

SCAPROC IMPACT ANALYSIS - PROVIDING QA AND COMPLIANCE INSIGHTS

SANDEEP JUNEJA / STIJN ROGIERIS / BEN BOCCHICCHIO



AGENDA

- SCAPROC
 - Introduction
 - Parsing SCAPROC information
- Aggregating impact analysis information
 - PC SAS
 - SAS® Life Science Analytics Framework (LSAF)
- Using workflow for impact analysis
- Graphical impact analysis through RDF

- Introduced in SAS 9.2
- SCAPROC procedure implements the SAS® Code Analyzer
- Captures information from a executing SAS program about
 - Inputs
 - Outputs
 - Use of macro symbols

SCAPROC SYNTAX

```
* Start SCAPROC;
proc scaproc; write; run;
proc scaproc; record "C:/AnalyzeIO/scafiles/sca.txt" EXPANDMACROS;
run;

* SAS PROGRAM CODE TO BE ANALYZED;
libname datalib "C:/AnalyzeIO";
data datalib.class;
    set sashelp.class;
run;

* End SCAPROC;
proc scaproc; write; run;
```

SCAPROC PARSING SCAPROC ANALYSIS

```
/* JOBSPLIT: JOBSTARTTIME 25APR2016:18:30:07.49 */  
/* JOBSPLIT: TASKSTARTTIME 25APR2016:18:30:07.51 */  
/* JOBSPLIT: DATASET INPUT SEQ #C00001.CLASS.DATA */  
/* JOBSPLIT: LIBNAME #C00001 V9 'C:\Program Files\SASHome\SASFoundation\9.4\core\sashelp' */  
/* JOBSPLIT: CONCATMEM #C00001 SASHELP */  
.....  
/* JOBSPLIT: DATASET OUTPUT SEQ DATALIB.CLASS.DATA */  
/* JOBSPLIT: LIBNAME DATALIB "C:/AnalyzeIO" */  
/* JOBSPLIT: FILE OUTPUT C:\AnalyzeIO\scafiles\sca.txt */  
/* JOBSPLIT: ELAPSED 31 */  
/* JOBSPLIT: SYSSCP WIN */  
/* JOBSPLIT: PROCNAME DATASTEP */  
/* JOBSPLIT: STEP SOURCE FOLLOWS */  
  
* SAS PROGRAM CODE TO BE ANALYZED;  
libname datalib "C:/AnalyzeIO";  
data datalib.class;  
  set sashelp.class;  
run;  
* End SCAPROC;  
/* JOBSPLIT: JOBENDTIME 25APR2016:18:30:07.54 */  
/* JOBSPLIT: END */
```

Start time of Execution

Input Data Set

Output Data Set

Libname Definition

End Execution Time

SCAPROC PARSING SCAPROC ANALYSIS

```
/* JOBSPLIT: JOBSTARTTIME 25APR2016:18:30:07.49 */
/* JOBSPLIT: TASKSTARTTIME 25APR2016:18:30:07.51 */
/* JOBSPLIT: DATASET INPUT SEQ #C00001.CLASS.DATA */
/* JOBSPLIT: LIBNAME #C00001 V9 'C:\Program Files\SASHome\SASFoundation\9.4\core\sashelp' */
/* JOBSPLIT: CONCATMEM #C00001 SASHELP */

.....
/* JOBSPLIT: DATASET OUTPUT SEQ DATALIB.CLASS.DATA */
/* JOBSPLIT: LIBNAME DATALIB "C:/AnalyzeIO" */
/* JOBSPLIT: FILE OUTPUT C:\AnalyzeIO\scafiles\sca.txt */
/* JOBSPLIT: ELAPSED 31 */
/* JOBSPLIT: SYSSCP WIN */
/* JOBSPLIT: PROCNAME DATASTEP */
/* JOBSPLIT: STEP SOURCE FOLLOWS */

* SAS PROGRAM CODE TO BE ANALYZED;
libname datalib "C:/AnalyzeIO";
data datalib.class;
  set sashelp.class;
run;
* End SCAPROC;
/* JOBSPLIT: JOBENDTIME 25APR2016:18:30:07.54 */
/* JOBSPLIT: END */

*Define patterns to extract different information;
if _n_=1 then do;
  type_pattern = prxparse("/(INPUT|OUTPUT|UPDATE LIBNAME|JOBSTARTTIME|JOBENDTIME)/i");
  obj_type_pattern = prxparse("/:s*\w+\s/i");
  * Change pattern to include all characters and spaces between quotes;
  lib_path_pattern = prxparse("/(\'|\\\"")+.*(\'|\\\"")+/i");
  xpt_libname_pattern = prxparse("/LIBNAME .* XPORT .*xpt/i");
  ttf_pattern = "/.*\.ttf/";
  * Start with digit 1 or more , followed by any characters till we '*' is encountered;
  time_pattern = prxparse("/\d+.*\*/i");
end;
```

IMPACT ANALYSIS SAS PROGRAM

- PC Execution

```
%include "C:/AnalyzeIO/macros/analyzeio.sas";
```

```
%AnalyzeIO (SASFile=%str(C:/AnalyzeIO/Study/Programs/ae.sas),  
            worklib=%str(C:/AnalyzeIO/worklib),  
            file_io=(Y),  
            master_io=%str(C:/AnalyzeIO/IO/master_io.sas7bdat)  
            );
```

SAS Program

Temp Folder

SAS Data Set

- Life Science Analytics Framework (LSAF) Execution

```
%include "&_sasws_/SAS/Files/AnalyzeIO/macro/analyzeio.sas";
```

```
%AnalyzeIO (SASFile=%str(&_sasws_/SAS/Files/Study/Programs/scan_log.sas),  
            worklib=%str(&_sasws_/SAS/Files/AnalyzeIO/worklib),  
            file_io=(Y),  
            master_io=%str(&_sasws_/SAS/Files/AnalyzeIO/IO/master_io.sas7bdat)  
            );
```

- SAS Macro can be downloaded from SAS Tech Support Link:

<http://support.sas.com/kb/58878>

IMPACT ANALYSIS SAS DATA SET WITH INPUTS/OUTPUTS (IO) INFORMATION

- PC Impact Analysis

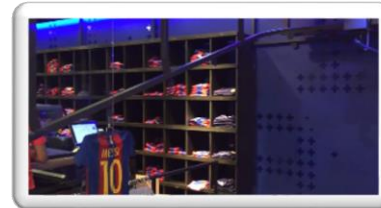
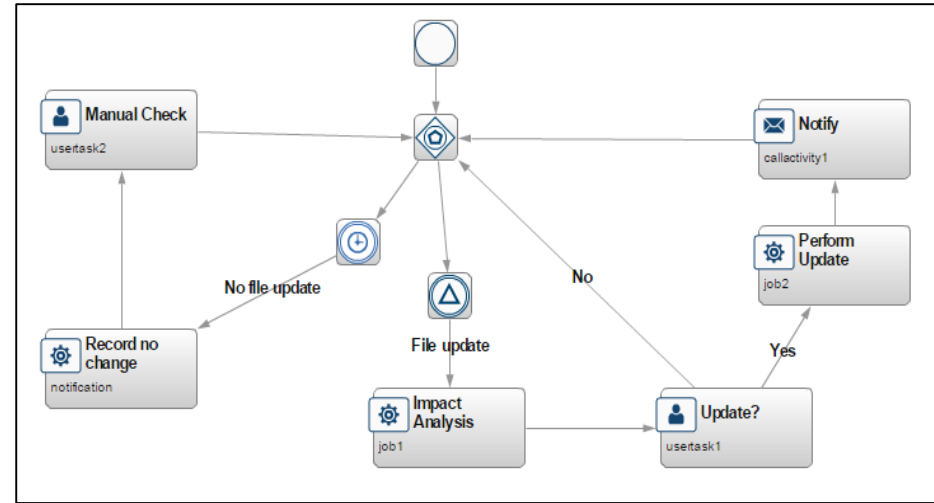
	sas_task	order_id	type	obj_type	file_path	timeinfo	createdDate	modifiedDate	fileSize
▶ 1	C:/AnalyzeIO/Study/Programs/ae.sas	1	JOBSTARTTIME	JOBSTARTTIME	C:/AnalyzeIO/Study/Programs/ae.sas	16AUG2016:08:04:23.60	25Apr2016:18:25:28	15Aug2016:20:14:50	756 Bytes
2	C:/AnalyzeIO/Study/Programs/ae.sas	2	JOBENDTIME	JOBENDTIME	C:/AnalyzeIO/Study/Programs/ae.sas	16AUG2016:08:04:23.77	25Apr2016:18:25:28	15Aug2016:20:14:50	756 Bytes
3	C:/AnalyzeIO/Study/Programs/ae.sas	3	INPUT	FILE	C:/AnalyzeIO/Study/Data/ae.sas7bdat		25Apr2016:14:43:50	16Aug2016:08:04:23	128 KB
4	C:/AnalyzeIO/Study/Programs/ae.sas	3	INPUT	FILE	C:/AnalyzeIO/Study/Data/formats.sas7bcat		25Apr2016:14:40:45	25Apr2016:14:42:53	17 KB
5	C:/AnalyzeIO/Study/Programs/ae.sas	3	INPUT	FILE	C:/AnalyzeIO/Study/macros/setup.sas		25Apr2016:14:40:45	25Apr2016:14:49:19	349 Bytes
6	C:/AnalyzeIO/Study/Programs/ae.sas	3	INPUT	FILE	C:/Program Files/SASHome/SASFoundation/9.4/core/sashelp/class.sas7bdat		03Dec2015:21:07:24	25Jun2015:01:56:11	128 KB
7	C:/AnalyzeIO/Study/Programs/ae.sas	4	OUTPUT	FILE	C:/AnalyzeIO/Study/Data/ae.sas7bdat		25Apr2016:14:43:50	16Aug2016:08:04:23	128 KB

- Life Science Analytics Framework (LSAF) Impact Analysis

Obs	sas_task	ord...	▲ type	obj_type	file_path	timeinfo	createdBy	createdDate	modifiedBy	modifiedDate	file Size
1	/SAS/Files/Study/Programs/api_test.sas	1	JOBSTARTTIME	JOBSTARTTIME	/SAS/Files/Study/Programs/api_test.sas	16AUG2016:12:07:04.14	sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	1 KB
2	/SAS/Files/Study/Programs/api_test.sas	2	JOBENDTIME	JOBENDTIME	/SAS/Files/Study/Programs/api_test.sas	16AUG2016:12:07:04.73	sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	1 KB
3	/SAS/Files/Study/Programs/api_test.sas	3	INPUT	FILE	/SAS/Files/Study/Programs/api_test.sas		sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	1 KB
4	/SAS/Files/Study/Programs/api_test.sas	3	INPUT	FILE	/SAS/Files/Study/Data/ae2.sas7bdat		sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	2 MB
5	/SAS/Files/Study/Programs/api_test.sas	3	INPUT	FILE	/SAS/Files/Study/Documents/CM_specs1.xlsxbak		sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	4 KB
6	/SAS/Files/Study/Programs/api_test.sas	3	INPUT	FILE	/SAS/Files/Study/Documents/CM_specs.xlsx		sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	8 KB
7	/SAS/Files/Study/Programs/api_test.sas	4	OUTPUT	FILE	/SAS/Files/Study/Data/ae4.sas7bdat		sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	2 MB
8	/SAS/Files/Study/Programs/api_test.sas	4	OUTPUT	FILE	/SAS/Files/Study/Documents/test.rtf		sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	25 KB
9	/SAS/Files/Study/Programs/api_test.sas	4	OUTPUT	FILE	/SAS/Files/Study/Documents/CM_specs1.xlsx		sajune	Sat Aug 13 22.4	sajune	Sat Aug 13 22.4	4 KB

BPMN-based workflow

- Set up event listener on folder
- Event trigger for file update
- Analyze impact of updated file as input to files
- Set up mechanism to
 - Notify owners of impacted file
 - Auto-execute impacted files



RDF RDF TURTLE

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix aio: <http://www.sas.com/sso/ontologies/analyzeio#>.
```

definition

```
<aio:#input> rdfs:range <aio:#ae.sas>.
<aio:#sasprg> rdfs:domain <aio:#ae.sas>.
<aio:#ae.sas> rdfs:domain <aio:#output>.
```

Program

```
<aio:#ae.sas> <aio:#Location> "C:/Study/Programs"^^<xs:string>;
  <aio:#createdby> "sajune"^^<xs:string>;
  <aio:#creationdate> "21JAN2016:21:10:51"^^<xs:dateTime>;
  <aio:#lastmodifiedby> "sajune"^^<xs:string>.
```

Input

```
<aio:#ae.sas7bdat> <aio:#Location> "C:/Study/Data"^^<xs:string>;
  <aio:#creationdate> "22JAN2016:10:00:00"^^<xs:dateTimeStamp>;
  <aio:#lastmodificationdate> "01FEB2016:11:00:02"^^<xs:dateTimeStamp>;
  rdfs:range <aio:#input>.
```

Output

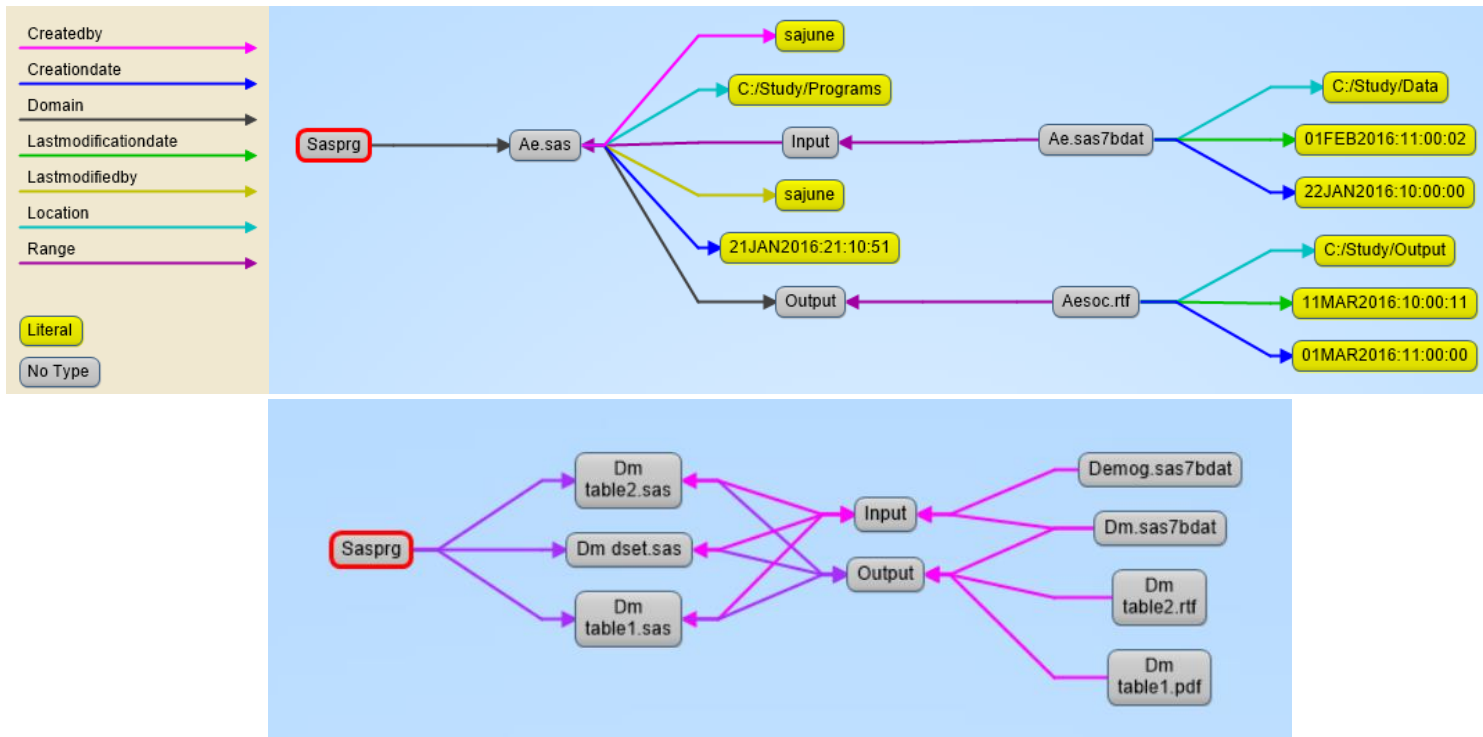
```
<aio:#aesoc.rtf> <aio:#Location> "C:/Study/Output"^^<xs:string>;
  <aio:#creationdate> "01MAR2016:11:00:00"^^<xs:dateTime>;
  <aio:#lastmodificationdate> "11MAR2016:10:00:11"^^<xs:dateTime>;
  rdfs:range <aio:#output>.
```

SAS Macro can be downloaded
from SAS Tech Support Link:

<http://support.sas.com/kb/58878>

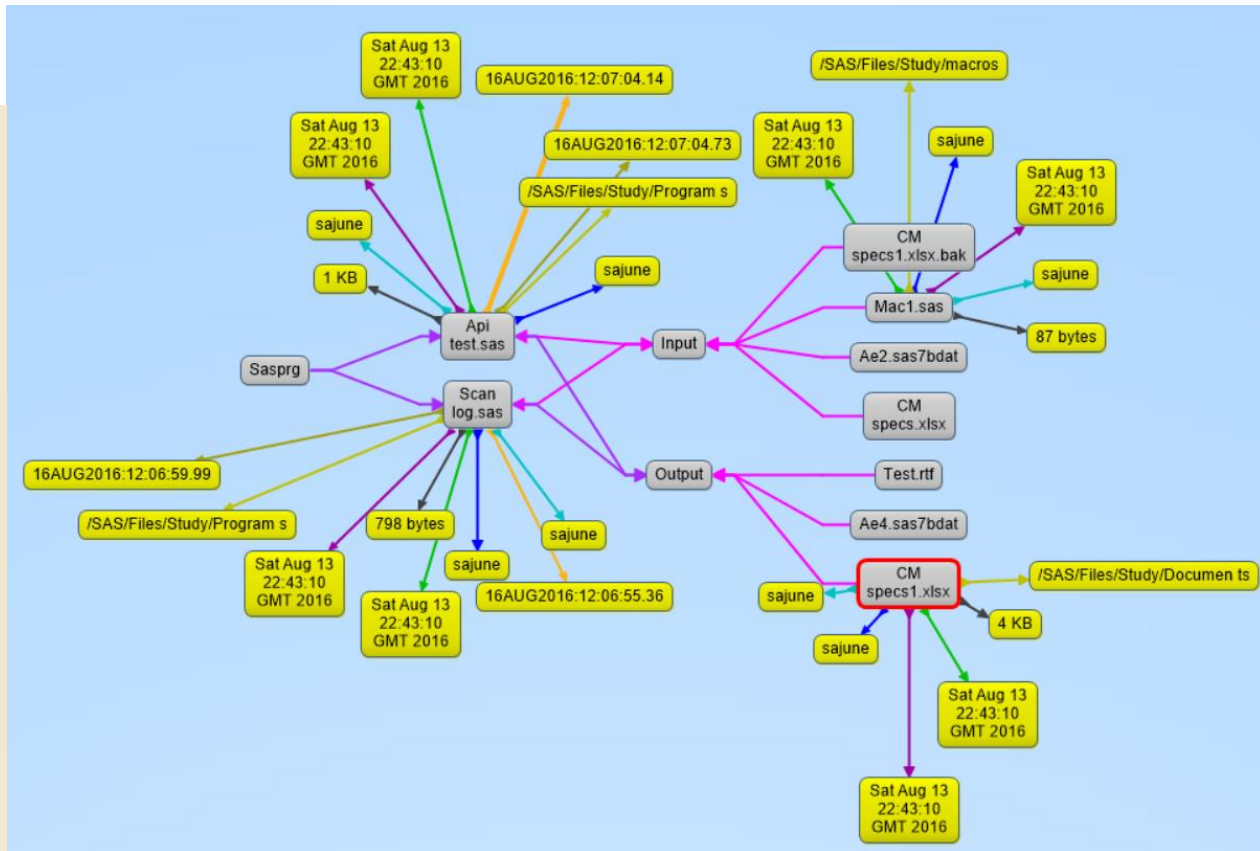
RDF GRAPHICAL IMPACT ANALYSIS

Using gruff-6.1.0-AG3.3-windows.64-ACL8.2 Tool



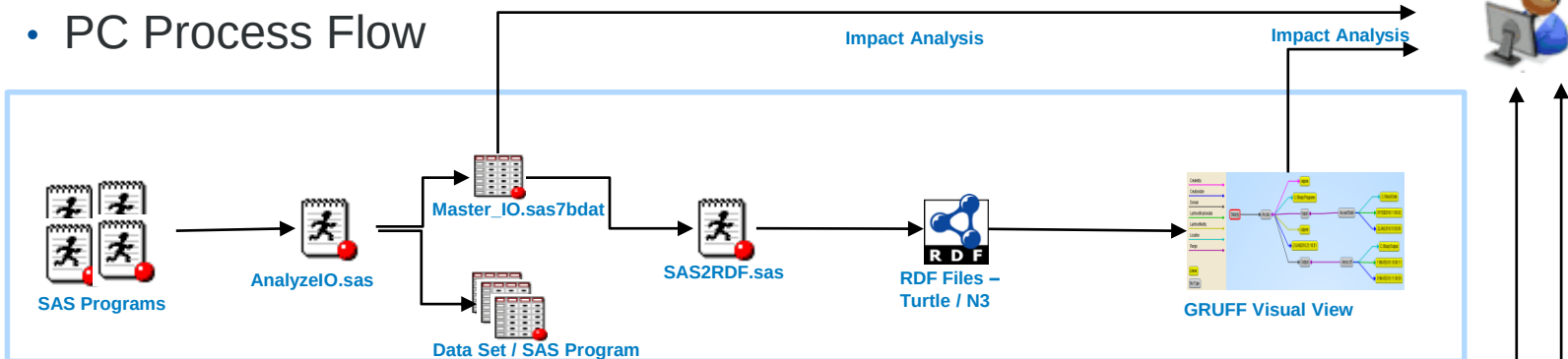
RDF GRAPHICAL IMPACT ANALYSIS

RDF Impact Analysis

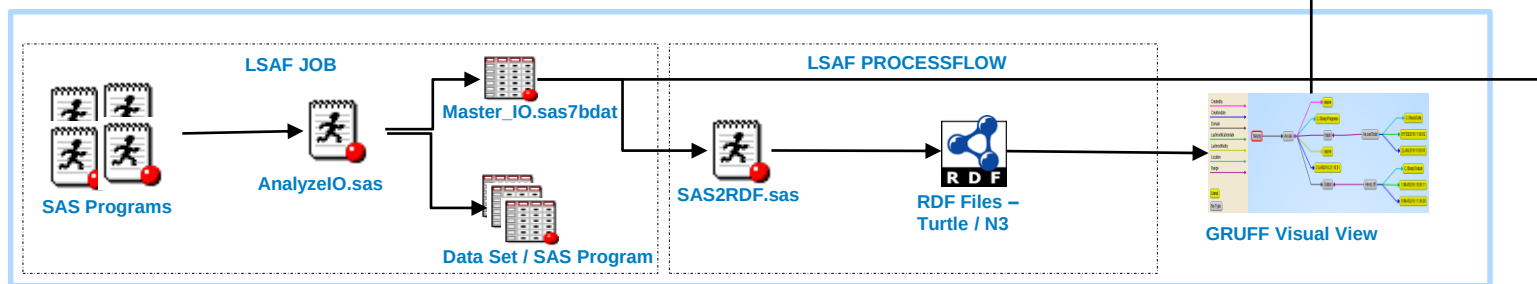


IMPACT ANALYSIS PROCESS FLOW

• PC Process Flow



• LSAF Process Flow



THANK YOU !



Sandeep.Juneja@sas.com
Ben.Bocchicchio@sas.com
Stijn.rogiers@sas.com