



Faster Access to Herschel Data

(using MyHSA)

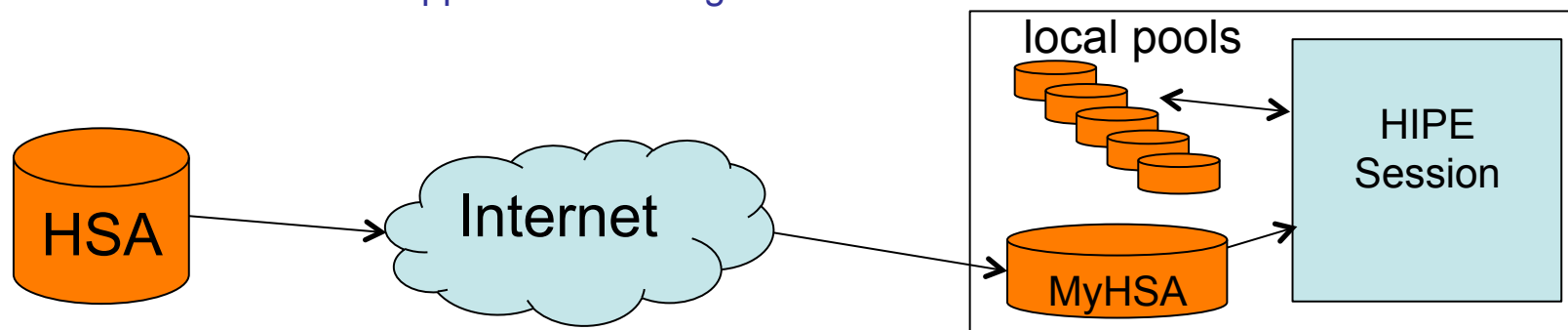
Bernhard Schulz

NHSC/IPAC

on behalf of the HSC and the NHSC

Introduction

- MyHSA was created to improve user access of Herschel data with HIPE.
- Main improvements:
 - Easier integration of HSA downloaded file structures.
 - Local buffer (proxy) memory.
 - Save option makes products persistent in MyHSA buffer.
 - More efficient product selection when loading observations.
 - Product identifiers (URNs) are identical to those used internally by HSA.
 - This causes the same product to be transferred only once.
 - Products common to more than one observation, like pointing history or calibration tree don't need to be loaded again.
 - Faster access mechanism to HSA.
 - Smaller memory footprint.
 - Data transfers happen in the background.



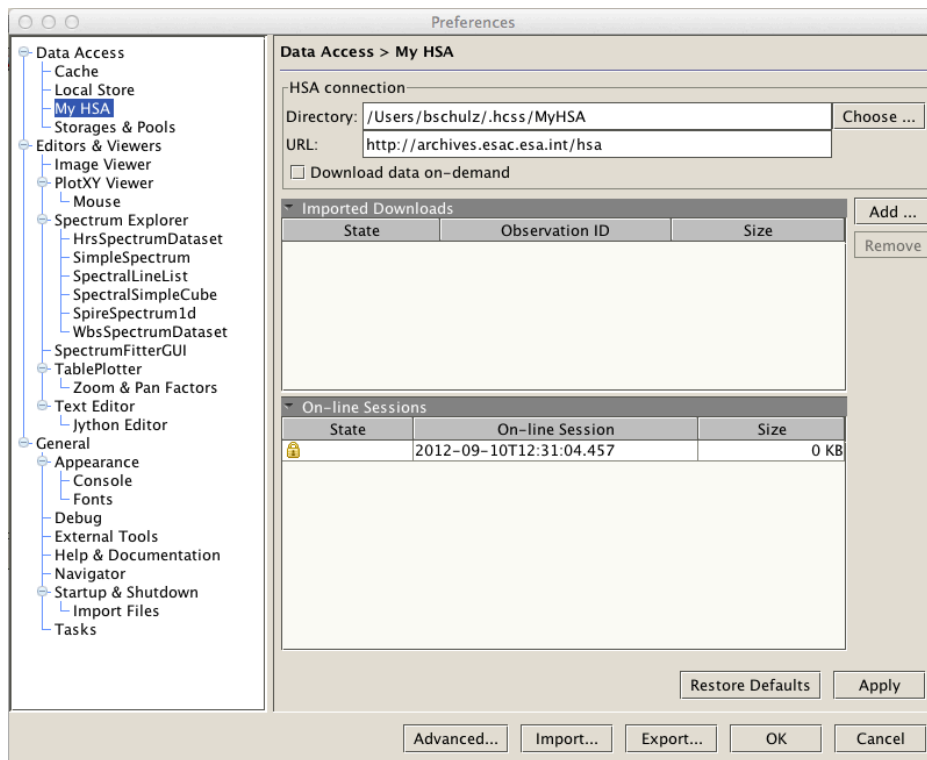


myHSA Configuration

- Edit->Preferences->Data Access->My HSA
- Set path for MyHSA directory on your computer to a place with a lot of disk space.
 - This is used for:
 - indexing HSA file structures
 - direct download of observations from the HSA through HIPE
- The “Add” button allows to include the data in HSA file structures that were downloaded from the archive.
 - The HSA file structures need to be unzipped.
 - The data is NOT relocated to the MyHSA directory! It is just indexed.
- Further you can:
 - Change the archive URL (don't do that!)
 - Activate the save data on-demand option
 - This will keep the products of observations directly loaded from the HSA permanently in the MyHSA buffer, after the session has ended.
 - The same effect can also be achieved by inserting the following line into the user.props or hipec.props file in the .hcss directory
 - `hcss.ia.pal.pool.hsa.on_demand = True`

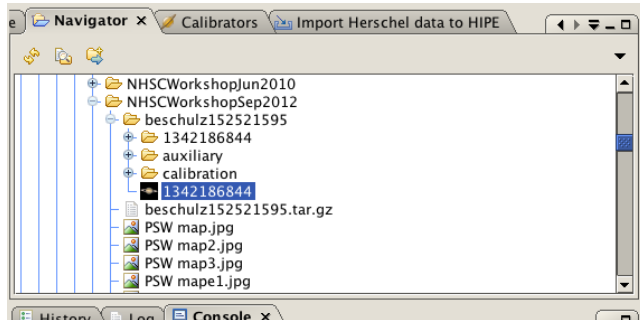


MyHSA Preferences Menu

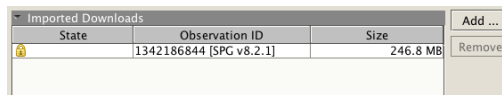
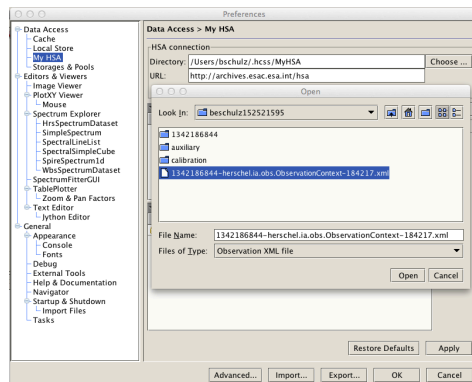


- Path to MyHSA buffer directory
 - Set to disk with most space. Herschel data have a lot of volume.
- Set URL of connection to Herschel archive.
 - You need a very good reason to change this.
- Make products of directly downloaded observations “stick” in MyHSA buffer on disk.
 - These products will be available in subsequent sessions of HIPE without the need to connect to the HSA again.
 - Note: If you only looked at a level 2 context but didn't look at the PSW map inside specifically, it was never loaded and won't be in your buffer. The next time you try to access the PSW map the system will complain until you allow HSA access in the Product Browser GUI.
- Include xml files of specific observations in HSA download file structures.
- View/manage data in MyHSA buffer by session.

Importing HSA File Structures



- Navigator
 - Use file Navigator to go to unzipped, untared, HSA data structure.
 - Select .xml file of observation and open (double-click).
- Preferences Menu
 - Use “Add” button to navigate to unzipped, untared, HSA data structure.
 - Select .xml file of observation and open.
 - Observation will appear in “Imported Downloads” panel.
- Observation will now be found in the MyHSA buffer.





Using a Script with MyHSA

- `obs=getObservation(obsid='1342186844')`
 - Retrieve observation from myHSA buffer.
- `obs=getObservation(obsid='1342186844',useHsa=True)`
 - Retrieve observation telling the system to explicitly go across the web to the HSA. Observation is not saved even if the `hsa.on_demand` property is set.
- `obs=getObservation(obsid='1342186844', useHsa=True, save=True)`
 - Retrieve observation from the HSA and keep in MyHSA buffer after end of the session. This will retrieve and save the **entire observation** in the background, i.e. no lazy loading in this case.
- Note regarding the SPIA tasks `spiaLoadObs` and `spiaCopyHsa`
 - (SPIRE only with SPIA plugin!)
 - The HSA and the myHSA buffer can be accessed with
 - `obs = spiaLoadObs(obsID=ObsID, Pool='hsa')` or
 - `obs = spiaCopyHsa(obsID=ObsID, Pool='outputPool')`
 - The HSA is reached only if “HSA” is selected in the Product Browser (this is a “feature”).
 - However, none will save the observation in the MyHSA buffer, even if the `hcss.ia.pal.pool.hsa.on_demand` property is set.



Using the Product Browser with MyHSA

Switch to allow going across the web to HSA.

Switch to search the local MyHSA buffer.

odNumber	obsid	tag	version	object	total size	instrument	aot	obsMode	startDate
168	1342186522		3	Neptune	2413228495	SPiRE	Photometer	Large Map	2009-10-29T03
168	1342186524		3	Neptune	2961230963	SPiRE	Photometer	Large Map	2009-10-29T10
275	1342190669		2	1 Ceres	299304483	SPiRE	Photometer	Large Map	2010-02-13T11
342	1342195298		2	Neptune	164763172	SPiRE	Photometer	Small Map	2010-04-20T23
341	1342195292		2	Gamma Dra	162214844	SPiRE	Photometer	Small Map	2010-04-20T19
359	1342195937		1	Neptune	167033517	SPiRE	Photometer	Small Map	2010-05-07T21
369	1342196677		2	Neptune	236384652	SPiRE	Photometer	Small Map	2010-05-18T17
528	1342207036		1	Hartley 2	458805757	SPiRE	Photometer	Small Map	2010-10-24T15
576	1342210941		1	BD +30 3639	160128659	SPiRE	Photometer	Small Map	2010-12-10T13
117	1342183479		4	GP 299_0	1863502246	SPiRE	Photometer	Large Map	2009-09-07T18
117	1342183475		5	NGC 5315	286208170	SPiRE	Photometer	Large Map	2009-09-07T16
808	1342224963		1	Gamma Dra	179077791	SPiRE	Photometer	Small Map	2011-07-31T08
701	1342218642		2	Field 22_0	1461932163	PACS	Parallel Mode	Parallel Mode	2011-04-15T04
701	1342218642		5	Field 22_0	6957376263	PACS	Parallel Mode	Parallel Mode	2011-04-15T04
880	1342230739		1	3C239	194239043	SPiRE	Photometer	Small Map	2011-10-10T15
404	1342198883		5	NGC3627	368608259	SPiRE	Photometer	Large Map	2010-06-21T17
181	1342186882		4	Neptune	424618455	SPiRE	Photometer	Large Map	2009-11-11T08

Select Data Source:

- HSA will allow accessing the archive across the internet.
- Local Pools selects all local pools. Individual ones can be selected as well.
- MyHSA selects the local buffer contents.

• Search parameters on the right are optional.

- Accessing the archive without a selection will give too many results.
- Selecting MyHSA only will show all observations in that buffer.

• Results are displayed in the panel below with one line per observation.

- Right click on column titles and selecting "Default Table Layout" will give nicer layout ("feature").

• Double clicking a result retrieves the observation context, not all products the observation consists of

- "Lazy Loading"
- Check Outline view
- If a product of an observation is not found because HSA is not selected, just switch it on and the product will be pulled automatically from the archive.

This tool is ideally used for interactive work, although it also produces command-line printouts that can be used to construct scripts.



I want to load and keep my data! What Are My Options?

Using MyHSA buffer

- Get FTP data structure from archive and register with MyHSA.
 - Full observation is in MyHSA buffer.
- Use getObservation with save and useHsa options.
 - Full observation is in MyHSA buffer.
- Use Product Browser while having hcss.ia.pal.pool.hsa.on_demand property set.
 - Only products that are actually needed are loaded and saved in MyHSA buffer.
 - Reprocess observation with SPG script and save to local pool.

Not using MyHSA

- Get FTP data structure from archive and import into new local pool with HIPE Import View.
 - not the fastest option
- Use Product Browser or getObservation and save observation context in local pool.
 - not the fastest option
- Exploit “lazy loading” because you are reprocessing your data anyway
 - Use Product Browser without hcss.ia.pal.pool.hsa.on_demand property set.
 - Use obs = getObservation(obsid=ObsID, useHsa=True) to access the archive.
 - Run SPG pipeline script from Level 0 and save resulting observation context to local pool.
 - This will only retrieve the Level 0 products, auxiliary and calibration products and save the time for loading everything else.
- Use obs = spiaCopyHsa(ObsID=ObsID, Pool='outputPool')
 - (SPIRE only with SPIA plug-in!)
 - In default mode it will download a minimal observation context containing only Level 0 and what is needed to reprocess the observation, and save this to a local pool.



Recommendation

- Use MyHSA with **save option** and **hcss.ia.pal.pool.hsa.on_demand** property set.
- Use local pools rather for **saving results** from own processing.