Load Balancing Techniques for a distributed WMS cluster based on WMSMonitor

**D. Cesini, D. Dongiovanni, E. Fattibene, T.Ferrari - INFN-CNAF, Bologna Italy - wms-support@cnaf.infn.it**

Overview

- ✓ **High throughput, scalability and high availability are critical requirements for a WMS/LB service**
- ✓ **A user transparent Load Balancing system of multiple instances allows to achieve those goals**
- ✓ **Given the complexity and variety of job submission patterns, load balancing must not only equally distribute jobs to all available WMS but also take into account each instance "health" status over time**
- ✓ **We propose a DNS based load balancing system which periodically exposes behind an alias the subset of WMS instances among those available considered functioning and less loaded**
- ✓ **The selection of proper instances for the subset is based on a service status metric and job submission test**
- ✓ **The load balancing system is implemented in WMSMonitor which collects information from WMS cluster, calculates the status metric and updates the alias defined in the DNS**
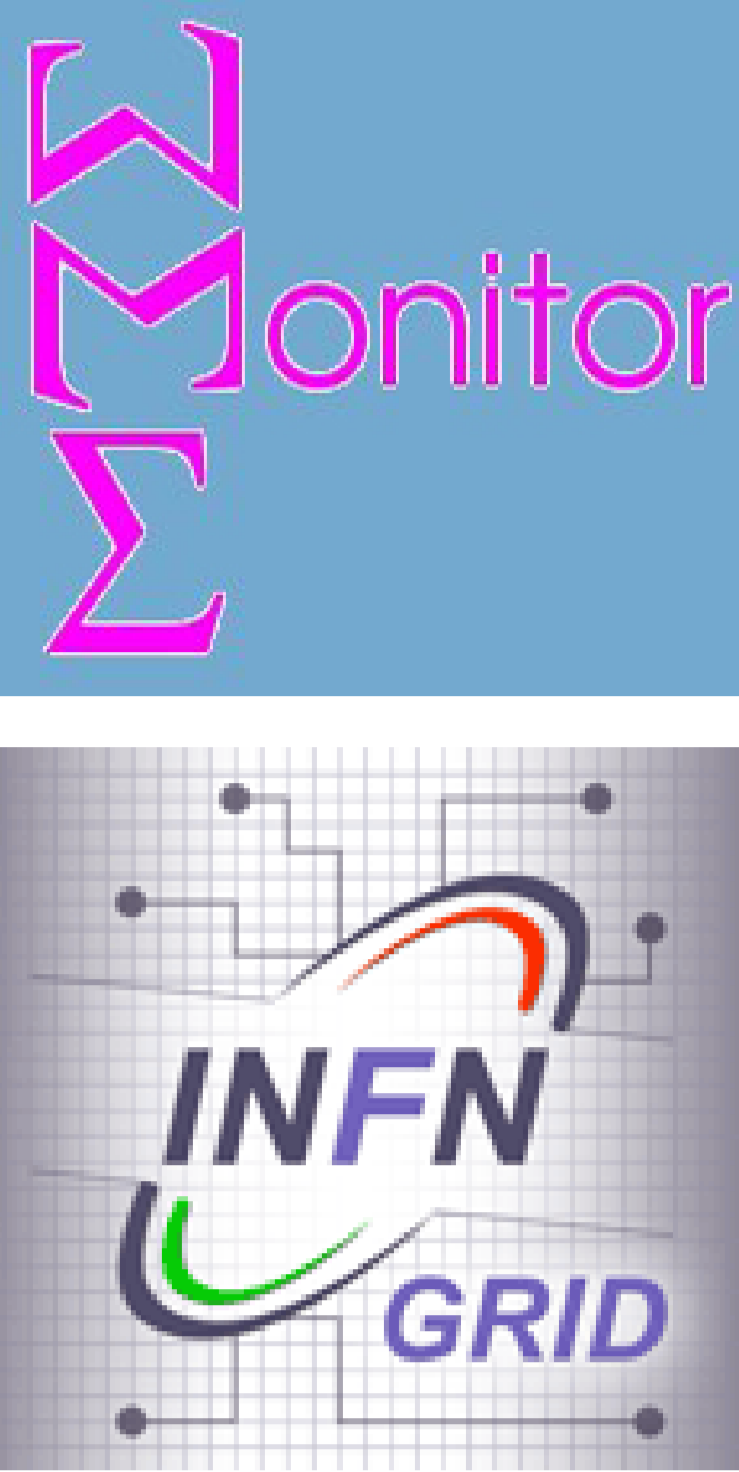
WMSMonitor Architecture & Functionalities

- ✓ **Monitors a pool of distributed gLite WMS/LB instances**
- ✓ **Publishes information through a user-friendly Web interface**
- ✓ **Detects failures affecting the services and supports administrators in fault prevention**
- ✓ **Untangles the complex WMS/LB internal dynamics**
- ✓ **Collects usage statistics aggregated per WMS and/or VO over configurable time intervals**
- ✓ **Displays Grid resource utilization statistics**
- ✓ **Computes activity statistics for each user**
- ✓ **Provides a overall status metric for each monitored WMS instance**
- ✓ **DB-Analyzer module periodically sends status notifications to the NAGIOS alarm system**
- ✓ **Arbiter module implements the DNS based load balancing system**

WMS view main page

WMS view detailed page

VO view main page

Resource / users pages
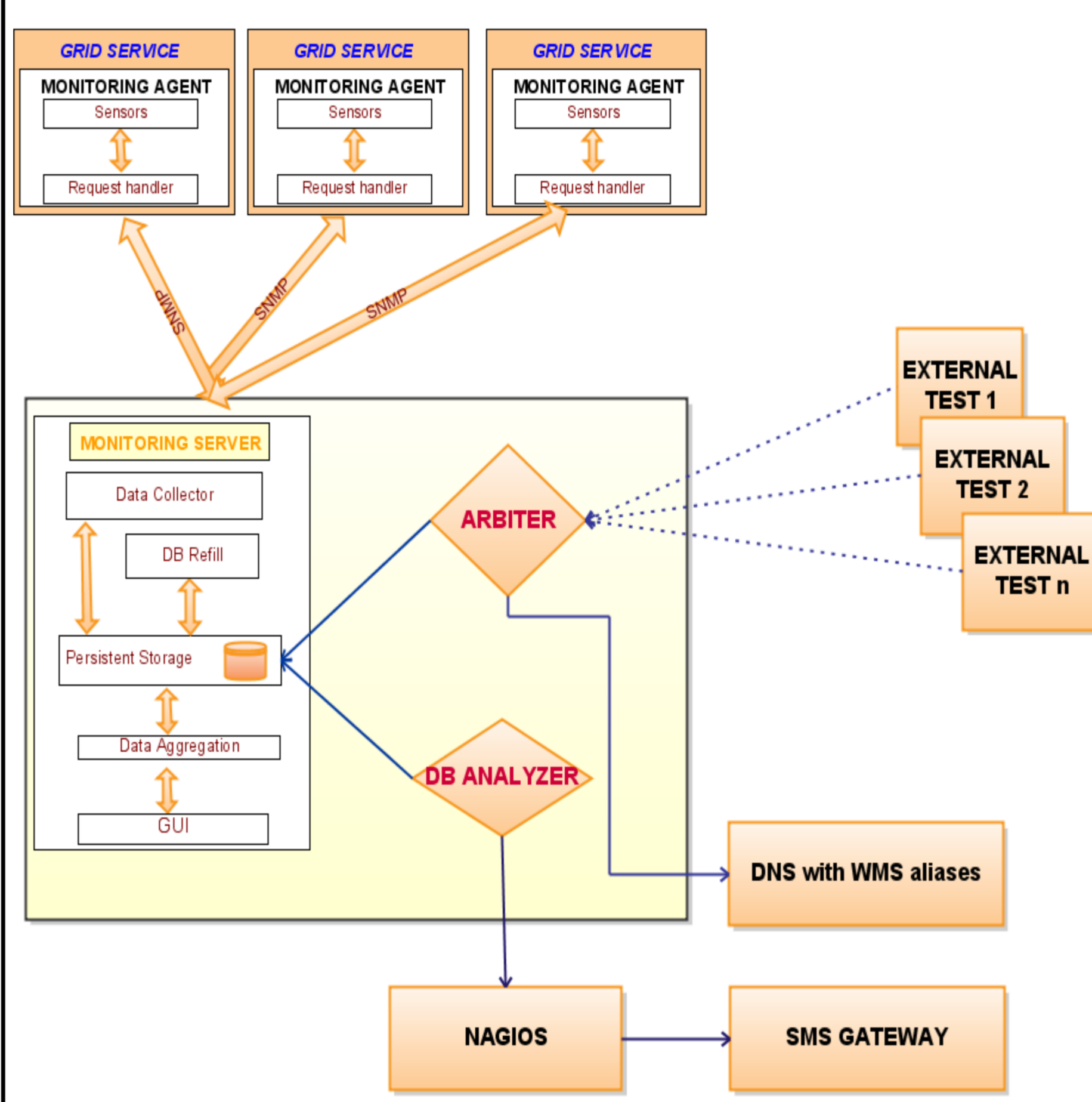
DNS based Load Balancing Method

PRE-REQUIREMENTS:

- • **A set of WMS alias defined in the DNS, for every VO or other suitable user grouping**
- • **A monitored set of WMS instances associated to each alias. The instances can be geographically distributed**

HOW THE ARBITER WORKS:

1. **Integrates WMS instances load-metric calculated by WMSMonitor with corresponding available external job-submission test**
2. **Selects for each alias the subset of functioning (i.e. daemons and submission test OK) and less loaded (i.e. lower metric value) associated WMS instances**
3. **Periodically updates the DNS alias accordingly (every 15 minutes)**

METRIC DESCRIPTION:

Given an instance i at time t:

$$\text{metric}_{i,t} = F(\text{daemons}, \text{drain}) * \{$$
$$L(\text{loadavg}, \text{loadavg limit}) +$$
$$Q_1(\text{WM input queue}, \text{Wm input queue limit}) +$$
$$Q_2(\text{JC input queue}, \text{JC input queue limit}) +$$
$$Q_3(\text{LB input queue}, \text{LB input queue limit}) +$$
$$M(\text{RAM usage}, \text{RAM usage limit}) +$$
$$D(\text{Disk space usage}, \text{Disk space usage limit}) \}$$

where:

$$F = \begin{cases} -1 : \text{if (logical OR)} \begin{cases} \text{daemons failure;} \\ \text{WMS in drain;} \\ \text{submission test failure} \end{cases} \\ 1 : \qquad \text{otherwise} \end{cases}$$

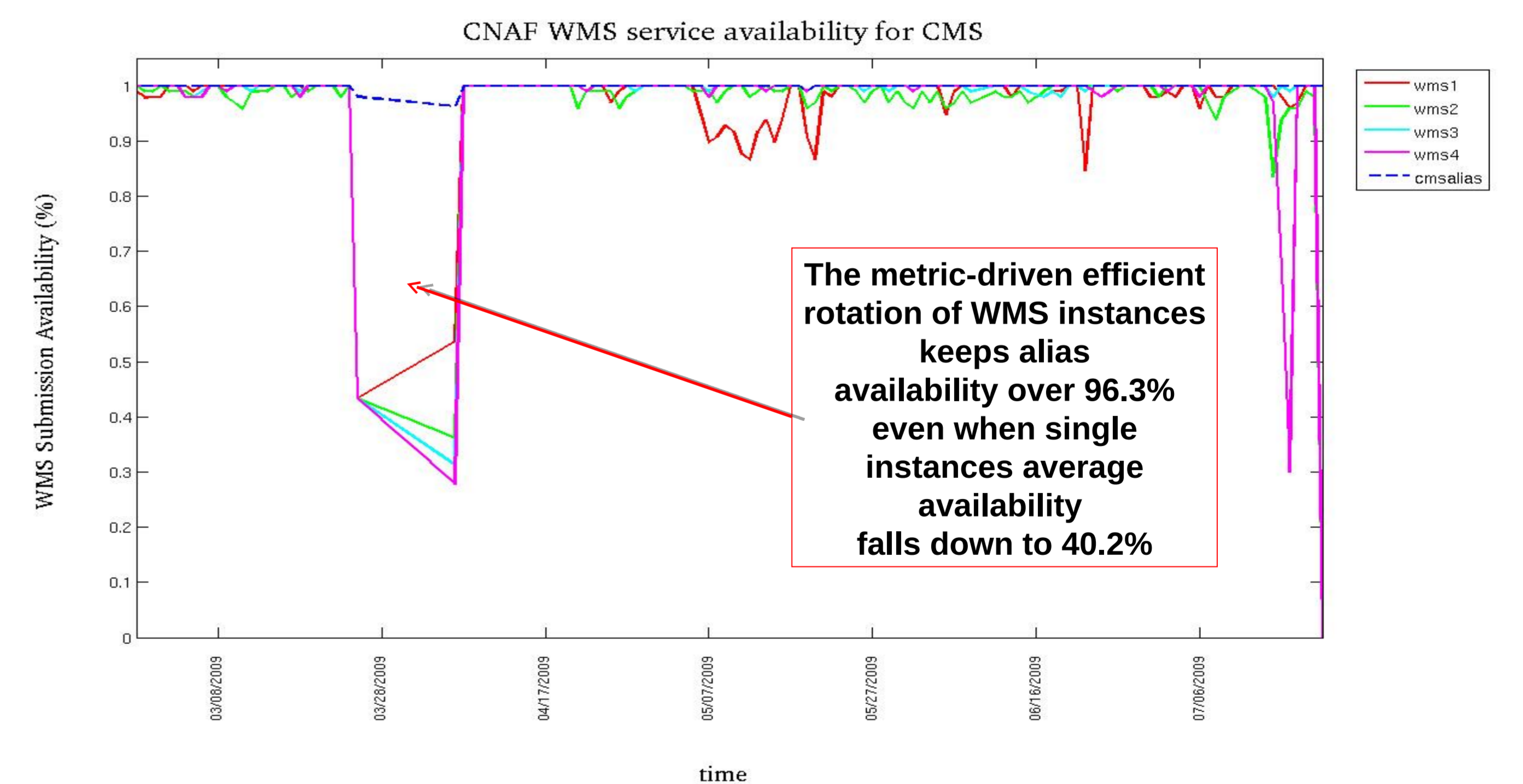$$L = \min\left\{1; \frac{\text{loadavg 15 minutes}}{\text{loadavg limit}}\right\}$$

$$Q_{1,2,3}, M, D(\text{var} \in N_0) = \min\left\{1; \frac{\text{var}}{\left|(\text{var\_limit}-1)(\text{var\_limit}-\text{var})\right|}\right\}$$
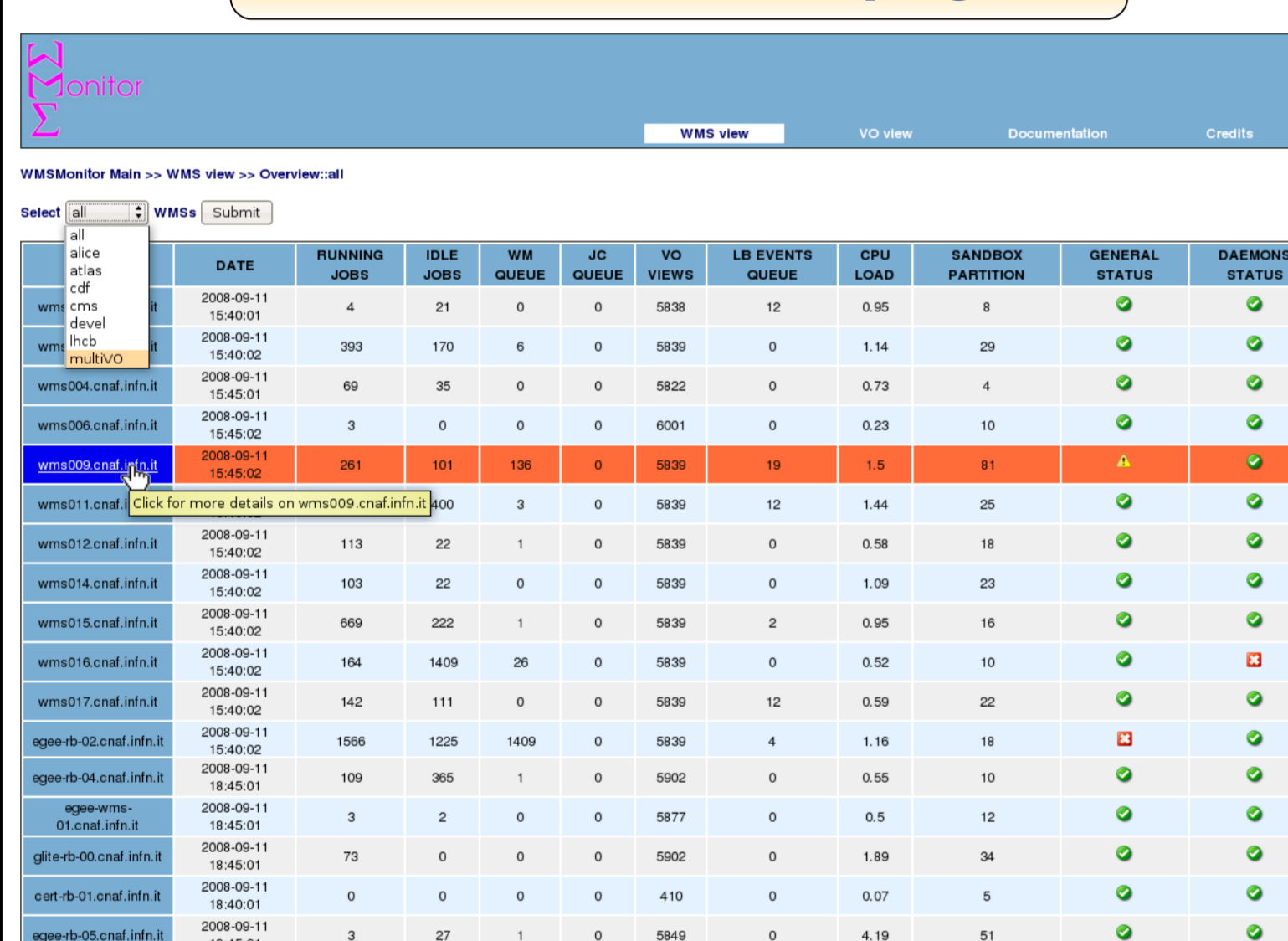
OBSERVATIONS:

- • Component F gives the functionality status of WMS instance: a negative value for F means WMS must be taken out of the alias
- • Other components summarize in a single positive monotonically increasing value the current instance usage status and load
- • While CPU load contributes linearly to the metric through function L, other variables like component queues, RAM and Disk memory usage, significantly contribute to the metric just when approaching the limit, i.e. only when becoming potentially dangerous for the service.

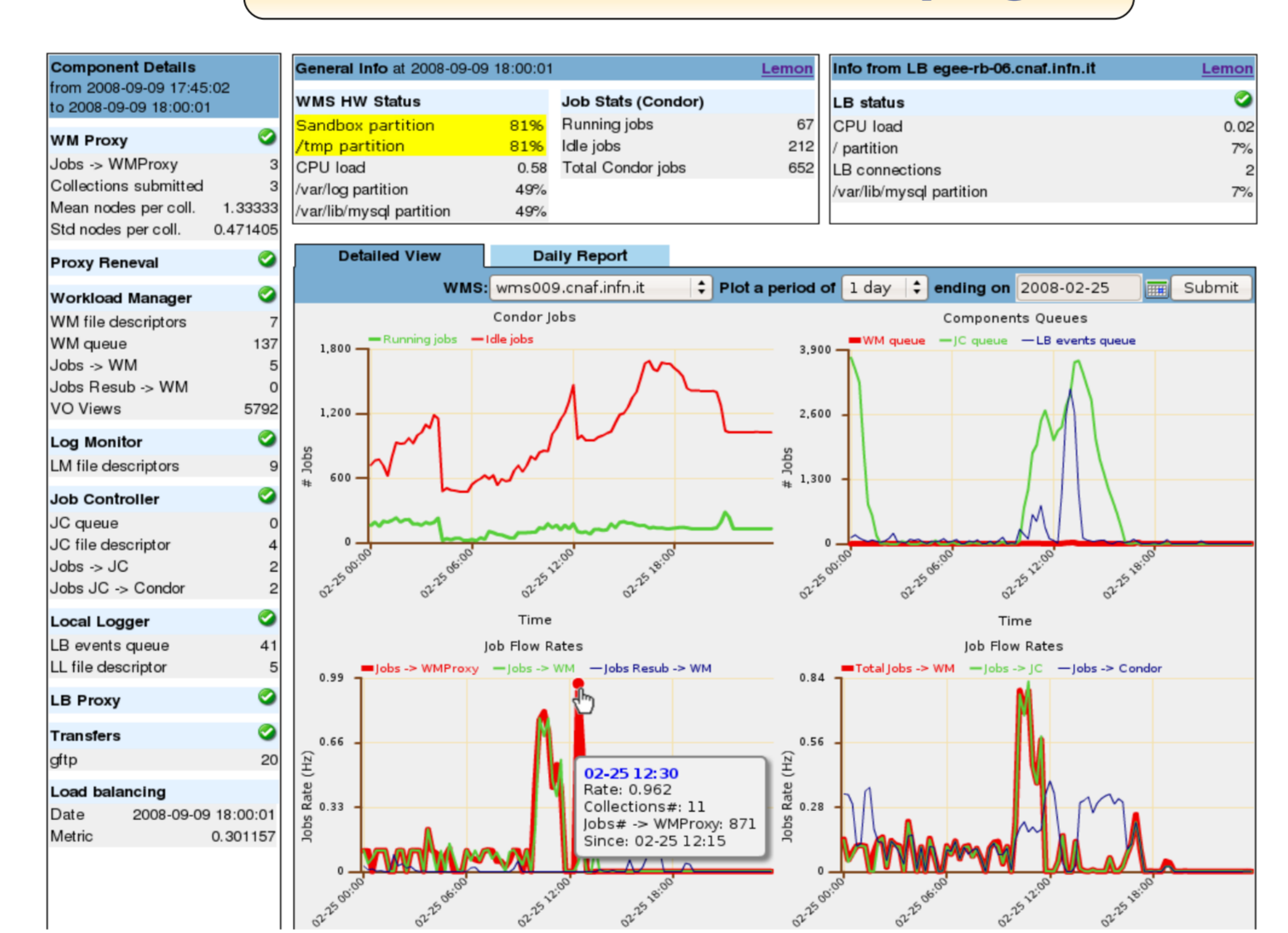## Improved WMS Service Availability

The metric-driven load balancing method, used for production WMS cluster at INFN-CNAF since early 2009, greatly improves service availability. Figure below shows WMS service daily availability for the CMS alias compared to that of each single instance. The alias reached a mean daily availability of 99.95% over the considered period of time (Feb-Jul 2009).
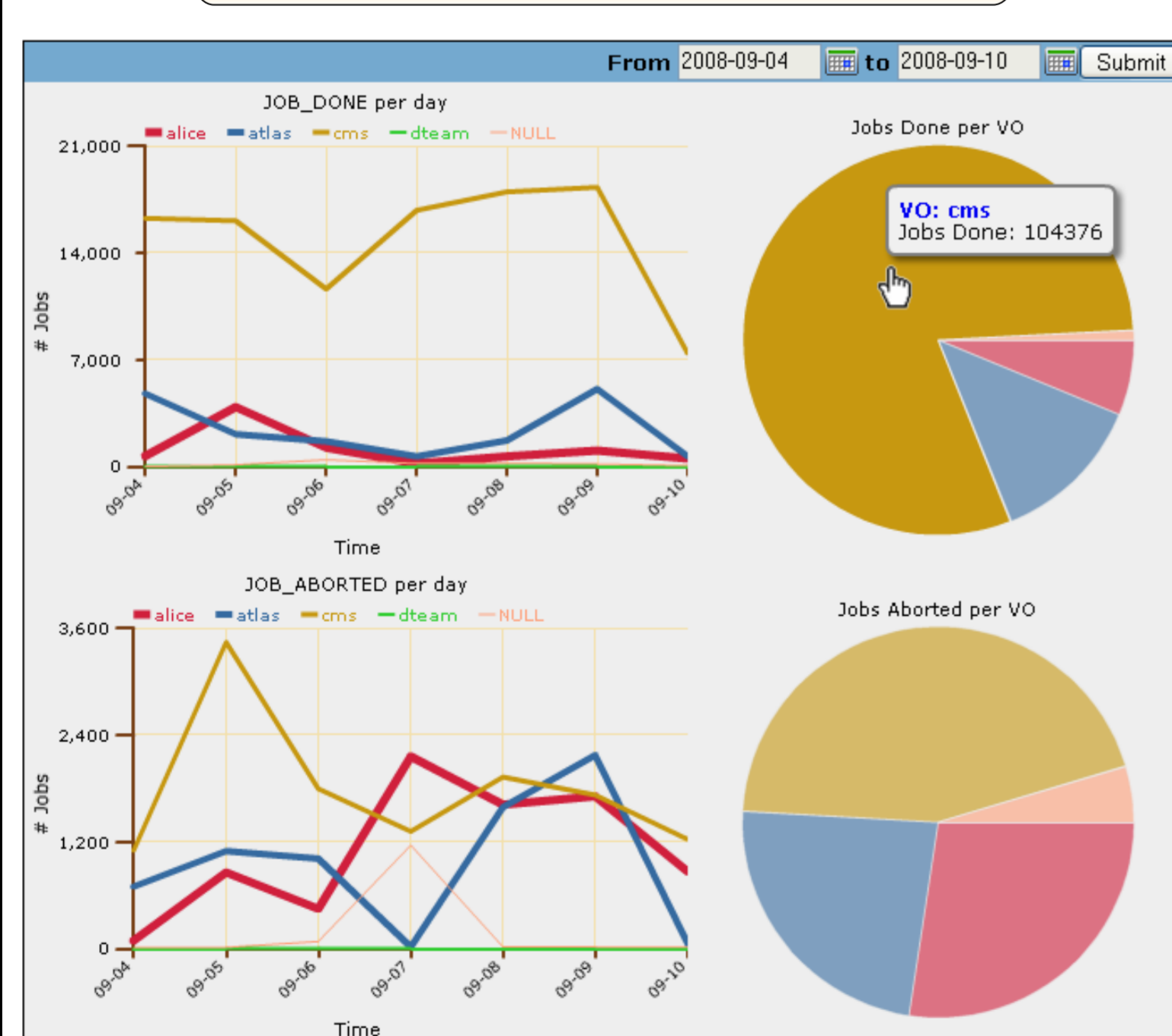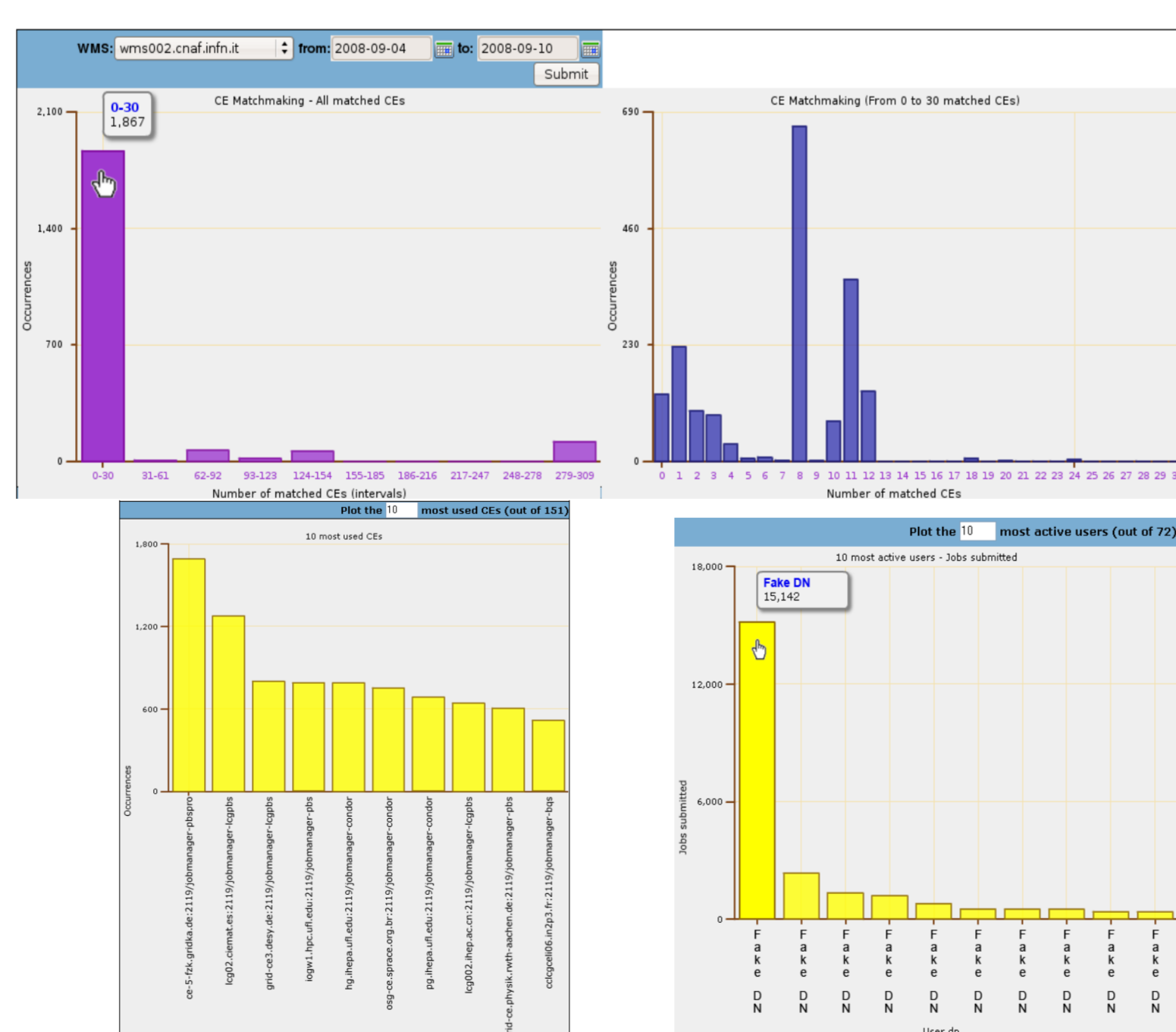

CNAF WMS service availability for CMS

The metric-driven efficient rotation of WMS instances keeps alias availability over 96.3% even when single instances average availability falls down to 40.2%

Conclusion & Future Work

- ✓ **The presented metric-driven load balancing system allows to efficiently distribute the job submission load maximizing availability of WMS cluster.**
- ✓ **It offers the possibility to implement an automatic geographically distributed WMS service failover**
- ✓ **Allows for transparent WMS instance maintenance without service interruption**

Working on:

- ➢ **Logging and Bookkeeping load balancing system**
- ➢ **Communication between different WMSMonitor collectors via ActiveMQ messaging system**
- ➢ **Monitor/arbiter/DNS redundancy**
- ➢ **Metric tuning**