CMS data quality monitoring: systems and experiences

L Tuura[1], **A Meyer**[2,3], **I Segoni**[3], **G Della Ricca**[4,5]

[1] Northeastern University, Boston, MA, USA
[2] DESY, Hamburg, Germany
[3] CERN, Geneva, Switzerland
[4] INFN Sezione di Trieste, Trieste, Italy
[5] Universitá di Trieste, Trieste, Italy

E-mail: lat@cern.ch, andreas.meyer@cern.ch, ilaria.segoni@cern.ch, giuseppe.della-ricca@ts.infn.it

Abstract. In the last two years the CMS experiment has commissioned a full end to end data quality monitoring system in tandem with progress in the detector commissioning. We present the data quality monitoring and certification systems in place, from online data taking to delivering certified data sets for physics analyses, release validation and offline re-reconstruction activities at Tier-1s. We discuss the main results and lessons learnt so far in the commissioning and early detector operation. We outline our practical operations arrangements and the key technical implementation aspects.

1. Overview

Data quality monitoring (DQM) is critically important for the detector and operation efficiency, and for the reliable certification of the recorded data for physics analyses. The CMS experiment at CERN's Large Hadron Collider [1] has standardised on a single end-to-end DQM chain (Fig. 1). The system comprises:

- tools for the creation, filling, transport and archival of histogram and scalar monitor elements, with standardised algorithms for performing automated quality and validity tests on value distributions;
- monitoring systems live online for the detector, the trigger, and the DAQ hardware status and data throughput, for the offline reconstruction and for validating calibration results, software releases and simulated data;
- visualisation of the monitoring results;
- certification of datasets and subsets thereof for physics analyses;
- retrieval of DQM quantities from the conditions database;
- standardisation and integration of DQM components in CMS software releases;
- organisation and operation of the activities, including shifts and tutorials.

The high-level goal of the system is to discover and pin-point errors early, with sufficient accuracy and clarity to reach good detector and operation efficiency. Toward this end, standardised high-level views distill the body of quality information into summaries with significant explaining power. Operationally CMS partitions the DQM activities in online and offline to *data processing*, *visualisation*, *certification*, and *signoff*, as illustrated in Fig. 2 and described further in subsequent sections. The CMS DQM supports mostly automated processes, but use of the tools is also foreseen for the interactive and semi-automated data processing at the CAF analysis facility [3].

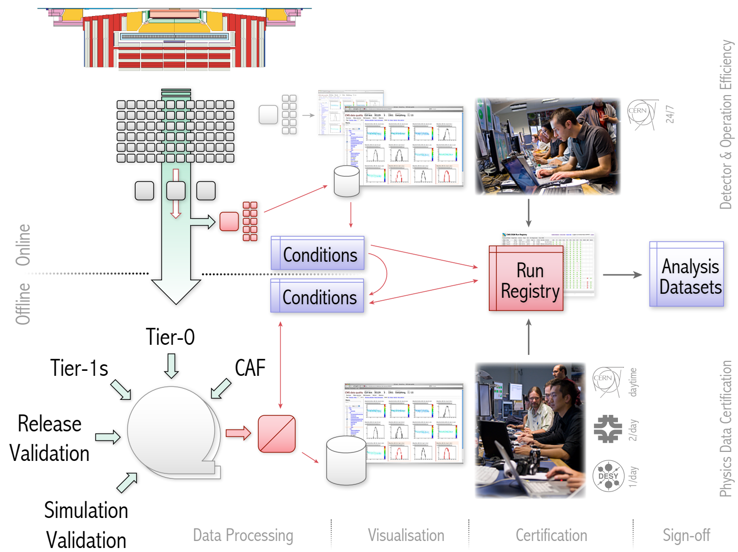
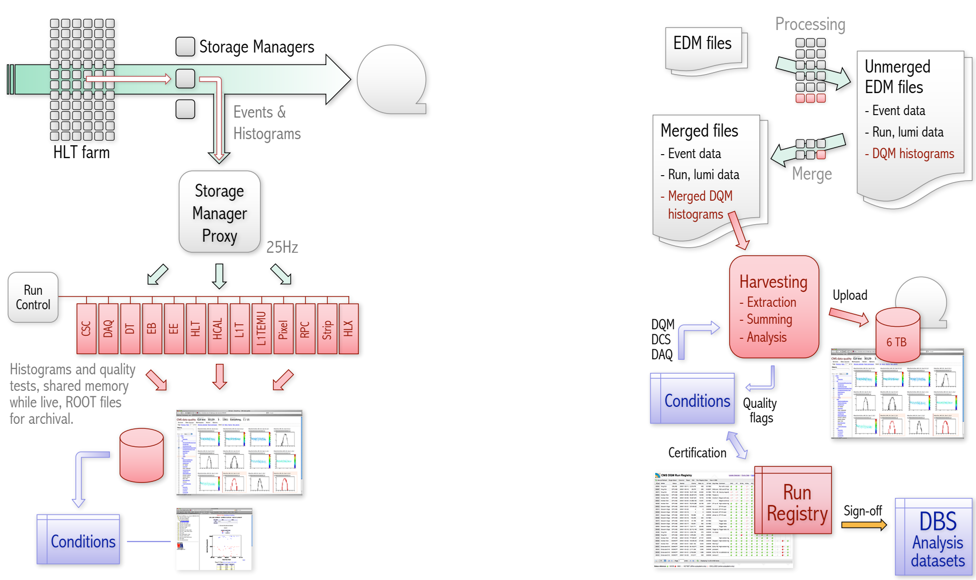


Figure 1. DQM system overview.



(a) Online DQM system.

(b) Offline DQM system.

Figure 2. DQM workflows.

2. Online DQM system

2.1. Data processing

As illustrated in Fig. 2(a), the DQM applications are operated in the online as an integral part of the rest of the event data processing at the cluster at CMS Point-5. DQM distributions are created at two different levels, *high-level trigger filter units* and *data quality monitoring applications*.

The high-level trigger filter units produce a limited number of histograms at up to 100 kHz. The histogram monitor elements are delivered from the filter units to the storage managers at the end of each luminosity section. Identical histograms across different filter units are summed together and sent to a storage manager proxy server, which saves the histograms to files and serves them to DQM consumer applications along with the events.

The data quality monitoring applications receive event data and trigger histograms from a DQM monitoring event stream from the storage manager proxy at the rate of about 10-15 Hz, usually one application per subsystem. Events are selected for the stream by applying DQM group specified trigger path selections. Each DQM application requests data specifying a subset of those paths as a further filter. There is no special special event sorting or handling, nor any guarantee to deliver different events to parallel DQM applications. The DQM stream provides only raw data products and only on explicit request additional high level trigger information.

Each application receives events from the storage manager proxy over HTTP and runs its choice of algorithms and analysis modules and generates its results in the form of *monitoring elements*, including meta data such as the run number and the time the last event was seen. The applications re-run reconstruction according to the monitoring needs. The monitor element output includes reference histograms and quality test results. The latter are defined using a generic standard quality testing module, and are configured via an XML file.

2.2. Visualisation

All the result monitor element data including alarm states based on quality test results is made available to a central DQM GUI for visualisation in real time [2], and stored to a ROOT file [4] from time to time during the run. At the end of the run the final archived results are uploaded to a large disk spool on the central GUI. There the files are merged to large ones and backed up to tape. The automatic certification summary from the online DQM step is extracted and uploaded to the run registry and on to the condition database (see section 4), where it can be analysed using another web-based monitoring tool, WBM [6]. Several months of recent DQM data is kept on disk availabe for archive web browsing.

2.3. Operation

Detector performance groups provide the application configurations to execute, with the choice of conditions, reference histograms and the quality test parameters to use and any code updates required. Reviewed configurations are deployed into a central replica *playback integration test system*, where they are first tested on recent data for about 24 hours. If no problems appear, the production configuration is upgraded. This practice allows CMS to maintain high quality standard at reasonable response time, free of artificial dead-lines.

The central DQM team invests significantly in three major areas: 1) to integrate and standardise DQM processes, in particular to define and enforce standard interfaces, naming conventions and look and feel of summary level information; 2) to organise shift activities, maintain sufficiently useful shift documentation, and train people taking shifts; and 3) to support and consult subsystem DQM responsibles and physicists using the DQM tools.

All the data processing components, including the storage manager proxy, the DQM applications and the event display, start and stop automatically under centralised CMS run control [5]. The DQM GUI and WBM web servers are long-lived server systems which operate independent of the run control. The file management on the DQM GUI server is increasingly automated.

3. Offline DQM systems

3.1. Data processing

As illustrated in Fig. 2(b), numerous offline workflows in CMS involve data quality monitoring: Tier-0 prompt reconstruction, re-reconstruction at the Tier-1s and the validation of the alignment and calibration results, the software releases and all the simulated data. These systems vary considerably in location, data content and timing, but as far as DQM is concerned, CMS has standardised on one standard two-step process for all these activities.

In the first step the histogram monitor elements are created and filled with event information, and stored as "run products" along with the processed events to the normal event data output files, called EDM files. When the CMS data processing systems merge output files together, the histograms are automatically summed together to form the first partial result.

In a second "harvesting" step, run at least once at the end of the data processing and sometimes periodically during the data processing, the histograms are extracted from the EDM files and summed together across the entire run to yield full event statistics on the entire dataset. The application also obtains detector control system (DCS, in particular high-voltage status) and data acquisition (DAQ) status information from the offline condition database, analyses these using detector-specific algorithms, and may create new histograms such as high-level detector or physics group summaries.

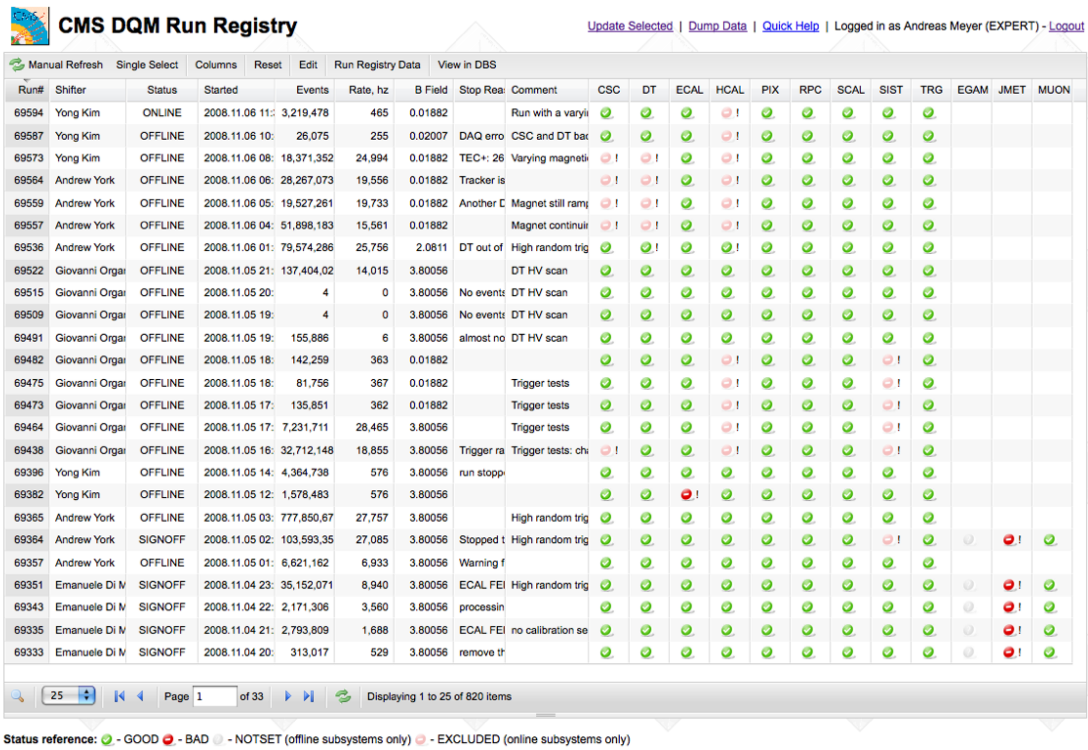
The final histograms are used to *calculate efficiencies* and *checked for quality*, in particular compared against reference distributions. The harvesting algorithms also compute the preliminary automatic *data certification decision*. The histograms, certification results and quality test results along with any alarms are output to a ROOT file, which is then uploaded to a central DQM GUI web server as on online.

The key differences between the various offline DQM processes are in content and timing. The Tier-0 and the Tier-1s re-determine the detector status on real data using full event statistics and full reconstruction, adding monitoring for physics objects. The Tier-0 does so at δt of a couple of a day or two whereas the Tier-1 re-processing takes from days to weeks. On CAF the alignment and calibration validation times vary from hours to days and the monitored quantities apply to these workflows; as the CAF activity is still very much seeking form, the DQM details are accordingly sketchy. The validation time range on simulated data is proportional to the sample production times, and can vary from anywhere from under a day for release validation to weeks on large monte carlo samples. The validation of simulated data differs from real data in that entire datasets are validated at once, rather than runs.

3.2. Visualisation

CMS provides one DQM GUI web server instance per offline activity, plus a test instance for DQM development. These servers are identical to the server used on online, and therefore provide a common look and feel for accessing all DQM data. From the user perspective, all the CMS DQM data is promptly available to the entire worldwide for inspection and analysis from one central location. While the GUI will offer all the capabilities needed for shift and expert use for all the DQM activities in its final form, it is nevertheless custom-built for the purpose of efficient interactive visualisation and navigation DQM results; it is not at all meant to be a completely general analysis tool.

As in the case of online, the DQM results from offline processing are uploaded to the central DQM GUI server with a large disk spool. There the result files are merged to larger size and backed up to the tape; recent data is kept cached on disk for several months. The automatic certification results from the harvesting, called "quality flags," are extracted and uploaded to



Run#	Shifter	Status	Started	Events	Rate, hz	B Field	Stop Rea:	Comment	CSC	DT	ECAL	HCAL	PIX	RPC	SCAL	SIST	TRG	EGAM	JMET	MUON
69594	Yong Kim	ONLINE	2008.11.06 11:	3,219,478	465	0.01882		Run with a varyi				!								
69587	Yong Kim	OFFLINE	2008.11.06 10:	26,075	255	0.02007	DAQ erro	CSC and DT bac				!								
69573	Yong Kim	OFFLINE	2008.11.06 08:	18,371,352	24,994	0.01882	TEC+: 26	Varying magneti	!	!		!								
69564	Andrew York	OFFLINE	2008.11.06 06:	28,267,073	19,556	0.01882	Tracker is		!	!		!								
69559	Andrew York	OFFLINE	2008.11.06 05:	19,527,261	19,733	0.01882	Another D	Magnet still ram	!	!		!								
69557	Andrew York	OFFLINE	2008.11.06 04:	51,898,183	15,561	0.01882		Magnet continui	!	!		!								
69536	Andrew York	OFFLINE	2008.11.06 01:	79,574,286	25,756	2.0811	DT out of	High random trig		!		!								
69522	Giovanni Orga	OFFLINE	2008.11.05 21:	137,404,02	14,015	3.80056		DT HV scan												
69515	Giovanni Orga	OFFLINE	2008.11.05 20:	4	0	3.80056	No events	DT HV scan												
69509	Giovanni Orga	OFFLINE	2008.11.05 19:	4	0	3.80056	No events	DT HV scan												
69491	Giovanni Orga	OFFLINE	2008.11.05 19:	155,886	6	3.80056	almost no	DT HV scan												
69482	Giovanni Orga	OFFLINE	2008.11.05 18:	142,259	363	0.01882						!				!				
69475	Giovanni Orga	OFFLINE	2008.11.05 18:	81,756	367	0.01882		Trigger tests				!				!				
69473	Giovanni Orga	OFFLINE	2008.11.05 17:	135,851	362	0.01882		Trigger tests				!				!				
69464	Giovanni Orga	OFFLINE	2008.11.05 17:	7,231,711	28,465	3.80056		Trigger tests				!				!				
69438	Giovanni Orga	OFFLINE	2008.11.05 16:	32,712,148	18,855	3.80056	Trigger ra	Trigger tests: ch	!			!				!				
69396	Yong Kim	OFFLINE	2008.11.05 14:	4,364,738	576	3.80056	run stopp													
69382	Yong Kim	OFFLINE	2008.11.05 12:	1,578,483	576	3.80056					!									
69365	Andrew York	OFFLINE	2008.11.05 03:	777,850,67	27,757	3.80056		High random trig												
69364	Andrew York	SIGNOFF	2008.11.05 02:	103,593,35	27,085	3.80056	Stopped t	High random trig								!			!	
69357	Andrew York	OFFLINE	2008.11.05 01:	6,621,162	6,933	3.80056	Warning f													
69351	Emanuele Di M	SIGNOFF	2008.11.04 23:	35,152,071	8,940	3.80056	ECAL FEI	High random trig											!	
69343	Emanuele Di M	SIGNOFF	2008.11.04 22:	2,171,306	3,560	3.80056	processin												!	
69335	Emanuele Di M	SIGNOFF	2008.11.04 21:	2,793,809	1,688	3.80056	ECAL FEI	no calibration se											!	
69333	Emanuele Di M	SIGNOFF	2008.11.04 20:	313,017	529	3.80056	remove th												!	

Figure 3. DQM run registry web interface.

the run registry. From there the values are propagated to the condition database and the dataset bookkeeping system DBS as described in section 4.

3.3. Operation

In all the offline processing the initial histogram production step is incorporated in the standard data processing workflows as an additional execution sequence. The harvesting step is currently technically implemented in a number of different ways for different activities. For Tier-0 the harvesting and upload to the GUI server are fully integrated into the processing system. Currently for other centralised data processing, as the processing progresses we send semi-manually custom analysis jobs with the CMS CRAB tool [7] to collect the results and upload the returned histogram file manually to the GUI server.

4. Certification and sign-off workflow

CMS uses a *run registry* as a central workflow tool to manage the creation of the final physics dataset certification result. The run registry is both a user interface managing the workflow (Fig. 3), and a persistent store of the information; technically speaking it is part of the online condition database and the web service is hosted as a part of the WBM system.

The certification process begins with the physicists on online and offline shift filling in the run registry with basic run information once the run control system has created the run, and adding any pertitent observations on the run during the shift. This information is then augmented with the automatic data certification results from the online, Tier-0 and Tier-1 data processing as described in the previous sections. This results in basic major detector segment level certification which accounts for DCS, DAQ and DQM online and offline metrics, and in future also potentially also power and cooling. For each detector segment and input one single boolean flag or a floating point value describes the final quality result. For the latter we define for each segment an appropriate threshold such that binary "good" or "bad" result is obtained. Were no number was calculated we assign the result "unknown."

Once the automatic certification results are known and uploaded to the run registry, the person on shift evaluates the detector and physics object quality following the shift instructions on histograms specifically tailored to catch relevant problems. This person adds any final observations to the run registry and where necessary adds a manual certification result which overrides the automatic result.

When the final combined quality result has been determined and communicated to the detector and physics object groups for confirmation, regular sign-off meetings collect the final verdict and deliver the agreed result to the experiment. At this stage the quality flags are copied to the offline condition database and to the dataset bookkeeping system (DBS) [8]. In the latter the quality flags are used to define convenience *analysis datasets*. The flags are also available for browsing and inspection via the CMS data discovery interface [9], a web interface to DBS. The quality flag values stored in the condition database are currently being made available as normal interval-of-validity based condition data for further use, e.g. in longer-term data quality evaluation and correlation with other entities such as temperature information.

For some time trivial trend plots of the key monitor element metrics have been generated automatically. We plan to extend this functionality to more comprehensive trend plotting of any selected histogram metric, and are working on common design for convenient access and presentation of trends over time.

5. Organisation and operation

On online the shifts take place mainly at the CMS Point-5 in Cessy, with remote assistance from the remote CMS centres [10]. Offline shifts are currently operated from the CMS centres at CERN Meyrin site, Fermilab and DESY. Standard shift instructions, as illustrated in Fig. 4, have been fully exercised. Perpetual effort to optimise histograms to maximise sensitivity to problems, to standardise the look and feel and to improve efficiency through better documentation.

6. Experience

CMS has commissioned a full end to end data quality monitoring system in tandem with the detector over the last two years. The online DQM system has now been in production for about a year and the full offline chain has just been commissioned. We have just recently completed the first full cycle of certification and sign-offs. DQM for the less structured alignment and calibration at CAF exists but a fair amount of work remains.

In our experience it takes approximately one year to commission a major component such as online DQM to production quality. Shift organisation, instructions, tutorials and supervision are major undertakings in their own right in particular. Significant amounts of effort are needed in various groups to develop the DQM algorithms, and centrally to standardise and integrate the workflows, procedures, code, systems and servers. While we find only modest amounts of code are needed for the DQM core systems, on the balance there is a perpetual effort to optimise histograms to maximise sensitivity to problems, to standardise the look and feel and to improve efficiency through better documentation, and against natural divergence in a collaboration as large as CMS toward sharing and using common code.

CMS has so far focused on commissioning a common first order DQM system through-out the entire experiment, with the aim of having an effective basic system ready for the first beam. We believe we have achieved this goal and will address second order features in due course.

CMS is very pleased with the DQM visualisation served using web technology and operating shifts from the CMS centres. Remote access to all the DQM information, especially offsite real-time live access to the online as been greatly beneficial and appreciated. Together the CMS

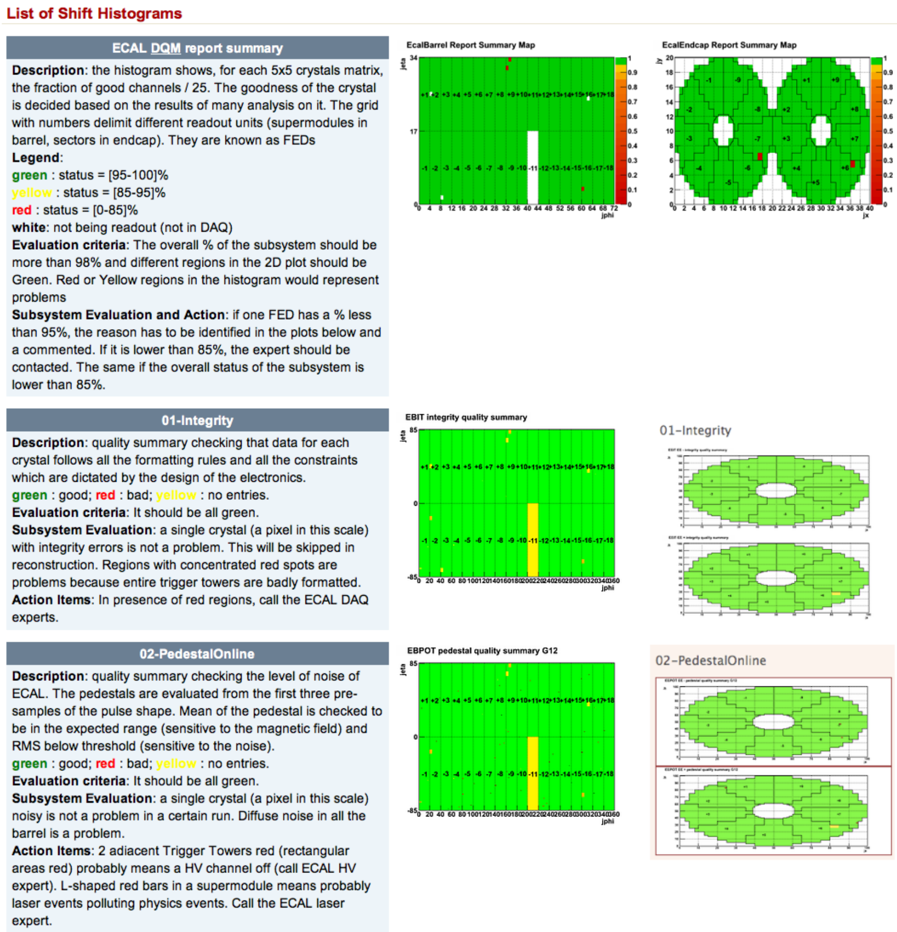


Figure 4. Example shift instructions.

centres and remote access have been essential and practical enabling factors to the experiment.

Acknowledgments

The authors thank the numerous members of CMS collaboration who have contributed the development and operation of the DQM system. Effective data quality monitoring is a truly collaborative effort involving a lot of people from several other projects: the trigger, detector subsystems, offline and physics software, production tools, operators, and so on.

References

[1] CMS Collaboration, 1994, *CERN/LHCC 94-38*, "Technical proposal" (Geneva, Switzerland)
[2] Tuura L, Eulisse G, Meyer A, 2009, *Proc. CHEP09, Computing in High Energy Physics*, "CMS data quality monitoring web service" (Prague, Czech Republic)
[3] FIXME: CAF.
[4] FIXME: ROOT.
[5] FIXME: Run control.
[6] FIXME: WBM.
[7] FIXME: CRAB.
[8] FIXME: DBS.
[9] FIXME: DBS discovery.
[10] FIXME: CMS centres.