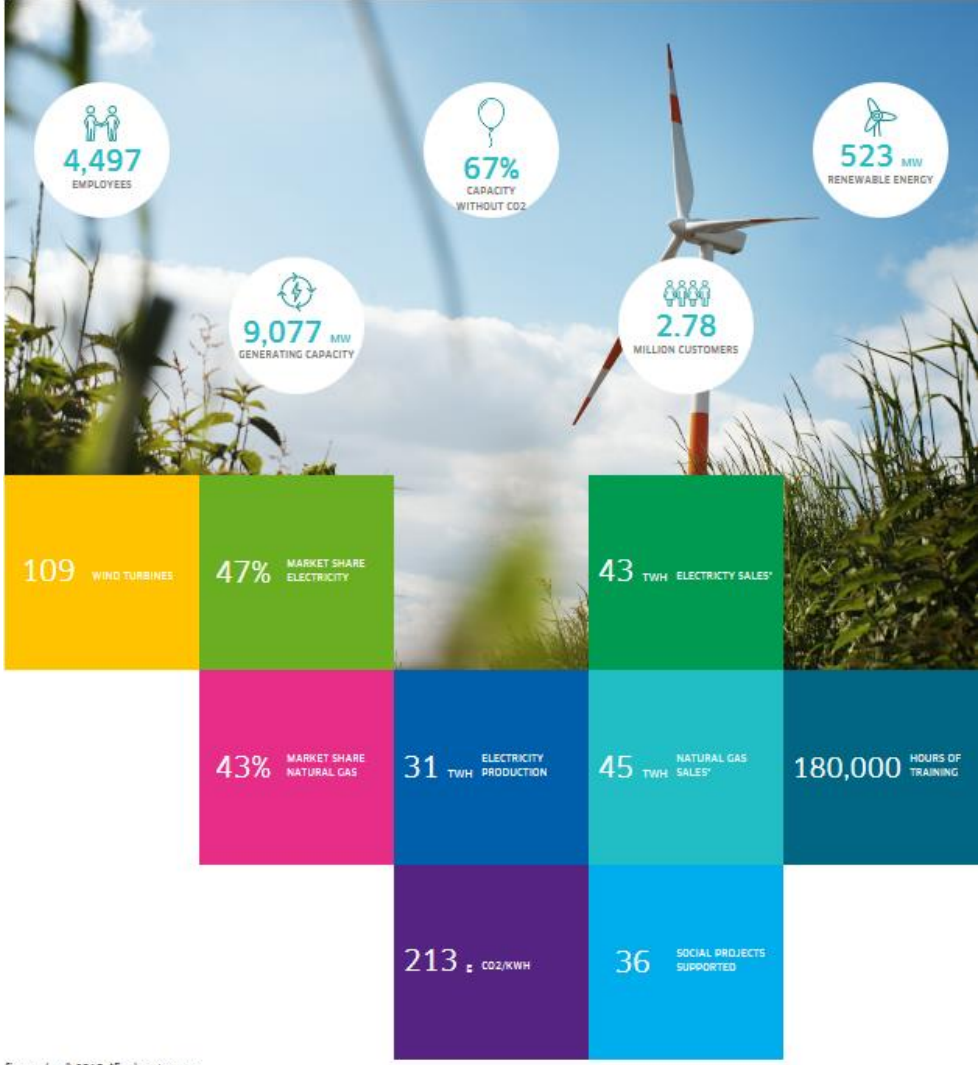
SAS BELUX FORUM 2016

SAS code optimization and best practices in Big Data context

KABACINSKI CHRISTOPHE – BEOPTIMIZED SPRL

ENGIE Electrabel is the number one producer and supplier of energy in **Belgium**.



figures (end) 2015; *final customers

CONTEXT | MIG6

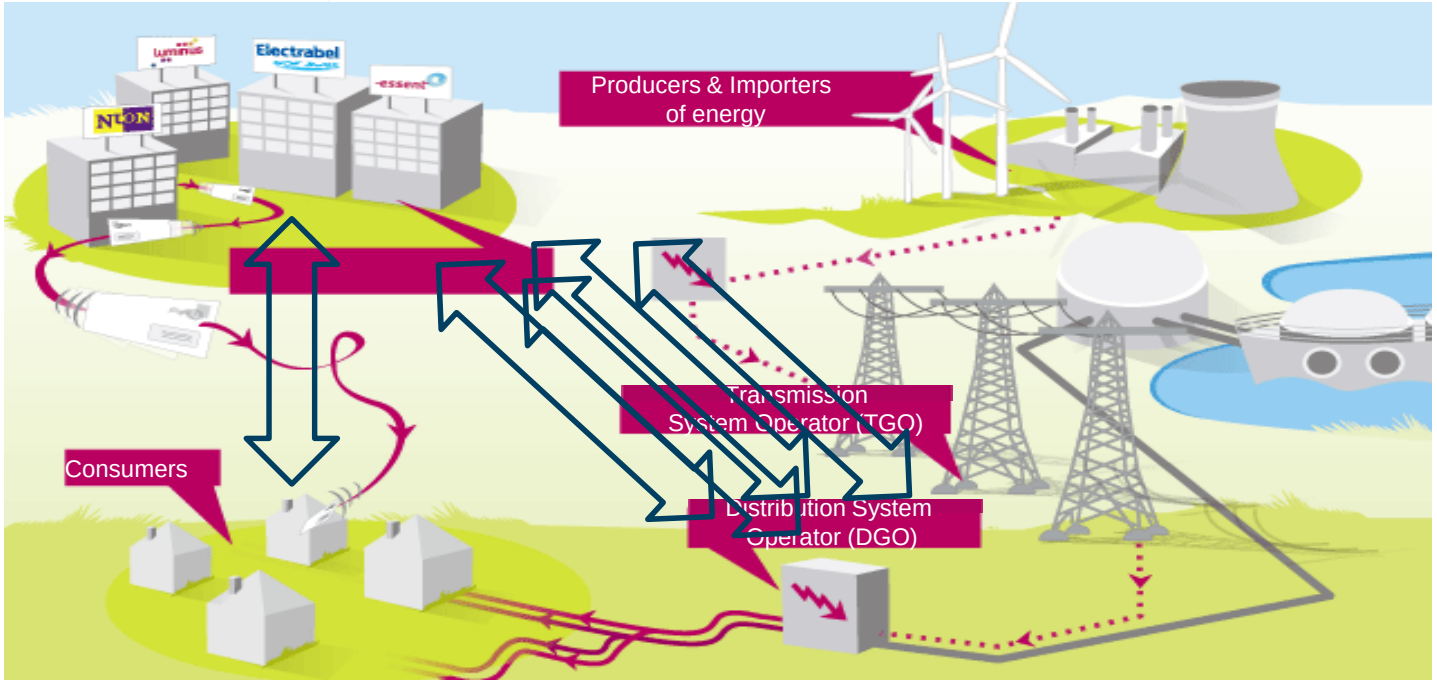
- The **MIG** (**M**essage **I**mplementation **G**uide) is the **regulatory-driven protocol** used to **standardize** Utility Market processes and associated communication between market parties.

- **MIG6: Introduction of ATRIAS, a clearing house,** acting as an intermediary between the energy suppliers and the DGOs on the one hand and the DGOs to TGOs on the other hand.

- **MIG6** represents the **largest evolution in** those market processes since the liberalization of the Belgian Energy Market. **Compliance with MIG6 is mandatory** for continued presence of ENGIE Electrabel as a Supplier and Balance Responsible. MIG6 is **due** in **January 2018** and **mandatory certification tests** are foreseen as from **May 2017.**

CONTEXT | ENERGY ACTORS

Lot of information are exchanged & validated between DGO's and energy suppliers…



Examples: TGO: Elia & Fluxys ; **DGO**: Sibelga, Ores, Infrax ,RESA, Eandis

CONTEXT | ENERGY ACTORS

In MIG6, messages will be grouped & standardized at ATRIAS.

Examples: TGO: Elia & Fluxys ; **DGO**: Sibelga, Ores, Infrax ,RESA, Eandis

CONTEXT | SAS AT THE M&S DEPARTMENT

- SAS used at Engie Electrabel Marketing & Sales department since 1993
  - Started on desktop, addition of ETL Server in 1998 and of Visual Analytics in 2013
- Operating System : Windows & Linux
- Database: SAS Data Warehouse – more than **7 Tb** of data
- Reports are then used in numerous departments

CHALLENGES

- Atrias regroup all DGO's in the market ➔ **BIG DATA Context**
- EAN will be replaced by SDP (Service Delivery Point) ➔ **new natural key** in all existing tables
- Web Service will replace EDI(EL) ➔ **redesign of all the input data**
- **Many other new processes (new calculations, new variables, European codes, Simplifications…)**
- Almost 11 000 existing SAS programs in the department (not all impacted by MIG6)

CHALLENGES

FEW CODES CREATED LONG TIME AGO NOT REALLY OPTIMIZED…

```
data edtar; set ed7; if DNB not in ('INTERENERGA','IVEG','INFRAX WEST','PBE VLA','PBE WAL','IVEG(ex-AGEM)')
                              then if factor in ('D_PUBLIC_SERVICE_MISSIONS_DAY','D_PUBLIC_SERVICE_MISSIONS_NIGHT')
                                        then factor='D_PUBLIC_SERVICE_MISSIONS';run;
proc sort noduplicates; by DNB Tariff MetMeth Factor From To Price Price_Unit; run;

/* Selectie TNB tarieven */
data et1; set elekTNB.tarieven_Elek; keep DNB Tariff Factor From To Price Price_Unit; run;
data et2; set et1; if Price_Unit not in (' ','%'); run;
data et3; set et2; if Factor not in ('D_CALC_ZERO');run;
data et4; set et3; if Price not in ('.');run;
data et5; set et4; metmeth=' '; run;
data et6; set et5; if DNB='GASELWEST VLA' then DNB='GASELWEST VL';run;
data et7; set et6; if DNB='IVEG' AND Tariff='30l' then Tariff='30L';if DNB='TECTEO' then Tariff='T' || Tariff; run;
data ettar; set et7;
if DNB='SIBELGAS' then DNB='SIBELGAS';
if DNB='DNBBA' AND factor='' AND Price_Unit='EUR/kwh' then Price=round(price,0.0000001);
if DNB in ('ORES (Namur)','ORES (Hainaut Electricité)','ORES (Verviers)','ORES (Brabant wallon)','ORES (Mouscron)',
                        'ORES (Luxembourg)','ORES (Est)','TECTEO ORES')
                        and factor='T_POWER_MONTH' then price=round(price,.0000001);
run;
proc sort noduplicates; by DNB Tariff MetMeth Factor From To Price Price_Unit; run;
/* Selectie GAS tarieven */
data g1; set gas.tarieven_gas; keep DNB Tariff Factor From To Price Price_Unit; run;
data g2; set g1; if Price_Unit not in (' ','%'); run;
data g3; set g2; if Factor not in ('D_CALC_ZERO');run;
data g4; set g3; if Price not in ('.');run;
data g5; set g4; format metmeth $3.; if factor in ('D_METERREADING_E13','D_METERREADING_B17','D_METERREADING_B18') then metmeth=substr(factor,16,3);run;
data g6; set g5; if factor in ('D_METERREADING_E13','D_METERREADING_B17','D_METERREADING_B18') then factor='D_METERREADING';run;
data g7; set g6; if DNB in ('GASELWEST WAL','IGH','SEDILEC','SIMOGEL','IDEG','TECTEO ORES') And Factor='D_MUNICIPAL_FEES' then Factor='D_TAXES_DE_VOIRIE';run;
data gtar; set g7;
If DNB in ('GASELWEST VL','GASELWEST WAL','IMEA', 'IMEWO', 'INTERGEM','IVEKA','IVERLEK','SIBELGAS')
                        and  Factor='D_POWER'  and Price_Unit='EUR/kW/jaar'     then Price=round(Price/12,.0000001);
If DNB in ('ORES (Namur)','ORES (Hainaut Gaz)','ORES (Luxembourg)','ORES (Brabant wallon)','ORES (Mouscron)')
                        and  Factor='D_POWER'  and Price_Unit='EUR/kW/jaar'     then Price=round(Price/12,.00001);
If DNB in ('SIBELGA')
                        and  Factor='D_POWER'  and Price_Unit='EUR/kW/jaar' then Price=round(Price/12,.000001);
If DNB in ('INFRAX WEST', 'INTERENERGA','IVEG')
                        and  Factor='D_POWER'  and Price_Unit='EUR/kW/jaar'     then Price=round(Price/12,.0000001);
/*If DNB in ('TECTEO') and  Factor='D_POWER'  and Price_Unit='EUR/kW/jaar' then Price=round(Price/12,.001);*/
if Factor='D_POWER' and Price_Unit='EUR/kW/jaar' then Factor='D_POWER_CAP';
run;
proc sort noduplicates; by DNB Tariff MetMeth Factor From To Price Price_Unit; run;

proc datasets library=base;  delete c8;quit;run;
```

THE PROCESS | ILLUSTRATION WITH THE GRIDFEE DETAIL

- **Gridfee detail:** message sent from the DGO (eg: Sibelga) to the supplier (eg: ENGIE Electrabel)
- **Content:** it contains the details of gridfee (ie: list of charges applied on the grid users for their grid usage) per Service Delivery Point (SDP)

THE MARKET MESSAGE

ILLUSTRATION WITH THE GRIDFEE DETAIL

- Contains gridfees for given periods (invoice & advance).
- XML converted in CSV
- Delivered **continuously**
- One line per SDP (gas or elec)
- Up to **6 000 000 lines**
- **5 300 Variables**
- DQ check **required** (missing, counts…) and verify the unitprice and quantity for each item
- Conditional aggregation

1.1.2.2. Exchanged Information Diagram

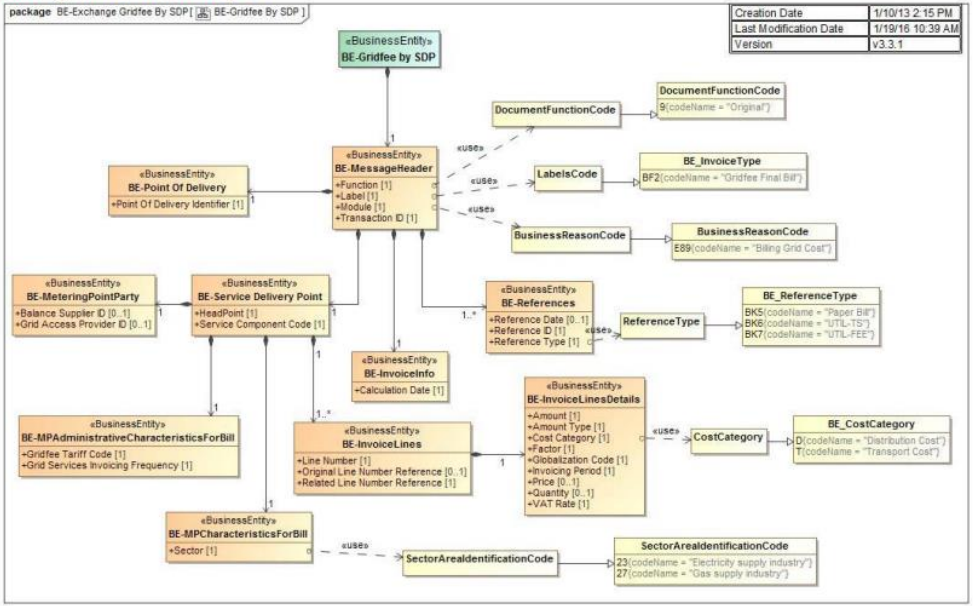


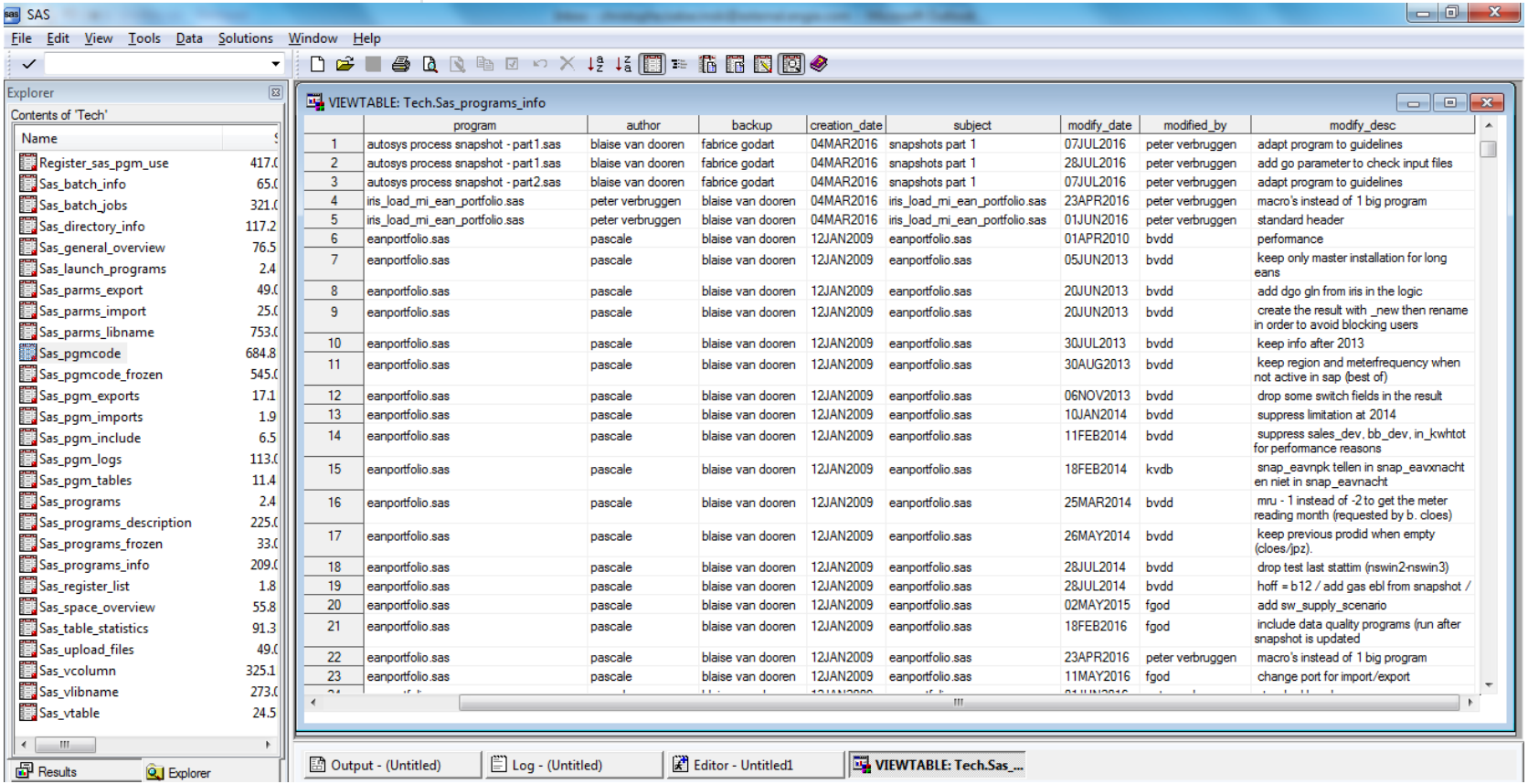
Figure 4 BE-Gridfee By SDP - *[ref: _17_0_2_2_b9402f1_1357823747764_547761_60823]*

COMMENTS

STANDARDIZATION OF THE PROGRAM HEADER AND COMMENTS IN THE CODES

```
/*************************************************************************************************************/
/* macro : Load_Detailed_Gridfee.sas                                                                         */
/*                                                                                                           */
/* Creation: 08/07/2016                                                                                      */
/*                                                                                                           */
/* Author: Kabacinski Christophe                                                                             */
/* Backup: Blaise Van Dooren                                                                                 */
/*                                                                                                           */
/* Subject: launch all detail gridfee data in SAS DWH                                                        */

/* Input:&folder_path47602.03_In\01_RawData\12_DGO\Gridfee_Detailed\Test\                                     */
/* Output:RDETGRID                                                                                           */

/* Steps :                                                                                                   */
/*                                                                                                           */
/*-----------------------------------------------------------------------------------------------------------*/
/* documentation start                                                                                       */
/* folder    :                                                                                               */
/* document  :                                                                                               */
/* reports   :                                                                                               */
/* description :                                                                                             */
/* documentation end                                                                                         */
/*-----------------------------------------------------------------------------------------------------------*/
/* modifications start                                                                                       */
/* DATE              | WHO                        | DESCRIPTION                                              */
/* ----              | ---                        | -----------                                              */
/*08/07/2016         | Christophe kabacinski      | creation for MIG6                                      */
/* modifications end                                                                                         */
/*************************************************************************************************************/
```

COMMENTS SCAN ALL PROGRAMS AND STORE METADATA IN SAS



	program	author	backup	creation_date	subject	modify_date	modified_by	modify_desc
1	autosys process snapshot - part1.sas	blaise van dooren	fabrice godart	04MAR2016	snapshots part 1	07JUL2016	peter verbruggen	adapt program to guidelines
2	autosys process snapshot - part1.sas	blaise van dooren	fabrice godart	04MAR2016	snapshots part 1	28JUL2016	peter verbruggen	add go parameter to check input files
3	autosys process snapshot - part2.sas	blaise van dooren	fabrice godart	04MAR2016	snapshots part 1	07JUL2016	peter verbruggen	adapt program to guidelines
4	iris_load_mi_ean_portfolio.sas	peter verbruggen	blaise van dooren	04MAR2016	iris_load_mi_ean_portfolio.sas	23APR2016	peter verbruggen	macro's instead of 1 big program
5	iris_load_mi_ean_portfolio.sas	peter verbruggen	blaise van dooren	04MAR2016	iris_load_mi_ean_portfolio.sas	01JUN2016	peter verbruggen	standard header
6	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	01APR2010	bvdd	performance
7	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	05JUN2013	bvdd	keep only master installation for long eans
8	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	20JUN2013	bvdd	add dgo gln from iris in the logic
9	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	20JUN2013	bvdd	create the result with _new then rename in order to avoid blocking users
10	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	30JUL2013	bvdd	keep info after 2013
11	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	30AUG2013	bvdd	keep region and meterfrequency when not active in sap (best of)
12	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	06NOV2013	bvdd	drop some switch fields in the result
13	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	10JAN2014	bvdd	suppress limitation at 2014
14	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	11FEB2014	bvdd	suppress sales_dev, bb_dev, in_kwhtot for performance reasons
15	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	18FEB2014	kvdb	snap_eavnpk tellen in snap_eavxnacht en niet in snap_eavnacht
16	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	25MAR2014	bvdd	mru - 1 instead of -2 to get the meter reading month (requested by b. cloes)
17	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	26MAY2014	bvdd	keep previous prodid when empty (cloes/jpz).
18	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	28JUL2014	bvdd	drop test last stattim (nswin2-nswin3)
19	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	28JUL2014	bvdd	hoff = b12 / add gas ebl from snapshot /
20	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	02MAY2015	fgod	add sw_supply_scenario
21	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	18FEB2016	fgod	include data quality programs (run after snapshot is updated
22	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	23APR2016	peter verbruggen	macro's instead of 1 big program
23	eanportfolio.sas	pascale	blaise van dooren	12JAN2009	eanportfolio.sas	11MAY2016	fgod	change port for import/export

DQ CHECK (PART 1) | CONTROL DQ 'ON THE FLY' …

```
* BASIC control;
*******************************************************************************;
    if _n_ eq 1 then do;
         errorCnt=0;
         missingCnt=0;
      end;

      if compress(sdp) eq '' then do;
         rejectMessage='Null Values';
         missingCnt+1;
         output    errors.&fileout._err                ;
      end;

      else if _error_ ne 0 then do;
         rejectMessage='Invalid Data';
         errorCnt+1;
         output   errors.&fileout._err;
      end;

      if lastRec then do;
         call symputx('recCnt',_n_);
         call symputx('errorCnt',errorCnt);
         call symputx('missingCnt', missingCnt);
      end;

/*we store raw data to the zip folder for possible further analysis*/
output tozip.&fileout;
run;

%ProcessMetric(macroname=&macro_main., start_time=&start_main. ,
    process_name= load &filein. , RecCnt=&RecCnt., ErrorCnt=&ErrorCnt.,
    MissingCnt=&missingCnt., dupCnt=);
```

DQ CHECK (PART 1) … AND STORE INFO IN DQ TABLE

	job_name	start_time	end_time	rc	type	us
814	SNAPCO_LOAD_FORECASTING	02AUG2016:15:23:52	02AUG2016:15:23:52	0	sub	CGM566
815	I_W_METERTYPE	02AUG2016:15:23:52	02AUG2016:15:23:53	0	sub	CGM566
816	REPORTDQSNAPSHOTEAVCHECK	02AUG2016:15:23:53	02AUG2016:15:23:53	0	sub	CGM566
817	LOAD_DETAILED_GRIDFEE	02AUG2016:15:31:36	.	.	sub	CGM566
818	DGO_UNIT_PRICE_CONTROL	02AUG2016:15:31:36	02AUG2016:15:31:40	0	sub	CGM566
819	LOAD_DETAILED_GRIDFEE	02AUG2016:15:34:18	.	.	sub	CGM566
820	DGO_UNIT_PRICE_CONTROL	02AUG2016:15:34:18	02AUG2016:15:34:21	0	sub	CGM566
821	LOAD_AGGREGATED_GRIDFEE	02AUG2016:15:49:12	02AUG2016:15:49:14	0	sub	CGM566
822	LOAD_AGGREGATED_GRIDFEE	02AUG2016:15:53:33	02AUG2016:15:53:36	0	sub	CGM566
823	LOAD_AGGREGATED_GRIDFEE	02AUG2016:15:55:32	02AUG2016:15:55:35	0	sub	CGM566
824	LOAD_AGGREGATED_GRIDFEE	02AUG2016:15:57:21				
825	LOAD_AGGREGATED_GRIDFEE	02AUG2016:16:12:15				
826	LOAD_AGGREGATED_GRIDFEE	02AUG2016:16:25:55				
827	LOAD_AGGREGATED_GRIDFEE	02AUG2016:16:28:33				
829	TEST_PAUL	02AUG2016:16:35:37				
830	AUTOSYSPROCESSSNAPSHOT_PART_1.SAS	03AUG2016:10:24:59				
831	LOAD_SNAPSHOT	03AUG2016:10:24:59				
832	REPORT_SNAPSHOT_EXTRACT	03AUG2016:10:25:03				
833	SNAPSHOT_TD	03AUG2016:10:25:05				
834	REPORT_DQ_DELTA_EAV	03AUG2016:10:25:10				
835	EMV_SNAP	03AUG2016:10:25:15				
836	AUTOSYSPROCESSSNAPSHOT_PART_2.SAS	03AUG2016:10:29:03				
837	SNAPCO_TD	03AUG2016:10:29:06				

	macroname_name	start_time	process_name	RecC	ErrorC	Missin	DupC
546	EMV_SNAP	06JUL2016:15:07:37	Tot Number of SDP in ddgo.emv_201607	4	.	.	.
547	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000707_5414489000508_23_21JUN2016142255.csv	.	.	.	.
548	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000707_5414489000607_27_21JUN2016142256.csv	.	.	.	.
549	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000707_5499762305006_27_21JUN2016142257.csv	.	.	.	.
550	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000707_5499764353302_23_21JUN2016142258.csv	.	.	.	.
551	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000905_5414489000508_23_21JUN2016142259.csv	.	.	.	.
552	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000905_5414489000607_27_21JUN2016142300.csv	.	.	.	.
553	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499762305006_27_21JUN2016142300.csv	.	.	.	.
554	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499762305006_27_21JUN2016142301.csv	.	.	.	.
555	LOAD_DETAILED_GRIDFEE	08JUL2016:14:59:39	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499764353302_23_21JUN2016142302.csv	.	.	.	.
556	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000707_5414489000508_23_21JUN2016142255.csv	2	0	0	.
557	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000707_5414489000607_27_21JUN2016142256.csv	2	0	0	.
558	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000707_5499762305006_27_21JUN2016142257.csv	2	0	0	.
559	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000707_5499764353302_23_21JUN2016142258.csv	4	0	0	.
560	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000905_5414489000508_23_21JUN2016142259.csv	2	0	0	.
561	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000905_5414489000607_27_21JUN2016142300.csv	2	0	0	.
562	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499762305006_27_21JUN2016142300.csv	1	0	0	.
563	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499762305006_27_21JUN2016142301.csv	2	0	0	.
564	LOAD_DETAILED_GRIDFEE	08JUL2016:16:03:20	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499764353302_23_21JUN2016142302.csv	4	0	0	.
565	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000707_5414489000508_23_21JUN2016142255.csv	2	0	0	.
566	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000707_5414489000607_27_21JUN2016142256.csv	2	0	0	.
567	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000707_5499762305006_27_21JUN2016142257.csv	2	0	0	.
568	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000707_5499764353302_23_21JUN2016142258.csv	4	0	0	.
569	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000905_5414489000508_23_21JUN2016142259.csv	2	0	0	.
570	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000905_5414489000607_27_21JUN2016142300.csv	2	0	0	.
571	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499762305006_27_21JUN2016142300.csv	1	0	0	.
572	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499762305006_27_21JUN2016142301.csv	2	0	0	.
573	LOAD_DETAILED_GRIDFEE	08JUL2016:16:11:59	load GridfeeDataByServiceDeliveryPoint_5414488000905_5499764353302_23_21JUN2016142302.csv	4	0	0	.
574	LOAD_DETAILED_GRIDFEE	08JUL2016:16:13:31	load GridfeeDataByServiceDeliveryPoint_5414488000707_5414489000508_23_21JUN2016142255.csv	2	0	0	.
575	LOAD_DETAILED_GRIDFEE	08JUL2016:16:13:31	load GridfeeDataByServiceDeliveryPoint_5414488000707_5414489000607_27_21JUN2016142256.csv	2	0	0	.
576	LOAD_DETAILED_GRIDFEE	08JUL2016:16:13:31	load GridfeeDataByServiceDeliveryPoint_5414488000707_5499762305006_27_21JUN2016142257.csv	2	0	0	.
577	LOAD_DETAILED_GRIDFEE	08JUL2016:16:13:31	load GridfeeDataByServiceDeliveryPoint_5414488000707_5499764353302_23_21JUN2016142258.csv	4	0	0	.

DQ CHECK (PART 2) | USE HASH OBJECT TO CHECK DATA QUALITY

```
data DGFDET.&fileout(keep=SDP S_Kwhhi S_Kwhlo S_Kwhlox S_Kwhtot S_Kwhtot_cred S_Kwhtot_deb  S_TRT_date  S_dpowe
                         S_eurfede S_eurtnbgreen  S_eurtot S_eurtrns S_grdtype S_invdatf S_invdatt S_market

        errors.&fileout._err
        tozip.&fileout(drop=db_:)

        &fileout._DQ(keep=filename SDP s_market S_eandgo payload_gridfee_tariff payload_meter_freq  item_price_
                     db_GLN db_DNB db_tariff db_factor db_metmeth db_from db_to db_price grid_end_date grid_s

length db_key $ 100 db_from db_to  db_price 8 db_tariff db_metmeth $3 db_GLN db_factor db_DNB $ 17;
if _n_=1 then do ;
declare hash prices(dataset:"all_tariffs");
    prices.definekey('db_key');
    prices.definedata('db_from','db_to','db_tariff','db_factor','db_metmeth','db_price','db_GLN','db_DNB');
    prices.definedone();
    call missing (db_from,db_to,db_tariff,db_factor,db_metmeth,db_price,db_GLN,db_DNB,db_key);
end;
```

```
rc=prices.find(key:gridkey);
if rc eq 0 then do;

  start=datepart(input(inv_start{i},E8601Dz25.));
  grid_start_date=mdy(month(start),1,year(start));
  end=datepart(input(inv_end{i},E8601Dz25.));
  grid_end_date=mdy(month(end),1,year(end));
```

GROUP VARIABLES

USE ARRAY AND VARIABLE LISTS TO PROCESS VARIABLES IN GROUP

```
length  SDP $ 30
        payload_function $ 1
        payload_ref_id $ 20
        payload_SDP_id $ 20
        payload_SDP_comp_id $ 10
        payload_meter_freq $ 3
        payload_gridfee_tariff $ 5
        payload_inv_start_occur_1 - payload_inv_start_occur_654 $ 25
        payload_inv_end_occur_1 - payload_inv_end_occur_654 $ 25
        payload_glob_code_1 - payload_glob_code_654 $ 4
        payload_amount_EAC_1 - payload_amount_EAC_654 8
        payload_quantity_1 - payload_quantity_654 8
        payload_credit_line_ind_1 - payload_credit_line_ind_654 $ 1(
        payload_item_price_EAC_1 - payload_item_price_EAC_654 8
        payload_add_trade_id_1 - payload_add_trade_id_654 $ 30;
```

```
input payload_function  payload_ref_id  payload_SDP_id  payload_SDP_comp_id payload_meter_freq  p
      payload_inv_start_occur_1  - payload_inv_start_occur_654
      payload_inv_end_occur_1 - payload_inv_end_occur_654
      payload_glob_code_1 - payload_glob_code_654
      payload_amount_EAC_1 - payload_amount_EAC_654
      payload_quantity_1 - payload_quantity_654
      payload_credit_line_ind_1 - payload_credit_line_ind_654
      payload_item_price_EAC_1 - payload_item_price_EAC_654
      payload_add_trade_id_1 - payload_add_trade_id_654 ;
      SDP=catt(payload_SDP_id,'_',payload_SDP_comp_id);

      S_Grdtype = 'A';
      S_pprref1=payload_ref_id;
      S_mtrmeth=payload_meter_freq;
      S_tarifid=payload_gridfee_tariff;

      array inv_start {654} payload_inv_start_occur_1  - payload_inv_start_occur_654 ;
      array inv_end {654} payload_inv_end_occur_1 - payload_inv_end_occur_654 ;
      array glob_code {654} payload_glob_code_1 - payload_glob_code_654  ;
      array amount {654} payload_amount_EAC_1 - payload_amount_EAC_654 ;
      array quantity {654} payload_quantity_1 - payload_quantity_654 ;
      array credit_line {654}     payload_credit_line_ind_1 - payload_credit_line_ind_654;
      array item_price {654} payload_item_price_EAC_1 - payload_item_price_EAC_654 ;
      array add_trade {654} payload_add_trade_id_1 - payload_add_trade_id_654 ;
```

USE MACROS | USE MACRO TO LIMIT CODE REPETITION

```
******************************** eurdist 100% ***********************************;

%Sum_amount_203_204 (  eurdist
                     , (credit_line{i}='+'   and
                        glob_code{i} in ('100','110','120','130','140','150','160','200','210','215','220','230'
                                        ,'240','300','310','320','330','800','810','811','812','813','820','830
                                        ,'840','850','860','890','891','250'
                                        )
                       )
                         OR
                        ( credit_line{i}='-'  and glob_code{i} ='211')

                      );

******************************** eurtrns 100% ***********************************;

%Sum_amount_203_204 (  eurtrns
                      ,( credit_line{i}='+' and
                        glob_code{i} in ('500','510','520','530','540','550','600','610','620','630','640','900
                                         '910','913','920','930','960','970','975','976','980'
                                        )
                       )
                        OR
                       (credit_line{i}='-' and glob_code{i} ='521')
                      );
```

USE MACROS

CREATE REUSABLE AUTOCALL MACROS WHICH CAN BE USED IN ALL CODES

```
/*****************************************************************************************
/* Description : START OF ARCHIVE
/*****************************************************************************************
    %Display_msg ( ! START OF ARCHIVE );

    data _null_;
    call symput ( "datzip" , put(Date() , yymmddn8.) );
    call symput ( "tijdzip" , compress(translate(put(time() , hhmm5.),'',':' )));
    run;

    %zipall (&file_path., &zip_path. , allocationresultsbyservicedeliverypoint*23 ) ;

    %delall  (&file_path., allocationresultsbyservicedeliverypoint*23 ) ;

    %let rc=&syserr.;
    %let end_time = %sysfunc ( datetime() , datetime20.);
    %sas_batch_info ( &macro_main. , &start_main., &end_time. , &rc. , sub ) ;

/*****************************************************************************************
/* Description : program END
/*****************************************************************************************
    %Display_msg ( ! END of &macro_main. );
```

CONTROL INPUT

PROCESS SPECIFIC FILES OF A DIRECTORY AND RETURN METADATA

```
FILENAME Filelist PIPE "DIR /B &RGRIDPATH.*.*";
DATA GRID_INV GRID_ADV   ;
  LENGTH path FileName  $ 200 key $ 20 id $ 7 process $ 2 recipient_id sender_id $ ;
  INFILE Filelist ;
  INPUT FileName;


    if upcase(scan( filename,-1,'.')) EQ "CSV";
    path="&RGRIDPATH";
    rundate=today();
    name=scan(filename,1,'.');
    msg_type=scan(name,1,'_');
    longdate=scan(name,5,'_');
    id=substr(longdate,9);
    recipient_id=scan(name,2,'_');
    sender_id=scan(name,3,'_');
    process=scan(name,4,'_');
    key=catt('GRID',put(md5(filename),hex16.));
    drop longdate;

    received_date=input(substr(longdate,1,9),date9.);
    format received_date rundate date9.  ;
if upcase(msg_type)='GRIDFEEDATABYSERVICEDELIVERYPOINT' then output  GRID_IN
else if upcase(msg_type)='GRIDFEEADVANCEBYSDP' then output GRID_ADV;
RUN;
```

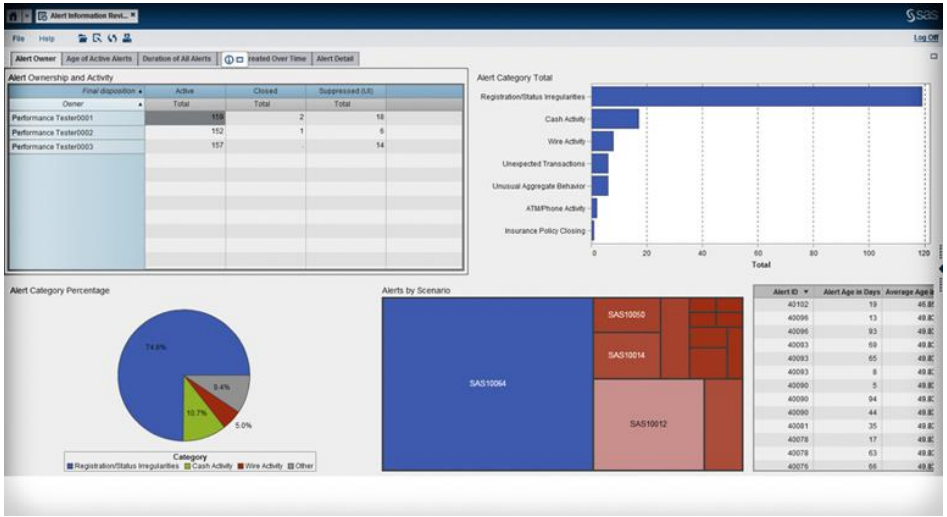
```
/*************************************************************************************
/* Description : load Gridfee Detailed Invoices
/*************************************************************************************
%Display_msg ( load Gridfee Detailed Invoices   );


    %do k=1 %to &numRESP_GRID_INV;
        %LoadGRID_INV(&&fileIN_GRID_INV&k,&&fileOUT_GRID_INV&k);
    %end;
/*augment the link table with number of sdp and S_pprref1*/

%do k=1 %to &numRESP_GRID_INV;
        proc sql noprint;
        select count(sdp) , S_pprref1, S_Grdtype into : sdp, : ref, : type from DGFDI
        update  grid_inv set num_sdp =&sdp, S_pprref1="&ref", S_Grdtype="&type" where
        quit;
    %end;
/*concatenation of the DQ error tables*/
    %do k=1 %to &numRESP_GRID_INV;
    %let table=&&fileOUT_GRID_INV&k;
    proc append base=grid_inv_tariff_err data=&table._dq force;
    run;
    %end;

proc sort data=grid_inv_tariff_err;
by rundate;
run;


   %if %sysfunc(exist(DGFDET.grid_inv_tariff_err)) %then %do;
    data DGFDET.grid_inv_tariff_err ;
```

OPTIMIZATIONS

USE SASFILE STATEMENT TO PUSH DATASET IN MEMORY

```
option compress=yes;
SASFILE all_tariffs LOAD;
%LoadallGRIDS;
SASFILE all_tariffs CLOSE;
```

NEXT STEP | REPORTING



Thanks to all the metadata we have created in all the processes (execution time, number of records, missing…), BI reports can be easily created in order to have excellent overview on all the processes…

CONCLUSIONS

- Creation of a SAS code template combining best practices, metadata capture and traceability,
- Application of Base SAS optimization technics in big data context
- Reporting to monitor daily processes and detect issues faster

CURRICULUM | KABACINSKI CHRISTOPHE – BEOPTIMIZED

christophe@beoptimized.be

ckabacinski

- Bio-engineer (UCLouvain),
- Master in Industrial Management, option: Environment & Energy (KULeuven)

9 years of SAS expertise and passion :

- 2008-2012 Training consultant at SAS
- 2012-2015 Sr Analytical consultant at SAS
- 2016 Creation of BeOptimized SPRL (Independent Consultant)
- 2016 Contractor via Maltem Consulting at ENGIE Electrabel

QUESTIONS ?

Christophe@BeOptimized.be