FUTURE INTERNET TESTBEDS
EXPERIMENTATION BETWEEN
BRAZIL AND EUROPE

Grant Agreement No.: 288356 (FP7) and 590022/2011-3 (CNPq)

FIBRE-EU

Future Internet testbeds/experimentation between BRazil and Europe – EU

Instrument**: *Collaborative Project***
Thematic Priority: ***[ICT-2011.10.1 EU-Brazil] Research and Development cooperation, topic c) Future Internet – experimental facilities***

D 4.4 - Report on the Federation Software Tools Deployment

Author: WP4
Revised by: WP3
Due date of the Deliverable: Month 28
Actual submission date: 21/10/2013
Start date of project: June 1st 2011 Duration: 34 months
version: v.1.0

Project co-funded by the European Commission in the 7th Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

FP7 Grant Agreement No.	288356
Project Name	Future Internet testbeds/experimentation between BRazil and Europe – EU
Document Name	FIBRE-EU D4.4
Document Title	Report on the Federation Software Tools Deployment
Work package	WP4
Authors	Adriano Spinola (UNIFACS) Igor Leonardo Eloy Macedo (UNIFACS) Igor Luiz Oliveira Souza (UNIFACS) Joberto Sergio Barbosa Martins (UNIFACS) José Augusto Suruagy Monteiro (UFPE) Marcelo Machado de Pinheiro (UNIFACS) Paulo Ricardo Rios Leite (UNIFACS) Raphael Augusto Dourado (UFPE) Roberto Monno (NXW) Thiago Souza Hohlenweger (UNIFACS)
Editor	Joberto Sergio Barbosa Martins (UNIFACS) José Augusto Suruagy Monteiro (UFPE)
Reviewers	WP3
Delivery Date	21/10/2013
Version	V1.0

Abstract

This document is part of the result of the task T4.3 – Federation Software Tools of WP4 of the FIBRE project. The objective of this task is to report specifically on the definition of federation monitoring framework and diagnosis tools to be used on the facility.

TABLE OF CONTENTS

TABLE OF FIGURES

1 Acronyms

AM	Aggregate Manager
API	Application Programming Interface
BWCTL	Bandwidth Test Controller
CMF	Control and Monitoring Framework
CNPq	Brazil's Council for Scientific and Technological Development
CPqD	Telecommunications Research and Development Centre
DNS	Domain Name System
ETOMIC	European Traffic Observatory Measurement InfrastruCture
EU	European Union
FI	Future Internet
FIBRE	Future Internet testbeds / experimentation between Brazil and Europe
FIBRE-BR	FIBRE's Brazilian Facility
FIBRE-EU	FIBRE's European Facility
FP7	Seventh Framework Programme
GÉANT	Pan-European Research and Education Network
GEMINI	A GENI Measurement and Instrumentation Infrastructure
GENI	Global Environment for Network Innovations
GPL	GNU General Public License
GMPLS	Generalized MultiProtocol Label Switching
GUI	Graphical User Interface
GW	Gateway
gLS	Global Lookup Service
hLS	Home Lookup Service
i2CAT	Fundació i2CAT
ICMP	Internet Control Message Protocol
ICT	Information and Communication Technologies
IP	Internet Protocol
iRODS	Integrate Rule Oriented Data System
I&M	Instrumentation and Measurement
JSON	JavaScript Object Notation
LAMP	Leveraging and Abstracting Measurements with perfSONAR
LDAP	Lightweight Directory Access Protocol
LS	Lookup Service
MA	Measurement Archive
MDIP	Measurement Data Integration Point
MI	Measurement Information
MS	Milestone
NITOS	Network Implementation Testbed using Open Source platforms
NMWG	OGF's Network Measurements Working Group

NOC	Network Operations Center
NTP	Network Time Protocol
NXW	Nextworks
OCF	OFELIA Control Framework
OF	OpenFlow
OFELIA	OpenFlow in Europe: Linking Infrastructure and Applications
OGF	Open Grid Forum
OMF	cOntrol, Management and Measurement Framework
OML	ORBIT Measurement Library
ORBIT	Open-Access Research Testbed for Next-Generation Wireless Networks
OS	Operation System
OWAMP	One-way Active Measurement Protocol
perfSONAR	PERFormance Service Oriented Network monitoring ARchitecture
ProtoGENI	Control Framework for GENI Cluster C
pSPT	perfSONAR Performance Toolkit
RFC	Request for Comments
RNP	Brazilian National Research and Education Network
RPC	Remote Procedure Call
RRD	Round Robin Database
RSpec	Resource Specification
SDN	Software Defined Network
SFA	Slice-based Facility Architecture
SNMP	Simple Network Management Protocol
SQL	Structured Query Language
SSH	Secure Shell
TDMI	TopHat Dedicated Measurement Infrastructure
UFF	Universidade Federal Fluminense
UFPE	Universidade Federal de Pernambuco
UFRJ	Universidade Federal do Rio de Janeiro
UNIFACS	Universidade Salvador
UPMC	Université Pierre et Marie Curie
USP	Universidade de São Paulo
VM	Virtual Machine
VT-AM	Virtual Technology Aggregate Manager
XML	eXtensible Markup Language
XMLRPC	XML Remote Procedure Call
WebUI	Web User Interface
WDM	Wavelength Division Multiplexing
WMI	Windows Management Instrumentation
WP	Work Package
WP2	Building and Operating the Brazilian Facility
WP4	Federation of Facilities
WP5	Development of technology pilots and showcases

2 Scope

This document is part of the result of the task T4.3 – Federation Software Tools of WP4 of the FIBRE project. The objective of this task is to report specifically on the definition of federation monitoring framework and diagnosis tools to be used on the facility.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [Bradner 1997].

3 Reference Documents

Abelém, A., Stanton, M., Machado, I., Salvador, M., Magalhaes, L, Fernandes, N., Correa, S., Cardoso, K., Marcondes, C., Martins, J., Monteiro, J., Carvalho, T., Rezende, J. (2013) "FIT@BR - a Future Internet Testbed in Brazil", In Proceedings of the APAN – Network Research Workshop.

Fernandes, N. C., Silva, E., Muchaluat-Saade, D., Magalhaes, L. (2013) "Identity Management in Brazilian Testbeds for Future Internet (in Portuguese)", In IV Workshop on Experimental Research on the Future Internet (WPEIF), Brasilia, Distrito Federal.

FIBRE. (2013). Future Internet Testbeds Experimentation Between Brasil and Europe [Online]. Available: http://www.fibre-ict.eu.

Hanemann, A., Boote, J. W., Boyd, E. L., Durand, J., Kudarimoti, L., Lapacz, R., Swany, D. M., Zurawski, J., Trocha, S., (2005) "PerfSONAR: A Service Oriented Architecture for Multi-Domain Network Monitoring", In "Proceedings of the Third International Conference on Service Oriented Computing", Springer Verlag, LNCS 3826, pp. 241-254, ACM SIGSOFT and SIGWEB, Amsterdam, The Netherlands.

iDrop. (2013). Available: https://code.renci.org/gf/project/irodsidrop/.

iRODS. (2013, Feb. 27). iRods: Data Grids, Digital Libraries, Persistent Archives and Real Time Data Systems [Online]. Available: https://www.irods.org/.

Matthew, P.. (2012, Jul. 31). Google SDN based inter-datacenter WAN case study [Online]. Available: http://www.sdncentral.com/sdn-blog/just-released-google-sdn-case-study/2012/07/

Monteiro, J. A. S. (2010), "Measurement Infrastructures for Future Internet Testbeds". First Workshop of the Brazilian Institute for Web Science Research. Rio de Janeiro.

Pinheiro, M. M., Macêdo, I. L. E., Souza, I. L. O., Hohlenweger, T .S., Leite, P. R. R., Spínola, A. L., Monteiro, H., Dourado, R. A., Sampaio, L. N., Monteiro, J. A. S, and Martins, J. S. B. (2012) "An Instrumentation and Measurement Architecture Supporting Multiple Control Monitoring Frameworks," In III Workshop on Experimental Research on the Future Internet (WPEIF), Ouro Preto, Minas Gerais.

RNP. (2013). Authentication infrastructure and federated authorization [Online]. Available: http://portal.rnp.br/web/servicos/cafe-en-how-to-cafe-works

Schroeder, W.. (2008, Nov. 19). Zone [Online]. Available: https://www.irods.org/index.php?title=Zone&oldid=3114.

Schroeder, W.. (2013, Apr. 16). PAM/LDAP Authentication / Authorization [Online]. Available: https://www.irods.org/index.php?title=PAM/LDAP_Authentication/Authorization&oldid=6840

Tooby, P.. (2009, Mar. 19-20). Overview of Data Management and IRODS [Online]. Available: https://www.irods.org/pubs/iRODS_Overview_0903.pdf.

Tooby, P.. (2010, Jun. 9). iRODS Architecture [Online]. Available: https://www.irods.org/index.php?title=iRODS_Architecture&oldid=4915.

White, J. (2010). A Tutorial Introduction to OML. GEC9. Available at http://omf.mytestbed.net/projects/oml.

4 Introduction – Measurement and Monitoring Tools

Providing the right information to the experimenter is key. In addition, granting access in an intelligent manner to the various tools that can be developed by the community is fundamental to ease the effort of the experimenter. Proper resource instrumentation is thus of utmost importance for experimenters to make sense of their results: do the results fit with the expectations? If not, what are the factors affecting them? This represents an advance towards a more scientific approach to experimental research (benchmarking is a current hot area of research investigating such issues).

We believe that the availability of such information as part of a testbed monitoring service has many advantages. As it relieves the users from having to perform such instrumentation by themselves, they can focus on developing the core of their experiments. Also, instrumentation is sometimes challenging, and a monitoring service will typically be realized by more experienced people, that will allow the user to benefit from best-of-breed tools. Finally, a monitoring service allows to easily make the information available to all users, since measurements might sometimes be costly to produce in terms of resources (network topology for example), or require special privileges (slice information for example).

4.1 Measurements in the experimental lifecycle

A close integration of measurements with the experimental lifecycle should help users to get the best out of the testbed resources.

4.1.1 Setting up an experiment: resource selection and booking

We can imagine that users willing to run an experiment will be directed to some sort of portal that will allow them to authenticate, discover the set of available resources through testbeds federation, and pick up and reserve some of them. At this step, monitoring information complementing the description of testbed resources will help a user choose the most appropriate resources for his/her experiment. The scope of this information can be very large, depending on the experimenters' needs, and will typically go beyond the scope of a single instrumentation platform.

As detailed, testbed information is abstracted into RSpecs (Resource Specifications). One possible way of dealing with measurements at this stage, would be to annotate those files by adding additional fields for each resource. Experiments RSpecs are critical to monitoring services due to the relevant information it contains about the reserved resources of a certain experiment.

4.1.2 Live infrastructure measurements

While an experiment is running, the interaction with measurements will occur through a set of queries (get measurements, etc.) or callbacks (periodic measurements, alerts when a change occurs, etc.). They will be used to inform the experiment about potential changes in the overlay, and trigger a response in the experimental plane (for logging purposes, reoptimizing the overlay, or, for instance, finding the closest node to an external user) or in the control plane (request new nodes to replace a failing one, etc.).

Here, the challenge will consist in collecting a large panel of information, both as a background task, or upon request of the user, and possibly in (near) real time. The ability for a user to plug its own measurements into the framework will be a plus.

Since monitoring will also consume resources – especially for real time – we might also want to consider monitoring data or infrastructure as a resource a user can reserve. For example, a node or a set of nodes in a slice could be associated to the corresponding measurement stream providing the user with periodic measurements or alerts when some events occur.

4.1.3 Historic measurements

It will also be convenient to access the archive of all measurements, performed during the timeframe of an experiment (both as a background task, or because it was requested by a user). This will be useful for post-processing the results, and understanding potential issues that may have affected the performance.

4.2 Review of existing monitoring infrastructures and tools

We review currently available monitoring infrastructures and tools useful for instrumenting testbeds and ensuring their interconnection. The list is not exhaustive but will highlight those that are for use in FIBRE.

4.2.1 Data integration tools

perfSONAR [http://www.perfsonar.net]

perfSONAR is an infrastructure for network performance monitoring, making it easier to solve end-to-end performance problems on paths crossing several networks. It contains a set of services delivering performance measurements in a federated environment. These services act as an intermediate layer, between the performance measurement tools and the diagnostic or visualization applications. This layer is aimed at making and exchanging performance measurements between networks, using well-defined protocols.

perfSONAR is considered one of the components of FIBRE´s monitoring and monitoring federation solution as discussed next.

TopHat and Manifold [http://www.top-hat.info]

TopHat is particular in the sense that it relies on its own measurements adapted to the PlanetLab[1] overlay TDMI (which will be presented in the next subsection), whose data is complemented by drawing upon excellent, proven third party services, notably the Dimes and Etomic measurement infrastructures, for specialized measurements thanks to the Manifold

[1] https://www.planet-lab.org/

framework. All data is served to the user through a unique consistent interface and data formats, which allow abstracting the set of tools used, or the origin of the data.

While the first integration was implemented by a set of ad hoc gateways, there is a current move towards a more robust integration framework, which aims at leveraging the added value of combining measurements. That means we can answer user queries by a combination of two or more platforms, which were not previously available through individual platforms.

This step is somehow related to the efforts in federating testbeds, and should allow for an ecosystem of measurement systems to emerge, and the various actors to seamlessly exchange information. This corresponds to the vision of a distributed federation of systems (as opposed to a single central entity), where the data can be accessed through different entry points (with their own strength and specificity) and still allowing access to the full range of available information.

TopHat currently provides access to a set of interconnected system of different types: (a) measurement systems (TMDI, Gulliver, Dimes, Etomic, SONoMA), (b) system-level monitoring infrastructure and testbed resources (CoMon, CoTOP, MyPLC, Monitor, SFA), (c) misc. information sources for topological and geographical data (MaxMind Geolite City, Team Cymru IP to ASN mapping, Georgia Tech AS characterization dataset). While new systems can be plugged into the existing infrastructure, it is also possible to deploy it independently (source code is available under the GPLv3 license[2]). In addition, UNIFACS and UPMC are developing within the FIBRE context a perfSONAR gateway allowing exposing perfSONAR data at the TopHat and MySlice level, thus allowing its exposition on the portal. This plugin allows for example the user to perform resource selection based on measured characteristics of the paths between a set of resources (one-way delay, bandwidth, loss rate).

TopHat is considered one of the components of FIBRE´s monitoring and monitoring federation solution as discussed next.

[2] Backend:
http://git.onelab.eu/?p=manifold.git;a=summary

Frontend:
http://git.onelab.eu/?p=myslice.git;a=summary

Development website:
http://trac.myslice.info/wiki/Manifold

Installation:
http://trac.myslice.info/wiki/Manifold/Install

4.2.2 Monitoring infrastructures and supporting tools

TDMI [http://www.top-hat.info]

The TopHat Dedicated Measurement Infrastructure (TDMI) is TopHat's own measurement infrastructure. It consists of modular probing agents that can be deployed on resources, and probe a set of destination (pre-determined and on-demand) in a distributed efficient manner. This agent wraps a set of tools such as Paris Traceroute to remove the artefacts arising from the presence of load balancers in the Internet. TDMI aims at providing the necessary information to users about the evolution of the overlay, and focuses on catching the dynamic aspects of the topology.

TDMI provides a set of reusable components such as (1) a common query interface based on Manifold, (2) the current support for new measurement tools, (3) a modular measurement agent and infrastructure that can be adapted and deployed for other needs, (4) the specific modules used in its original PlanetLab context (XMLRPC interface, scheduler, measurement campaigns, etc.), as well as (5) server-side support (all source is free software under the GPL license).

OML [http://oml.mytestbed.net/projects/oml]

OML is a measurement library that allows application writers to define customizable measurement points inside applications. Experimenters running the applications can then direct the measurement streams from these measurement points to storage in a remote measurement database (OML Server). OML was originally conceived to provide measurement facilities within the OMF Testbed Management Framework, but can also be run independently of OMF. Both OMF and OML are available under free software licenses, and can be adapted to fit general user needs.

OML is considered one of the components of FIBRE´s monitoring and monitoring federation solution as discussed next.

ZenOSS [http://community.zenoss.org/index.jspa]

Zenoss (Zenoss Core) is an open-source application, server, and network management platform based on the Zope application server. Released under the GNU General Public License (GPL) version 2, Zenoss Core provides a web interface that allows system administrators to monitor availability, inventory/configuration, performance, and events.

ZenOSS comprises various components, the most important are:

- Collector, which runs ping, snmp, ssh processes, and collects statistics from the devices,
- Database, which stores the collected statistics,
- WebUI, to configure and present the status of the infrastructure.

These enhanced capabilities, features and community support of ZenOSS makes it a versatile infrastructure monitoring tool.

ZenOSS is considered one of the components of FIBRE´s monitoring and monitoring federation solution as discussed next.

5 FIBRE Monitoring and Monitoring Federation Context

The monitoring and monitoring federation context comprises the set of experiments deployed at the FIBRE network facility.

The FIBRE facility is currently composed by a set of 13 "islands" (10 in Brazil and 3 in Europe) with a simplified topologic structure illustrated in Figure 5-1. This topologic structure is, for the purpose of the monitoring and monitoring federation discussion, split in two set of islands with different technical and implementation characteristics:

- FIBRE-BR islands: composed by the islands physically located in Brazil which are interconnected by an overlay backbone network implemented over the RNP network and mentioned from now on as FIBRE-BR;
- FIBRE-EU islands: composed by the islands physically located in Europe which are interconnected by a specific overlay backbone network implemented over GÉANT and mentioned from now on as FIBRE-EU.
- FIBRE-BR and FIBRE-EU islands are physically interconnected through Internet2 and RedClara**.**

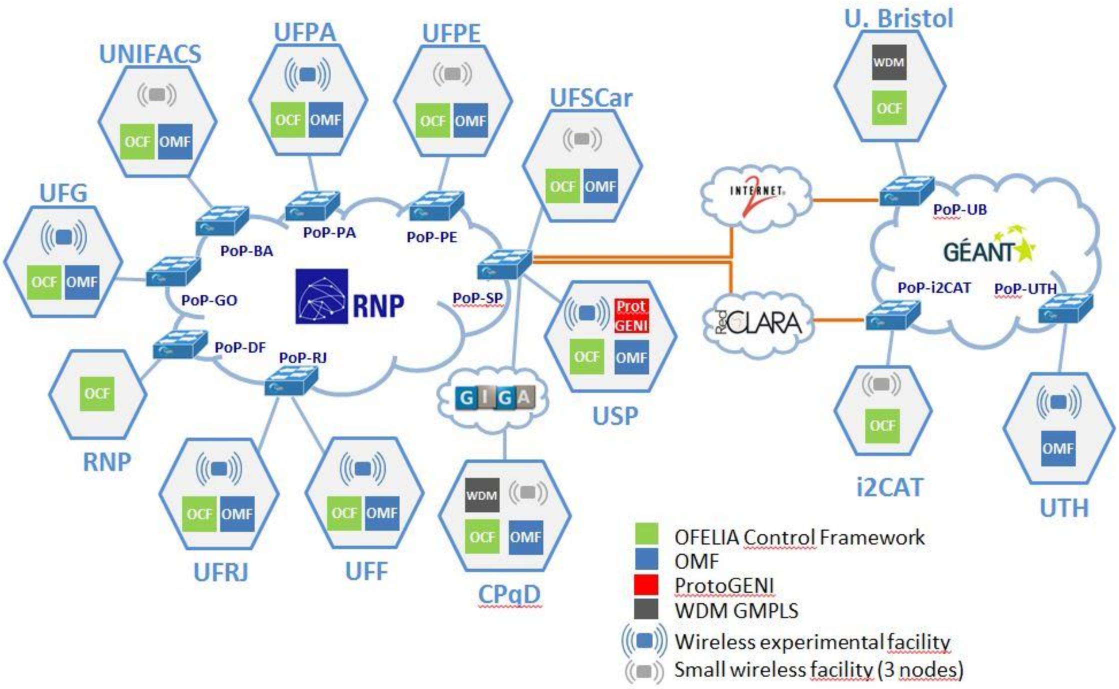

Figure 5-1 – FIBRE Islands

The overall context of the FIBRE facility comprises then the set of resources deployed over islands in Brazil and Europe that will support experiments with various possible deployments.

FIBRE islands adopt and support different Control and Monitoring Frameworks (CMFs) organized in different ways by islands in both FIBRE-BR and FIBRE-EU. The CMFs adopted in FIBRE are:

- OFELIA Control Framework (OCF)
- OMF
- ProtoGENI

FIBRE-BR islands, by definition, may adopt any combination of CMFs according to the focus and research interest of the institution implementing the island in the FIBRE project. FIBRE-EU islands, at this moment, adopt different CMFs (OCF and OMF) but remain homogeneous in terms of each island individually.

Monitoring Federation may then occur in the FIBRE facility context in various ways depending on how the experiment is effectively deployed over the network. The basic experiment deployment alternatives in terms of adopted CMFs are:

- An experiment is deployed over an island and uses a single control and monitoring framework (OCF, OMF, or ProtoGENI);
- An experiment is deployed over an island and uses multiple CMFs (OCF ←→OMF; OCF ← → ProtoGENI, other);
- An experiment is deployed over multiple islands and uses a single CMF; and
- An experiment is deployed over multiple islands and uses different combinations of CMFs.

As such, the experiment deployments and consequently monitoring and monitoring federation scenarios are multiple and the defined Instrumentation and Monitoring Architecture (I&M) discussed next considers these scenarios.

6 Instrumentation and Monitoring Architecture and Federation

The Instrumentation and Monitoring (I&M) Architecture supporting FIBRE islands and corresponding CMFs presents a set of architectural requirements as follows:

- I&M has necessarily to deal with a variety of CMFs which, in turn, have their own monitoring tools and facilities;
- I&M should provide a way to instrument the experiment;
- I&M should provide a way to collect, optionally store (persistent data) and visualize the monitored data;
- I&M should allow monitoring federation considering the control frameworks adopted in FIBER-BR and FIBRE-EU.

The generic strategic solution adopted by the I&M Architecture in order to support monitoring and monitoring federation is illustrated in Figure 6-1 and, in brief, uses a set of gateways and a portal entity.

The gateways have the purpose of conforming CMF´s specific monitoring tools and corresponding monitoring data (standards) to a FIBRE monitoring data format standard, in case, OGF's NMWG schema[3] (PerfSONAR-like standard).

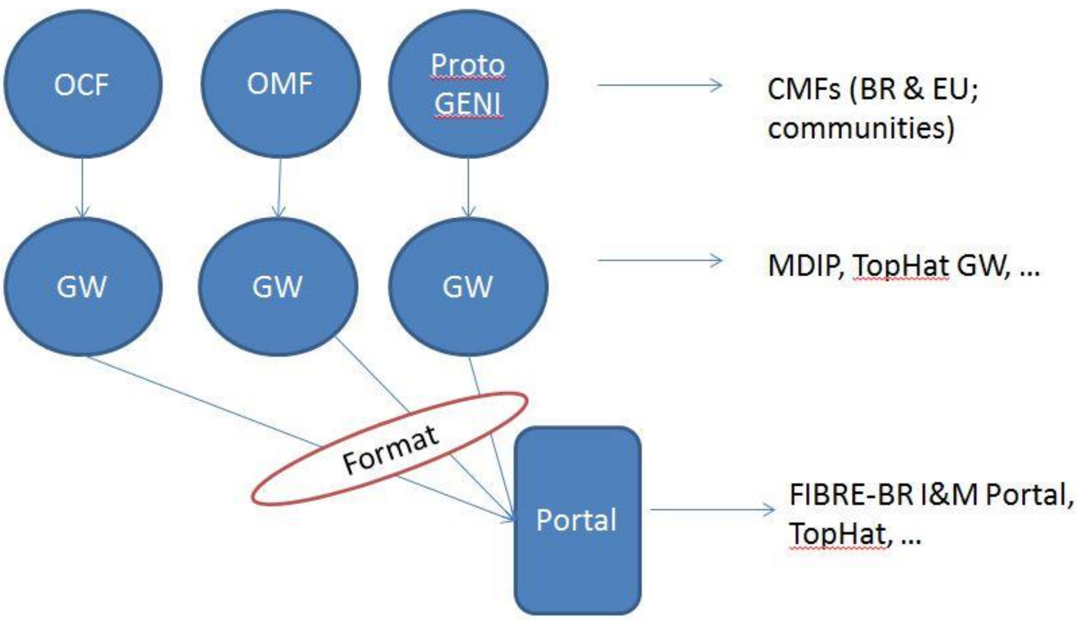

Figure 6-1 – Generic View of FIBRE's I&M Architecture

[3] http://redmine.ogf.org/projects/nm-wg/repository/show/schema

The indicated portal entity assumes the role of displaying relevant monitoring data and, also, supports the inherent instrumentation process necessary to setup and configure the monitoring tools and facilities involved.

6.1 Instrumentation and Monitoring Architecture

A major FIBRE goal is to provide Instrumentation and Measurement (I&M) facilities in order to help experimenters, network administrators and researchers to define experiments and collect infrastructure and/or experiment specific data.

The basic requirement for the I&M Architecture is the capability to configure, monitor, collect, and display both infrastructure and experiment specific data for distinct federated or individual CMF aggregates. Besides that, the architecture also includes or considers the set of requirements adopted by the FIBRE experimental facility which includes infrastructure measurements, experiment measurements, privacy of measured data, link measurements and persistent storage.

The building blocks of I&M architecture leverage from current CMFs' measurement capabilities. They create a gateway facility that conveys the measured data to other I&M services through a common format standard, according to its policy and experimental network requirements.

As such, the perfSONAR based Instrumentation and Monitoring (I&M) Architecture, proposed by the Brazilian partners, is based on a set of gateways (GW) referred to as MDIP (Measurement Data Integration Point) whose main functionalities and purpose are as follows:

- To collect monitoring raw data from CMF´s native monitoring tools;
- To conform monitored data to NMWG standard;
- To interface with the FIBRE monitoring portal entity through web services in order to support monitoring data visualization and access from experimenters (FIBRE users).

The Measurement Data Integration Point (MDIP) conforms the collected data from the available CMFs to perfSONAR standard format, representation, and distribution (including visualization). This service includes all measurement data processing related aspects such as, message format, message transport protocol and/or service, access privileges, and common data storage or on-the-fly data distribution.

The overall perfSONAR based I&M architecture, with MDIPs supporting heterogeneous CMFs installed in the islands, is illustrated in Figure 6-2.

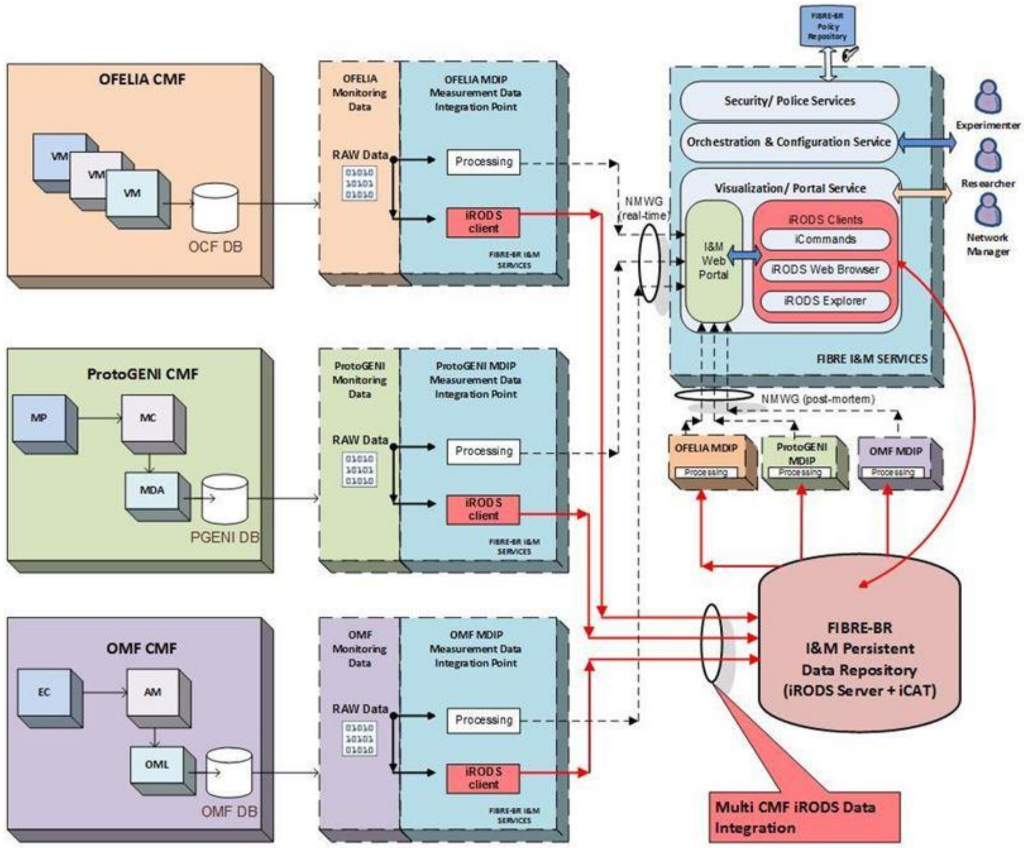

Figure 6-2 – perfSONAR Based I&M Architecture

In the context of FIBRE-BR islands, three MDIPs were proposed:

- ProtoGENI MDIP
- OMF MDIP
- Ofelia (OCF) MDIP

ProtoGENI MDIP benefits from the LAMP (Leveraging and Abstracting Measurements with perfSONAR) project[4] architecture, which is based on perfSONAR. In this specific case, MDIP data manipulation is much simpler due to the fact that perfSONAR format and schema have been adopted also for the FIBRE I&M architecture.

OMF MDIP interfaces with the OML Server and sends the requested data to FIBRE portal entity and I&M Persistent Data Repository. In this case, experimenters should create their measurement points based on some guidelines in order to facilitate the conversion of the resulting data to the NMWG standard.

OCF MDIP is the gateway between OCF and FIBRE portal providing data collection and formatting according to NMWG (perfSONAR standard).

[4] http://groups.geni.net/geni/wiki/LAMP

FIBRE I&M Architecture is intended to be an evolutionary I&M architecture in the sense that, firstly, it evolves from integrating aggregates belonging to homogeneous CMFs and, secondly, it integrates federated aggregates in multiple CMFs.

6.2 Monitoring Federation

The monitoring solution for FIBRE I&M considers that we may effectively adopt two independent monitoring solutions, one based on TopHat and the other based on perfSONAR (NMWG) standards, and defines a monitoring federation strategy as illustrated in Figure 6-3.

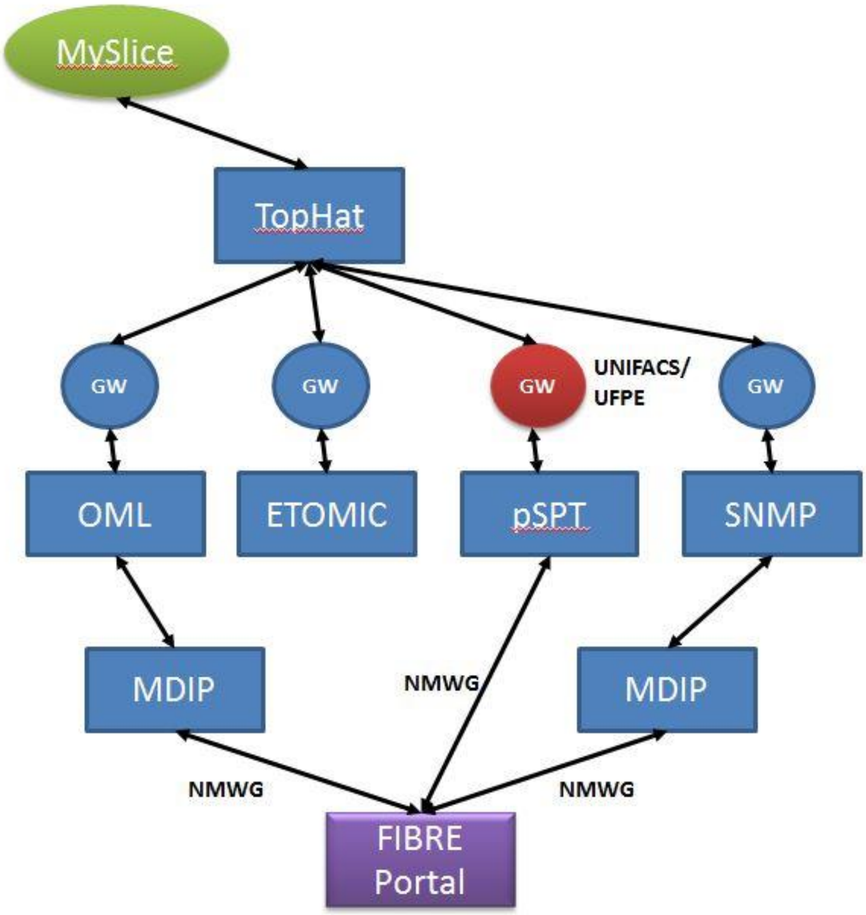

Figure 6-3 – Monitoring and Monitoring Federation in FIBRE

In terms of FIBRE monitoring federation, two independent solutions are being deployed and converge in FIBRE:

- MDIPs are the basic building block of the perfSONAR based I&M and a portal (FIBRE Portal) considers the perfSONAR-like standard adopted and, beyond that, considers the specificities of the FIBRE-BR setup (multiple islands, multiple CMFs, other specific requirements);

- MySlice/ TopHat are the basic building block for federating FIBRE-EU and FIBRE-BR islands.

The referred convergence in terms of the monitoring federation is achieved through the TopHat GW implementation which, in brief, will expose perfSONAR-like measurements to current TopHat users.

As such, infrastructure and/or experiment (slice) measurements will potentially be instrumented and visualized through both solutions and the convergence for the user might be seen as the possibility to acquire the required measurements through both facilities (perfSONAR based Portal and/or TopHat/MySlice setups).

7 Monitoring Federation Use Cases

This section describes the monitoring use cases identified for the FIBRE multi-CMF federated environment. We divided these use cases into two groups:

1) *Infrastructure Measurements*, which are the measurements automatically performed among the various islands of the testbed in order to provide connectivity information; and

2) *Experiment/Slice Measurements*, which are the measurements that the experimenter can perform within his/her slice.

The following subsections describe these two groups in detail.

7.1 Infrastructure Measurements

7.1.1 Use case 1: Active Measurements

This use case comprises the active measurements that will be periodically performed among FIBRE islands. The results of these tests will aid both experimenters and the testbed operators: the first in choosing resources that will better serve his/her experiment requirements, and the latter in providing an overview of the testbed connectivity as well as helping in troubleshooting performance issues.

The implementation setup for this use case is presented in Section 8.3.2.

7.1.2 Use case 2: Passive Measurements

This use case comprises the collection of passive measurements results (interface utilization, loss and errors rates, uptime/downtime, etc.) from the servers and services on each island of the FIBRE testbed.

The implementation setup for this use case is presented in Section 8.3.1.

7.2 Experiment/Slice Measurements

This section deals with use cases related to measurements performed *within* a slice. We divide these use cases in “intra-CMF” and “inter-CMF”. An *intra-CMF* measurement is one performed between nodes from a single CMF (e.g., two OCF nodes). An *inter-CMF* measurement is one performed between nodes from two different CMFs (e.g., an OCF node and a ProtoGENI one). Figure 7-1 illustrates these two concepts in a “single-CMF slice” and a “multi-CMF slice”.

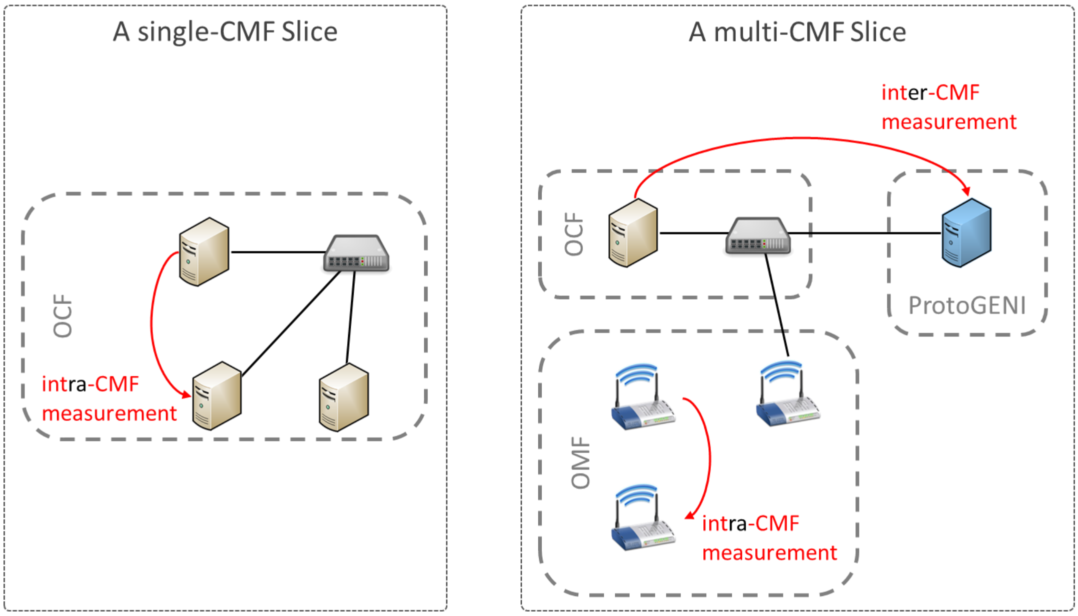

Figure 7-1 – Difference between "intra-CMF" and "inter-CMF" measurements

The following subsections describe the use cases that arise from the scenarios depicted in Figure 7-1.

7.2.1 Use case 3: Intra-CMF

In this scenario the experimenter uses the CMF's own I&M software (e.g.: GEMINI[5] on ProtoGENI) to configure and manage his measurements and, as such, I&M Portals will *not* be needed to *configure* intra-CMF measurements, only to *visualize* the results.

In order to make the measurement results available to the Portals, I&M Architecture will install gateways within the slice whenever needed. These gateways should be able to:

- Identify measurements being performed within the slice;
- Register the measurement data on a Measurement Information (MI) Service, to be presented in Section 8.2, so external clients such as Portals can discover this data;
- Provide an API for data retrieving, so external clients such as Portals can fetch the measurement data for visualization and backup;

Figure 7-2 shows a scenario where a gateway is deployed inside an OCF slice. In this scenario, the steps for visualizing the measurement results are:

1) The gateway will automatically identify the measurements being performed by the CMF's I&M software and register them on the MI Service;
2) By querying the MI Service, any Portal will be able to discover the slice's intra-CMF measurements;

[5] http://groups.geni.net/geni/wiki/GEMINI

3) Using the gateway within the slice, the Portals can fetch the desired data and present it graphically to the user;
4) Once the experiment is finished, any Portal can fetch all measurements stored within the slice (via the gateway API) and save them to some Persistent Storage System.

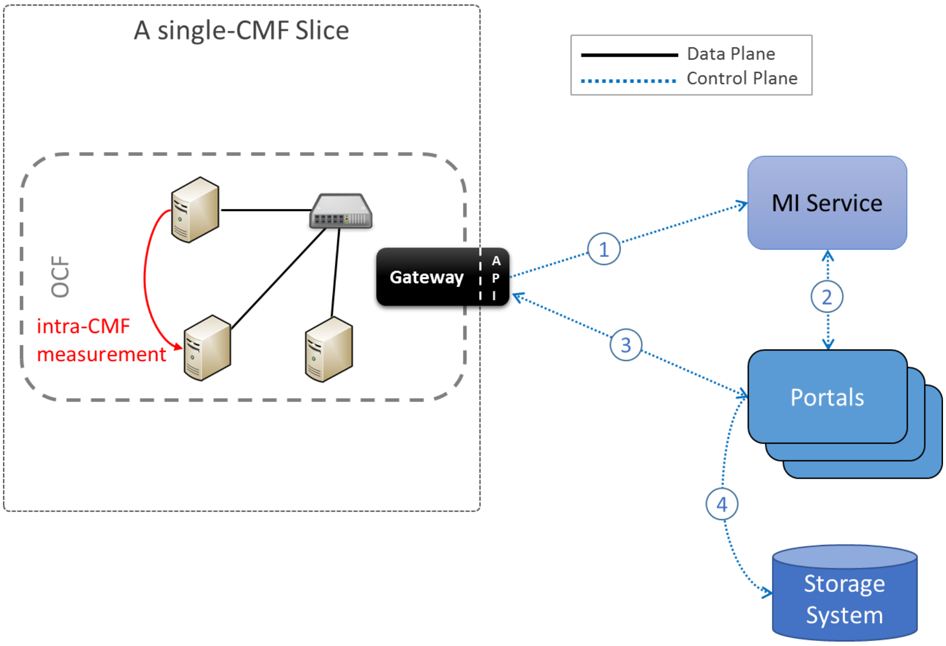

Figure 7-2 – Intra-CMF measurements

Note that this strategy works both with active and passive metrics. Figure 7-3 shows a scenario involving a multi-CMF slice, where each gateway exposes their passive measurement data stored inside the slice through its API.

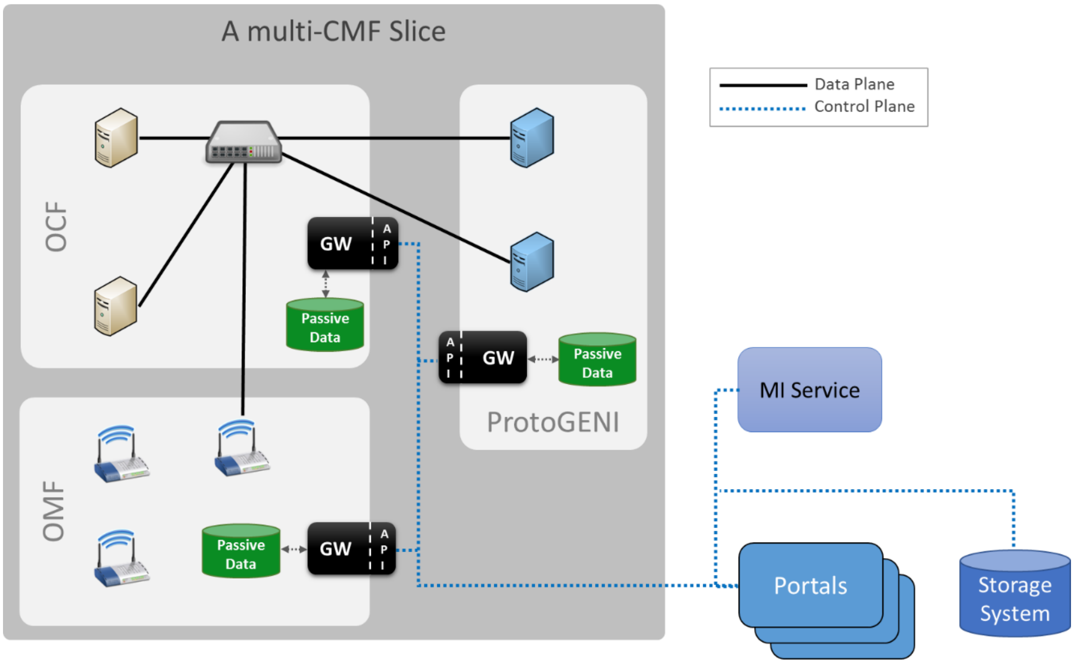

Figure 7-3 – Intra-CMF access to passive measurement results in a multi-CMF slice

7.2.2 Use case 4: Inter-CMF

As depicted in Figure 7-1, in a multi-CMF slice an experimenter may need to perform an inter-CMF measurement. This use case deals with this scenario.

Since each CMF has its own I&M software, the first challenge is to instrument all nodes with a common measurement tool, so any node can perform tests towards any other node, regardless of its CMF.

Figure 7-4 depicts how this proposal can be applied to a multi-CMF slice by installing on every node a common measurement tool (M. Tool). Note that for every CMF there is a gateway, which is responsible for configuring the measurement on behalf of the Portals and exposing the results in a standardized data format.

Also note that while in Use case 3 the experimenter would use the CMF's I&M Portal to configure the measurements, here we propose that FIBRE Portals will be used instead.

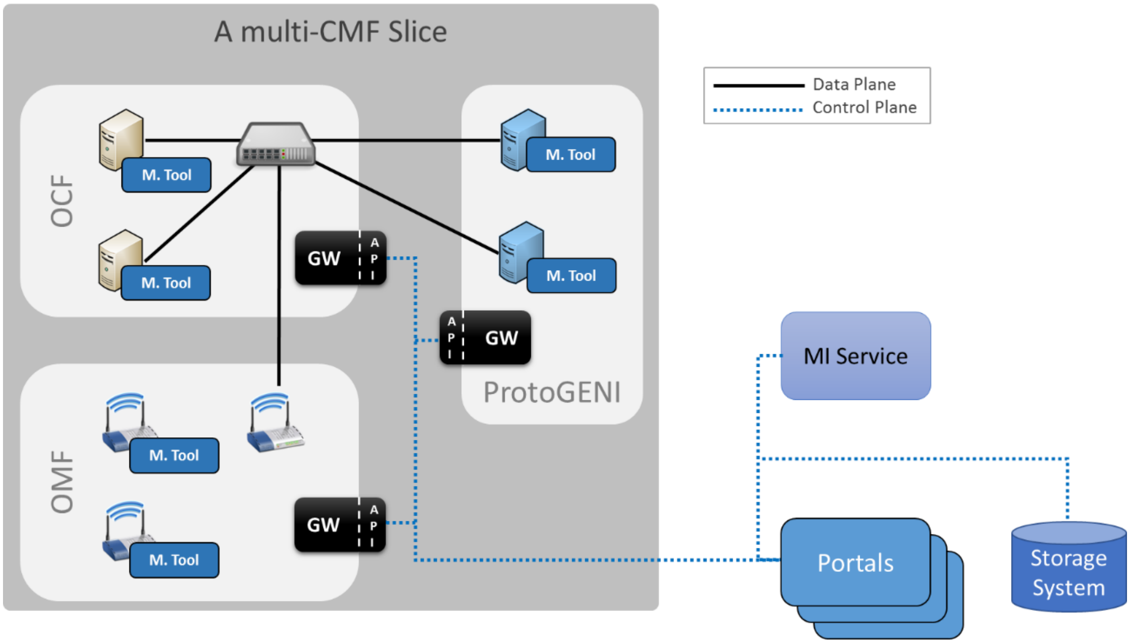

Figure 7-4 – Inter-CMF I&M solution in a multi-CMF slice

8 Monitoring Federation Building Blocks

In this section the basic building blocks composing the proposed Instrumentation and Monitoring (I&M) Architecture such as portals, gateways and agents are further detailed.

8.1 Monitoring Portals

8.1.1 Interconnecting monitoring systems with TopHat and Manifold

It is often the case that a testbed has one or several monitoring and measurement systems, with heterogeneous data and interfaces. TopHat (http://www.top-hat.info) and MySlice use an interconnection framework (named Manifold) allowing the different platforms to expose their information in a consistent way to the federation, and it helps mapping the different measurements to the corresponding Slice-based Facility Architecture (SFA) resources identifiers.

The Manifold information framework is able to answer user queries by a combination of two or more platforms, and thus provide information that was not previously available through individual platforms. MySlice uses it heavily to expose monitoring information to its users, and to allow for RSpec annotation.

More specifically, Manifold proposes a simple an efficient **query language** to request several sources of data and simple processing across interconnected platform, with many similarities to the capabilities offered by SQL:

- an engine performing efficient **query dispatching and result aggregation**;
- a **standard interface** for platforms to expose measurement information (similar to tables in a relational database), alongside a metadata syntax to describe them;
- a set of **gateways** to support platforms not complying to this interface. Those gateways are responsible for handling authorization/authentication with the platforms, and interpreting the transport of data (e.g., XMLRPC), their format (e.g., XML), and the ontology used for representing information;
- a modular interface for the visualisation (used by MySlice, for example).

Our work towards the interconnection of measurement systems is tightly linked to the efforts in federating testbeds, and should allow for an ecosystem of measurement systems to emerge, and the various actors to seamlessly exchange information. This corresponds to the vision of a distributed federation of systems (as opposed to a single central entity), where the data can be accessed through different entry points (with their own strength and specificity) and still allowing access to the full range of available information.

8.1.2 FIBRE I&M Portal

FIBRE I&M Portal has been modelled and developed with the objective to support the researcher/experimenter on the analysis of the results of their experiments deployed over the different CMFs in FIBRE. To make it possible, a portal framework (Figure 8-1) is proposed in order to facilitate the integration of heterogeneous platforms.

The FIBRE portal will be initially deployed with a simple web interface (module 2), handling the visualization of the retrieved data from the experiments through the MDIP/MA communication interface (module 3). The communication interface, in brief, collects experiment and infrastructure data form the experiment in a standard format (NMWG) through the MDIPs gateways.

Another focus of the FIBRE-BR portal is the persistent storage interface (module 3). The integration of the portal with the persistent storage system iRODS (a.k.a, Idrop web) adopted in FIBRE-BR, will allow the researcher to handle raw data. An iRODs web interface is being developed in order to facilitate the integration of iRODs persistent storage (Section 8.5) with FIBRE's Portal.

Figure 8-1 presents the basic building blocks of FIBRE's portal structure.

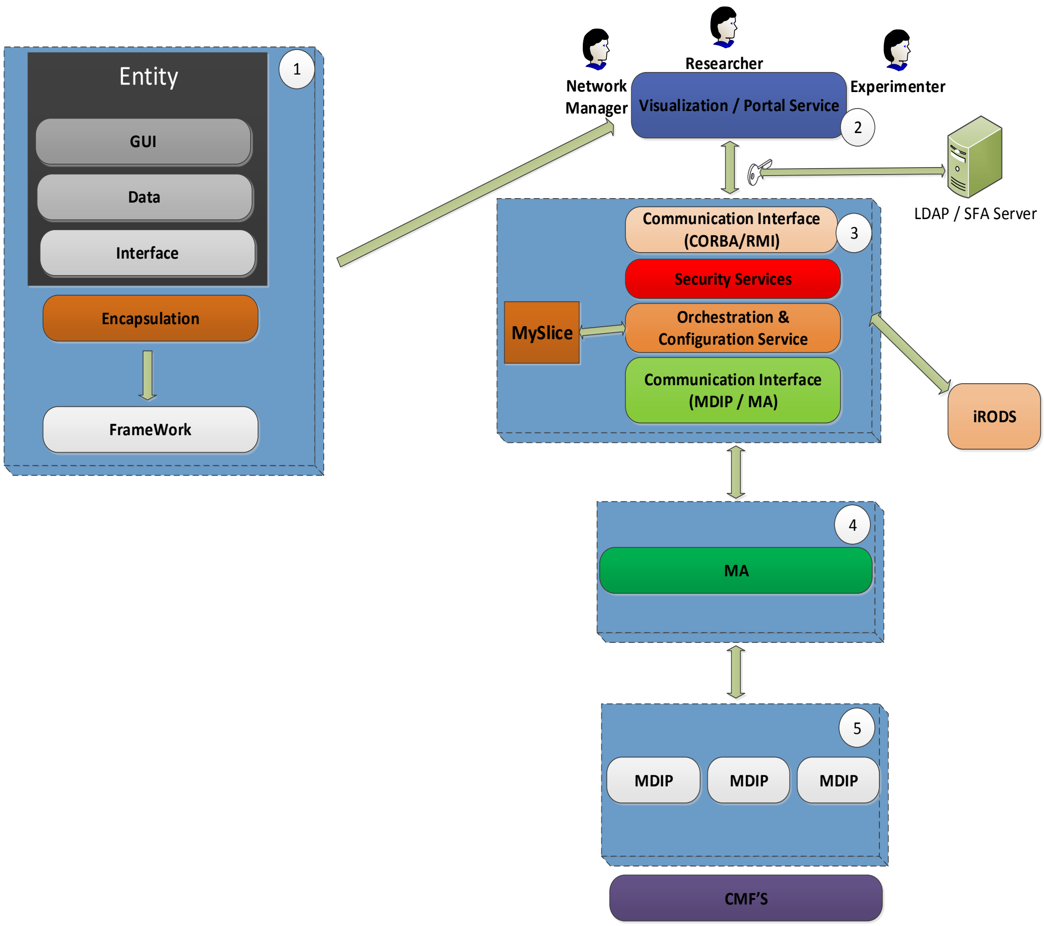

Figure 8-1 – FIBRE Portal Components Building Blocks

8.2 MI Service & Monitoring

The Measurement Information (MI) Service will act as a distributed metadata directory, providing FIBRE data clients (like I&M Portal) information such as: which measurements are being performed within each slice, the addresses of the GWs that can provide measurement data, etc. It should also provide advanced query mechanisms in order to facilitate the data discovery process for the clients.

Based on these requirements, we propose the use of the perfSONAR PS Lookup Service (pS-LS)[6], a distributed information service that meets the needs of the I&M Architecture. The deployment of the pS-LS is similar to a traditional DNS service, with the Home Lookup Service (hLS) acting as a local DNS server, and the Global Lookup Server (gLS) acting as a DNS root server.

Figure 8-2 shows our proposed deployment for the MI Service in the FIBRE I&M Architecture. Each island will have an hLS server where all MDIPs must register themselves and the measurement data they can provide. The Brazilian NOC will host a gLS cluster with two or more root servers, where each hLS will automatically synchronize its data (this process is represented by the dotted lines in the Figure).

Figure 8-2 also shows how the clients can use the MI Service to discover measurement data. Following the steps depicted in the Figure: (1) the client query the gLS Cluster asking which hLS has a given dataset, and receives the address of the appropriate hLS; (2) the client query the hLS asking which MDIP has the desired dataset; and (3) the client directly contacts the appropriate MDIP to retrieve the desired data.

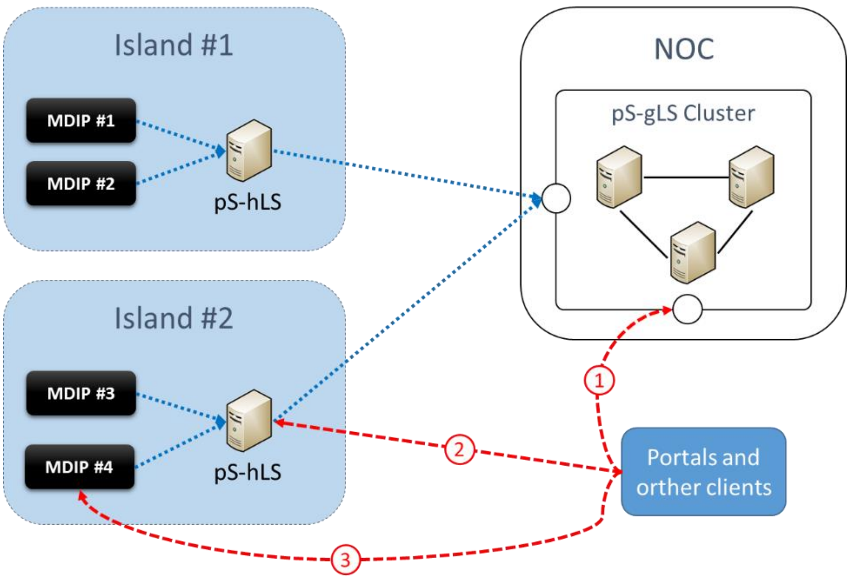

Figure 8-2 – MI Service Deployment Architecture and Usage

6 http://psps.perfsonar.net/lookup/index.html

8.3 Infrastructure Measurements

In this section the monitoring infrastructure supporting infrastructure measurements is presented.

8.3.1 ZenOSS

Infrastructure monitoring in each island is supported by ZenOSS[7]. ZenOSS is a tool capable of managing a set of IT resources of an organization. It monitors the network, servers (physical or virtual), power consumption, and even applications. ZenOSS is organized in 4 layers as shown in Figure 8-3.

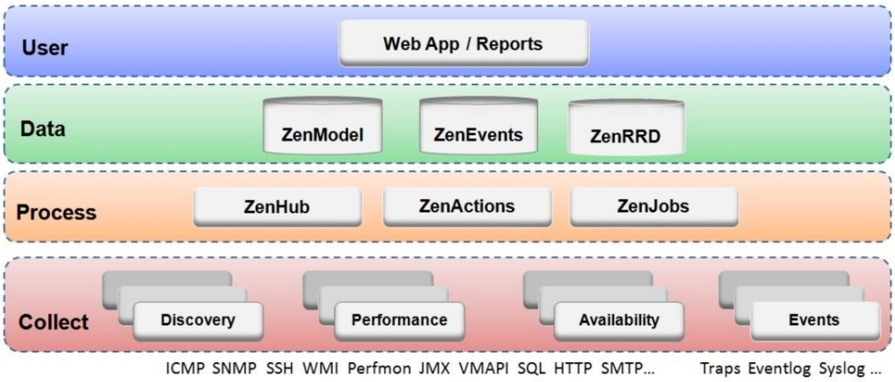

Figure 8-3 – ZenOSS Functional Structure

The user layer is a web portal that gives access to the managed devices, to the dashboard, generated events, user's management, and to reports. This layer interacts with the data layer that is responsible for information visualization.

The data layer stores the information in three different databases: ZenModel and ZenEvents stores devices, monitoring, and events information in a MySQL database. ZenRRD stores measurement data in a Round Robin database (RRD). Each collector owns its specific measurement object, as an example, the CPU measurement generates its own RRDs archives.

The process layer has the responsibility to intermediate the communication between the data and collect layer. This communication occurs through remote procedure calls (RPCs) between the diversity of services that run to support all ZenOSS functionalities.

The Collect layer is responsible to manage the services that collect and feeds the data layer. Data Collection process can be done through a variety of protocols, such as: SNMP, ICMP, SSH, and WMI (for specific Windows process). This layer has a variety of plug-ins that "translates" the raw data to "understandable" data to the data layer.

ZenOSS monitoring process model follows the workflow presented in Figure 8-4.

[7] http://www.zenoss.com/

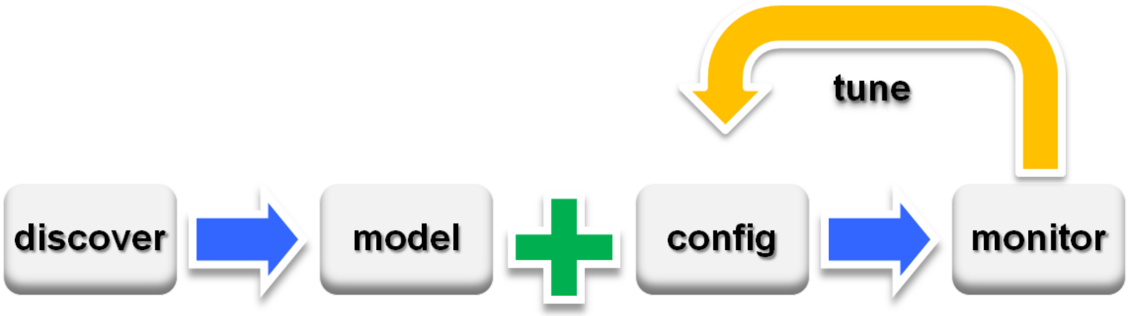

Figure 8-4 – ZenOSS Monitoring Process Workflow

The first step is to discover the resources/devices over the network, either manually or automatically. It is followed by the definition of the configuration model, through a range of available templates or using extensions called ZenPacks. Once it is applied, the monitoring starts. If it is necessary to make adjustments, they should be done in the configuration step to make it reflect in monitoring.

ZenOSS has a concept of class devices that is an organizer of used metrics, by default, accordingly to the device classification. For example, for a Linux server, it has a specific class with a variety of items that are common monitored in Linux servers. Naturally, this class of devices can be adjusted regarding the objective of monitoring. This process is done at the configuration step, just after the resource Discovery, as it can be seen in Figure 8-4.

ZenOSS' available monitoring parameters may not suit all the user needs. Therefore, there are additional packages, called Zenpacks (ZenOSS extensions). These additional packages allow enlarging the range of monitoring possibilities. It is possible to develop a customized Zenpack or import an existing one from the ZenOSS repository.

In FIBRE, two Zenpack are imported: to monitor Xen Servers and Network Interface Cards (NICs) of the devices in each island.

The specific monitoring parameters currently monitored in FIBRE with the ZenOSS tool, are:

- CPU consumption
- Workload SO
- Memory utilization
- Data disk consumption
- I/O operation
- Process and applications (databases, webservers, NIC status)
- NIC interfaces
- Virtual NICs of the OXA server environment

Beyond those monitoring metrics already discussed, ZenOSS has a set of alerts and alarms for critical events that notifies the sysadmin through e-mails with detailed information.

8.3.2 perfSONAR Toolkit and Dashboard

Figure 8-5 shows the strategy adopted for *Active Infrastructure Measurements.* Each island on FIBRE-BR (mandatory) and FIBRE-EU (optional) will be equipped with two perfSONAR PS Toolkits[8] servers for latency and bandwidth measurements. This setup will allow active measurements among FIBRE-BR and FIBRE-EU islands.

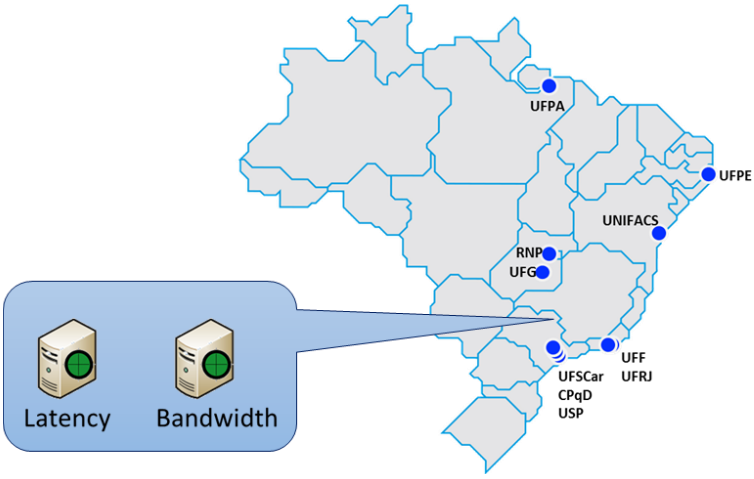

Figure 8-5 Strategy for Active Infrastructure Measurements

Each perfSONAR Toolkit server periodically perform measurements among FIBRE islands and saves the results locally. The users can visualize these results using each server's own web interface.

This is certainly helpful for independent islands operation, one specific requirement of the FIBRE project, but monitoring federation requires a more general approach. As such, in order to provide a global view of the entire facilities performance, a NOC (Network Operating Center) structure is adopted in FIBRE. In terms of the Brazilian islands (FIBRE-BR), the NOC will host a web-based Dashboard where experimenters and operators can visualize measurements from all islands in a single interface. Figure 8-6 shows some screenshots of a prototype deployment of this interface.

[8] http://psps.perfsonar.net/toolkit/

Figure 8-6 - Screenshots from a prototype deployment of the FIBRE-BR perfSONAR Dashboard

In order to make these measurement data available for TopHat, we are in the process of developing a gateway that will translate perfSONAR data to TopHat's format. The details about this gateway are discussed in Section 8.4.1.

8.4 Adopted Gateways

8.4.1 TopHat Gateway

TopHat's gateways acts as an adaptation layer between different monitoring platforms and its framework structure (MANIFOLD). Taking into account that some of the adopted monitoring tools are perfSONAR-based, it was defined that a gateway of this platform to TopHat would be helpful to FIBRE testbed users.

In the proposed infrastructure measurements using the perfSONAR Toolkits we are actively measuring both achievable throughput (BWCTL – Bandwidth Test Controller), and one-way delay and losses (using OWAMP – One-way Active Measurement Protocol) between islands. The results of such measurements are stored in a database and made available through Measurement Archive (MA) services. Considering the TopHat GW, each service has a set of characteristics that must be defined to MANIFOLD's metadata. Currently the metadata description of BWCTL is under development in order to implement the query functions to allow MANIFOLD to make perfSONAR queries to the MA service.

The integration of perfSONAR and TopHat through the GW can be glimpsed in Figure 8-7.

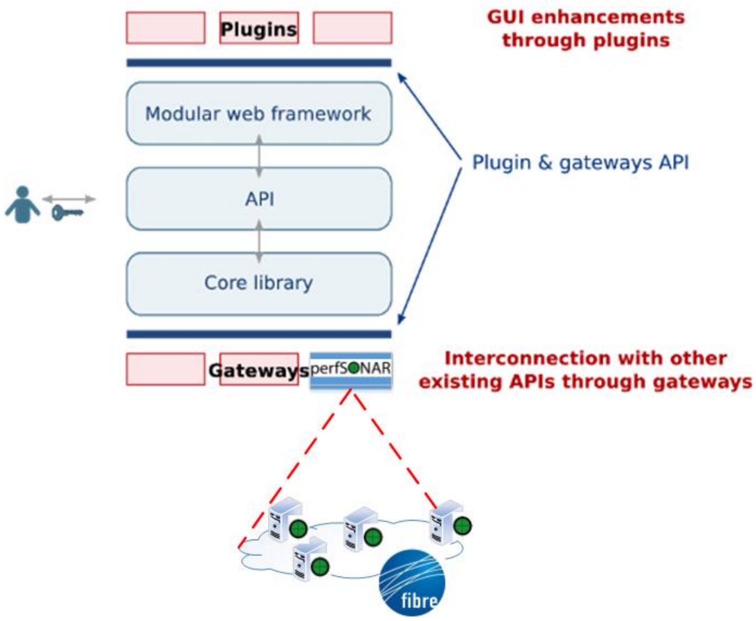

Figure 8-7 – TopHat's perfSONAR Gateway

8.4.2 OCF enhancements with metering agents

Figure 8-8 depicts an overview of the architecture proposed for enhancing OCF to support the collection of statistics parameters.

The architecture is based on a centralized module, the OCF-collector, able to retrieve and collect measurements coming from several monitoring agents. Each agent will be specialized to manage specific information, e.g. VT-AM agent for computer and storage meters, Optin-AM agent for network parameters, etc.

To support the federation environment, the OCF-collector will expose a northbound interface fully compliant to SFA specifications.

It is important to note that the proposed architecture is flexible enough to integrate other "specific" agents that could manage and monitor other particular devices or aggregates.

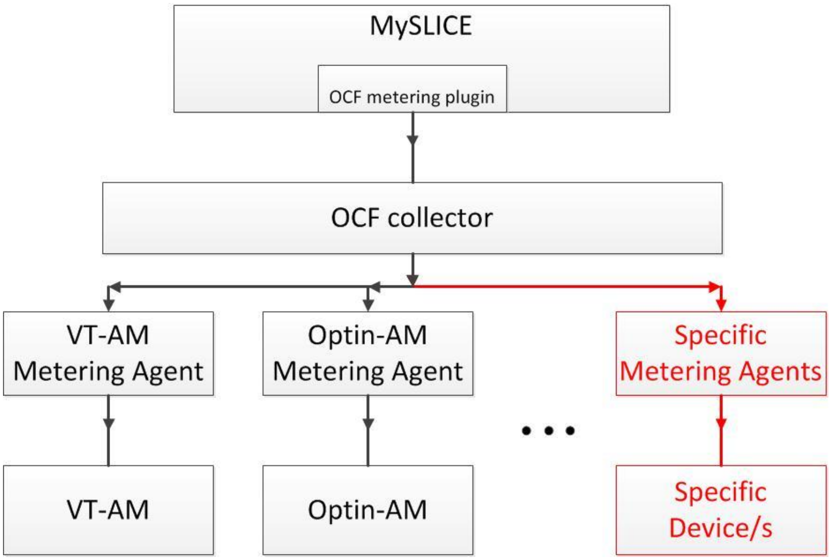

Figure 8-8 – Overview of OCF Metering Architecture

The following sections will give a brief overview of the different building blocks.

8.4.2.1 *VT-AM metering agent*

The Virtual Technology Aggregate Manager (VT-AM) metering agent will be responsible for retrieving computing and storage information from VT-AM and to periodically send them to the OCF collector. This kind of information could be related to power consumption, memory utilization, available RAM or disk space, number of active VMs, etc.

The VT-AM agent will use southbound dedicated XML-RPC methods exposed by VT-AM to retrieve meters and a standard (e.g., OML Measurement Stream Protocol) northbound interface to expose them to the OCF collector.

8.4.2.2 *Optin-AM metering agent*

The Optin-AM metering agent will be specialized to manage network information and OpenFlow statistic parameters related to OF devices and per-slice information. It could retrieve data like average or instant delay, number of packets received/transmitted/dropped, etc.

The Optin-AM agent will communicate to the Flowvisor instance that manage OF physical devices, using its JSON-RPC monitoring methods. On the other hand, it will use the same standard northbound interface, already introduced in Subsection 8.4.2.1, to communicate to the OCF collector.

8.4.2.3 *OCF collector*

The OCF collector will be the centralized entity that will store all the received metering data and expose them in a standard format (i.e. XML-based) compliant to SFA specifications.

The OCF collector will receive data streams from the metering agents and will store them in a database backend (e.g. SQLite or PostgreSQL based). This will give the possibility to perform any kind of data analysis for discovering potential issues, network failures or congestions, etc.

8.4.2.4 *Zenoss interfacing with OML:*

One of the selected and deployed facility & infrastructure monitoring for FIBRE OCF islands is Zenoss. Zenoss daemon collects and stores performance data from network equipment, links, computers and virtual machines using various interfaces SSH, SNMP, Telnet etc. It supports ZenPack which is a plug-in or extension to Zenoss. Zenoss has many (more than 200) ZenPacks available for various network & IT devices. Typically, ZenPacks add the ability to monitor new types of devices, but can also be used to add new capabilities to the software itself, add new reports, etc.

Its default storage backend relies on RRDTool, from which time-based graphs can be generated from external tools. However, a writer plugin is needed which makes the data available through OML. The UNIVBRIS FIBRE facility has developed a plugin which extracts all the Zenoss monitoring data and exposes it via an OML interface. The measurement data is packaged in a common format: the OML stream, which the experimenters can use. This OML measurement databases can be made accessible to the OCF collector which can further expose it Myslice monitoring interface.

8.4.2.5 *MySLICE plugin*

The MySLICE OCF metering plugin will be responsible to retrieve and manage OCF measurements and to graphically show them to the end users.

The MySLICE OCF metering plugin will communicate to the OCF collector using the SFA control framework. Due to the complexity and the large amount of available data, the modelling of metering resources (RSpecs based) will be a crucial point.

8.4.3 ProtoGENI MDIP

The ProtoGENI I&M solution (GEMINI) stores active measurements distributed among the nodes (the node that initiates the test saves the results) and passive measurements at a special node dedicated to monitoring activities, called Global Node, which has access to all nodes within the slice through the Control Plane.

To take advantage of this architecture, the ProtoGENI MDIP was developed to work within the Global Node as an additional service, providing Measurement Archive (MA) interfaces that enable external clients to retrieve the measurement data. Figure 8-9 shows the overall MDIP architecture.

In order to offer a simplified interface for retrieving active measurement data, we developed a collector that caches the distributed data in a temporary database (Cache DB), which is attached to a MA service based on pS-BUOY[9]. Besides offering an interface for external access, this MA also registers its data on the MI Service.

On the other hand, passive measurements are already stored inside the Global Node as RRD files. To expose these data to external clients, we are in the process of adapting the pS-SNMP-MA Service[10] to expose the already generated RRD files through an MA interface.

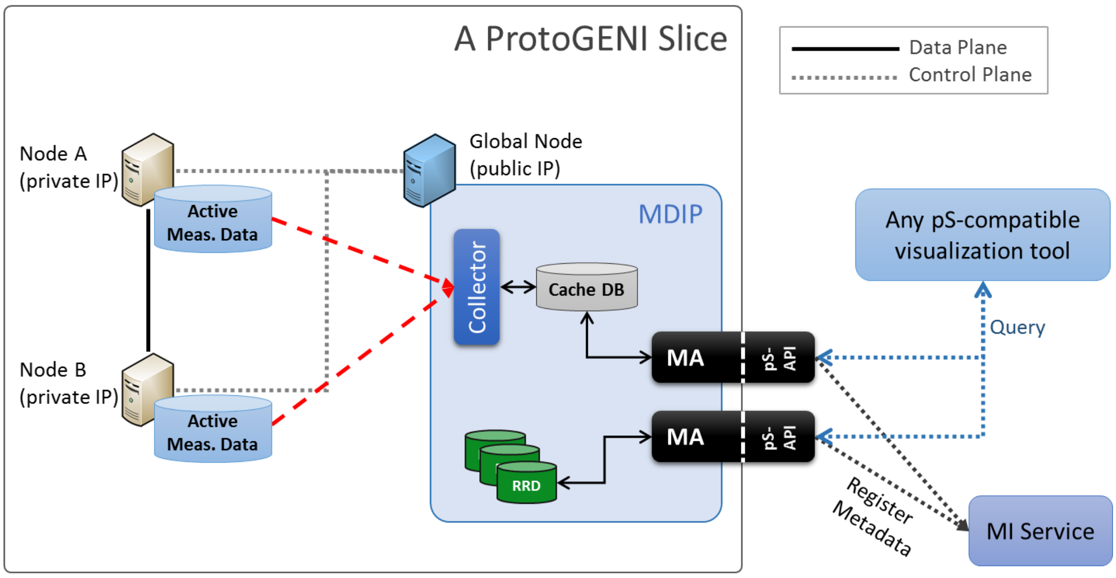

Figure 8-9 – ProtoGENI MDIP Architecture

8.4.4 OCF MDIP

The OCF MDIP is based on an inherent measurement node transparently instantiated on every slice in OCF CMF. This node is mainly composed by a collector module, a pool of measured data, and two modules for exposing the measurements to the external environment, interfacing with I&M Portal and with the Persistent Data server.

The collector is responsible for gathering the measurement data from SNMP services on hosts and OpenFlow devices in the experiment. Those data are stored in RRD files, keeping historical registers of the metrics. These data can be accessed in real time through a MA (perfSONAR SNMP-MA) converted to the NMWG standard format. The gateway still has the capability to export the measured data to the Persistent Data server through a client/server interface. Figure 8-10 shows the overall OCF MDIP architecture.

[9] http://psps.perfsonar.net/psb/

[10] http://psps.perfsonar.net/snmpma/index.html

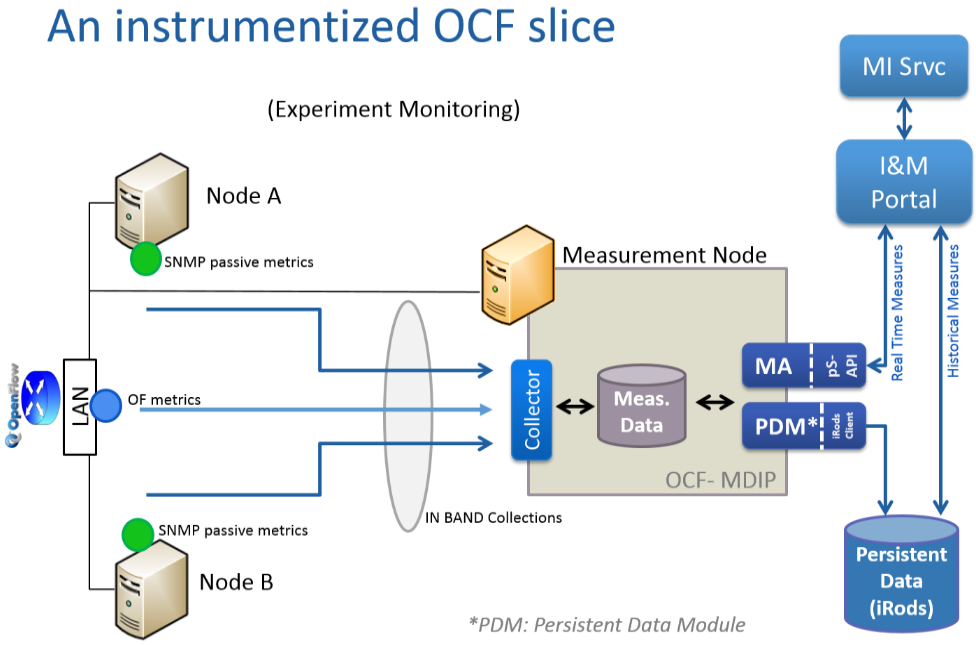

Figure 8-10 – OCF MDIP Architecture

8.4.5 OML MDIP

OML MDIP consists of a perfSONAR tool “plugged” to the OML Server. The OML MDIP is capable of access OML Server database to query for data when requested.

As shown in Figure 8-11, the OML MDIP is composed by a Web Service module that listens for client requests, always made by using the perfSONAR communication protocol (using NMWG schema), process these and then passes all valid requests to the DB Module, responsible for retrieving data requested by the other modules.

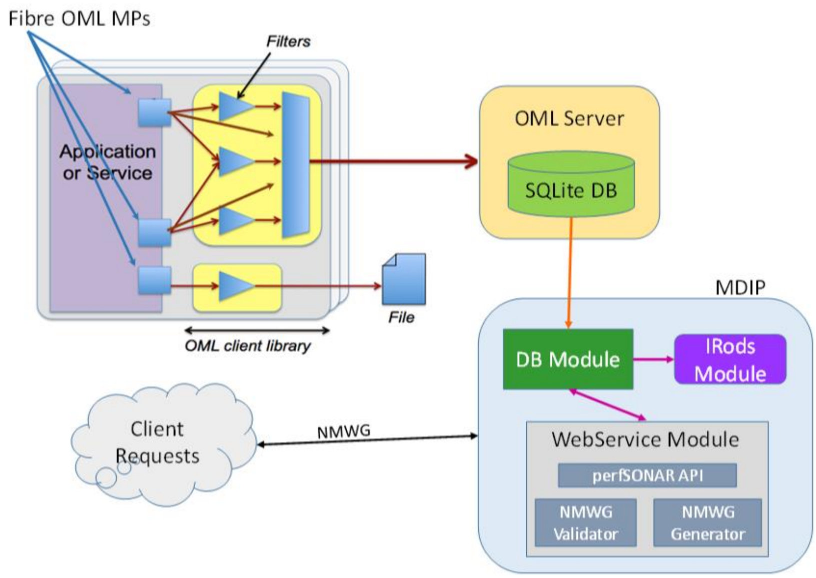

Figure 8-11 – OML MDIP Architecture

The Web Service module is based on a perfSONAR SQL Measurement Archive.

8.5 Monitoring Persistent Data & iRODS

iRODS is a data grid system that aims at consolidating data systems. The adoption of iRODS in the FIBRE I&M Architecture was motivated by the fact that its components and functionalities meet most of the FIBRE I&M requirements for persistent data storage.

This software provides a way to manage, organize, share, protect, and preserve data among users. Based on a hierarchical structure, it is possible to build a transparent storage solution (centralized or distributed) that looks from the final user perspective just as a regular directory.

The adopted strategy is defined by an instance of MDIP with iRODS client (icommand) that will send data collected of the experiment environment to an iRODS icat server in each island/zone. These servers are interconnected through iRODS-federation allowing a user (experimenter, researcher, other) to store data in a remote zone. In order to make it possible, it will be necessary to comply with user's permissions. The access to the monitored data will be allowed through the web portal (iDROP-Web) that, after authenticating the user with the same credentials stored at the LDAP DB, will enable the user to delete an archive, insert metadata, share data and download the data to his/her local machine. Figure 8-12 illustrates the adopted scheme with three federated islands: island 1, island 2 and island 3. Each island will store the experiments data of their testbeds and will catalog them on its own iRODS iCAT.

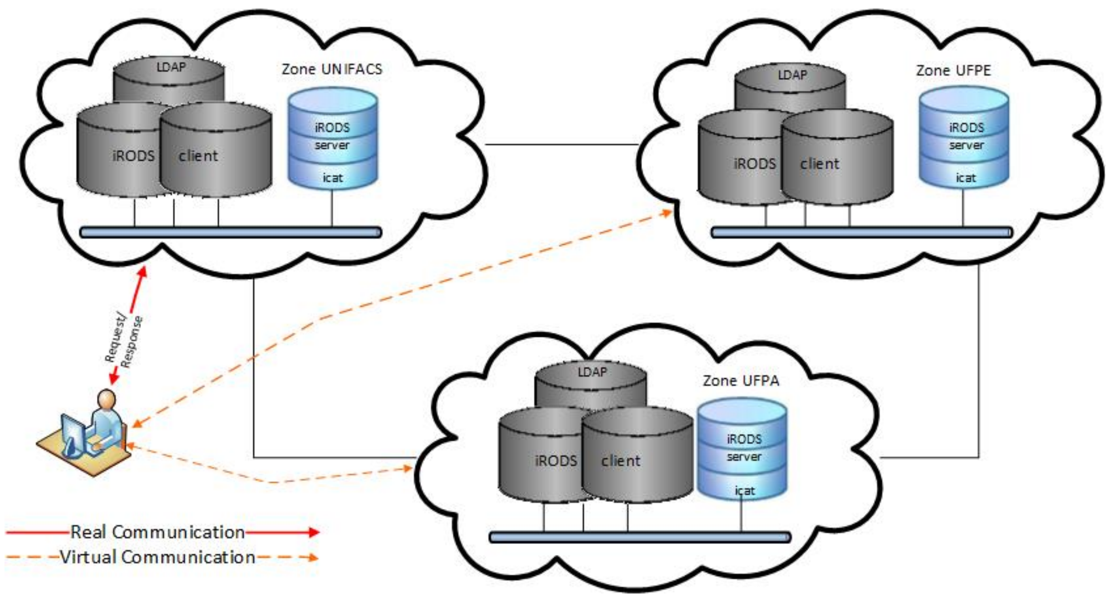

Figure 8-12 - iRODS Federation with Distinct CMFs

Zone access privileges, in general, will be managed by the federated LDAP. Accessing data at the original zone is represented in Figure 8-12 as the real communication. Clients will also be able to query for data stored in other trusted zones, through a virtual communication making use of iRODS mechanism. If the data is not stored at its original zone, distinct iRODS iCATs will communicate among them to locate it. Regarding monitoring data visualization, the iRODS clients will also access data repositories in local or remote islands/zones as needed.

9 Conclusion

The objective of this document was to present the definition of the federation monitoring framework and diagnosis tools to be used on the FIBRE facility.

The proposed FIBRE I&M Architecture and Monitoring Federation were presented. There are basically two approaches. One based on TopHat and the other based on perfSONAR. All the monitoring building blocks for both approaches were presented.

All the alternative approaches are under development and expected to be completed by the end of the project.

"*This work makes use of results produced by the FIBRE project, co-funded by the Brazilian Council* for Scientific and Technological Development (CNPq) and by the *European Commission within its Seventh Framework Programme."*

END OF DOCUMENT