

PREP: Production and Reprocessing management tool for CMS

F. Cossutti ^{a)}, P. Lenzi ^{b)}, N. Naziridis ^{b)}, D. Samyn ^{b)}, F. Stöckli ^{c)}
on behalf of the CMS Collaboration

^{a)} INFN Trieste, Italy ^{b)} CERN Geneva, Switzerland ^{c)} MIT Cambridge, USA

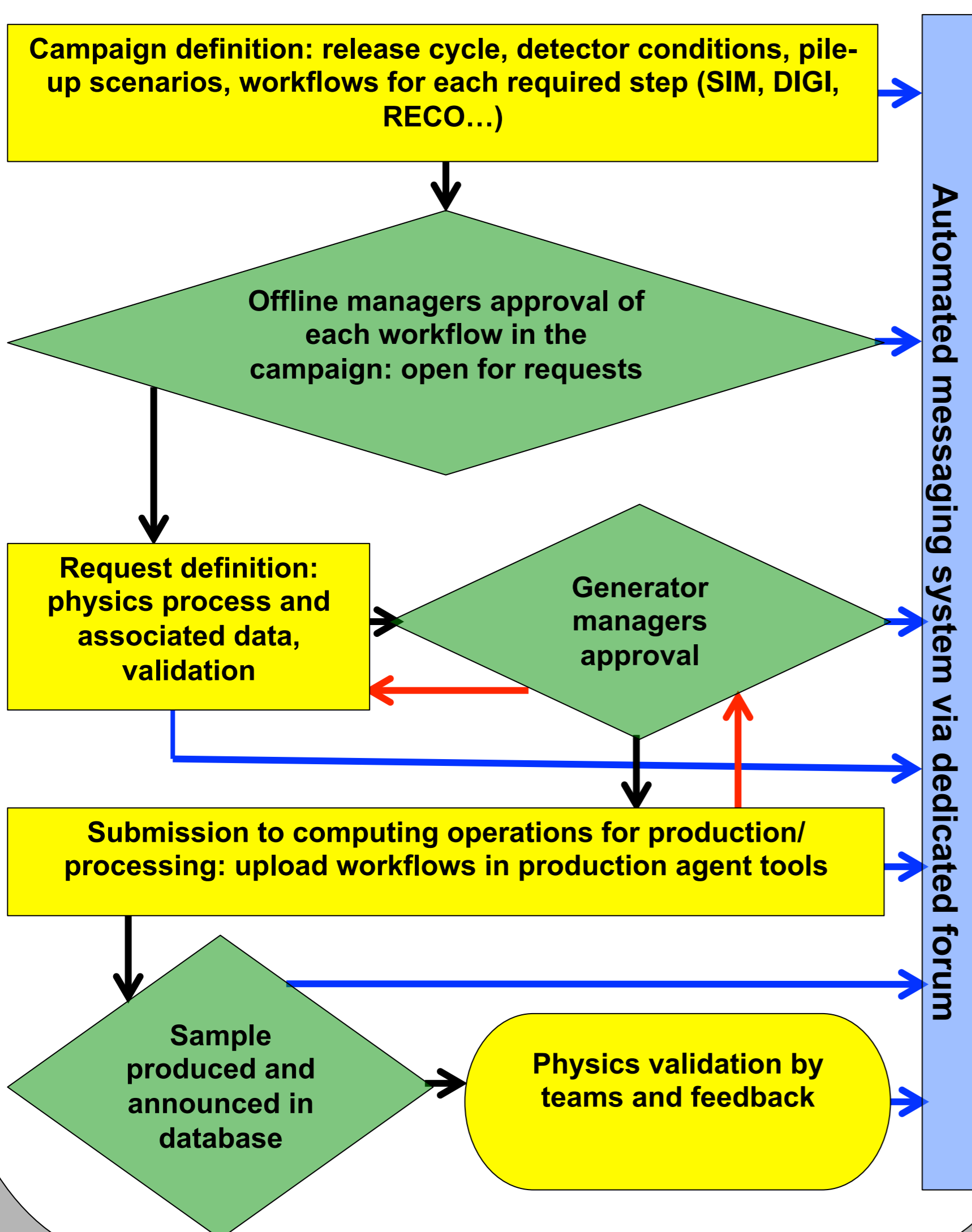


Simulation production and reprocessing for LHC

The needs of the simulation of physics samples for a general purpose LHC experiment like CMS are driven by the large number and complexity of the physics processes involved in pp collisions. This implies the generation of many different final states, with possible different phase space selections, and the usage of a variety of calculations and models based on different assumptions and levels of approximation in the physics description. Since the simulated events are fully reconstructed up to analysis level objects, re-processing of data using improved reconstruction algorithms, detector calibrations and alignment usually implies a similar re-processing of the corresponding simulated data. From the computational point of view all these activities translate into the need of managing several thousands of different workflows in order to satisfy the simulation needs related a single period of data taking. The necessity of effectively managing the flow of actions implied by each requested sample and of storing of the corresponding information in a user friendly format has suggested the development of a dedicated tool, PREP.

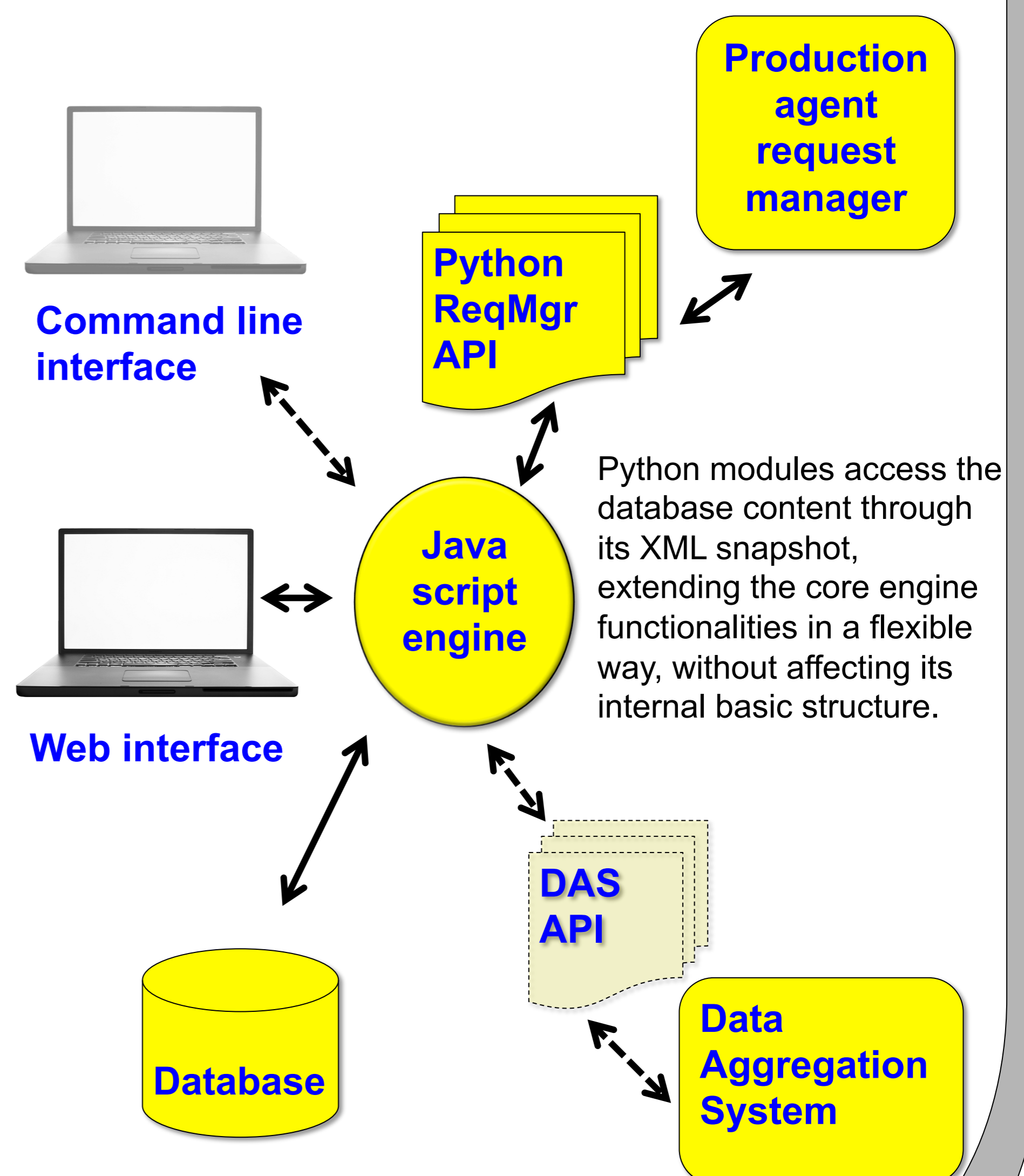
Logical flow of actions

Campaign: a collection of samples produced with the same release cycle and a homogeneous set of conditions. Simulation campaigns can be either production ones (starting from event generation and detector simulation) or re-processing ones (pile-up and electronics simulation and/or reconstruction).
Request: a single sample corresponding to a user defined process/generator/model/conditions belonging to a specified campaign.



Basic architecture

The tool is based on a Javascript core which allows information to be stored in a MySQL database through various interfaces (dashed lines indicate components under construction), and allows interactions with the CMS production environment through dedicated python APIs. The database content is exported using a standard XML format.



User interfaces: production management and data retrieval

Campaign Management view

Request Management view

Actions	WaitingFor	Code	Status	P	Approvals	Contact	CrossSection(ab)	FilterEff	MCDBId	DataSetName	SWRelease	N	Type	Link	EventContent	Pvt
cmsDriver [1]	Done	QCD-Summer12-00001	Done	30000	show 2 steps	Luisa Marchioni (ANTWERPEN)	7.294999E10	1.0	-1	MinBias_Tune4C_8TeV-pythia8	CMS5W_5_0_0_patch2	25000000	Prod	DAS	SIM	
show Resources (4 versions)	1	show Options (4 versions)	1	show Comments(11)												
cmsDriver [1]	Done	QCD-Summer12-00002	Done	100000	show 3 steps	Luisa Marchioni (ANTWERPEN)	7.270000E10	1.0	-1	MinBias_TuneZzstar_8TeV-pythia8	CMS5W_5_0_0_patch2	50000000	Prod	DAS	SIM	
show Resources (3 versions)	1	show Options (4 versions)	1	show Comments(12)												
cmsDriver [3]	Done	QCD-Summer12-00003	SUBMIT	40000	show 3 steps	Pierluigi Lenzi (CERN)	1.036E7	1.0	5475	QCD_HT-100To250_TuneZzstar_8TeV-madgraph-pythia	CMS5W_5_1_3	52000000	Prod	DAS	SIM	
show Resources (6 versions)	1	show Options (1 version)	1	show Comments(11)												

There are two main views accessible in PREP: **Campaign Management** and **Request Management**. The former provides production managers with a tool to define the characteristics of a given production campaign, in terms of release and condition versions to be used, workflow list to be run, and ensure all the concerned software coordinators can approve the chosen settings. At the same time it provides a reference for any user of that campaign about its general settings. The Request Management view provides the list of sample definitions, it is used to manage the whole life cycle of each request, from its submission by the physics groups till its completion and storage in the catalogue. At the same time it is the main source of information for analysts who want to know both the status of the request, and to retrieve all the physics parameters associated to it.