The Data Explosion

Here it is, not even halfway through the year and your IT people are warning that you're running out of disk space on the server again - even though you added space just last year. It seems there is never enough - and nobody is bothering to remove old documents, drawings and files from project directories. When a project is complete, the project team wants to keep the files on the server. They want to be able to reference information for other projects.

I've got to agree with the project teams. As I've mentioned in this column before (December, 2000); an A/E firm's only resource is knowledge, and it's only product is information. Pulling valuable project information from your file storage makes it unavailable to your teams - and your knowledge collects dust in a fireproof safe with the rest of your backup tapes.

Most businesses see a 50% data storage growth every year. If you're using digital photographs in your proposals and field reports, you're probably seeing closer to 100% growth per year in your data storage needs.

MasterCard knows how valuable information can be - and how quickly data can grow. They have a physical warehouse full of disk drives holding about 7 **Terabytes** of information about our purchasing habits, which they sell to manufacturers and marketers. When MasterCard's Data Warehouse was created in 1995, it took 350 hard drives to contain 1 Terabyte of data. A TeraByte is a thousand Gigabytes (or a million Megabytes).

But you keep running out of space, and often getting more space means getting another server to manage it! How can we hope to cope with the vast amount of knowledge that keeps accumulating?

If you have not already done so, dissociate your hard drive storage from your servers. Inexpensive NAS devices are easy to plug into your network with no server upgrades and no downtime. Second, buy the storage you need. It's always cheaper than pruning project files, only to restore the ones you will need again. Third, if you **do** need to prune files, prune annually based upon the recommendations in the sidebar.

Of course, the biggest factor in this equation is the limited capacity of each storage device. The industry is approaching the maximum capacity of traditional magnetic-media disk drives, and we're all looking for the next storage technology that will increase storage capacities by at least an order of magnitude.

Around 1877, Edison invented the phonograph. It captures data on the surface of a spinning platter. Since that time, we have simplified the signal by digitizing it and have replaced the needle with a laser, but we're still using devices very much like the early phonographs to listen to our music - and to record our every-growing mass of business data.

> If you **do** need to prune files, annually move files that have not been accessed for 24-36 months to tapes (3 copies), and maintain an active file on the network listing pruned files which identifies which tapes contain each.
>
> It's important to realize that a file **access** date is different from the file **creation** date. Never judge file use based upon the creation date.

Inside the hard drive are several 2 1/2" phonograph records with 2 tonearm 'needles' for each platter - one for each platter side. These 'needles' don't touch the surface of the platter, but use magnetism to record and read the pattern of data hovering a few microns away on platters rotating at 7200 rpm or more. The geometry hasn't really changed much since the phonograph. Each point on the surface of each platter records one distinct piece of information. Each platter side can hold as much as 7 Gigabytes or so, so higher capacity hard drives simply stuff more platters and heads inside the casing.

The most immediately promising alternate technology is Holographic Data Storage. Holography was theorized in 1947 by Dennis Gabor while trying to improve the resolution of Electron Microscopes.

Unfortunately for him, the process requires a bright light of a single wavelength to work, and the ideal light source for it - the laser - was not invented until 13 years later.

Using Holography, you can record different images at different depths in the medium, or you can alter the wavelength or angle of light to record many different images in the same place.

Holographic data can be retrieved at the speed of light from a stationary medium the size of a sugar cube, or 150 times current DVD speeds. It's estimated that holographic techniques can pack trillions of bytes of data into a standard CD. A sugar cube a centimeter square should theoretically hold a terabyte of data.

Holography has many attributes which makes it an ideal data storage method. Like Edison's Light Bulb, everything is worked out except for the perfect medium. In Edison's case, it was the ideal filament that turned the light bulb from a curiosity to a useful invention. For holographic data storage, it will be a recording medium (probably crystalline) that is clear and visually stable enough to keep the recorded information safe over time.

Until Holographic Data Storage hits the shelves, how are you dealing with the data explosion in your office? E-mail me to let me know your tips and methods.

Michael Hogan, AIA - head chiphead at Ideate, provides custom web solutions and provides consulting services to the AEC industry in Chicago. He welcomes comments by e-mail at mhogan@id-8.com