

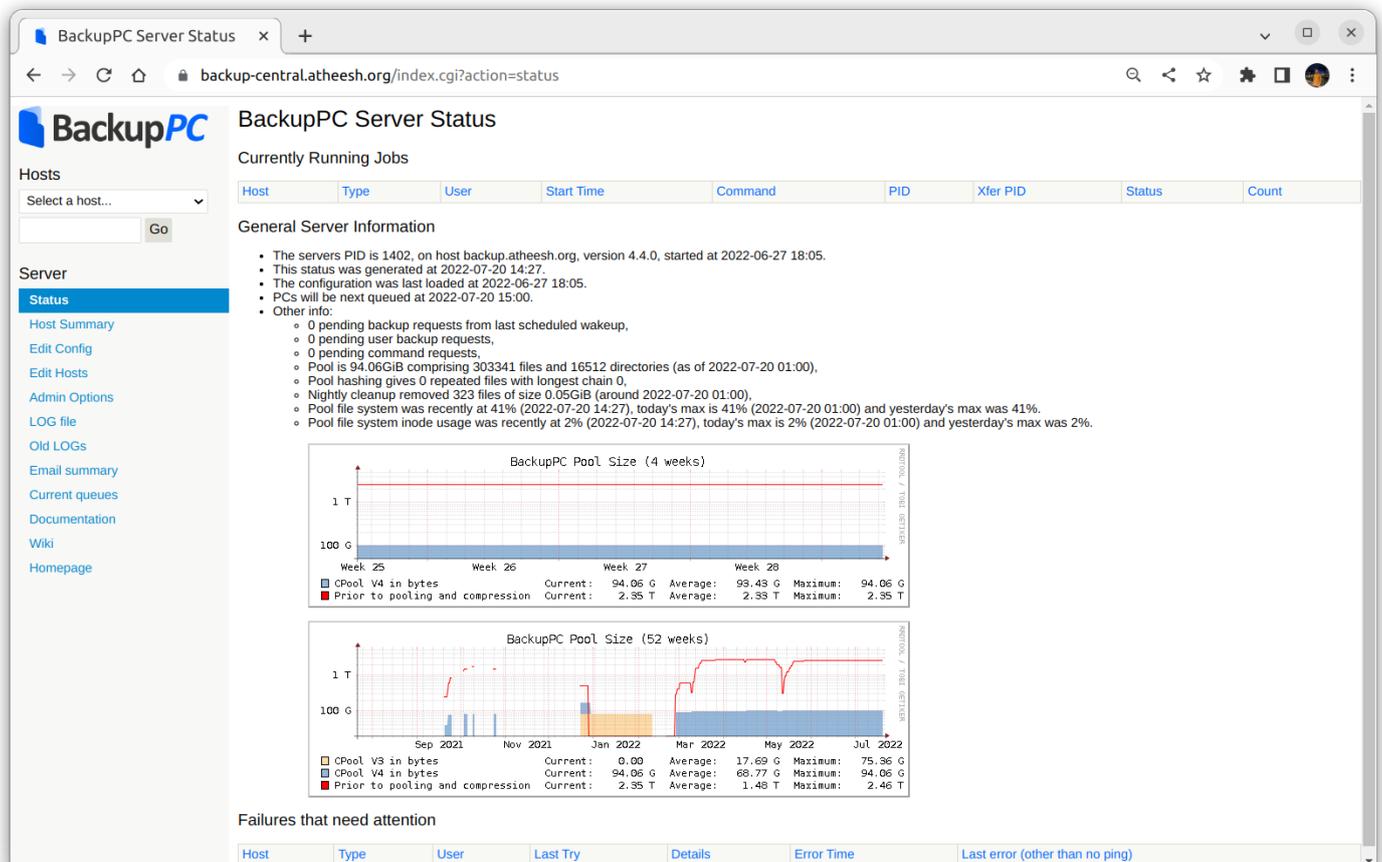
The Auto Backup Service

?Whoops!, I just lost all my data...

Just Kidding! Data Backups are really really essential. For me they're Docker Volumes, For my Family they're Photos and for Chartered Accountants in my family they're accounting software data.

Accounting Software Data in one of my servers accounts for 500GB of data on my primary Hard Drive. Manually Backing Up such huge amounts of data is next to impossible. BackupPC proves as the best open-source, self-hosted solution for this.

Deploying a BackupPC Instance



Detailed information and steps on deploying a BackupPC Instance can be found on [the official documentation](#)

Deploying BackupPC involves checking Filesystem compatibility between the clients and the server. For Example, Windows `NTFS` has case-insensitive filenames, but in Linux's `ext4`, they aren't. I Deployed the BackupPC Server instance on Linux (Ubuntu) with Windows Clients.

Further calculations include Disk Space Requirements on the sever. I usually have twice to thrice the total data which needs to be backed up from all the clients. This makes sure that **multiple incremental and full backup** copies are present **providing a better file history in case of loss of data**

BackupPC works on the principle of pulling data instead of pushing. By this I mean, the Backup Client provides access to data stored in it via `SSH`, `rsyncd`, `SMB`, etc.

RAID? No, GlusterFS...

Ha, Ha! Redundant Array of Inexpensive Disks just proves to be expensive for me. In the chartered accountants' office, around 10 employees have computers with 1TB HDDs but, they use only 100GB (added 75G as buffer) as all of them store their data in the central server.

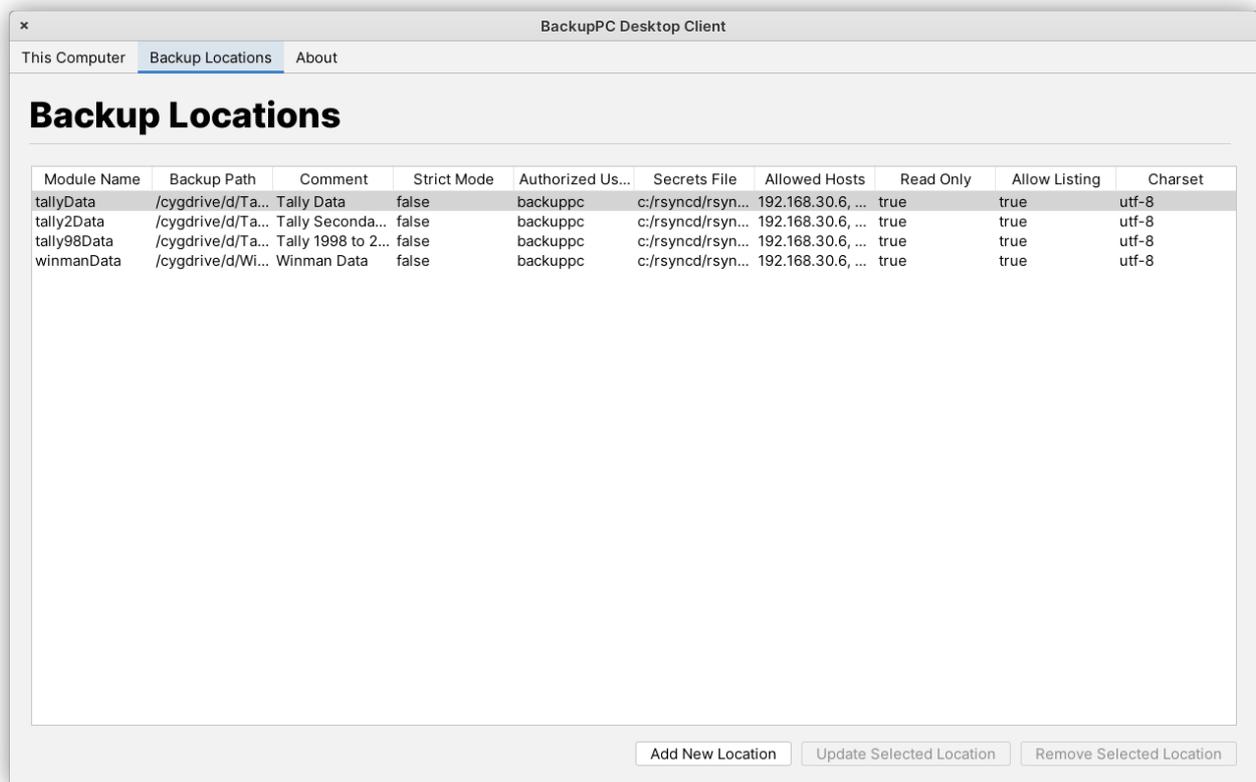
So, the rest 900G is for me. **Instead of buying HDDs and a computer with really good hardware to setup a real fast RAID, all I did was to create a scalable network filesystem (resembles to a NAS but using the TCP/IP Protocol Suite) using GlusterFS.**

GlusterFS is a scalable network filesystem suitable for data-intensive tasks such as cloud storage and media streaming. GlusterFS is free and open source software and can utilize common off-the-shelf hardware. To learn more, [please see the Gluster project home page](#).

Source: <https://docs.gluster.org/en/latest/>

Now, GlusterFS works only on Linux and on the FUSE Architecture. FUSE stands for **Filesystem in Userspace**. So, in a bigger picture my setup using GlusterFS involves using 10 computers with 900G Hard Disks (accounted only space available for GlusterFS) under a **Distributed Dispersed GlusterFS Volume** architecture. In RAID terms, this is a **configuration where two independent sets of stripped volumes are mirrored**. This configuration provides me with around 4TB of Disk Space with high redundancy (mirroring of stripped disks)

Configuring a Windows BackupPC Client



Configuring `rsync-bpc` (a `cygwin` wrapper for `rsync` by BackupPC) sometimes proves to be difficult. To streamline this process, I've created my own Java Based Configuration File Generator for Windows BackupPC Clients as I had to deploy it for around 10 to 15 computers easily. Read more about it here.

Revision #12

Created 5 July 2022 15:22:16 by Atheesh

Updated 17 August 2022 01:10:21 by Atheesh