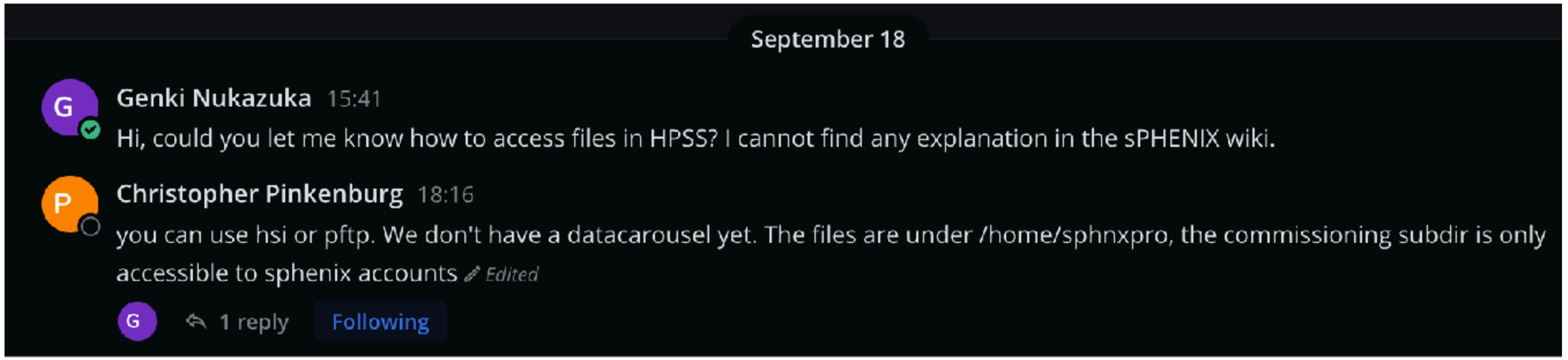
About HPSS

G. Nukazuka (RIKEN/RBRC)

Data in HPSS

Mass data storage (~200 PB?) in SDCC using tape (NOT disc) devices.
Our data in the buffer box is (automatically?) transferred to HPSS.
We need to know HPSS more.
Information on HPSS is not so much. No description is in the sPHENIX wiki.

This is a used tape storage in CERN, by the way (~10 CHF).



SDCC suggests using *pftp* or *hsi* commands to access data (cf. hpss-archiving-tools).
I tried pftp with Chris yesterday but cannot access data somehow. hsi worked well.

Accessing Data in HPSS using *hsi*

To launch the interactive shell(?)

```
[nukazuka@sphnx05 02:38:10 ~] $ hsi
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]
? pwd
pwd0: /home/nukazuka
```

But the user experience was terrible. Executing command through hsi command is even better.

```
[nukazuka@sphnx05 02:36:35 ~] $ hsi pwd
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]
pwd0: /home/nukazuka
[nukazuka@sphnx05 02:36:38 ~] $ hsi ls
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]
[nukazuka@sphnx05 02:36:42 ~] $ hsi ls /home/sphnxpro
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]

/home/sphnxpro:
00190000000000000C9E068   commissioning/            data_test/               test/                    test3file
```

Let’s see what’s in HPSS

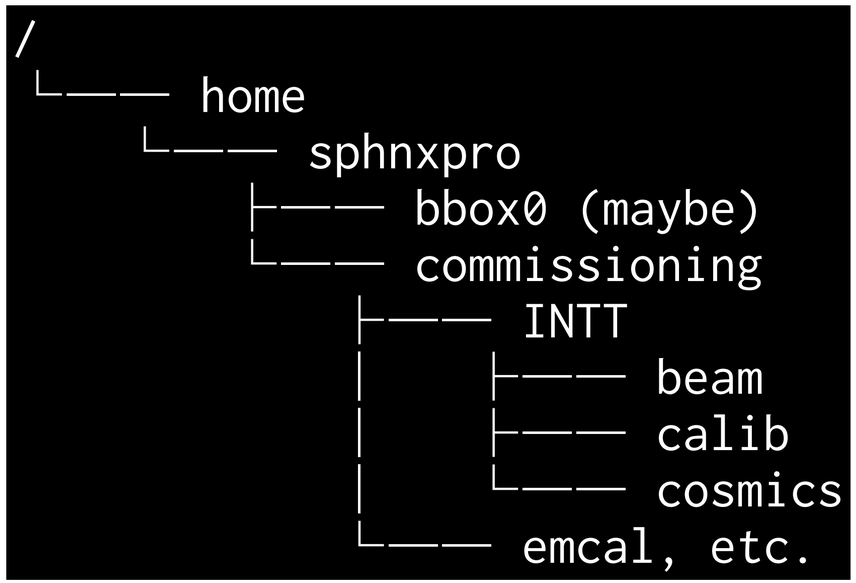
```
[nukazuka@sphnx05 02:40:33 ~] $ hsi ls /home/sphnxpro/commissioning/
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]

/home/sphnxpro/commissioning/:
emcal/  GL1/    HCal/   INTT/   LL1/    mbd/    MVTX/   tpc/    TPOT/   ZDC/
[nukazuka@sphnx05 02:40:58 ~] $ hsi ls /home/sphnxpro/commissioning/INTT
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]

/home/sphnxpro/commissioning/INTT:
beam/     calib/    cosmics/
```

Accessing Data in HPSS using *hsi*

The directory structure looks like



The directory structure is very similar to the buffer box. According to Martin's explanation (not documented) and Chis's words, bbox0, bbox1, …, and bbox5 must be under the sphnxpro directory. Those directories should contain beam data in the future, just like directories with the same name in the buffer box. But those can be accessed only with sphnxpro account (probably).

Let's see what's inside the beam, calib, and cosmics directories.

```
[nukazuka@sphnx05 02:51:06 temp] $ hsi ls /home/sphnxpro/commissioning/INTT/cosmics
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]

/home/sphnxpro/commissioning/INTT/cosmics:
cosmics_intt0-00025145-0000.evt   cosmics_intt1-00026078-0000.evt   cosmics_intt3-00025360-0000.evt   cosmics_intt4-00026385-0000.evt   cosmics_intt6-00025950-0000.evt
cosmics_intt0-00025146-0000.evt   cosmics_intt1-00026081-0000.evt   cosmics_intt3-00025361-0000.evt   cosmics_intt4-00026386-0000.evt   cosmics_intt6-00025951-0000.evt
cosmics_intt0-00025147-0000.evt   cosmics_intt1-00026082-0000.evt   cosmics_intt3-00025362-0000.evt   cosmics_intt4-00026387-0000.evt   cosmics_intt6-00025952-0000.evt
cosmics_intt0-00025148-0000.evt   cosmics_intt1-00026083-0000.evt   cosmics_intt3-00025383-0000.evt   cosmics_intt4-00026388-0000.evt   cosmics_intt6-00025954-0000.evt
cosmics_intt0-00025149-0000.evt   cosmics_intt1-00026085-0000.evt   cosmics_intt3-00025386-0000.evt   cosmics_intt4-00026389-0000.evt   cosmics_intt6-00025956-0000.evt
cosmics_intt0-00025150-0000.evt   cosmics_intt1-00026090-0000.evt   cosmics_intt3-00025388-0000.evt   cosmics_intt4-00026390-0000.evt   cosmics_intt6-00025961-0000.evt
```

The last data (checked at Sep/27 3am) is cosmics_intt7-00026978-0000.evt, which was taken on Sep/25.
Yes, this is the latest data!

Data

/bbox/commissioning/INTT or
/bbox/bbox[0-5]/INTT

/sphenix/lustre01/sphnxpro/commissioning/INTT or
/sphenix/lustre01/sphnxpro/bbox[0-7]/INTT

Updated numbers are shown in red.

	All evt	evt in HPSS /home/sphnxpro/ commissioning/INTT	Evt in the buffer box	Evt in SDCC	ROOT files in SDCC	Event-base ROOT files in SDCC	Merged event-base ROOT files in SDCC
Beam		12385	0 24	13012 0	6704	6675	32 (6675/8~784)
Cosmics		3260	983 1208	5350 752	5025	4895	611 (4895/8~612)
Pedestal	70?	NA	92	94 320	40	40	0
Junk	196?	NA	96	291 8	144	144	0
Calib	705	0	0	705 0	?	?	0

Probably, all beam data is in HPSS.
But HPSS doesn't have all cosmic data.
Also, the directories where files from the buffer box had been transferred are getting chaotic.

Accessing Data in HPSS using *hsi*

Let's get the latest cosmic data as a demonstration.

```
[nukazuka@sphnx05 03:18:46 temp] $ hsi get /home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt7-00026978-0000.evt
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]
get  '/direct/sphenix+u/nukazuka/temp/cosmics_intt7-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt7-00026978-0000.evt'
 (2023/09/26 16:31:48 5775360 bytes, 357517.8 KBS )
[nukazuka@sphnx05 03:18:59 temp] $ ls
cosmics_intt7-00026978-0000.evt
```

Does "?" in the regular expression work?

```
[nukazuka@sphnx05 03:19:01 temp] $ hsi get /home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt?-00026978-0000.evt
Username: nukazuka  UID: 100532  Acct: 100532(100532) Copies: 1 COS: 0 Firewall: off [hsi.8.3.0.p0 Thu Jun 10 07:40:52 EDT 2021]
get  'cosmics_intt6-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt6-00026978-0000.evt' (1969/12/31 19:00:00 0 bytes, 0
.0 KBS )
get  '/direct/sphenix+u/nukazuka/temp/cosmics_intt2-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt2-00026978-0000.evt'
 (2023/09/26 16:24:01 5300224 bytes, 153487.0 KBS )
get  '/direct/sphenix+u/nukazuka/temp/cosmics_intt4-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt4-00026978-0000.evt'
 (2023/09/26 16:58:44 5308416 bytes, 90847.6 KBS )
get  '/direct/sphenix+u/nukazuka/temp/cosmics_intt1-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt1-00026978-0000.evt'
 (2023/09/26 16:31:51 5816320 bytes, 153639.0 KBS )
get  '/direct/sphenix+u/nukazuka/temp/cosmics_intt0-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt0-00026978-0000.evt'
 (2023/09/26 16:19:30 5185536 bytes, 108052.7 KBS )
get  '/direct/sphenix+u/nukazuka/temp/cosmics_intt3-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt3-00026978-0000.evt'
 (2023/09/26 16:29:29 5160960 bytes, 90618.0 KBS )
get  '/direct/sphenix+u/nukazuka/temp/cosmics_intt7-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt7-00026978-0000.evt'
 (2023/09/26 16:31:48 5775360 bytes, 163571.4 KBS )
get  '/direct/sphenix+u/nukazuka/temp/cosmics_intt5-00026978-0000.evt' : '/home/sphnxpro/commissioning/INTT/cosmics/cosmics_intt5-00026978-0000.evt'
 (2023/09/26 16:37:47 5627904 bytes, 37327.5 KBS )
[nukazuka@sphnx05 03:20:21 temp] $ ls -1
cosmics_intt0-00026978-0000.evt
cosmics_intt1-00026978-0000.evt
cosmics_intt2-00026978-0000.evt
cosmics_intt3-00026978-0000.evt
cosmics_intt4-00026978-0000.evt
cosmics_intt5-00026978-0000.evt
cosmics_intt6-00026978-0000.evt
cosmics_intt7-00026978-0000.evt
```

YES, it works!

More about *hsi*

```
$ man hsi
no manual
```

Note: Output of *hsi ls* is to stderr. So
`$ hsi ls | wc -l`
doesn't work. You need to do
`$ hsi ls 2>&1 | wc -l`
to count the number of line from *hsi ls*.

I found the HSI manual (https://www.racf.bnl.gov/Facility/HPSS/Documentation/HSI/)

Table of Contents

Data

/bbox/commissioning/INTT or
/bbox/bbox[0-5]/INTT

/sphenix/lustre01/sphnxpro/commissioning/INTT or
/sphenix/lustre01/sphnxpro/bbox[0-7]/INTT

	All evt	evt in HPSS /home/sphnxpro/ commissioning/INTT	Evt in the buffer box	Evt in SDCC
Beam		12385	0 24	13012 0
Cosmics		3260	983 1208	5350 752

In the beginning of the run2023, all data taken by RCDAQ was stored in
`/bbox/commissioning/INTT`
Data in the directory was transferred to
`/sphenix/lustre01/sphnxpro/commissioning/INTT`
About a month ago, the directory to store data in the buffer box for new data was changed
`/bbox/bbox[0-7]/INTT`
The transfer script keeps the directory structure so data was sent to
`/sphenix/lustre01/sphnxpro/bbox[0-7]/INTT`
Yesterday, I found data new data in
`/bbox/bbox[0-7]/INTT`
was transferred to
`/sphenix/lustre01/sphnxpro/commissioning/INTT`
The situation is confusing and may not be stable yet. 🤮