

Re: too many files....

Re: too many files....

Source: <http://linux.derkeiler.com/Newsgroups/comp.os.linux.misc/2006-08/msg00528.html>

- *From:* Dances With Crows <danSPANceswitTRAPhcrows@xxxxxxxxxx>
 - *Date:* Mon, 24 Jul 2006 08:59:00 -0500
-

On Fri, 21 Jul 2006 16:22:14 -0700, CptDondo staggered into the Black Sun and said:

Dances With Crows wrote:

CptDondo wrote:

Dividing the files into groups and directories is not an option; they are used in a website and all of the picture URLs are stored in an SQL database. Dividing them now would be a huge nightmare.....

You screwed up. Fix it now, or fix it later; the pain will be less if you fix it now. It's **never** a good idea to try to pack every file into 1 directory if the number of files is going to grow and

I'm a victim of my own success.... I never thought there'd be more than about 1K images altogether. It's not until the website started slowing down that I actually did a count....

I'm going to try and create directories for new images, but there's an awful lot of scripts written around the images and realistically it will be a few months (and maybe another thousand or two pics) before I can devote the time to that.

This is why it's a good idea to think through the design **first**, before starting a large project.

[Does] part of the problem of slow access involve some SQL lookup?

I would guess this is part of the problem. Make sure you're not doing something dumb in the queries being used and the server-side whatever (PHP? Ruby? CGI-Perl? **yuck** Java?) you're using.

Re: too many files....

Re: too many files....

PHP + SQL. And I know next to nothing about either one.... So yes, most likely it is SQL dragging the whole thing down.

Right... IIRC, there's a comp.lang.php* group or heirarchy, so you might want to post relevant snippets of your PHP code there. I'd take a look at them myself, but the last time I used PHP seriously was when PHP3 was the big thing. The database queries themselves would be a good thing to post. There's probably a subset of comp.lang.* that deals with SQL, but I forget what exactly that is.

The SQL DB is stored on the same physical devices as the images, and that could also be a part of the problem.

I doubt it. It's more probable that a query is doing something suboptimal somewhere, and you're getting tons of results when you want 5 (or something). At one point, a co-worker managed to enter a Bugzilla query that retrieved all 4000 bugs in the database[0] and he was ... less than happy about it.[1] A more sane query ("all bugs not Fixed or Closed in project N") returned in half a second.

I assume that I could do some sort of SQL magic to optimize the indexes to make retrieval faster....

This depends greatly on which database you're using. Are you using My, Postgres, Firebird, DB2, or something else? Anyway, HTH,

[0] roughly 3300 of them were marked fixed or closed, TYVM.
[1] Bugzilla returned the random quote "The request could not be completed because of Error 0x0032 : Parachuting Weasels" and sat there retrieving all bugs and arranging them into a <table>. His browser displayed the header and the quote, and no bugs, for ~3 minutes. This confused him. Oh well....

--

Matt G|There is no Darkness in Eternity/But only Light too dim for us to see
Brainbench MVP for Linux Admin / mail: TRAP + SPAN don't belong
<http://www.brainbench.com> / "He is a rhythmic movement of the
-----/ penguins, is Tux." --MegaHAL

Re: too many files....