gCube

gCube Grid Service

Pedro Andrade
CERN

- Introduction

- What is gCube

- Functionality and Services

- Infrastructure Set-up

- DILIGENT & D4Science VREs

gCube is a framework to host, define and manage dynamic Virtual Research Environments (VREs) to satisfy the collaboration needs of distributed Virtual Organisations (VOs)

gCube was developed by the DILIGENT project and is now being maintained and further enhanced by the D4Science project

VREs are collaborative environments created on-demand through remote sharing of resources to support the research activities of distributed communities

A VRE consists of several resources

- Processing and storage resources
- Heterogeneous collections of data
  - reports, maps, sensor data, statistical data
- Applications for retrieving and accessing data, processing it and produce new knowledge
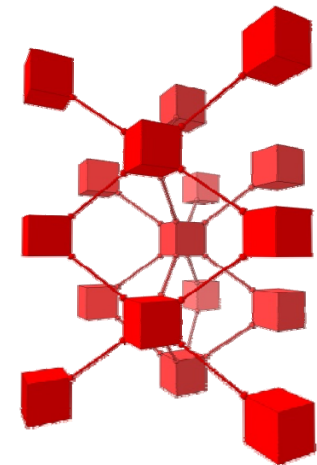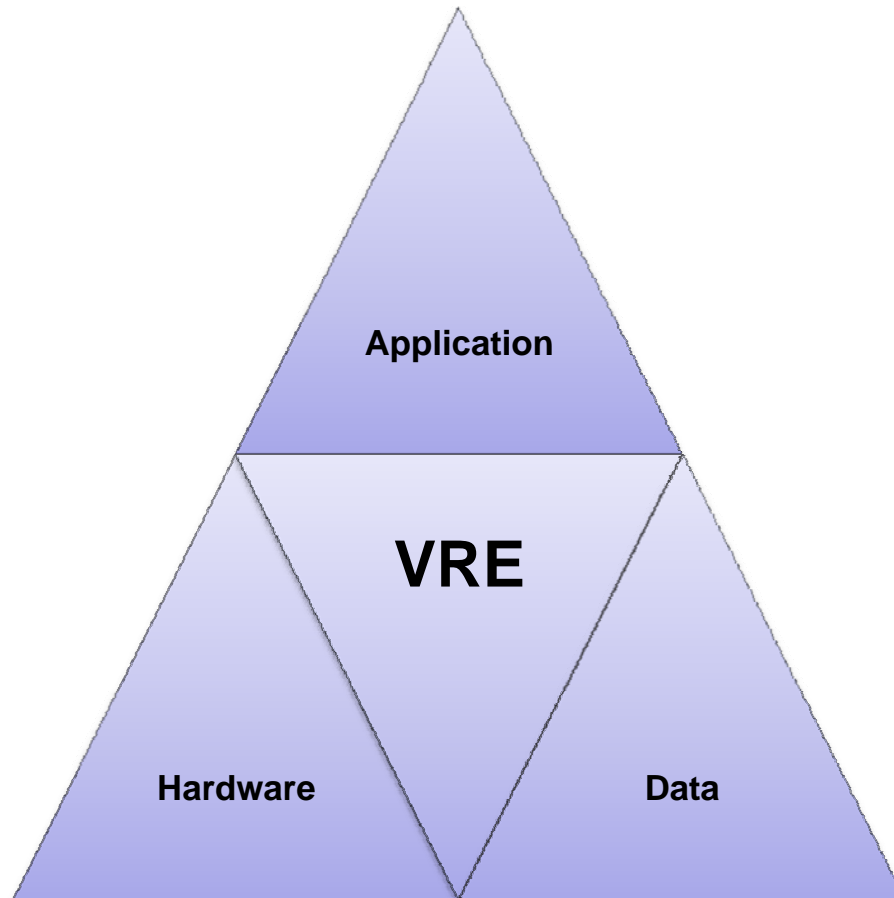
- gCube is a distributed system of middleware and domain services to host, define and manage VREs

- Workspaces in which users collaborate through controlled sharing of hardware, data, and application resources
  - 3rd generation of the Grid vision of resource sharing

A framework for the development of services that can be outsourced to a grid-enabled infrastructure

- Easy development of gCube services

An advanced container of web services (remotely managed)

- Easy deployment and run-time of gCube services

A runtime environment for executing VRE build-in services

- Provision of information about shared resources
- Workflow management
- Replication and partition of federated content
- Indexing, selection, fusion, extraction, and annotation

gCube

gCube Grid Services
13 February 2008, Clermont-Ferrand (France)

www.gcube-system.org

Content management highlights

- Storage of inter-related information and data by exploiting database and grid storage capabilities
    - Logical grouping of content in collections
    - Logical sharing of content among several collections
    - Replication and partition of content
    - Importing/linking pre-existing content
    - Bulk upload and update

VRE Specific

- Storage of structured heterogeneous metadata compliant with different formats and schemas

- Programmatic and manual annotation of data sources via text and images

Information retrieval services highlights

- Distributed search engine composed by different, autonomous, and pluggable elements
  - XML processing (joiners, sorters, transformers, filterers)
  - Lookups (FT indices, XML indices, geo indices, external sources)
  - Combining of results (fusion / merging)
- Supported search types
  - Structured data (fielded search / xml search)
  - Semi structured data (xml search)
  - Geospatial / temporal data (R-Tree)
  - Content based search (full text search, image similarity search)
- Data transport mechanism (ResultSet)

VRE Specific

9
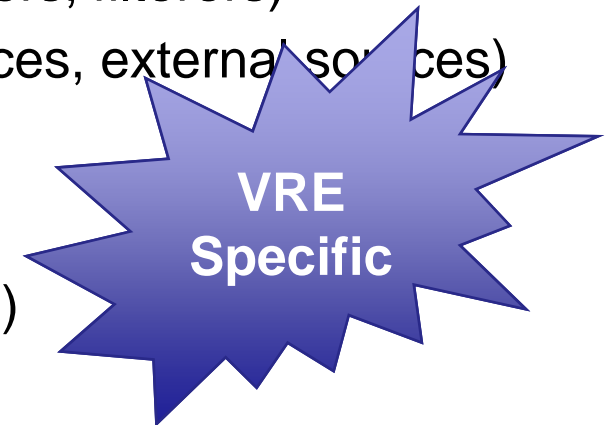
Workflow management services highlights

- Fully distributed workflow engine
  - Based on a subscription/notification mechanism
- Support of different operators
  - Loops, iterations, conditional execution
- Re-execution of workflows
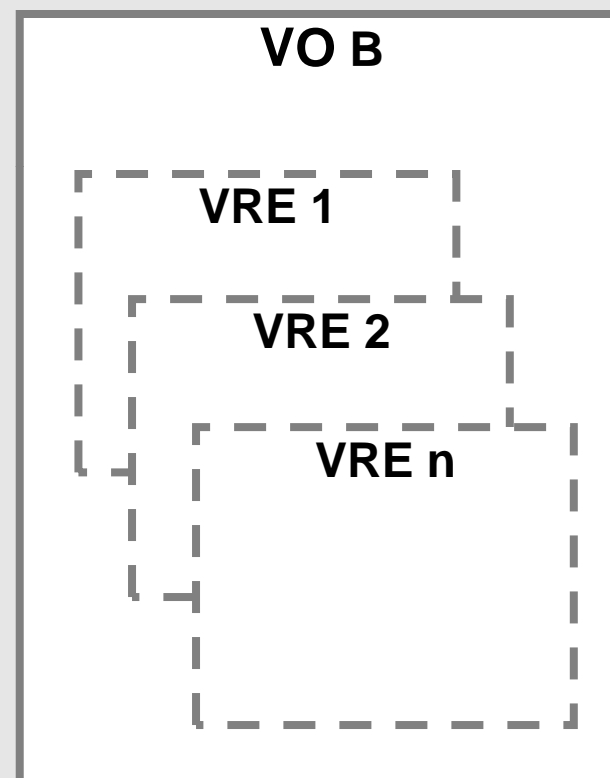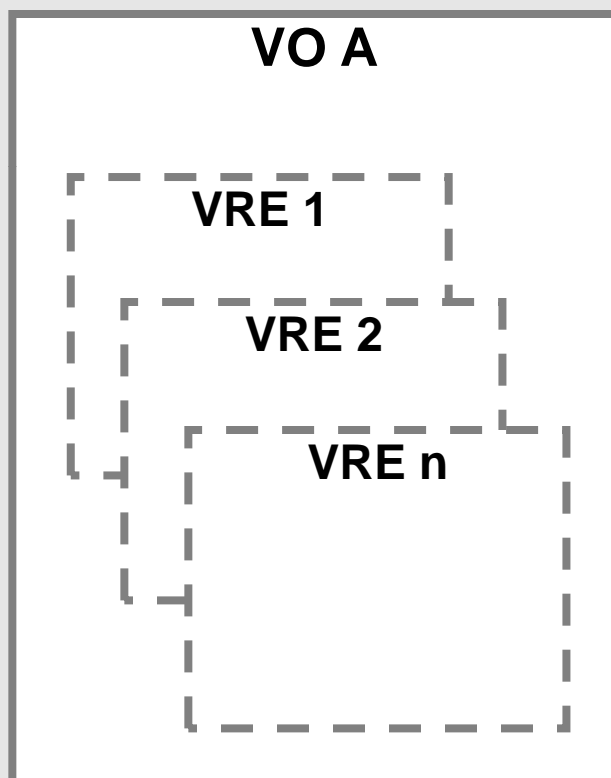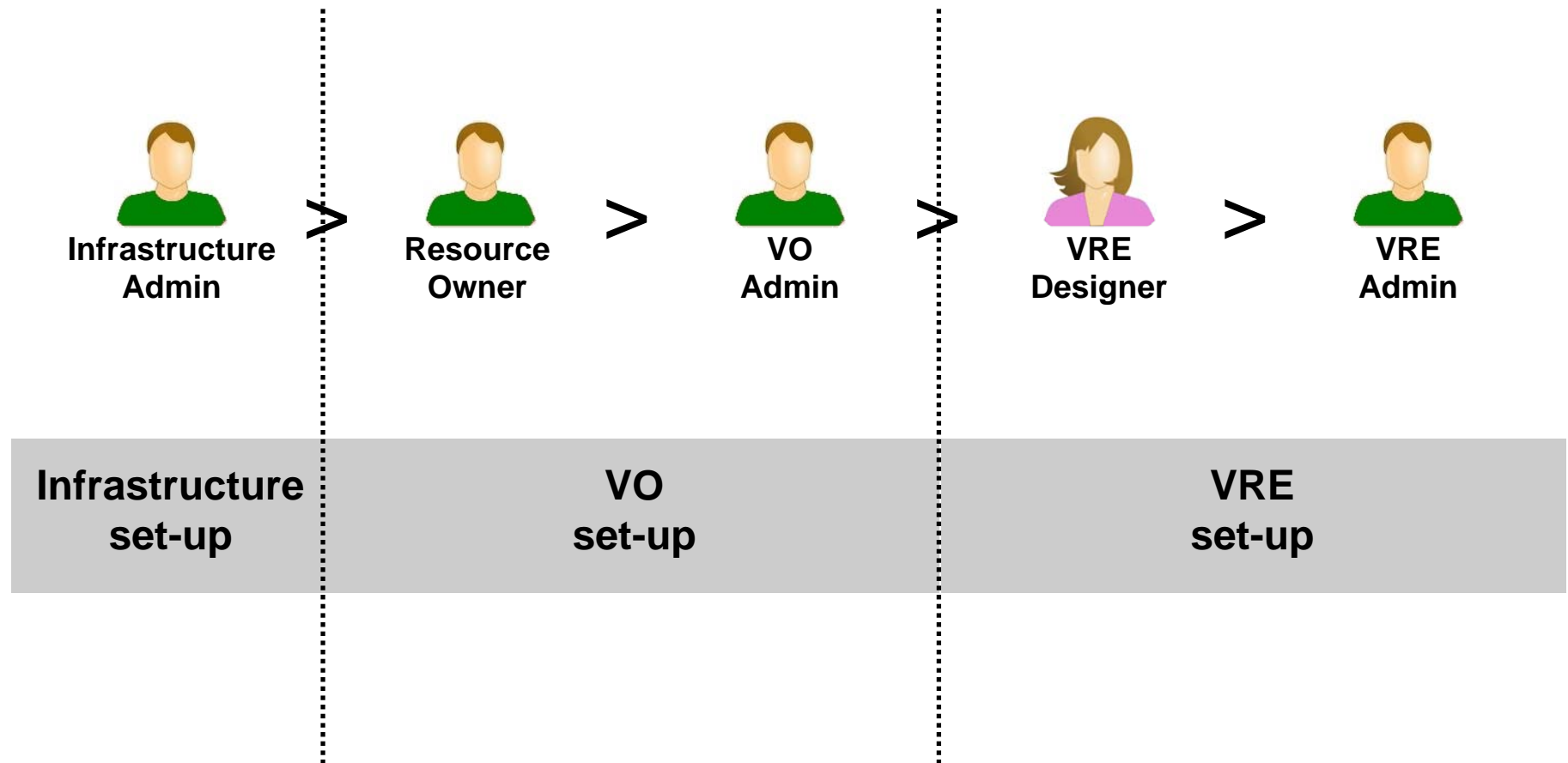- BPEL complaint

VRE Specific

VRE management services highlights

- Management of gCube resources
- Remote deployment
  - no manual deployment overheads
- Environment configuration
- Lifetime management
- Service provision continuity
- Optimal placement of services

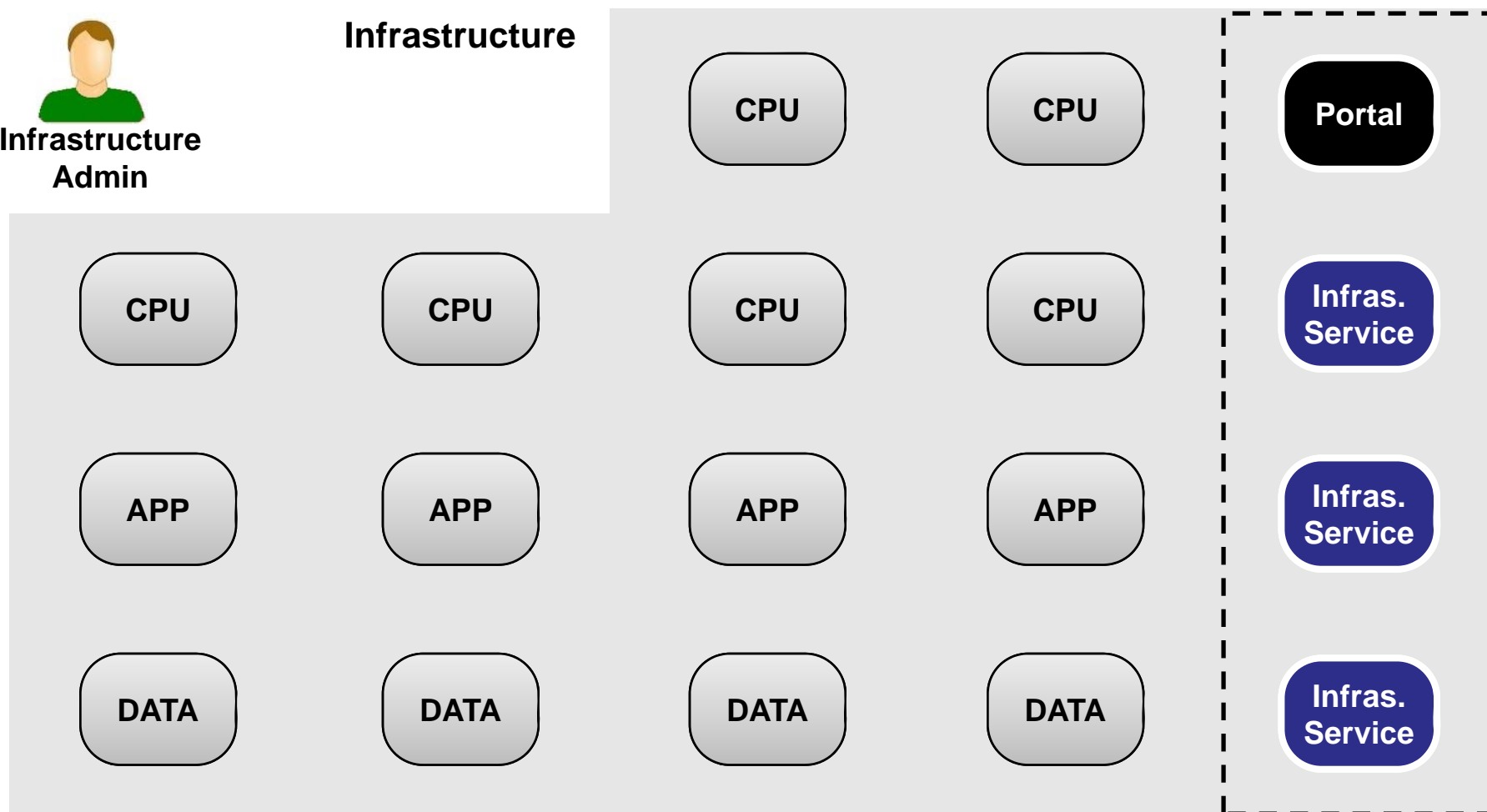**Infras. Specific**

gCube Infrastructure

VO A

VRE 1

VRE 2

VRE n

VO B

VRE 1

VRE 2

VRE n

Infrastructure Admin > Resource Owner > VO Admin > VRE Designer > VRE Admin

Infrastructure set-up	VO set-up	VRE set-up

Infrastructure

Infrastructure Admin

		CPU	CPU	Portal
CPU	CPU	CPU	CPU	Infras. Service
APP	APP	APP	APP	Infras. Service
DATA	DATA	DATA	DATA	Infras. Service

gCube

Virtual
Organisation

Resource
Admin

VO
Admin

CPU

CPU

APP

APP

DATA

DATA

gHN

gHN

Portal

gLite

gLite

Infras.
Service

APP

APP

Infras.
Service

DATA

DATA

Infras.
Service

gCube

Virtual Research Environment

VRE Designer / VRE Admin

CPU CPU

APP APP

DATA DATA

VRE Service VRE Service **Portal**

gLite gLite **Infras. Service**

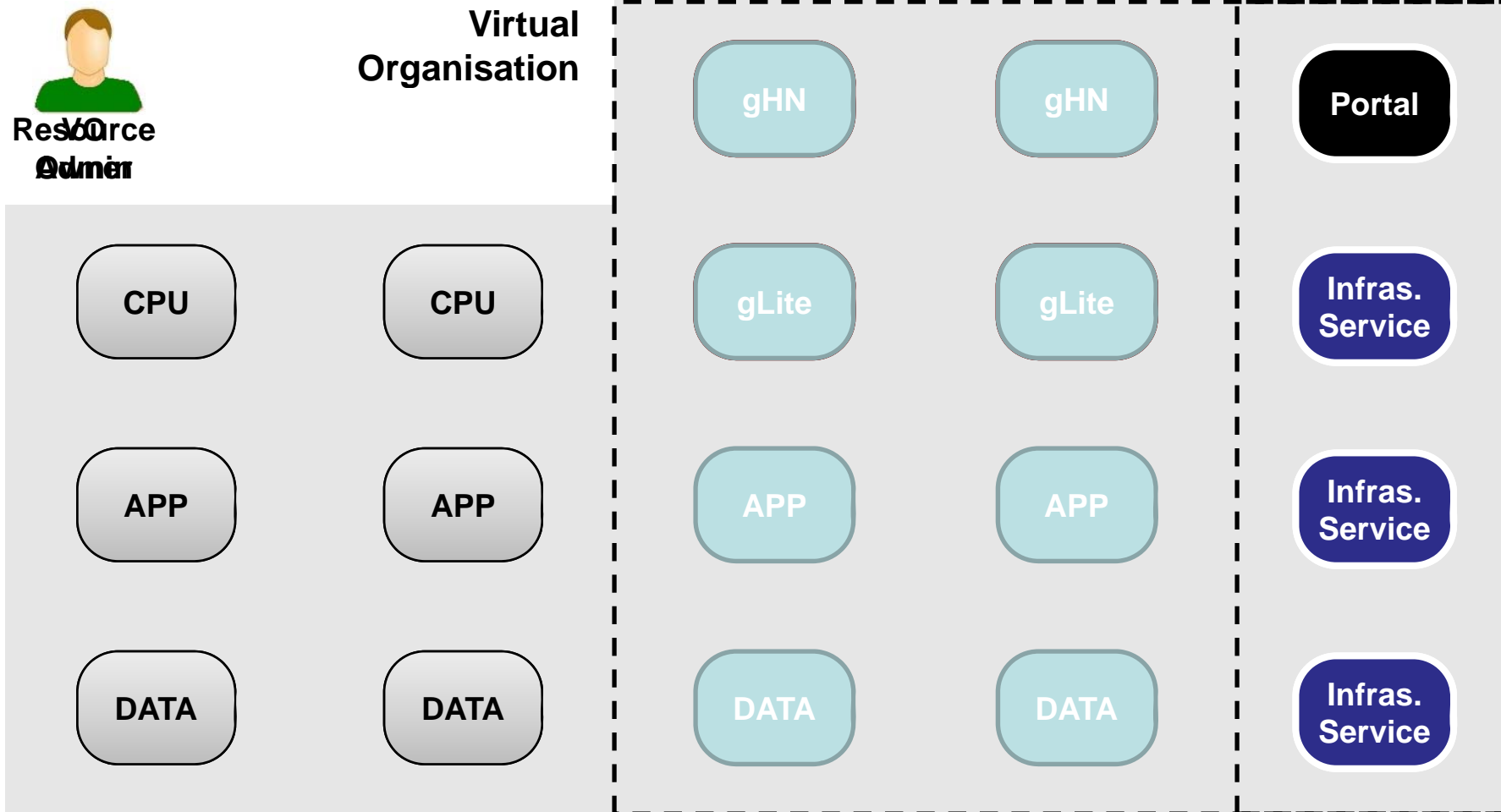APP APP **Infras. Service**
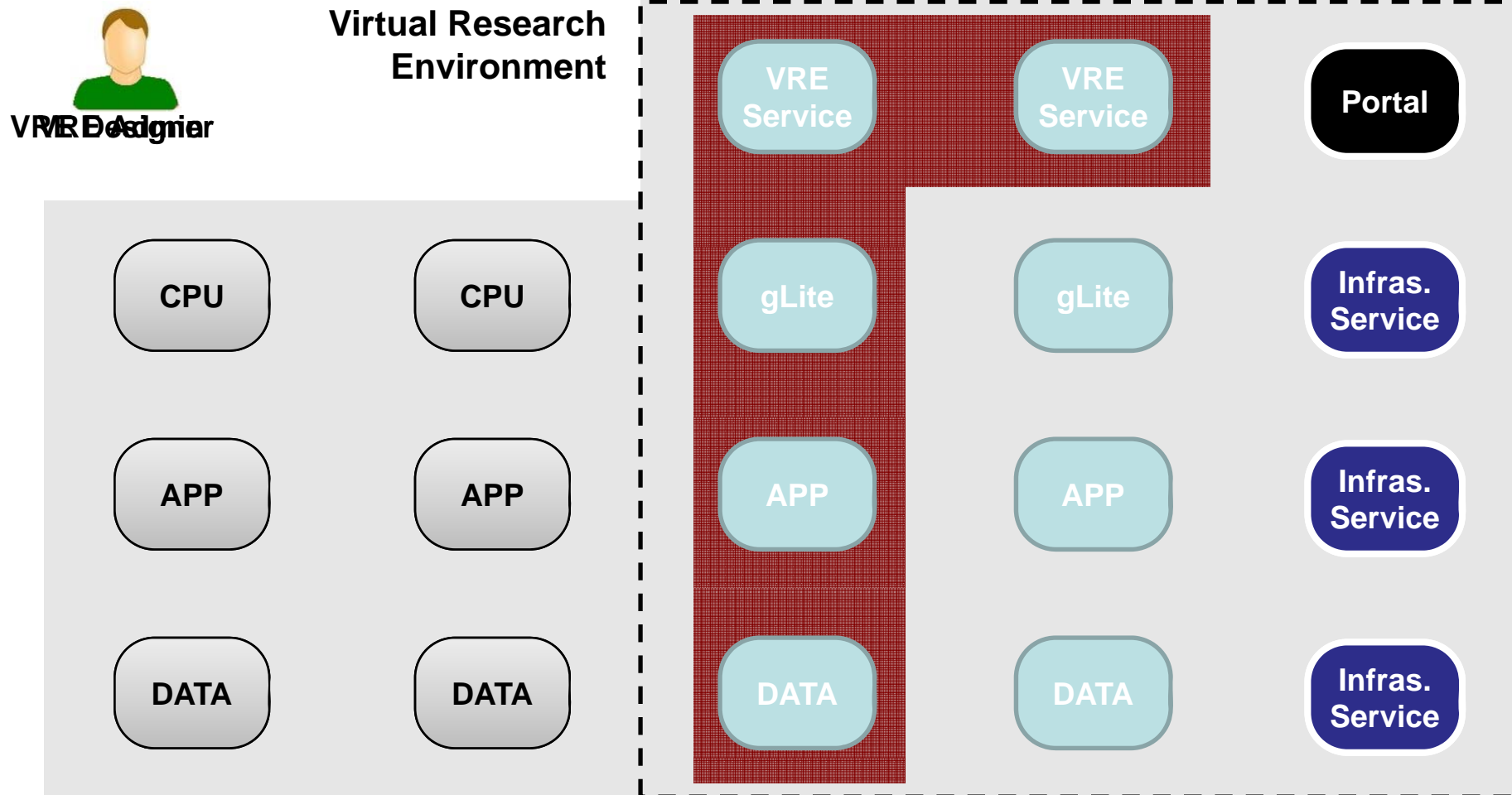
DATA DATA **Infras. Service**

DILIGENT
pre-prod
infrastructure
example

VO
ImpECt
(from ESA)

VRE
Vegetation

* gLite PPS sites
already available

SCOPE	ACTION	TIME
Infrastructure	install collective layer	< 1 day
	install portal	< 1 day
VO	install 1 DHN	< 10 min
	register resources (DHN, data)	< 1 min
	approve resource (DHN, data)	< 1 min
	data import (metadata, indexes)	hours
	manage users	< 10 min
VRE	define VRE	< 10 min
	approve VRE	< 1 min
	deploy VRE	< 2 hour
	modify VRE	< 1 hour

D4Science objective is to "deploy, consolidate and expand the e-Infrastructures built so far by the EGEE and DILIGENT projects to address the needs of several new scientific communities affiliated with the disciplines of Environmental Monitoring and Fishery Resources Management"
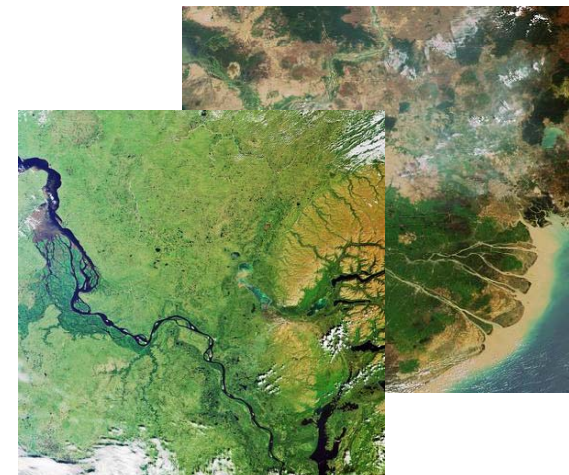
Environmental Monitoring

- European Space Agency
- Provide political and technological solutions to global environmental issues (forest ecosystem, air quality, etc.)
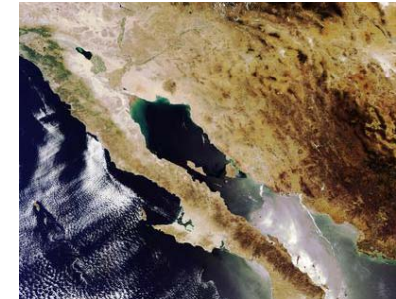

Fishery Resources Management

- FAO Fishery Dep. and WorldFish Centre
- Facilitate and secure the long term sustainable development and utilization of the world's fisheries and aquaculture

gCube

- Duration: 24 months
  - from January 08 to December 09
- Budget: 3 916 735 €
  - 3 150 000 € EC-funded
- Consortium: 11 partners
  - from 7 countries

- http://www.d4science.eu
- http://www.gcube-system.org
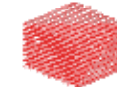
Virtual Research Environments are powerful collaborative workspaces. gCube provides the tools to easily create them.

gCube provides a vast number of VRE services for Content Management, Workflow Management, Information Retrieval, VREs Management, etc.

The creation, management and exploitation of VREs using gCube has been successfully tested in DILIGENT

- A new VO can join the infrastructure in less then 1 day
- A new VRE can be deployed in less then 2 hours

gCube

THANK YOU !