

# The Old New Thing

## You'd think that with the name scratch, people wouldn't expect it to be around for a long time

7 Jun 2011 7:00 AM

61

There is a server run by the Windows team called scratch. Its purpose is to act as a file server for storing files temporarily. For example, if you want to send somebody a crash dump, you can copy it to the scratch server and send the person a link. The file server is never backed up and is not designed to be used as a permanent solution for anything.

The Windows team likes to use the server to test various file server features. For example, the scratch server uses hierarchical storage management and migrates files to tape relatively aggressively, so that the HSM development team can get real-world usage of their feature.

The file system team will occasionally wipe the hard drives on the server and reformat them with a new version of NTFS, so that they can put the new file system driver through its paces in a situation where it is under heavy load.

When these sort of "mass extinction events" takes place, you can count on somebody sending out email saying, "Hey, what happened to the sprocket degreaser? It was on \\scratch\\temp\\sprocket\_degreaser, but now I can't find it. I have an automated test that relies on the sprocket degreaser as well as some data files on \\scratch\\temp\\foobar\_test\_data, and they're all gone!"

— Um, that's a scratch machine. Why would you put important stuff on it?

"Um, well..." (scratches forehead)

— Okay, well before we reformatted the hard drive, we copied the data to \\scratch2\\save, so try looking there. But remember, the scratch server is for temporary file storage and comes with no service level agreement.

"Oh, phew, thanks."

You'd think that with the name `scratch`, people wouldn't expect it to be around for a long time. Maybe they could call it `can_be_reformatted_at_any_time`.

[Raymond is currently on his way to sunny Hawaii; this message was pre-recorded. Mahalo.]

### Blog - Comment List MSDN TechNet

#### Comments



p

7 Jun 2011 7:05 AM

#

We have a file share named "temp" on our network which is used for important order data that shouldn't be deleted. As in, that's actually the point of the folder. \*sigh\*

**nathan\_works**

7 Jun 2011 7:17 AM

#

Hope the trip is fun, hike a volcano or learn to surf..

**henke37**

7 Jun 2011 7:33 AM

#

Sounds like these are the same kind of people who stores their documents in the bitbucket.

**Nick**

7 Jun 2011 7:37 AM

#

When I saw the title I thought it was about the scratch program.

**Adam Rosenfield**

7 Jun 2011 7:40 AM

#

@Nick: Same here. But good ol' scratchy's still kicking around after over 6 years!  
[blogs.msdn.com/.../410773.aspx](http://blogs.msdn.com/.../410773.aspx)

**Joshua Ganes**

7 Jun 2011 7:54 AM

#

This reminds me of a place I used to work. One department started saving all their files to the "backup" folder. It seems they wanted to make sure the files would be backed up.

**Bryan**

7 Jun 2011 8:21 AM

#

We have a scratch server that suffers from the problem that it gets non-periodically deleted with no warning. More than a few times I've seen situations where someone puts up a file on scratch and leaves for the night after sending an e-mail to say something is on your scratch and arrives the next morning to find it was deleted a mere hour later.

**Maurits [MSFT]**



7 Jun 2011 8:30 AM

#

We had this problem at the last place I worked. The design was a file server with various different shares, each with different permissions; fine so far.

Everybody had a personal share they could use for their stuff; no-one else could see it. This worked well.

Every department also had a share of their own; everybody in the department could see it, no-one else. This worked too.

Then there was a global share that everybody could see. This had the problem described above.

People had a tendency to put *\*everything\** in the global share. Over time, things accumulated. No-one would know whether a given thing was still needed.

I solved the problem by enforcing /periodic/ (weekly) wipes of the global share. On the weekend, everything is moved into a `.\lastweek` folder (stuff in the `.\lastweek` folder from before is deleted.) The periodicity ensured that everyone abided by the intended usage for the share.

This had the side effect of shaking out a lot of legitimate needs for interdepartmental file sharing. "Accounting and Sales need a long-lived share for budget files." "What did you do before?" "We stuck the budget files on the global share and used Excel's 'Save with password' feature." When these came up I would create a special-purpose share (in this case, a "Budget" share) and grant access only to the departments that needed it.



**Simon Farnsworth**

7 Jun 2011 8:34 AM

#

The trouble is that it doesn't matter what your rules are - someone will get annoyed when the scratch server loses things.

I worked somewhere where the rules were simple:

1. Anything created more than 90 days ago will be deleted.
2. Anything not accessed in the last 30 days will be deleted.
3. If the scratch server has less than 1GB space free, the oldest file will be deleted to make room. This will be repeated until the scratch server has at least 5GB free.

Nice and clear, and yet people struggled with the idea that something they put on the scratch server 6 months ago might not be there now - or worse, when someone puts a big file on the server, their 14 day old file might disappear. In the end, IT gave in and removed the scratch server completely - it was too much hassle to support the idiots, and cheaper to just expand the main filestores to have room for big files, so long as people cleaned up after themselves.



**Martin Cowen**

7 Jun 2011 8:56 AM

#

Reminds me of the story of the band Erasure who lost a master tape due to it being left in the studio labelled up correctly and someone decided it was an instruction instead of a name.

**Jason**

7 Jun 2011 9:07 AM

#

We had the same problem with a network share named 'temp'. After a couple of complaints about missing files, the share was renamed 'temp\_DAMMIT' and all was well. Just kidding... some people still didn't get it and complained bitterly when their files would disappear.

**JamesW**

7 Jun 2011 9:13 AM

#

The lesson is obviously to store all your critical files on //scratch2.

**Ross**

7 Jun 2011 9:21 AM

#

I've known people who stored files in the recycle bin because they thought it was a cute directory name.

**Rick C**

7 Jun 2011 9:29 AM

#

I was curious to see what the overlay looked like so I wrote a trivial program to set the bit. On Win7, instead of a black clock, you get a gray X.

This isn't intended to be any kind of nitpicking--I'd never seen the overlay before, so I ran the program on Win7, and saw the X, so then I ran it on an XP machine to see the clock. The file properties dialog on XP does not have any way of showing the offline flag, so you can't find out what the overlay means, other than the file size being in parentheses (as mentioned) in a command prompt directory listing. In Windows 7, there's an Attributes tab on the dialog that shows the file has the O attribute. The dir command doesn't appear to have a switch that shows attributes.

Of course, as Raymond has pointed out, you wouldn't normally see the overlay so it wouldn't be an issue.

**ias88**

7 Jun 2011 9:34 AM

#

It's a very common problem with editing attached files, too - the (non-MS) email client we use at work presently will stick an attachment in %TEMP% and open that file in the appropriate app. Make a load of changes, save, exit, go to send the updated file back ... whoops, it was automatically deleted from %TEMP% as soon as the application exited, since you were no long using that temporary file. With a bit of luck, that should change with our switch to Office 365, saving my users some hair loss and wasted time.

**Ray Trent**

7 Jun 2011 10:06 AM

#

Hey, is that where you mount your scratch monkey? ([www.catb.org/.../scratch-monkey.html](http://www.catb.org/.../scratch-monkey.html))

**James**

7 Jun 2011 11:08 AM

#

Do non-technical people understand what "scratch" means? "Temp" seems like it'd be better, although obviously still not foolproof given some of the other comments.

**James Schend**

7 Jun 2011 11:35 AM

#

@Other James: It's not 1995 anymore, how about naming it "temporary"? "temp" could stand for a dozen things, most of which do not imply the files could disappear. "Temperature?" "Temper?" "Temple?" "Tempura?"

I still like the suggestion of "these\_files\_could\_be\_deleted\_at\_any\_time" though, it's impossible to interpret that wrong.

**640k**

7 Jun 2011 11:49 AM

#

In my experience, when admins didn't get funds to their pet projects, they usually delete users' files for fun, or to harass users. Or both. Admins always complain the isn't enough disk space. Come on! Disk is \*cheap\*. Often a magnitude or two more cheap than letting people spending hours on deleting small old files.

**Wag**

7 Jun 2011 12:01 PM

#

While disk space is cheap, installing said disk space is not cheap in time, money or hassles.

**pc**

7 Jun 2011 12:06 PM

#

It does seem that with the cheapness of space now, it's cheaper to buy more space than spend anybody's time looking for files to delete or coming up with automated ways of deleting them. Perhaps sad, but probably true.

**scld**

7 Jun 2011 12:15 PM

#

@James Why should technical people know what 'scratch' means? I would assume that only sports people know what scratch means! :-p

**Ben Voigt**

7 Jun 2011 12:50 PM

#

@James, @scld: I'm pretty sure the meaning of "scratch paper" is widely understood, and not just by technical people.

**Joseph Le Brech**

7 Jun 2011 1:01 PM

#

maybe 'scratch' should be renamed to 'current' or 'thisWeek', then a script would move 'thisWeek' to 'Week##'

**DJ FABULOUS**

7 Jun 2011 1:37 PM

#

EVBODY NOES SCRATCH IS DA SWEET SOUND OF MA NEEDLE BUSTIN DAT BEAT

**KH**

7 Jun 2011 1:49 PM

#

Name the share \\PleaseDeleteThese\\Files

**edgar**

7 Jun 2011 2:56 PM

#

Raymond, STOP. Don't move.

Hawaii is... scratched !

**Andrew Pennebaker**

7 Jun 2011 3:23 PM

#

"scratch" is an emacs term, but that would not have clued me to the fact that the server is regularly erased. Please rename it something like TMP, which people do associate with temporary storage. Microsoft has a legacy of bad naming conventions.

**GrumpyYoungMan**

7 Jun 2011 3:46 PM

#

@Andrew Pennebaker

>"scratch" is an emacs term

I really don't know whether to laugh or cry.

**Marcel**

7 Jun 2011 3:54 PM

#

Scratch of course means pocketing the white ball in pool! Which is a foul... and the white ball is gone... taking all files with it! It's obvious, isn't it :-)

Also: Itchy & Scratchy. Scratchy always loses. In this case files.

**Mark**

7 Jun 2011 4:27 PM

#

>can\_be\_reformatted\_at\_any\_time

Which, btw, is not a valid ietf name. If you did name it that, then put a web server on it and directed IE at it, it would silently disable cookies.

That one took a long time to figure out. Don't put underscores in your hostnames if you want to use them anytime after the 1990s.



**Maurits [MSFT]**

7 Jun 2011 4:59 PM

#

To clarify my earlier comment: this in no way stopped people from complaining that their files were deleted. People still occasionally complained.

I said this "solved" the problem because:

a) the global share took up a reasonably small amount of space now - so much so that we could back it up

b) the frequency of the purging ensured that everybody was aware of it, and they were able to self-medicate by dragging from the .\lastweek folder to the main folder.



**Ens**

7 Jun 2011 5:16 PM

#

@Ben Voigt

I'm not familiar with scratch paper. I'm familiar with scratch pad. I would not necessarily expect by that analogy that I couldn't put something there in the long term. I'm from Canada, for what it's worth.

According to this: [forum.wordreference.com/showthread.php](http://forum.wordreference.com/showthread.php), "scratch paper" may be regionalised to just certain parts of the U.S. (there's another Canadian who agrees with my usage, and some Americans that claim never to have seen that use before).



**JustSomeGuy**

7 Jun 2011 6:37 PM

#

My question is this: WHY is this machine being used for such disparate purposes? Surely the NTFS team could buy their own machine for testing new features. Surely the aggressively-stash-on-tape team can have their own server as well. What happens when the former buggers up the environment for the latter? I think the problem here is that this machine was tasked for multiple (and very much, potentially incompatible) things. Social engineering will damn well guarantee that someone will attempt to use it for source control at some point, simply because it's there :-)

*[In other words, "Dogfooding is a stupid idea. Lab testing should be good enough. And*



*integration testing is totally unnecessary." -Raymond]*



**Cheong**

7 Jun 2011 6:54 PM

#

Actually I'd think that at some point people will start putting funny recordings on the share, knowing that someone will delete them later so you can put it there, send links to others and then safely forget it.



**v.u.**

7 Jun 2011 6:57 PM

#

'— Um, that's a scratch machine. Why would you put important stuff on it?

"Um, well..." (scratches forehead)

He stores his forehead there too??



**Chaz**

7 Jun 2011 8:00 PM

#

I have a colleague who used to retain all of her important e-mails in her Deleted Items folder. I have no idea how she thought this was the best answer to a problem.



**Charlie Kindel**

7 Jun 2011 8:06 PM

#

@JustSomeRandomGuy - The NTFS team does have their own test servers (many of them). However, servers like \\scratch are used in all sorts of ad-hoc scenarios that are more 'real-world' than those in a lab. And there's value in testing systems in this way.

FWIW, my favorite server/share at Microsoft was \\paddy\giveadogabone.



**Dave**

7 Jun 2011 10:15 PM

#

>Raymond, STOP. Don't move.

>

>Hawaii is... scratched !

No, this record is scratch-ed. I will not buy it.



**Drak**

7 Jun 2011 10:43 PM

#

@640K: Disk may be cheap, but backing-up more disk takes more time or newer equipment, and that is not cheap.



**Ooh**

8 Jun 2011 3:39 AM

#

@JustSomeRandomGuy - I'm pretty sure the NTFS and HSM guys do have their own test servers, but even a good test environment misses bugs because there is this class of problems which only occur in production usage. Raymond wrote about it some time ago in another pre-recorded post called "The Microsoft corporate network: 1.7 times worse than hell":

[blogs.msdn.com/.../416846.aspx](http://blogs.msdn.com/.../416846.aspx)



**Alex Grigoriev**

8 Jun 2011 6:30 AM

#

@640k:

"Disk is \*cheap\*."

Not if it is a first-tier brand's array. Additional disks cost about 100 times more per GB than a retail SATA disk. While this f-ing array still feels like the world's slowest SMB filer. Not kidding. Used to keep SourceSafe "database" on it, and it was incredibly slow.



**Voo**

8 Jun 2011 9:03 AM

#

@Drak: Huh? Why would it make any difference if your backup script/program/whatever copies from 5 servers or 10? Or backups 100gb instead of 80 per week? Sure the backup process may run a bit longer and you may need more backup storage but then space is cheap and I hope most network admins don't have to babysit something as regular as backing up data.

And I mean most user data we're talking about here is in the lower mb range I'd wager (at least I don't see many valid reasons to store gb of videos/pictures for work) - so how much could this possibly be? One upper management guy wasting an hour searching/restoring his files will pay for a lot of 1TB disks. Not to think of the time wasted by supporting guys having lost their data - which I'm sure will happen even if you

give the server the most wonderful and plausible name ;)



**James Schend**

8 Jun 2011 9:42 AM

#

All of you people debating on the cost of disk space need to look into cloud computing services like Amazon's S3. Disk is dirt cheap, if you don't keep it in-house.



**kog**

8 Jun 2011 10:14 AM

#

I think this whole disks are cheap debate is missing the point. This server was created as a temporary storage server to test NTFS changes, HSM etc. I would have to presume that MS employees have another server available to them for long term storage. If you going to make changes to the underlying NTFS file system there is a good chance that you could lose all data or may need to reformat the drives. The problem isn't the MS can't afford additional hard drives it that the IT warned employees that all their data could disappear at any time Then when it does the employees are stocked and angry. Maybe IT wasn't clear enough about the purpose of this server, or maybe end users just never read the emails, but this has nothing to do with the price of storage.

Also

"Okay, well before we reformatted the hard drive, we copied the data to scratch2"

While I understand the logic behind this and as a sys admin I would have done this myself. You just removed all consequences from the user. They can now happy continue ignoring all of IT's emails and using systems however they want knowing that they can get their files recovered with a simple email to IT.



**Cesar**

8 Jun 2011 11:28 AM

#

Is the share writable by everyone? If it is, what prevents someone else from simply erasing the file you put there? If nothing prevents it, why do you expect anything you put in the share to stay there?

*["If your office door doesn't have a lock, what prevents somebody else from simply walking into your office and stealing the photos of your family? If nothing prevents it, why do you expect anything you put in your office to stay there?" -Raymond]*



**640k**

8 Jun 2011 12:00 PM

#

>>can\_be\_reformatted\_at\_any\_time

>Which, btw, is not a valid ietf name.

Neither is it a valid NETBIOS name.



**640k**

8 Jun 2011 12:43 PM

#

>> "Disk is \*cheap\*."

> Not if it is a first-tier brand's array. Additional disks cost about 100 times more per GB than a retail SATA disk. While this f-ing array still feels like the world's slowest SMB filer. Not kidding. Used to keep SourceSafe "database" on it, and it was incredibly slow.

Secure backups are not expensive if done right. It sounds like you're still trapped in thinking of last millennium with expensive tape robots and backup software which requires a whole staff to operate. SSDs is cheaper than "first-tier brand's" sas/fc disks, an several times faster.



**Alex Grigoriev**

8 Jun 2011 1:57 PM

#

@640k:

I don't get why you started to talk about secure backup.



**640k**

8 Jun 2011 5:35 PM

#

Storage in enterprise is more than a simple disk, and --> SSDs is cheaper than "first-tier brand's" sas/fc disks, and several times faster.



**Alex Grigoriev**

8 Jun 2011 7:03 PM

#

Whether SSD is cheaper or faster or not does not matter, if the SSD is not on the array's vendor's list.



**Tony**

8 Jun 2011 9:06 PM

#

There's a Scratch disk here at University, with a BIG notice that says "this disk is

temporary storage and everything will be deleted every 24 hours".

Still doesn't stop some students of mine putting their entire essays on there and there comes crying to me when they couldn't find it.



**Falcon**

8 Jun 2011 10:31 PM

#

There's a compromise solution: back up the "temporary" storage, but tell the complaining users that the data MAY be archived somewhere, then retrieve it for them after a random period of a few weeks or months, or sometimes not at all. That way, there are still consequences for them, even if they ultimately get their files back - it's not just a "simple email to IT" any more!



**Drak**

8 Jun 2011 10:34 PM

#

@Voo: It starts to matter when you backups start taking more than the time between the last person leaving the office and the first person entering the office. Some (smaller) companies just don't have the cash and personnel needed to 'simply' upgrade from their current backup strategy to something faster.



**Drak**

8 Jun 2011 10:35 PM

#

@Voo (addendum): At least, that's what IT keeps saying to us when we complain of too little disk space on our build server.



**Worf**

8 Jun 2011 11:18 PM

#

Anyone realize that with Raymond's huge queue of posts (years long), if he suddenly quit and stays in Hawaii, we won't know?

If we take his Windows 2000 comment at that PDC2008 thing, it means he can disappear for 3 years before we'd notice. Like the sun can blink out of existence and we'd see everything as normal for 8 minutes...



**James Bray**

9 Jun 2011 3:52 AM

#

We have almost exactly the same setup - a share called "scratch-1" on our Linux Samba server.

We even put a file in the root directory called "THIS\_DISK\_WILL\_NOT\_BE\_BACKED\_UP" :-)

James



**Alex Grigoriev**

9 Jun 2011 6:42 AM

#

@Drak:

"It starts to matter when you backups start taking more than the time between the last person leaving the office and the first person entering the office"

Not if you're using shadow copy (volume snapshot) backup. Which Windows backup was doing since XP.



**Simon Farnsworth**

10 Jun 2011 4:25 AM

#

@Alex Grigoriev:

Even with shadow copy, it matters; while the server is backing up, it's using some resources to handle the backup. If backup takes so long that it's still running when people come in, the server appears "slower".

The solution is to spend more on hardware so that the backup doesn't push you over the edge in terms of I/O performance; for example, use the LAN to copy across at high speed to a dedicated backup system, which does the slow copy to removable media or a service on the WAN (you do push your backups off-site, right?). At this point, you've just turned \$100 of new disk into \$1,000 of server + disk + new procedures - not so cheap any more...



**Gabe**

10 Jun 2011 10:43 PM

#

Saying that disk space is cheap is like saying a feature is cheap. Sure you can buy a 2TB drive for under \$100, but that doesn't pay for installing it, the downtime to install it, backing it up, etc.

In the same vein, some features take under an hour to code but that neglects the specs, the testing, fixing the bugs, the documentation, translating the documentation into 24 languages, supporting that feature for the next 15 years, and so on.

**Dave**

12 Jun 2011 11:11 PM

#

You'd also think that with a directory name like %TEMP%, Visual Studio wouldn't use it for persistent file storage, but it does.

(VS 2010, as part of its process of reinstalling half the OS on install, needs to reboot during the install process. It writes the install file that it needs in order to continue to %TEMP% rather than the VS 2010 install location. Since these are student machines that magically fill themselves up with crap, %TEMP% gets wiped on reboot. Result: A lab full of lobotomised machines stuck halfway through the VS 2010 install, can't move backwards, can't move forwards. ARGGHHH!).