfortable with the idea – while there isn't anyone I'd be comfortable handing my data over to now – I can see that at some point, that might change. It's a hassle to keep my own servers running and maintained and backed up and safe.

And it's not a big leap to move to the service of letting someone else manage my data. After all, I let someone else generate my electricity, grow my food, and provide telephone service, right? Of course, those utilities are rather mature.... Eventually, data hosting utilities could become just as mature, right?

The key requirements for third party data hosting are two-fold – permanent connections and high bandwidth. Remember back ten or fifteen years ago – the old call, "Network is down!" Nowadays, it's really fairly rare for a network to go down – sure, it happens occasionally, but the days of servers being down for days at a time on a regular basis are behind us. We take the existence of a network connection to be as assured as the existence of a mouse (or pointing device, ahem!) next to the keyboard.

It's getting like that with a connection to the Internet (it doesn't surprise you that I'm assuming access to this supposed data hosting utility would be done through the Internet, does it?) Most everyone has one, and more and more, we get nonplussed when the hotel we're staying at doesn't have a broadband port of some sort or another.

So we're getting there...

Remember when I said that two things stunned me about my friends' attitude about their mail being stored on AOL servers? The second one – and the more startling one, actually – was that they didn't know – or care! It was totally transparent to them – whether their data was on their machine or on a *nix box in Colorado or Virginia or the Philippines – the question never even entered their minds!

And that attitude is what's going to make data hosting possible – the idea that people will be able to get to their data from anywhere, and they won't care where it actually "is" as long as it's accessible.

I wrote about the two terabyte server I was planning to install in my basement one of these days.... That's a lot of space – I'm wondering if maybe it'd be easier and more foolproof to have someone else host it – so I can store whatever it is people keep in their basement when they don't have a half-dozen computers there.

I suspect this approach is going to overtake us faster than we think – when 4 pound notebooks are sporting 50 and 100 GB hard disks – can the days of personal terabyte servers be far off? Think about this as you're thinking about your next few applications. The apps you build now could well still be used in ten years. Where's the data going to be in ten years? Will your users even care?