## Database development and usage for prototype and pre-series testing

#### **Ali Al Kadhim**

HGCAL Week June 21, 2022







### Outline

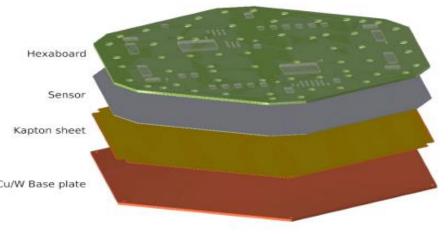
- Motivation
- Personal involvement: spreadsheets to database (DB)
- CMS HGCAL DB overview
- DB Graphical User Interface (GUI) and XML schemas
- Connecting to the DB with SQLDeveloper tutorial
- Proposed changes to DB and HGCAL community involvement

### Motivation

- Phase-2 upgrade of the CMS HGCAL will encompass tens of thousands of new sensors.
- Extensive campaigns have been conducted by CMS to determine suitable material and layout parameters for operation in HL-LHC environment.
- Sensor Quality Control (SQC) institutions routinely run tests on these sensors to determine if they pass these requirements.
  - Successful HGCAL construction will require collaboration between institutions/users on sensors and their results -> Sensor results must be stored in a central database (DB)!

### Motivation

- HGCAL requires ≈ 26,000 Si detector modules (5% are tested):
  - CE-E module (electromagnetic) or a CE-H module (hadronic).
- A unified HGCAL DB allows users from all institutions access to test results, status, location, etc. Cu/W Base plate
- When assembling these modules, users must have easy access to sensor info/results from a unified database.



**CE-E Si module** 

(PCB)

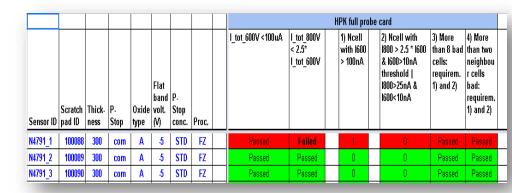
Si 8" sensor

For electrical insulation from base plate

For placement onto cassettes

### Spreadsheets to Database

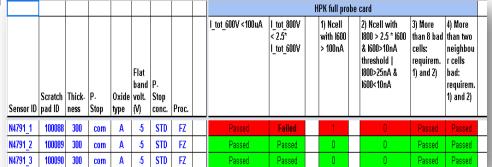
- Summary results of Si sensor analyses.
   have been entered manually in spreadsheets
   (which grew organically), such as:
  - This is very error-prone and time-consuming!
  - Not query-able.
  - Doesn't capture one-to-many relationships.

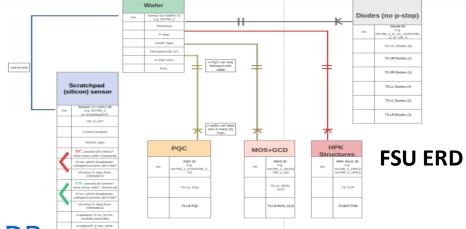


## Spreadsheets to Database

Summary results of Si sensor analyses.
 have been entered manually in spreadsheets
 (which grew organically), such as:

- This is very error-prone and time-consuming!
- Not query-able.
- Doesn't capture one-to-many relationships.
- We (FSU) constructed a relational (SQL) database that stores and updates sensor attributes, summary results, history, location, etc.
- Code: <a href="https://github.com/AliAlkadhim/FSU\_HGCAL\_DB">https://github.com/AliAlkadhim/FSU\_HGCAL\_DB</a>
- More on FSU HGCAL DB here: <u>https://indico.cern.ch/event/1164937/contributions/4891850/attachments/245144</u> <u>3/4200827/HGCAL\_DB\_May25\_Alkadhim.pdf</u>





#### CMS HGCAL DB

- Global DB for all HGCAL parts, tests, modules, etc. from all testing, production, etc. institutions.
- There are 2 instances of HGCAL Database (DB)
  - HGCAL Development DB INT2R Database
  - HGCAL Production DB CMSR Database
- To learn more about the different schemas in the HGCAL DB, see <a href="https://indico.cern.ch/event/1164937/contributions/4891851/attachments/2451351/4200663/XML%20Templates%20HGCAL%20Si%20Sensors.pdf">https://indico.cern.ch/event/1164937/contributions/4891851/attachments/2451351/4200663/XML%20Templates%20HGCAL%20Si%20Sensors.pdf</a>
- The DB (I will discuss INT2R, but the two are nearly identical), is based on XML schemas.
- XM has tag-based syntax (like HTML), all it is: <TAG\_NAME> ......DATA............ <TAG\_NAME>

  Open Tag

  Information, including other tags

Ali Al Kadhim - HGCAL DB

#### **HGCAL DB**

HGCAL DB consists of the following inter-connected schemas (accounts), each used to store different data types:

Tables in these schemas are related to each other!

CMS\_HGC\_CORE\_ATTRIBUTE

to store attribute information related to parts and data types

CMS\_HGC\_CORE\_COND

to store meta data of user-generated data (conditions data).

CMS\_HGC\_HGCAL\_COND

To store **detailed tables** for specific user-generated test data.

• CMS\_HGC\_CORE\_CONSTRUCT to store detectors parts.

CMS\_HGC\_CORE\_IOV\_MGMNT

to store data IOVs (interval of validity) of user generated data.

CMS\_HGC\_CORE\_MANAGEMNT

to store "management" data type, e.g. locations, institutions, etc.

• Every part has a *kind of Part* and a *serial Number,* and is uniquely identified in the DB by this pair of parameters.

### Accessing the Database with SQLDeveloper

• SQLDeveloper is the preferred way to access and query this XML database.

 For more instructions on how to access and view the database, see the backup slides!

 Once connected, you can see the CMS\_HGC\_\* schemas previously mentioned.

You can also do any SQL queries (also in backup).



Sample ERD

Sample SQL query

### Connecting to the DB with SQLDeveloper Tutorial

 Now put the following text in a file named tnsnames.ora and put it in your /etc/ directory.

Now ssh to lxplus

```
ssh -XY -L 10131:itrac1609-
v.cern.ch:10121 -L 10132:itrac1601-
v.cern.ch:10121
```

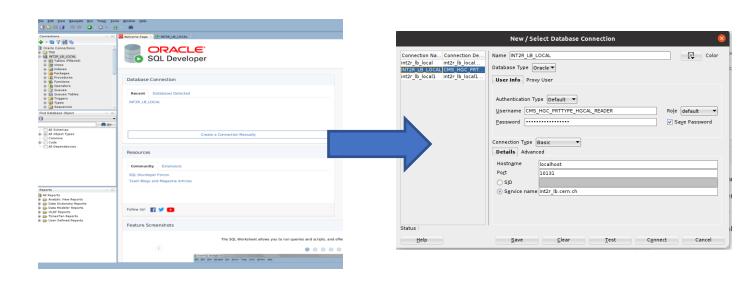
Where username is your lxplus username.

username@lxplus.cern.ch

#### tnsnames.ora

## Connecting to the DB with SQLDeveloper Tutorial

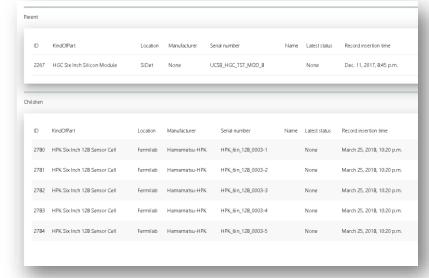
- Launch SQLDeveloper (on unix by typing sqldeveloper).
- Press "Open a Connection Manually" in the front page.
- On the "New / Select Database Connection" Dialogue box, fill:
  - Name: INT2R\_LB\_LOCAL
  - Username: CMS\_HGC\_PRTTYPE\_HGCAL\_READER
  - Password: HGCAL\_Reader\_2016
    - Check "save password" if you want.
  - Hostname: localhost
  - Port: 10131
  - Check on "Service name"
  - Service name: int2r\_lb.cern.ch
  - Press "Connect"
  - You're connected!

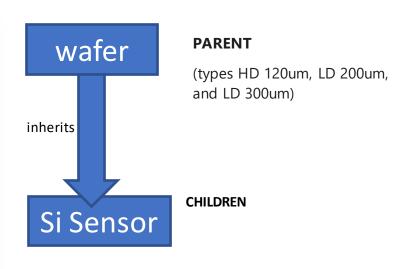


#### GUIs! OMS and DBA

 Production Detector Construction Application (DCA: https://cmsdca.cern.ch/hgc\_cmsr/construct/parts/) GUI

High level Information on detector assembly parts: mainly for current status and tracking (location)





- To view OMS, you need to set up an ssh tunnel with a proxy. Instructions on how to do that can be found here:
  - https://twiki.cern.ch/twiki/pub/CMS/HGCALDataBase/ssh-tunnel-for-development-oms-dca-dec-2019.pdf

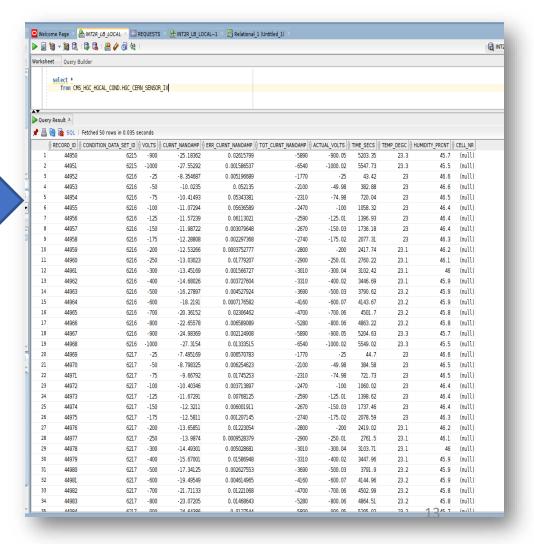
### The DB is XML schema

#### XML Template for HGC\_CERN\_SENSOR\_IV table

Ali Al Kadhim - HGCAL DB

```
<ROOT>
   <HEADER>
       <TYPE>
           <!-- Comment: Table Name -->
           <EXTENSION_TABLE_NAME>HGC_CERN_SENSOR_IV
           <!-- Kind of condition name -->
           <NAME> HGC CERN Sensor IV</NAME>
       </TYPE>
       <RUN>
           <RUN NAME>Your Run Name
           <!-- Enter your timestamp -->
           <RUN BEGIN TIMESTAMP>2018-05-14 00:00:00/RUN BEGIN TIMESTAMP>
           <RUN END TIMESTAMP>2018-05-14 00:00:00/RUN END TIMESTAMP>
           <INITIATED BY USER>Your Name</INITIATED BY USER>
           <LOCATION>CERN</LOCATION>
           <COMMENT DESCRIPTION>Your Comments//COMMENT DESCRIPTION>
       </RUN>
   </HEADER>
   <DATA SET>
       <PART>
          <!-- Kind of Part name -->
          <KIND OF PART>HGC Sensor</kind OF PART>
           <!-- Sensor serial number -->
           </PART>
       <DATA>
          <!-- table columns with data -->
          <VOLTS>-25</VOLTS>
           <CURNT NANOAMP>-7.609905</CURNT NANOAMP>
           <ERR CURNT NANOAMP>0.01653122/ CURNT NANOAMP>
           <TOT CURNT NANOAMP>-2000</TOT CURNT NANOAMP>
           <ACTUAL VOLTS>-25</ACTUAL VOLTS>
           <TIME SECS>10</TIME SECS>
           <TEMP DEGC>23</TEMP DEGC>
           <HUMIDITY_PRCNT>46.6/HUMIDITY_PRCNT>
           <CELL NR>YYYYY</CELL NR>
       </DATA>
```

#### Viewed in SQLDeveloper



# Converting .txt results to .xml DB files (using the XML layout)

- HGCAL sensor db interface:
  - https://gitlab.cern.ch/CLICdp/HGCAL/hgcal\_sensor\_db\_interface
  - Example Usage
    - Convert HPK data to .txt:
      - python db\_convert\_hpk\_to\_txt.py -i
         '/home/data/cms/results/hamamatsu/HPK\_8in\_432ch\_120um/200805\_S1559103(ES1) inspection sheet.xlsx'
    - Convert .txt data to .xml:
      - python db\_register\_data\_iv.py -i '/home/data/cms/results/hamamatsu/8-432-120E-00-N1308-00001/8-432-120E-00-N1308-00001\_IV.txt' --serial '8-432-120E-00-N1308-00001-SE-1' --run 'IV Test 01' --location 'CERN' --operator 'Florian Pitters' --comment 'Nothing special.'
- Similar operations can also be found here
  - https://github.com/AliAlkadhim/FSU HGCAL DB/tree/master/CMS HGCAL DB/XML TABLES

### Uploading to the DB (CMSR or INT2R)

- Once the XML files for the data that you want to store in the DB have been generated (based on the templates mentioned previously <u>here</u> and <u>here</u>), you can upload to the DB using DBLoader.
  - DBloader code: <a href="https://github.com/valdasraps/cmsdbldr">https://github.com/valdasraps/cmsdbldr</a>
- More instructions to upload are in backup slides.
- To upload to the INT2R (development) database, do
  - scp <file.xml> xml joshi@dbloader-hgcal.cern.ch:/home/dbspool/spool/hgcal/int2r
- To upload to the CMSR (production) databse, do
  - scp <file.xml> xml joshi@dbloader-hgcal.cern.ch:/home/dbspool/spool/hgcal/int2r

## We Need More Involvement & Input from HGCAL!

- Everyone from HGCAL (testing, module assembly, etc.) is encouraged to upload to the DB!
- If you need permissions to upload to HGCAL DB, please contact Imran (<a href="Muhammad.Imran@cern.ch">Muhammad.Imran@cern.ch</a>), Rao (<a href="muhammad.atif.shad.rao@cern.ch">muhammad.atif.shad.rao@cern.ch</a>), or Umesh (<a href="mujoshi@fnal.gov">joshi@fnal.gov</a>).
- Attend an HGCAL DB meeting <a href="https://indico.cern.ch/category/11190/">https://indico.cern.ch/category/11190/</a>.
- Do we want to add extra information (tables or attributes) to the DB? We are working on changes, but we need your valued opinions!
  - E.g. HGC\_CERN\_SENSOR\_IV\_SUMRY has summary of grading criteria for sensors. Do we want to add more information to this summary?
  - Should the sensor tables include extra information, e.g. current location?
  - Perhaps sensors should also have a tag, whether they have been tested or not (only 5% of sensors undergo testing).

### Backup

### SQL Developer: installation

- Follow the instructions on <a href="https://docs.oracle.com/cd/E39885">https://docs.oracle.com/cd/E39885</a> 01/doc.40/e38928/install.htm#RPTIG12
   for installation instructions on any system.
- Important: SQL Developer requires JDK 7 or above. Download from <a href="https://www.oracle.com/java/technologies/downloads/#java8">https://www.oracle.com/java/technologies/downloads/#java8</a> and make sure your default Java in your system points to this **Oracle JKD.** 
  - This means setting the JAVA\_HOME and JKD\_HOME environment variables are pointing to the Oracle Java.
  - Doing java -version should result in "java version ..." not "openjkd version ..."
- SQLDeveloper could now be launched.
- On Unix, launch with sqldeveloper to see that it works.

## Connecting to the Database with SQLDeveloper

(1)

tnsnames.ora

- Now put the following text in a file named tnsnames.ora and put it in your /etc/ directory.
- Now ssh to lxplus

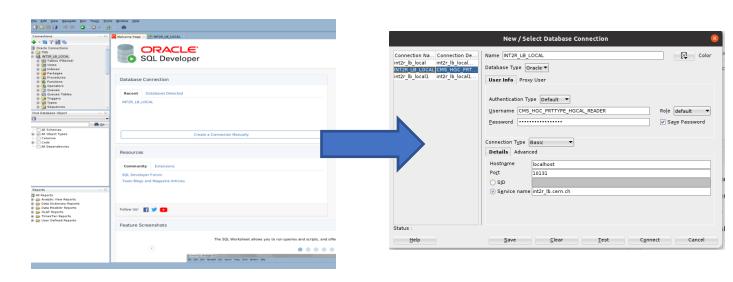
```
ssh -XY -L 10131:itrac1609-
v.cern.ch:10121 -L
10132:itrac1601-v.cern.ch:10121
username@lxplus.cern.ch
```

```
int2r lb local=(
    DESCRIPTION=
    (ADDRESS= (PROTOCOL=TCP)
                               (HOST=localhost)
                                                 (PORT=10131))
     (ADDRESS= (PROTOCOL=TCP)
                               (HOST=localhost)
                                                 (PORT=10132))
    (LOAD BALANCE=yes)
    (CONNECT DATA=
     (SERVER=DEDICATED)
    (SERVICE NAME=int2r lb.cern.ch)
     (FAILOVER MODE=
             (TYPE=SELECT)
             (METHOD=BASIC)
             (RETRIES=180)
             (DELAY=5)
```

Where username is your lxplus username.

# Connecting to the Database with SQLDeveloper (2)

- Launch SQLDeveloper.
- Press "Open a Connection Manually" in the front page.
- On the "New / Select Database Connection" Dialogue box, fill:
  - Name: INT2R\_LB\_LOCAL
  - Username: CMS\_HGC\_PRTTYPE\_HGCAL\_READER
  - Password: HGCAL\_Reader\_2016
    - Check "save password" if you want.
  - Hostname: localhost
  - Port: 10131
  - Check on "Service name"
  - Service name: int2r\_lb.cern.ch
  - Press "Connect"



# Connecting to the Database with SQLDeveloper (3)

- Click View at the top of the screen.
- Click on "Find DB Object"
- Click on this database icon and select INT2R\_LB\_LOCAL
- Click on the + sign on the left of "All Schemas"
- Scroll down to entries that start with "CMS\_HGC\_\*" and check one (or more) that you want to view.
  - For us the relevant tables are actually ones like "HGC\_CERN\_SENSOR\_IV"
- Then click on "Go" (which has binoculars next to it).

### Uploading to the HGCAL DB (1/2)

- Generate XML files for your data based on the templates
  - Templates from Umesh: <a href="https://indico.cern.ch/event/1164937/contributions/4891851/attachments/2451351/4200663/XML%20Templates%20HGCAL%20Si%20Sensors.pdf">https://indico.cern.ch/event/1164937/contributions/4891851/attachments/2451351/4200663/XML%20Templates%20HGCAL%20Si%20Sensors.pdf</a>
  - Find the relevant templates from the documentation: <u>https://twiki.cern.ch/twiki/bin/view/CMS/ConstructionDB</u>
  - Scripts that can help in conversion to xml can be found <u>here</u> or <u>here</u> (or write your own scripts).
- You can upload to the DB using DBLoader.
  - DBloader code: <a href="https://github.com/valdasraps/cmsdbldr">https://github.com/valdasraps/cmsdbldr</a>

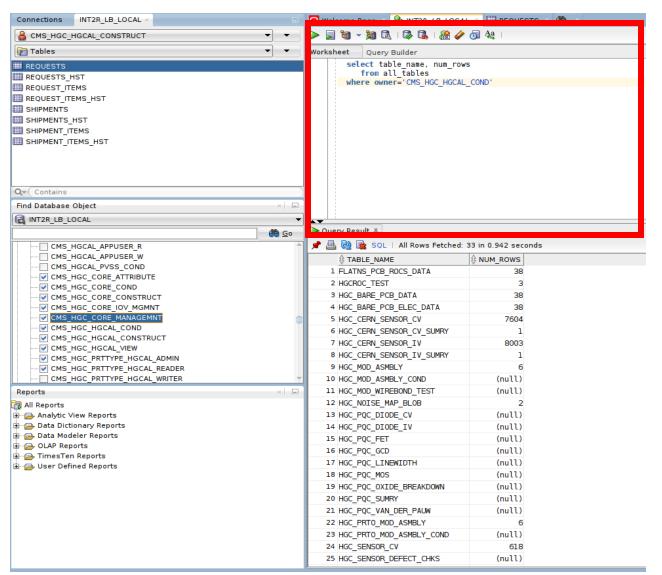
### Uploading to the HGCAL DB (2/2)

- To upload to the INT2R (development) database, do
  - scp <file.xml> xml <u>joshi@dbloader-hgcal.cern.ch:/home/dbspool/spool/hgcal/int2r</u>
- To upload to the CMSR (production) databse, do
  - scp <file.xml> xml joshi@dbloader-hgcal.cern.ch:/home/dbspool/spool/hgcal/int2r
- Check the status of the job on /home/dbspool/state/hgcal/int2r/filename.xml
  - E.g. joshi@dbloader-hgcal.cern.ch:/home/dbspool/state/hgcal/int2r/filename.xml
- Check the log file /home/dbspool/logs/hgcal/int2r/filename.xml for log error information.

#### Queries in SQLDeveloper Example 1

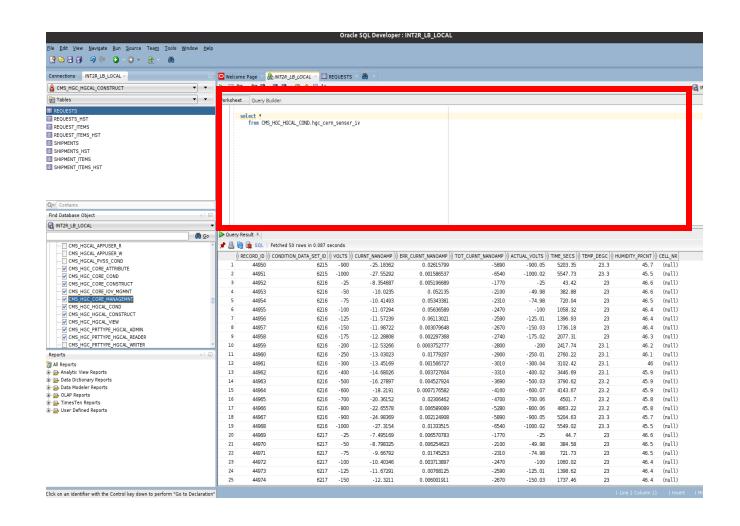
- On "worksheet" tab, write any SQL query
- Press ctrl+enter to execute.

(you don't have to check any tables, since you Have access to all of INT2R\_LB\_LOCAL.)

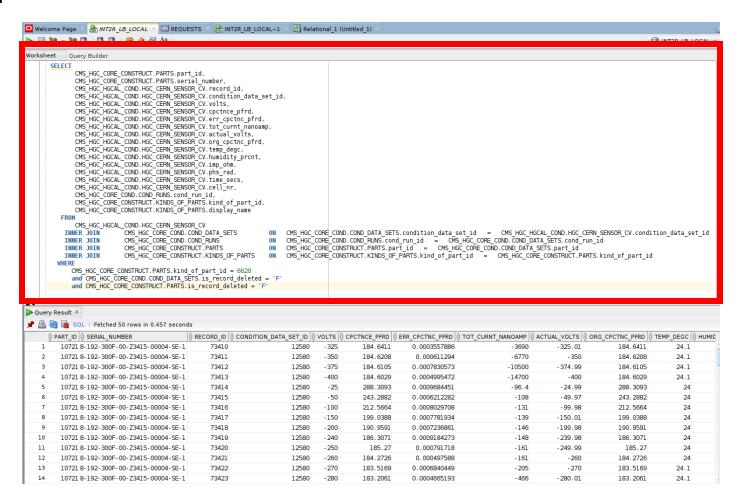


### Queries in SQLDeveloper Example 2

 See everything in the "HGC\_CERN\_SEN SOR\_IV" table.



# You can make SQL Queries infinitely complex!



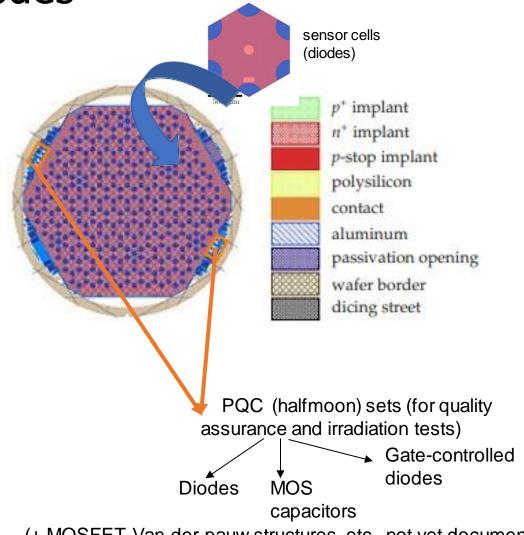
# Want to Generate ERDs? You can do that too!

- Click File, then Data Modeler, then Import, then Data Dictionary
- Select DB Connection (e.g. INT2R\_LB\_LOCAL), and click Next
- Select a schema (e.g. CMS\_HGC\_HGCAL\_COND), and click Next
- Select all the objects in the schema, and click Next
- Resulting ERD:



SQC, PQC and their Analysis Codes

- HexDaq is used for the output of raw data test results.
- SQC (sensor Quality Control) tests (such as IV and CV) tests the actual Si sensors. Analysis code for the results is <u>lcd hgcal analysis workflow</u>
   S.
- PQC (Process Quality Control) is performed on test structures. Analysis codes is <u>analysis-pqc</u>.



(+ MOSFET, Van-der-pauw structures, etc., not yet documented)