



Interoperable execution of e-Science applications on Grids and Clouds through open standards

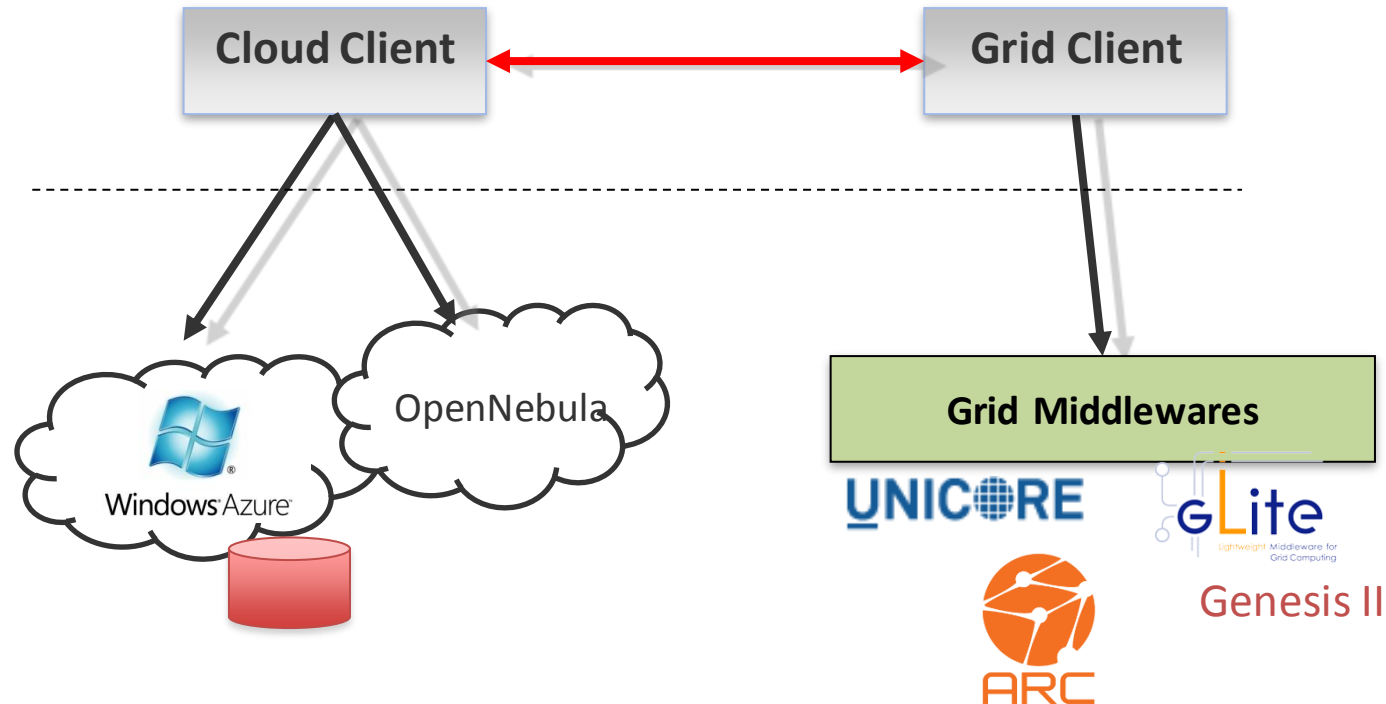
UNICORE Summit 2012, Dresden

31. Mai 2012 **Daniele Lezzi, Shahbaz Memon, Roger Rafanell, Hakan Soncu, Morris Riedel, Rosa M. Badia**

Outline

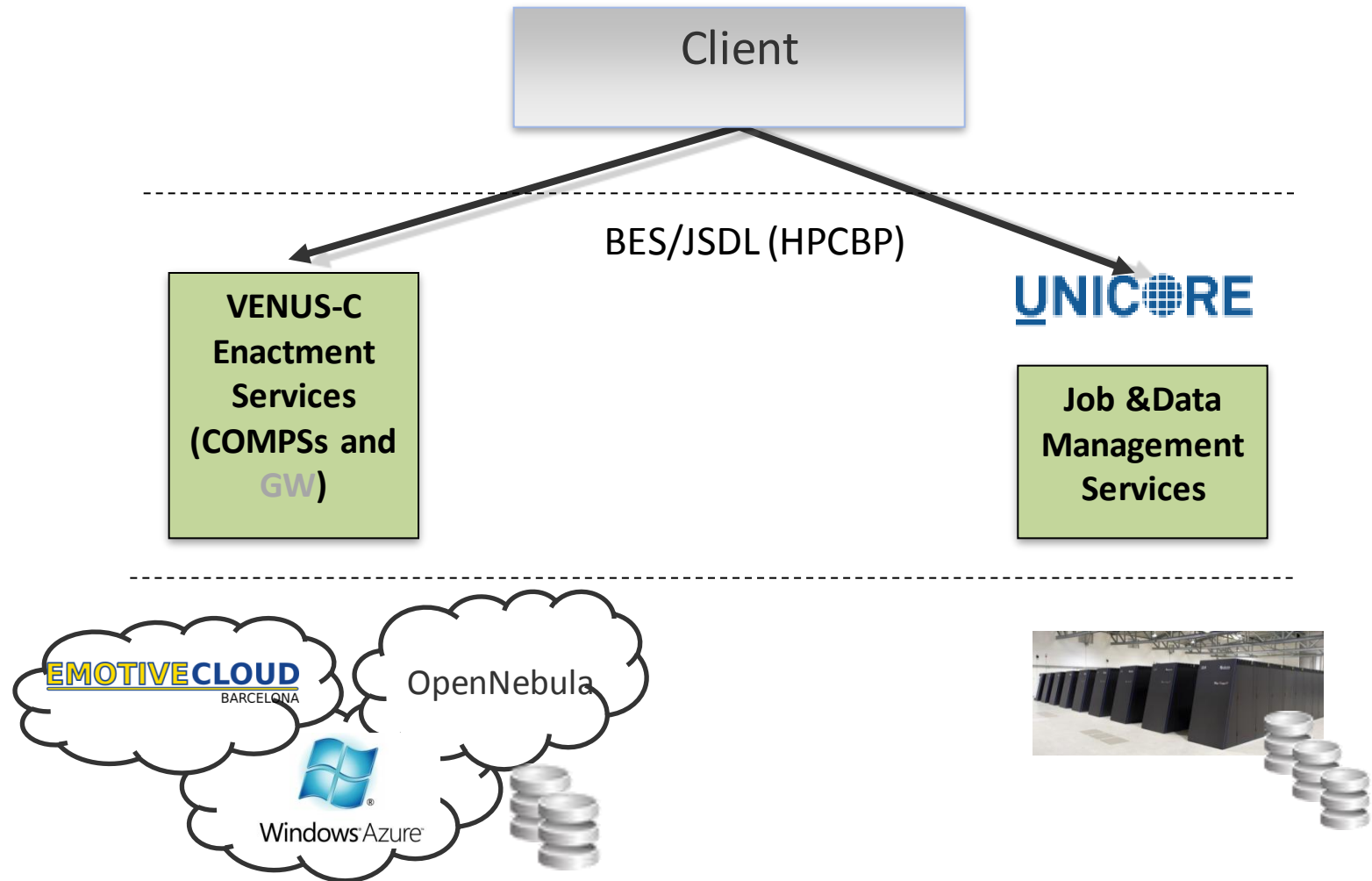
- Motivation
- Approach
- Scenario
- Implementations
 - UNICORE-BES Implementation
 - The Venus-C Platform
 - The Venus-C Enactment Service (BES & JSDL)
 - Generic Worker Service
- Challenges
- Conclusion

Motivation



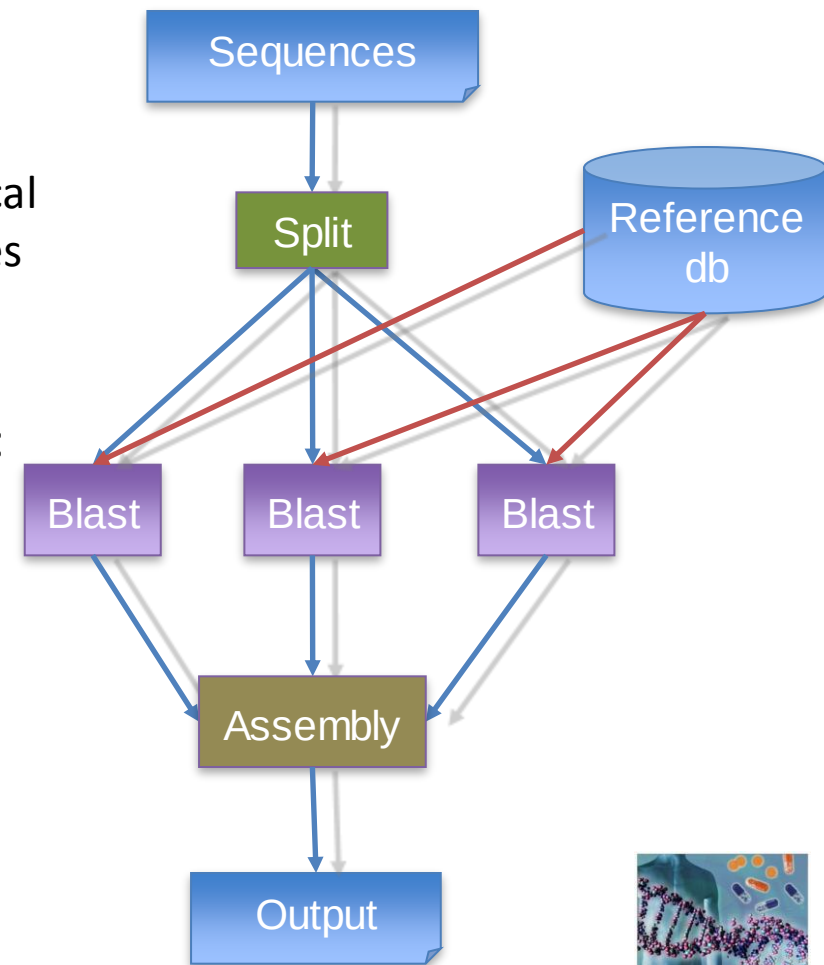
Need to introduce adaptors for each service type!

Using Standard Interface

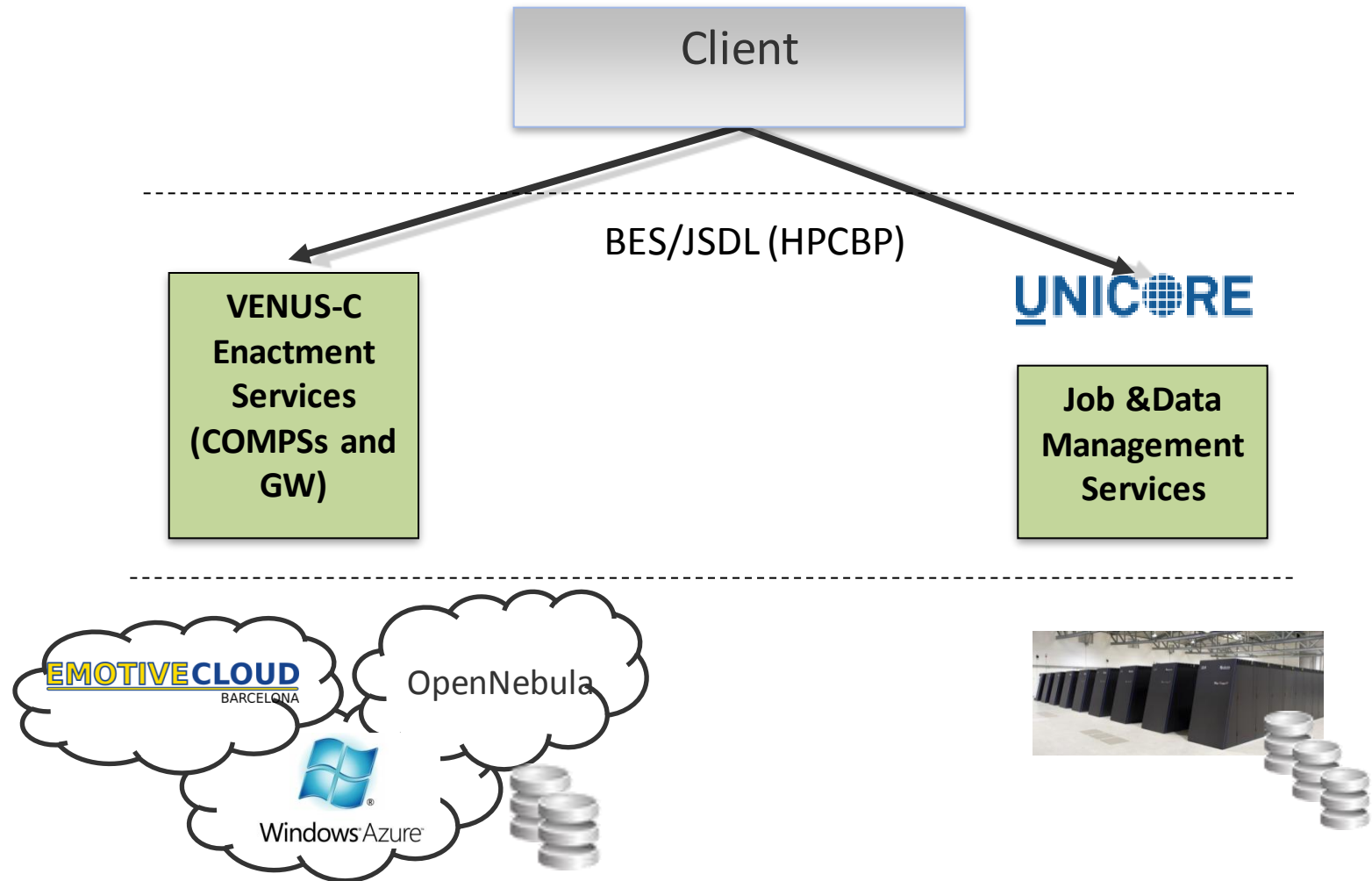


Bioinformatics Scenario: BLAST

- BLAST (Basic Local Alignment Search Tool):
 - An algorithm for comparing primary biological sequence information (amino-acid sequences of different proteins, nucleotides of DNA sequences)
- Scenario: Alignment of 1000 Metagenomes:
 - Input file: Saragasso's Sea metagenome
 - Reference: SwissProt (258 MB)
 - Algorithm: blastx
 - eValue: 0,1
 - Output: 14MB
 - COMPS workflow implementation

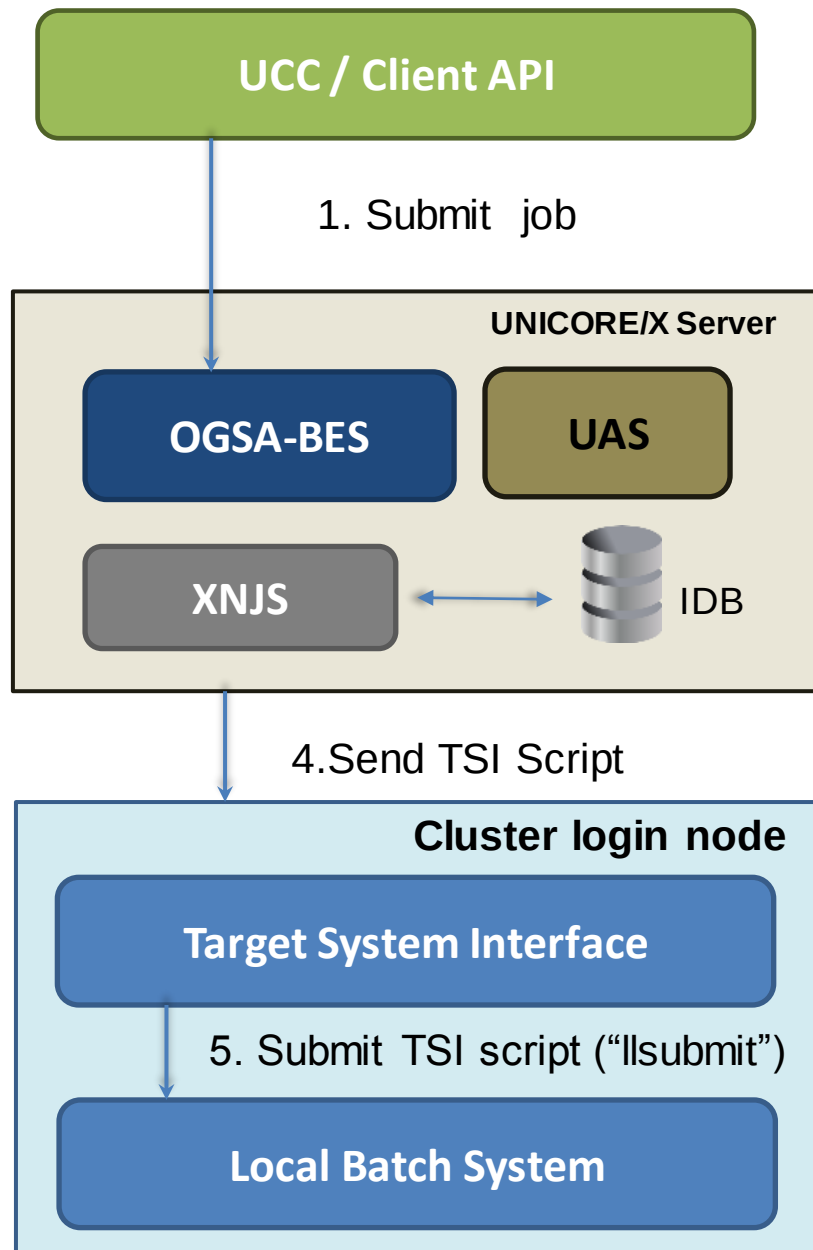


Using Standard Interface



UNICORE-BES

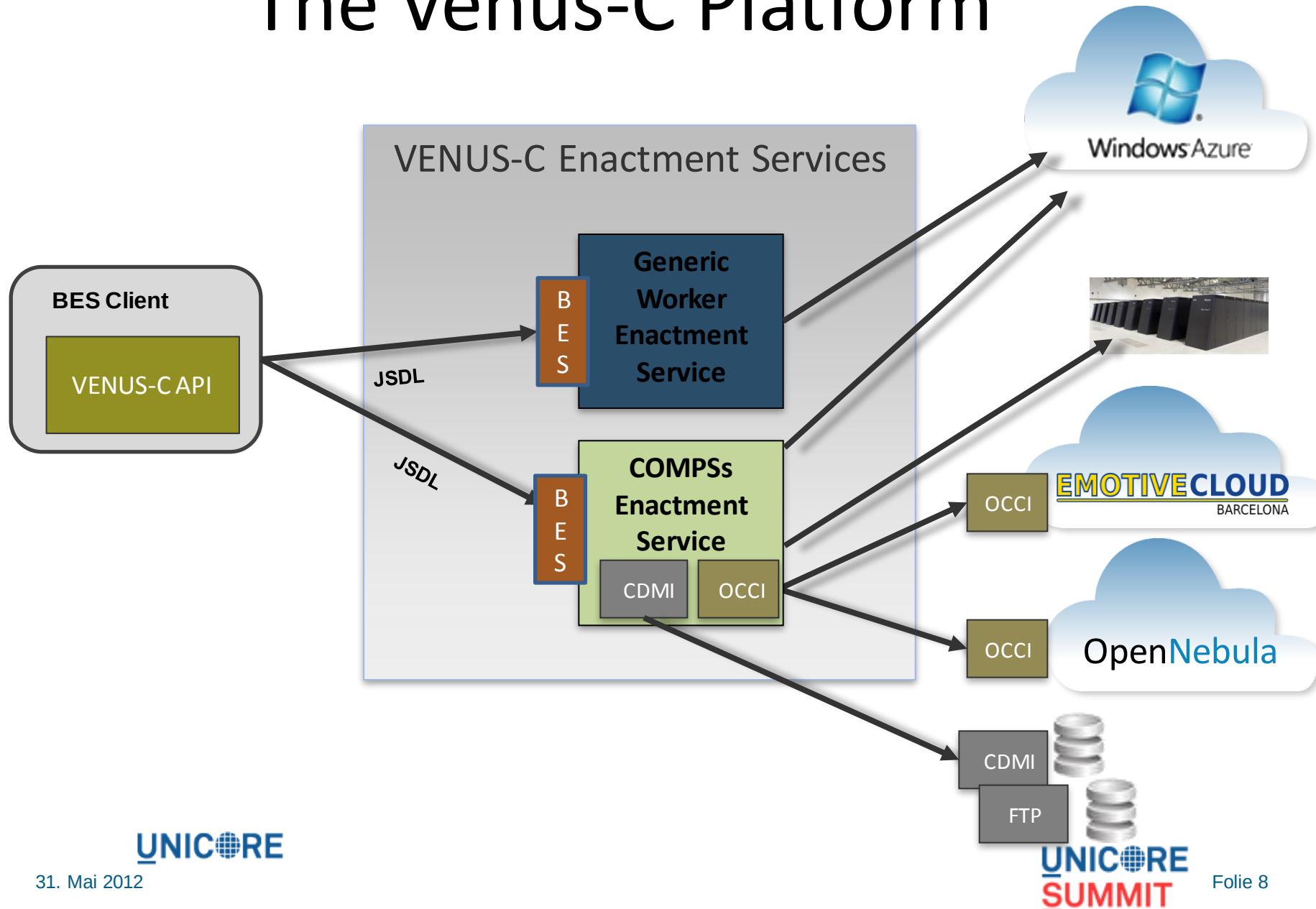
UNICORE



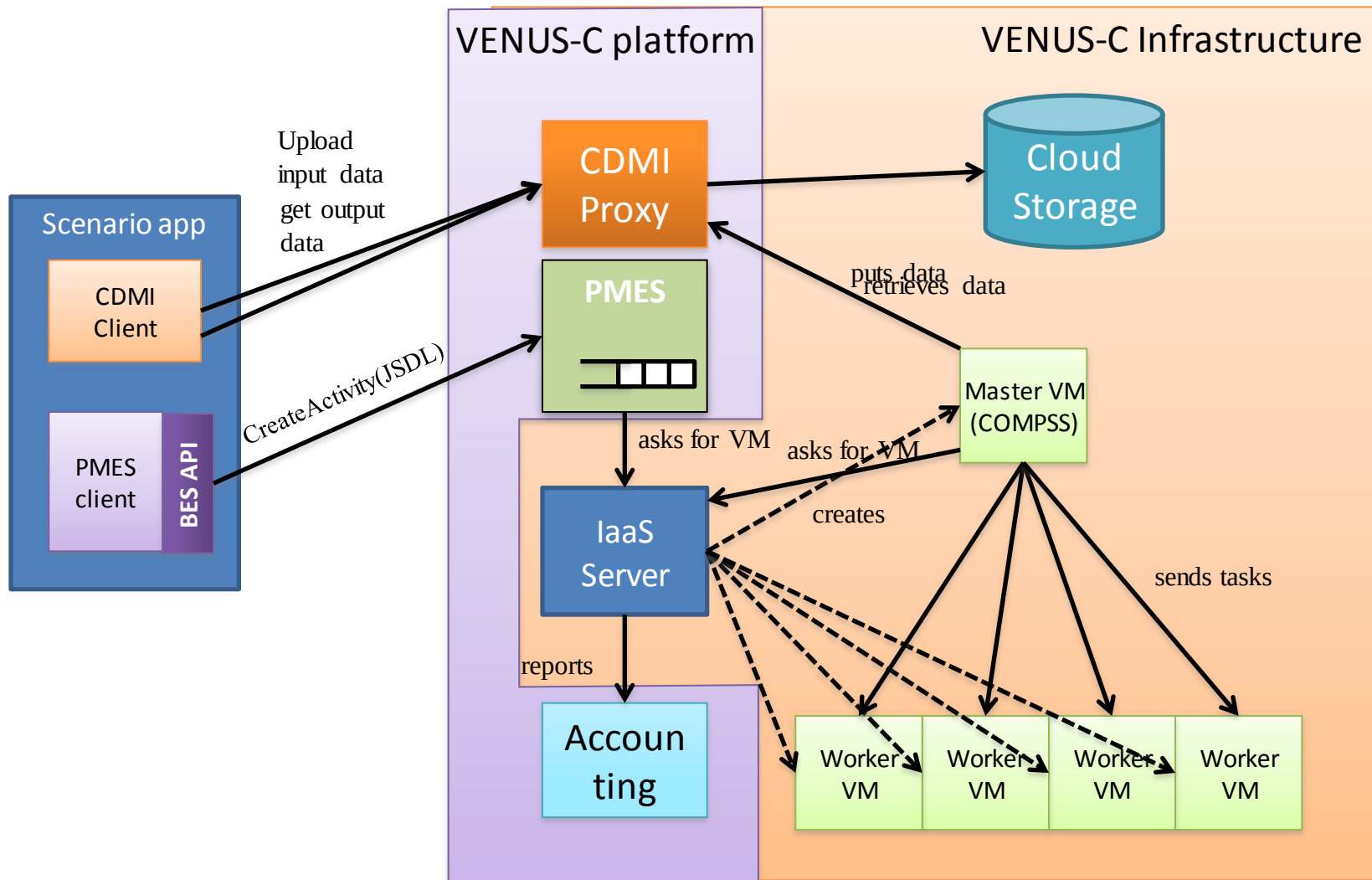
2. Remote request received by OGSA-BES, UAS

3. XNJS generates a TSI script from the job description and local configuration stored in IDB file

The Venus-C Platform

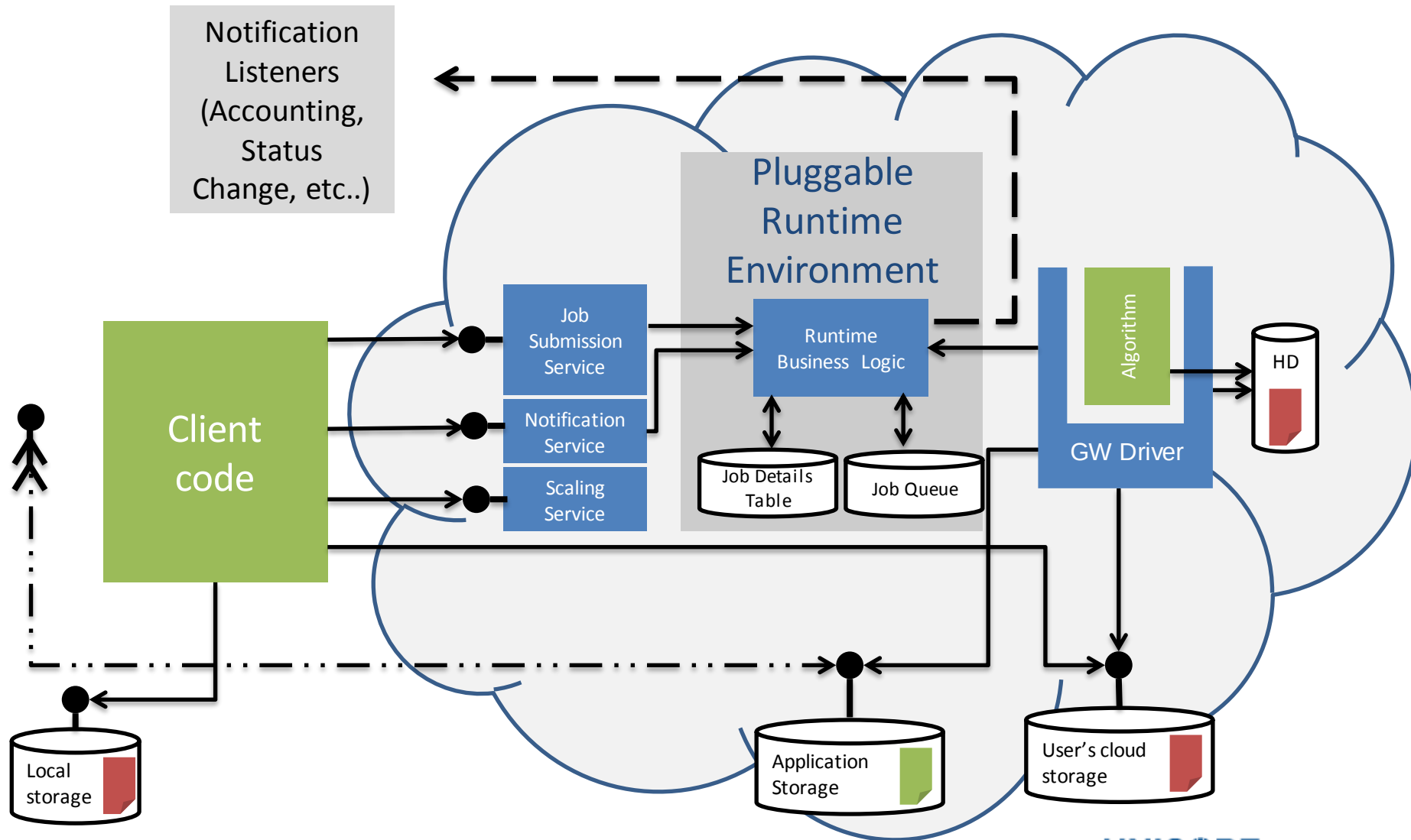


COMPSs Enactment Service



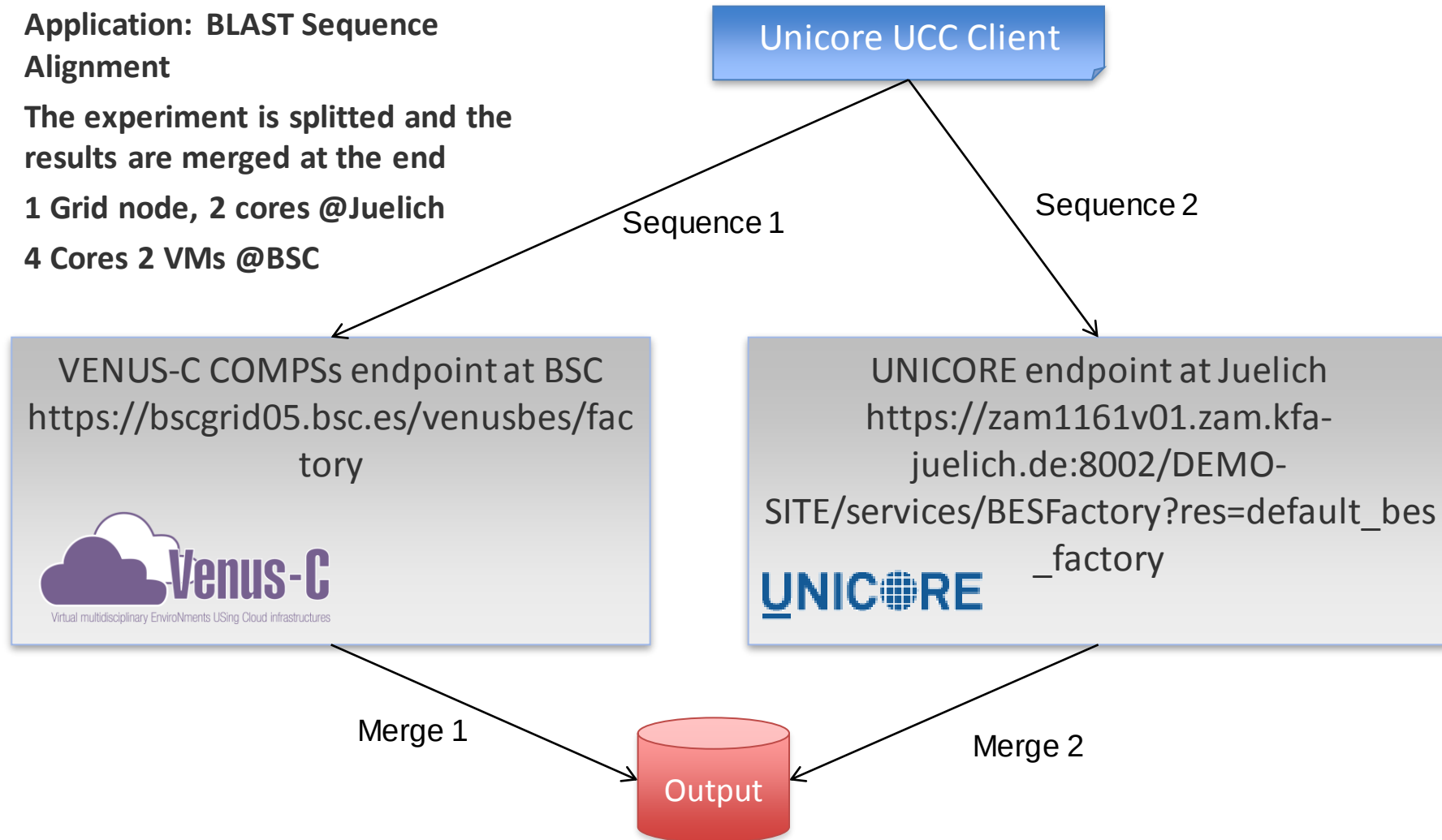
Generic Worker

Researcher



Application Run

- ❑ **Application: BLAST Sequence Alignment**
- ❑ **The experiment is splitted and the results are merged at the end**
- ❑ **1 Grid node, 2 cores @Juelich**
- ❑ **4 Cores 2 VMs @BSC**



JSDL Instance - HPCProfile

```
<HPCProfileApplication >  
  <Executable>runcompss</Executable>  
  <Argument>blast.Blast</Argument>  
  <Argument>true</Argument>  
  <Argument>blastall</Argument>  
  <Argument>COMPSs/databases/swissprot</Argument>  
  <Argument>sargasso_1MB_small.fasta</Argument>  
  <Argument>2</Argument>  
  <Argument>blast</Argument>  
  <Argument>out-unicore.txt</Argument>  
  <Argument>-v 10</Argument>  
  <Argument>-b 10</Argument>  
  <Argument>-e 1e-10</Argument>  
  <Output>stdout-unicore</Output>  
  <Error>stderr-unicore</Error>  
</HPCProfileApplication>
```

OGSA-BES Extensions

Name	GetJobs
Description	Retrieves the jobs of a specific owner
Method Signature	public List<EndPointReferenceType> GetJobs(String owner)
Name	GetAllJobs
Description	Retrieves list of all jobs in the system
Method Signature	List<EndPointReferenceType> GetAllJobs()
Name	GetAllJobsByGroup
Description	Retrieves a list of all jobs belonging to a given group of a specific owner
Method Signature	List<EndPointReferenceType> GetJobsByGroup(string owner, string groupName)
Name	GetRoot
Description	Retrieves the root of the job hierarchy
Method Signature	EndPointReferenceType GetRoot(EndPointReferenceType job)
Name	GetHierarchy
Description	Retrieves list of all jobs which are below the given job in the hierarchy; if the root of the hierarchy is given, all the jobs in the hierarchy will be retrieved
Method Signature	List<EndPointReferenceType> GetHierarchy (EndPointReferenceType root)

Notification and Scaling Methods

Name	CreateSubscription
Description	Creates a notification subscription for all statuses
Method Signature	CreateSubscription(EndpointReferenceType job, List<KeyValuePair<string, string>> pluginConfig)
Name	CreateSubscriptionForStatuses
Description	Creates a notification subscription for a list of statuses
Method Signature	CreateSubscription(EndpointReferenceType job, List<JobStatus> statuses, List<KeyValuePair<string, string>> pluginConfig)
Name	Unsubscribe
Description	Remove a notification subscription from the system
Method Signature	Unsubscribe(EndPointReferenceType job)
Name	ListDeployments
Description	Get a list of Deployments (pairs of CloudProviderID, InstanceCount)
Method Signature	List<DeploymentSize> ListDeployments()
Name	UpdateDeployment
Description	Modify the number of instances on a specified cloud provider
Method Signature	void UpdateDeployment(DeploymentSize newDeploymentSize)

Challenges – Cross Interoperability

- Application deployment and configuration
- SOAP version incompatibility (1.1 vs 1.2)
- Issues with JSDL FSP (File Staging Profile) interpretation
- Grid Job stagings from Cloud storage interface (CDMI)

Conclusions

- Demonstrated the use of standards in executing scientific application on heterogeneous platforms
- BLAST application run on Venus-C's COMPSs (Cloud) and UNICORE-BES (Grid) environments
- Identified set of challenges during the interoperability excursions
- Resolve Generic Worker (Windows Azure) issues with SOAP version issues
- Extend BLAST scenario with automated split and merge instead of manual submits i.e. in a workflow
- Open further opportunities for integrating Grid and Cloud technologies
 - Outsourcing, On-demand, Customized Exec. Env., Programming Models..



Contacts

Daniele Lezzi (daniele.lezzi@bsc.es)

Roger Rafanell

Rosa M. Badia

Barcelona Supercomputing Center



Shahbaz Memon

Morris Riedel

Juelich Supercomputing Centre



Hakan Soncu

European Microsoft Innovation Center (EMIC) Aachen



Microsoft
Research

