

THE SEN2AGRI SYSTEM DATABASE

WHO?

- Laurențiu Nicola, CS Romania
- Sentinel-2 for Agriculture developer
- laurentiu.nicola@c-s.ro, lnicola@c-s.ro
- <http://forum.esa-sen2agri.org/>

WHAT?

- We'll cover the most important database tables
- You shouldn't need to look into it
- You might still need to

PLATFORM

- PostgreSQL 9.4 (not in CentOS 7 / EPEL)
- Should be compatible with future Postgres versions
- PostGIS
- Some headaches about binary compatibility

SITES

- *Sites* are areas of interest
- MULTIPOLYGON
- You can split a site to get around system limitations
- Download and processing are per-site
- Beware of overlaps

SITES

```
create table site(  
  id smallserial not null primary key,  
  name character varying not null,  
  short_name character varying,  
  geog geography(multipolygon) not null,  
  enabled boolean not null  
);
```

- name used for display
- short_name used for disk paths
- Disabling a site stops all future processing

SEASONS

```
create table season(  
  id smallserial not null primary key,  
  site_id smallserial not null,  
  name text not null,  
  start_date date not null,  
  end_date date not null,  
  mid_date date not null,  
  enabled boolean not null,  
  unique (site_id, name)  
);
```

- name is used for display
- Dates are used to scheduling the processing
- Download starts in advance (for L2A initialization)
- Beware of overlaps

CONFIGURATION

```
create table config_metadata(  
    key character varying not null primary key,  
    is_advanced boolean not null  
    -- [...]  
);  
create table config(  
    id serial not null primary key,  
    key character varying not null,  
    site_id smallint,  
    value character varying not null  
    -- [...]  
);
```

- Simple key-value mapping
- Per-site settings, fallback to global if not configured
- Use `sen2agri-config`
- *Advanced* parameters are read-only

TILE DEFINITIONS

```
create table shape_tiles_s2(  
  tile_id character(5) not null,  
  geog geography not null  
  -- [...]  
);  
create table shape_tiles_l8(  
  pr int,  
  geog geography not null  
  -- [...]  
);
```

- Dumps of S2 and L8 tile definitions
- Used to determine site tiles
- Beware of overlaps
- Beware of changes

SITE TILES

```
create table site_tiles(  
    site_id smallint not null,  
    satellite_id int not null references satellite(id),  
    tiles text[]  
);  
create table satellite(  
    id int not null primary key,  
    satellite_name varchar not null,  
);
```

- You *want* to manually customize the tile list
- Not available in the GUI, unfortunately

DOWNLOADS

```
create table downloader_history(  
  id serial not null primary key,  
  site_id smallint not null,  
  satellite_id smallint not null,  
  product_name character varying not null,  
  full_path character varying not null,  
  created_timestamp timestamp with time zone not null,  
  status_id smallint not null,  
  no_of_retries smallint not null,  
  product_date timestamp with time zone not null,  
  orbit_id int not null  
);
```

- Shows L1 download and processing status
- Useful to trigger re-download or reprocessing

PRODUCTS

```
create table product(  
  id serial not null primary key,  
  product_type_id smallint not null,  
  processor_id smallint not null,  
  site_id smallint not null,  
  full_path character varying not null,  
  name character varying(512),  
  job_id integer references job(id),  
  geog geography,  
  satellite_id integer,  
  tiles character varying[] not null  
  -- [...]  
);
```

- Contains L2 to L4 products

JOBS

```
create table job(  
  id serial not null primary key,  
  processor_id smallint not null,  
  site_id smallint not null,  
  start_type_id smallint not null,  
  parameters json,  
  submit_timestamp timestamp with time zone not null,  
  start_timestamp timestamp with time zone,  
  end_timestamp timestamp with time zone,  
  status_id smallint not null  
    references activity_status(id),  
  status_timestamp timestamp with time zone not null  
)
```

- A job is a (non- L2A) processor invocation
- Jobs are submitted when runnable (not scheduled)
- Jobs may create multiple output products

TASKS, STEPS

- These are artefacts of our parallelization method
- Jobs are heterogenous, have one or more tasks
- Tasks are homogenous, have one or more steps
- Steps only differ in input parameters (e.g. tiles)
- You'll hopefully never deal with these

STEP RESOURCE USAGE

```
create table step_resource_log(  
  step_name character varying not null,  
  task_id int not null,  
  node_name character varying not null,  
  entry_timestamp timestamp with time zone not null,  
  duration_ms bigint,  
  user_cpu_ms bigint,  
  stdout_text character varying not null,  
  stderr_text character varying not null,  
  primary key (step_name, task_id)  
  -- [...]  
);
```

- System resource usage for each step
- Includes console messages

NODE RESOURCE USAGE

```
create table node_resource_log(  
  node_name character varying not null,  
  cpu_user smallint not null,  
  mem_used_kb int not null,  
  swap_used_kb int not null,  
  load_avg_1m int not null,  
  disk_used_bytes bigint not null,  
  "timestamp" timestamp without time zone not null  
  -- [...]  
);
```

- Node (computer) resource usage
- Displayed in the GUI
- Consider using an off-the-shelf monitoring solution

SCHEDULED TASKS

```
create table scheduled_task(  
  id smallserial not null primary key,  
  name character varying not null,  
  processor_id smallint not null,  
  site_id smallint not null,  
  season_id smallint not null,  
  processor_params character varying,  
  repeat_type smallint,  
  first_run_time character varying  
  -- [...]  
);
```

- Automatic (non-custom) processor executions
- Configured in the GUI

SCHEDULED TASKS

```
create table scheduled_task_status(  
  id smallserial not null primary key,  
  task_id smallint not null,  
  next_schedule character varying,  
  last_scheduled_run character varying,  
  last_run_timestamp character varying,  
  last_retry_timestamp character varying,  
  estimated_next_run_time character varying  
);
```

- Status of scheduled tasks
- The columns above are timestamps

USERS

```
create table "user"(  
  id smallint not null primary key,  
  login character varying(50) not null,  
  email character varying(100) not null,  
  role_id smallint not null references role(id),  
  site_id smallint references site(id),  
  password text not null,  
  unique (login)  
);
```

- User accounts for the GUI
- These *should* be editable directly

WHY?

- Change system settings
 - `site_tiles`
 - `config` (consider using `sen2agri-config`)
 - `site`
- Monitor/alter downloads : `downloader_history`
- Check product list : `product`
- Debug job execution : `job`, `task`, `step`, `event`
- Some examples in the FAQ document

SEN2AGRI - CONFIG

- Use `sen2agri-config` to change system settings
- Must be executed locally or over X11 forwarding
- Some settings will be read-only
- Some settings will not be available
 - tile filtering for sites
 - user management
 - site footprints
- See user manual for parameter descriptions

Sen2Agri Configurator

General

L3A Composite

L3B Vegetation Status

L4A Crop Mask

L4B Crop Type

Archiver

Executor

Site [Global]

Bands mapping file for S2

are/sen2agri/bands_mapping_s2.txt

...

Output L3A as Cloud Optimized Geotiff

Specifies if composite for S2 20M resolution should be generated

1

Half synthesis interval in days

25

L3A LUT file path

/usr/share/sen2agri/composite.map

...

Scattering coefficients file for S2 10 m

/sen2agri/scattering_coeffs_10m.txt

...

Scattering coefficients file for S2 20 m

/sen2agri/scattering_coeffs_20m.txt

...

Difference in days between the scheduled and the synthesis date

30

Maximum value of the linear range for weights w.r.t. AOT

0.8

Maximum weight depending on AOT

1

Minimum weight depending on AOT

0.33

Coarse resolution for quicker convolution

240

Standard deviation of gaussian filter for distance to large clouds

10

Standard deviation of gaussian filter for distance to small clouds

2

Minimum weight at edge of the synthesis time window

0.5

Save

Close

OTHER STUFF

- Migrations on updates (version in the meta table)
- You may want to truncate the resource log tables
- System log (journal) also tends to grows
- SLURM needs a database, we install MariaDB
- We are switching to a new downloader in 1.8
 - Heads-up: things might work differently

QUESTIONS?

- Or ask on the [forum](#)
- Or write us an [email](#)