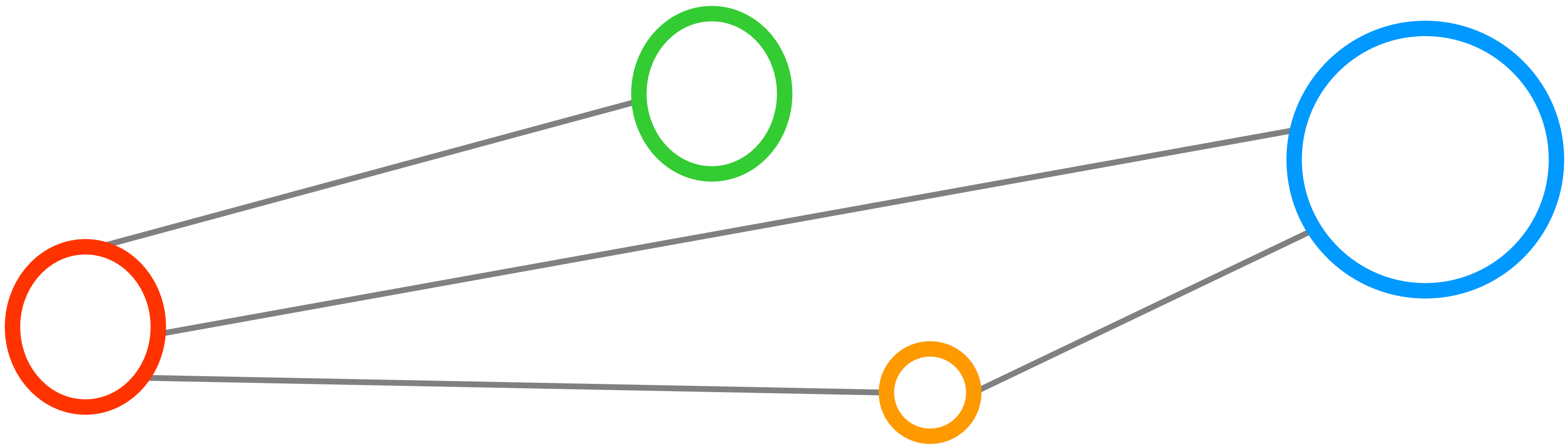
# The Seven Steps towards Interoperability for e-Science

## Lessons learned from standardization in the context of HPC & HTC

Morris Riedel (Jülich Supercomputing Centre & EMI) et al.
Group Co-Chair Grid Interoperation Now & Production Grid Infrastructure

*"…separating the e-science technology & standard hypes from e-science production infrastructure reality today and tomorrow…"*

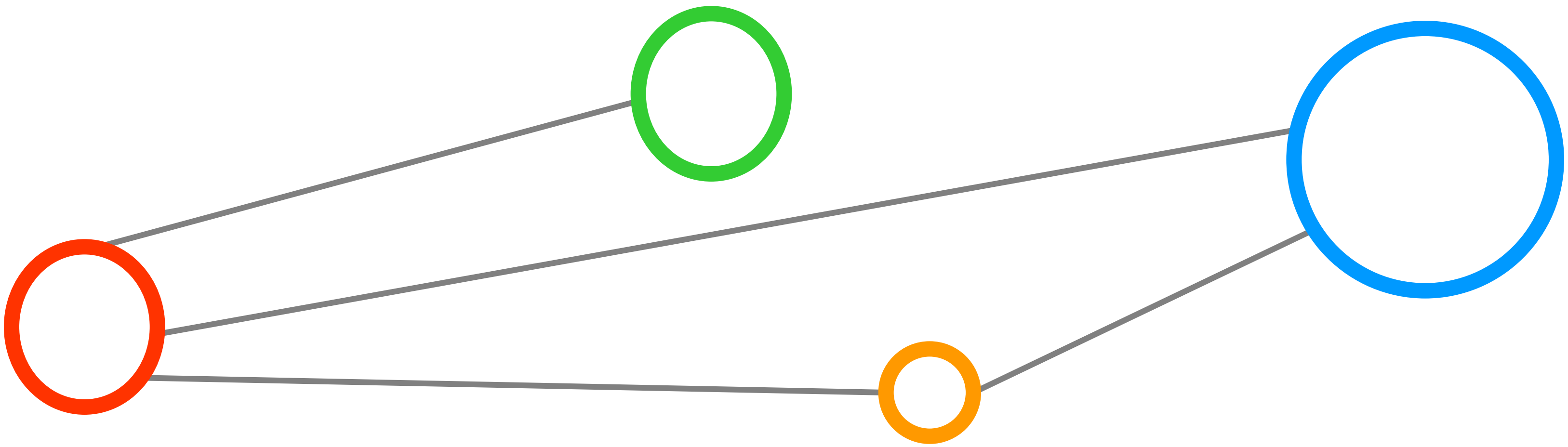
# Outline



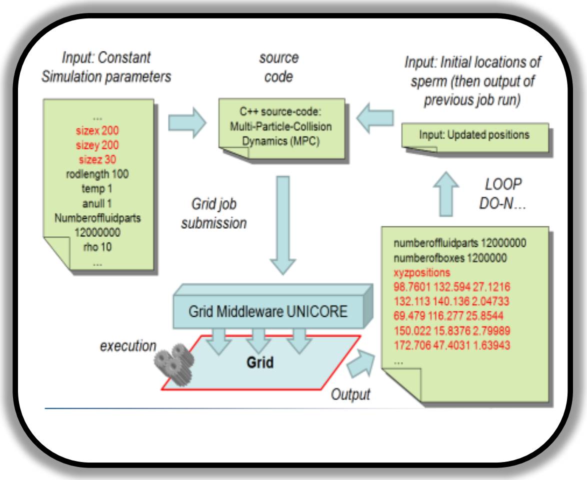
# Outline

- A Design Pattern in e-science
- Interoperability Approaches
- The Seven Steps
- Summary
- References

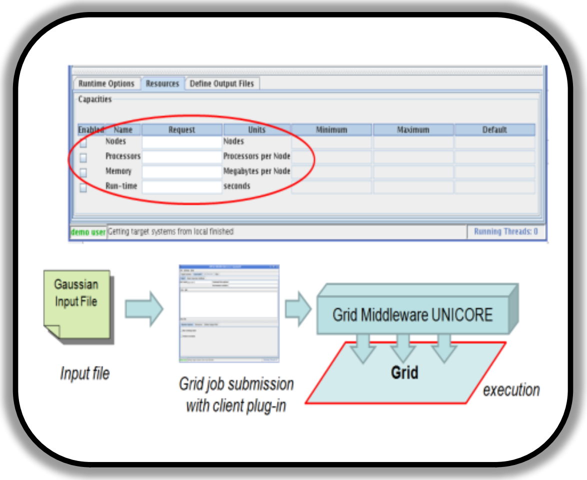
# A Design Pattern in e-Science



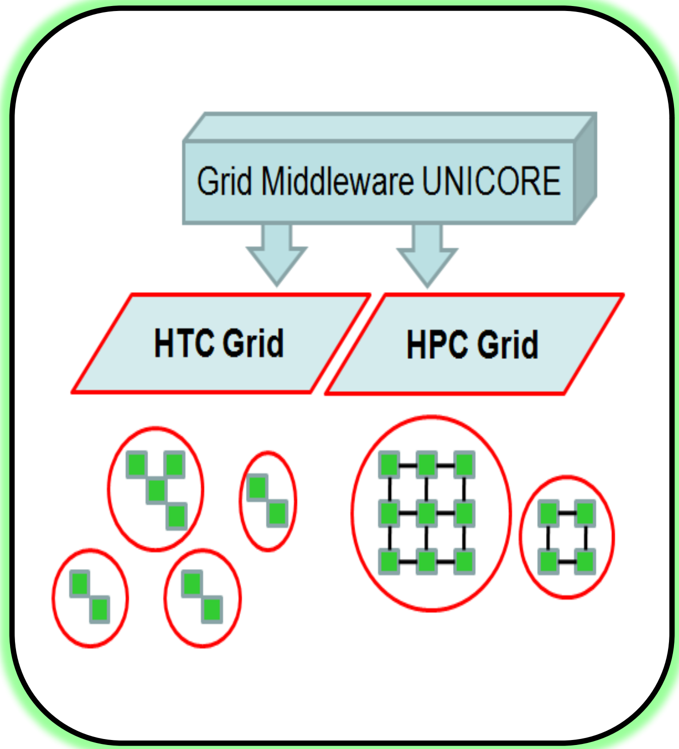
# Different Approaches for e-Science



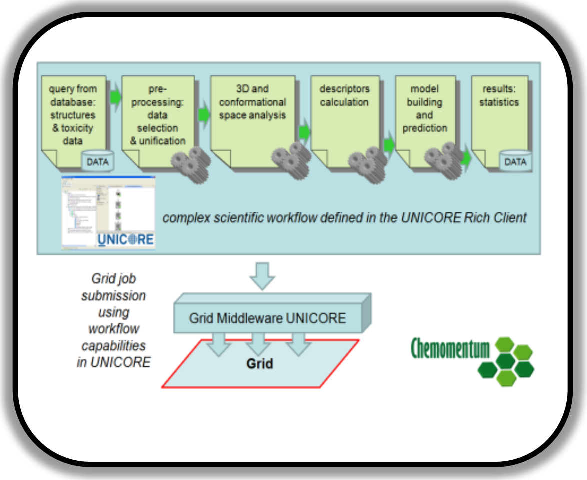

**Simple Scripts & Control**



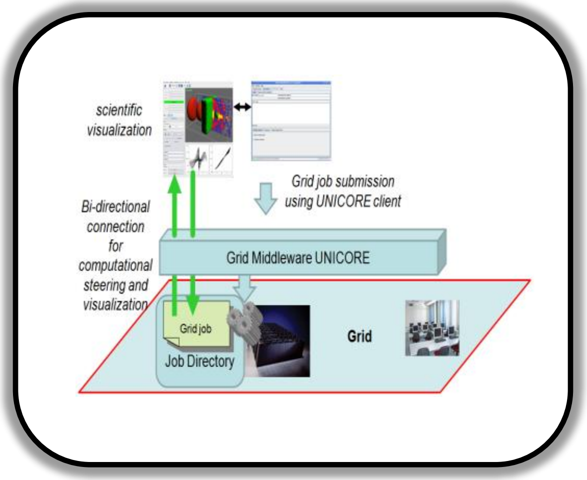

**Application Plug-ins**




**Interoperability of HTC & HPC Infrastructures**




**Complex Workflows**




**Interactive Access**

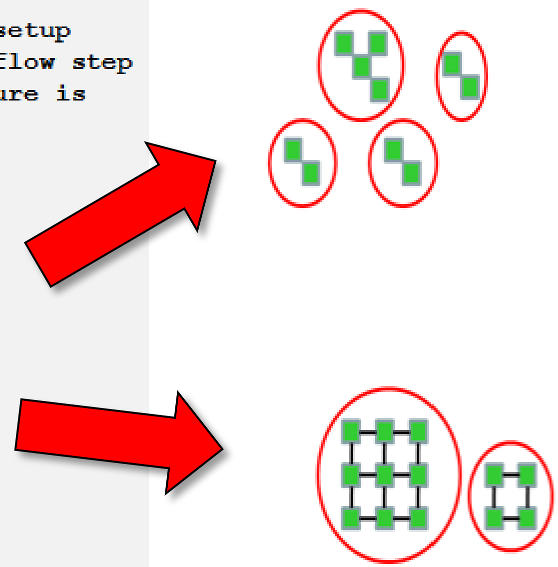
**[3] Riedel and Kranzlmueller et al.,**
***'Classification of Different Approaches for e-Science Applications in Next Generation Computing Infrastructures '***

# A Design Pattern in e-Science

OpenGridForum

```
Begin

1  Begin GridInformationProvisioning
     Grid Information Providers (GIPs) publish pieces of
     information about infrastructures (HPC and HTC resources)
   End

   scienceworkflowfinished = false

   WHILE (scienceworkflowfinished)
     Begin Brokering
       End-user uses client technology (CT) and performs application setup
2        and defines HPC or HTC requirements for next scientific workflow step
       Compute resource (CR) of corresponding HPC and HTC infrastructure is
         found based on the information exposed by GIPs
     End

     Begin JobSubmitToResource
       If CR.type is HTC then
         End-user of CT submits HTC job to a HTC resource
         using middleware MA of the corresponding infrastructure IA
3      End If
       If CR.type is HPC then
         End-user of CT submits HPC job to a HPC resource
         using middleware MB of the corresponding infrastructure IB
       End If
     End

     Begin AnalysisScienceComplete
4      If end-user need no further computing then
         scienceworkflowfinished = true
     End

   End While
```

a new 'toolset' is given…



[1] Riedel et al. *'E-Science Infrastructure Interoperability Guide'*

# Two Case Studies / Use Cases

OpenGridForum

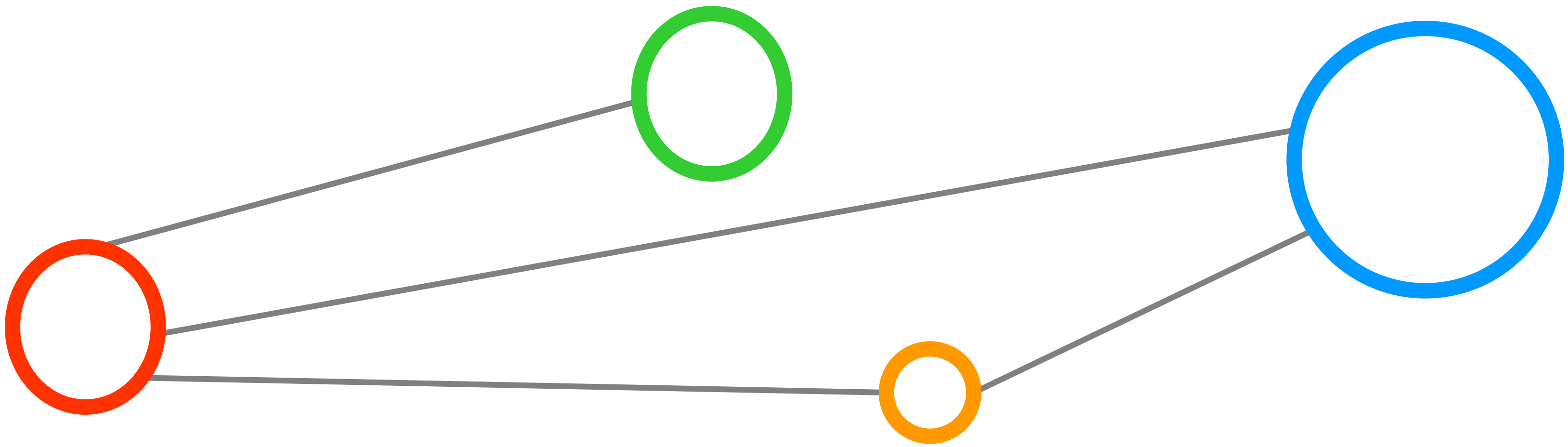
- WISDOM workflow implements design pattern
  - Case study of the bio-informatics domain
  - e-Scientists use HTC infrastructures (EGEE/EGI)
  - … and HPC infrastructure (DEISA)
  - Molecular docking on HTC first, then molecular dynamics on HPC

**[5] Riedel et al. *'E Improving e-Science with Interoperability of the e-Infrastructures EGEE and DEISA'***

- EUFORIA workflow implements design pattern
  - Case study of the ITER fusion domain
  - e-Scientists use HTC infrastructure (EGEE/EGI)
  - …and HPC infrastructure (DEISA)
  - Massive parallel fusion app. on HPC, other fusion app. on HTC

**[4] Memon and Riedel et al. *'Lessons learned from jointly using HTC- and HPC-driven e-science infrastructures in Fusion Science'***

# Interoperability Approaches



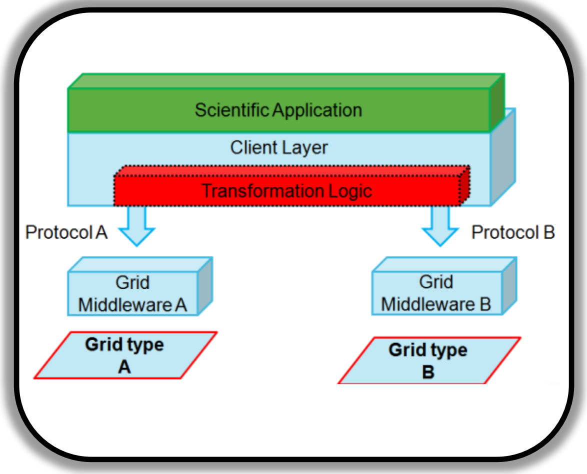
# Approaches: Reference Models

- No reference model exists addressing all relevant factors

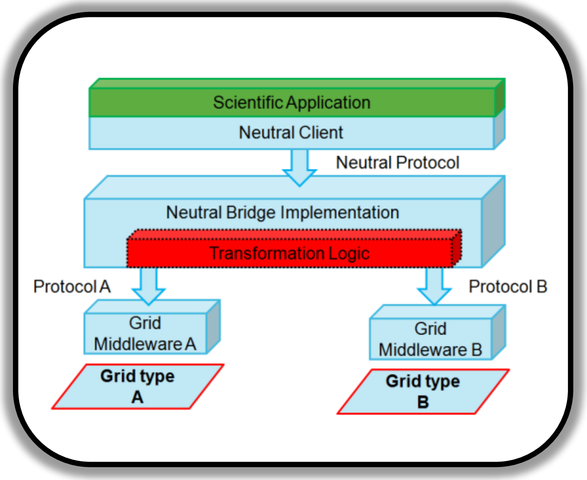
| Relevant Factors | OGSA | EGA | CAA | CSA | CPN |
|---|---|---|---|---|---|
| Service based | yes | yes | yes | yes | no |
| e-Science Context | yes | no | yes | no | yes |
| Detailed enough | no | no | yes | yes | no |
| Realistically implementable | no | no | yes | no | no |
| Standards based | yes | yes | yes | yes | no |
| Adoption in e-science production technologies | no | no | no | no | no |
| Relationships between func-tional areas | no | yes | no | yes | no |

**[1] Riedel et al. *'E-Science Infrastructure Interoperability Guide'***

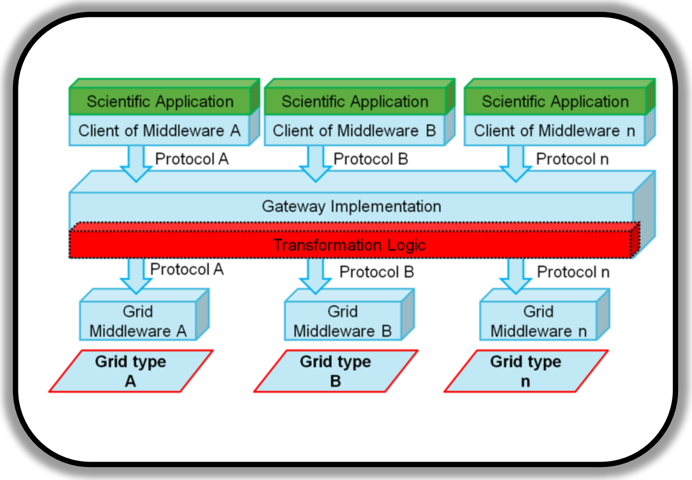
# Approaches: Transformation Logic



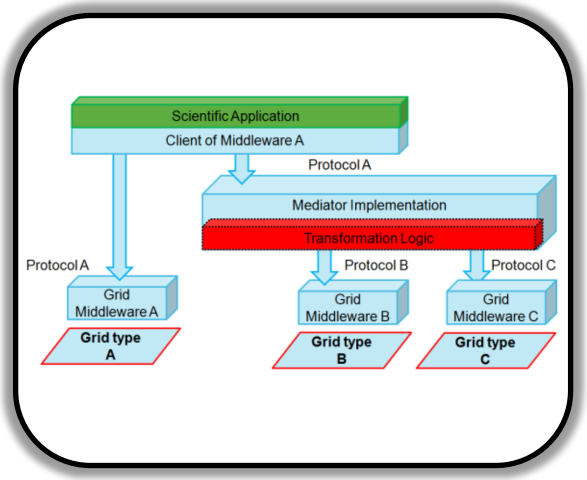

**Client Layer Approach**



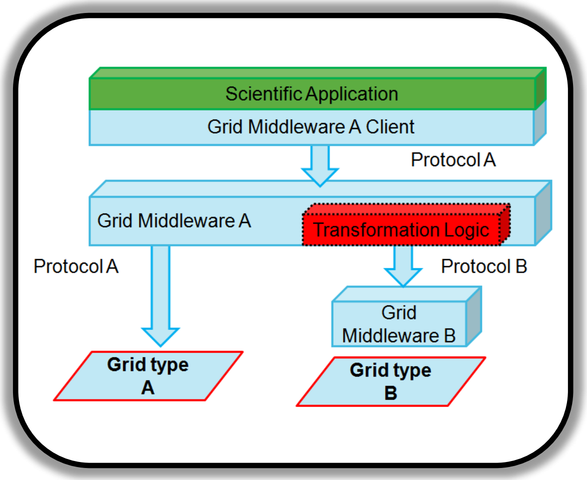

**Neutral Bridge Approach**



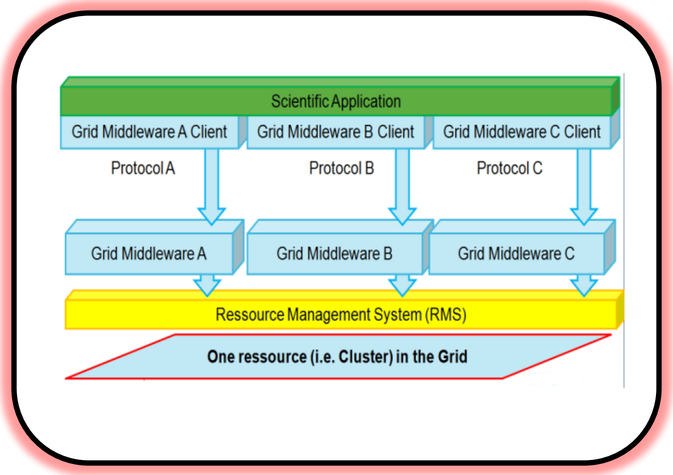

**Gateway Approach**




**Mediator Approach**



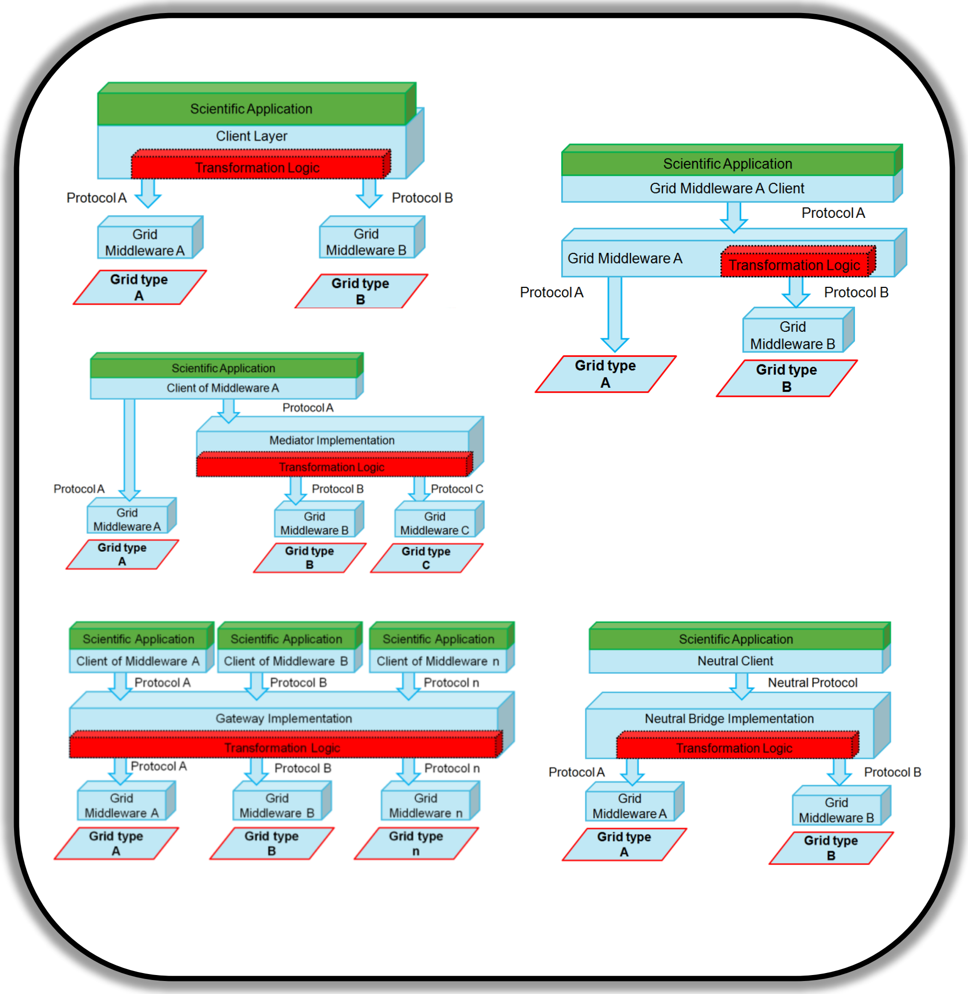

**Adapter Approach**




**Middleware Co-Existence**

**[2] Riedel et al., '*Research Advances by using Interoperable e-Science Infrastructures – The Infrastructure Interoperability Reference Model applied in e-Science* '**

# Transformation Logic




**Approaches that require transformation logic**

protocol A or schema A

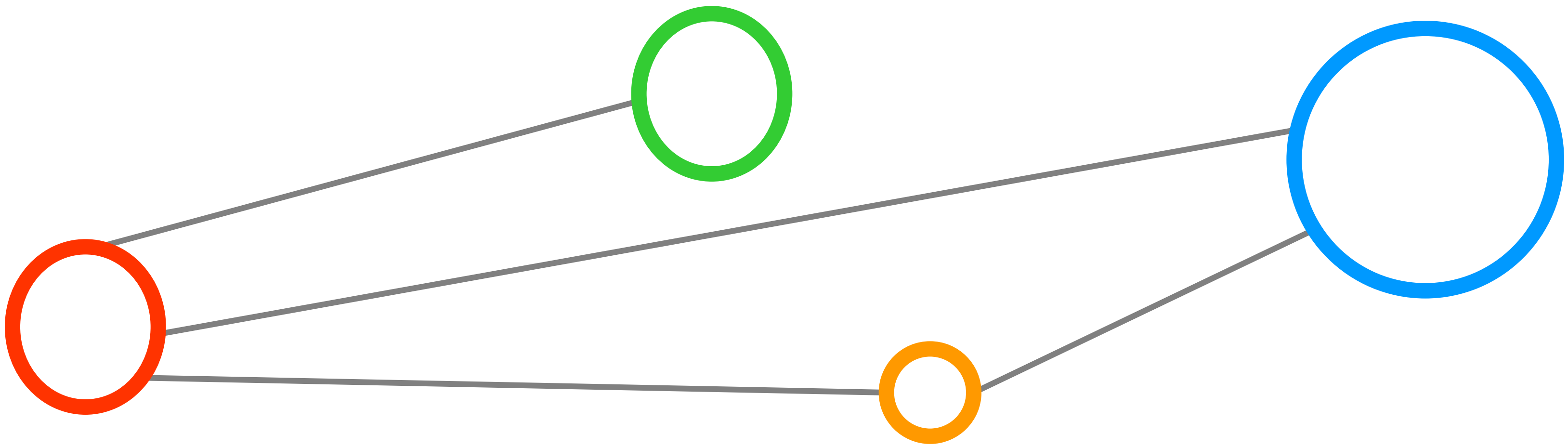
**Transformation Logic**

- Time consuming & error-prone
- Difficult to maintain (n*n versions)
- Production deployment dimension
- Protocol B‘ implies result is often only a subset of protocol B

protocol B‘ or schema B‘

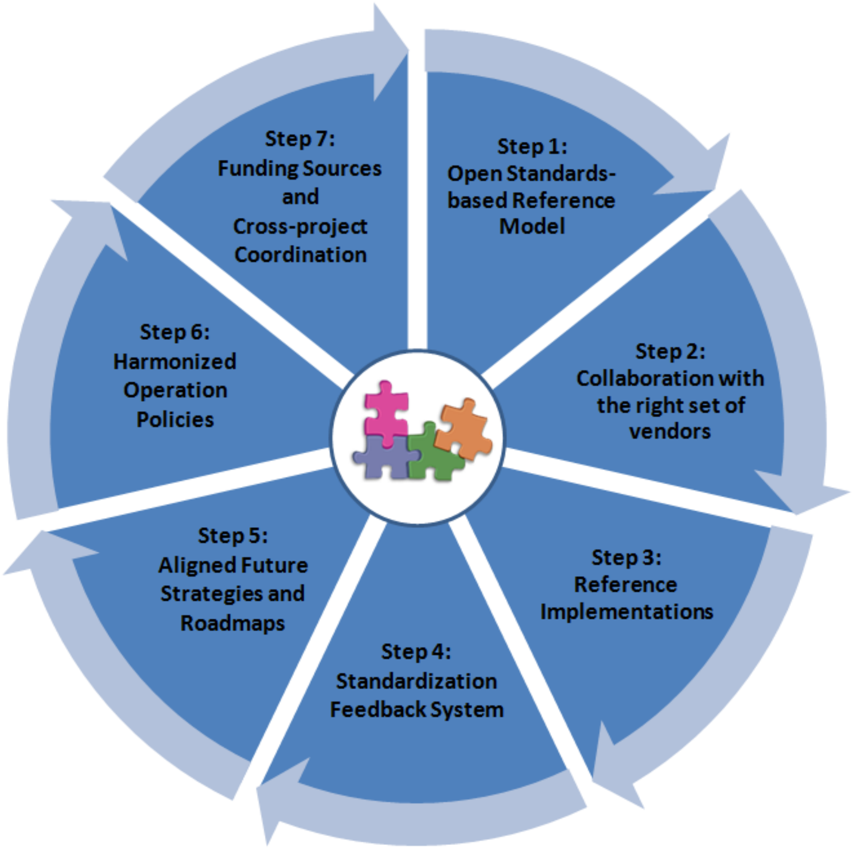

**[2] Riedel et al., ‘*Research Advances by using Interoperable e-Science Infrastructures – The Infrastructure Interoperability Reference Model applied in e-Science* ‘**

# The Seven Steps



# The Seven Steps Process

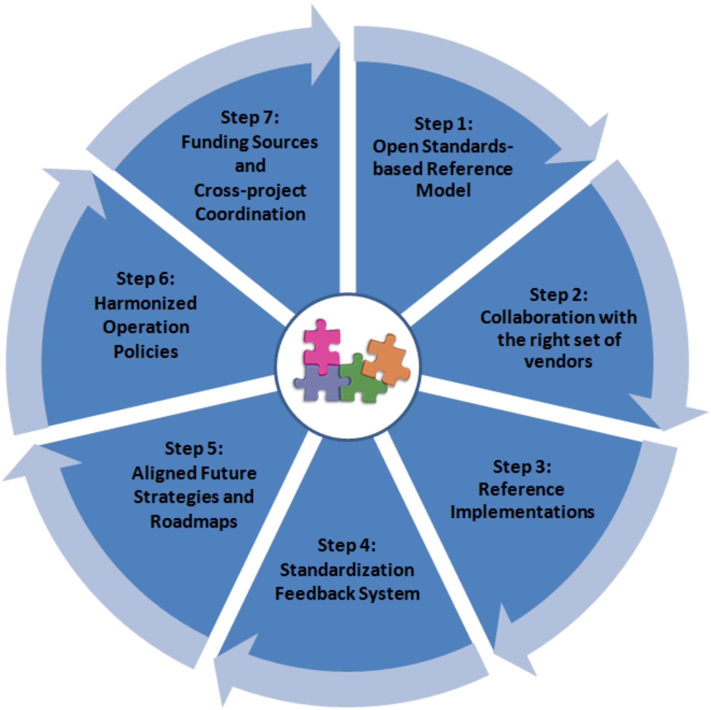
- ‚Seven Steps to e-Science Infrastructure Interoperability‘
- Standards are key to success – but not enough

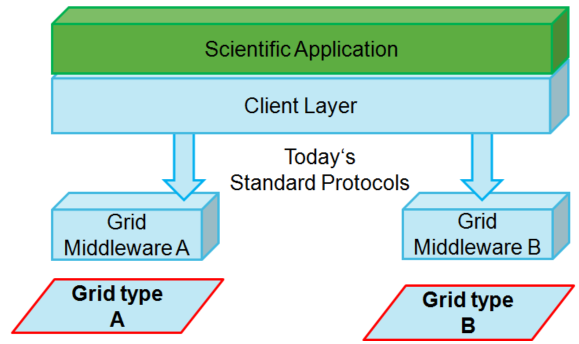



a new ‘mindset’ is required…

[1] Riedel et al. ***‘E-Science Infrastructure Interoperability Guide’***

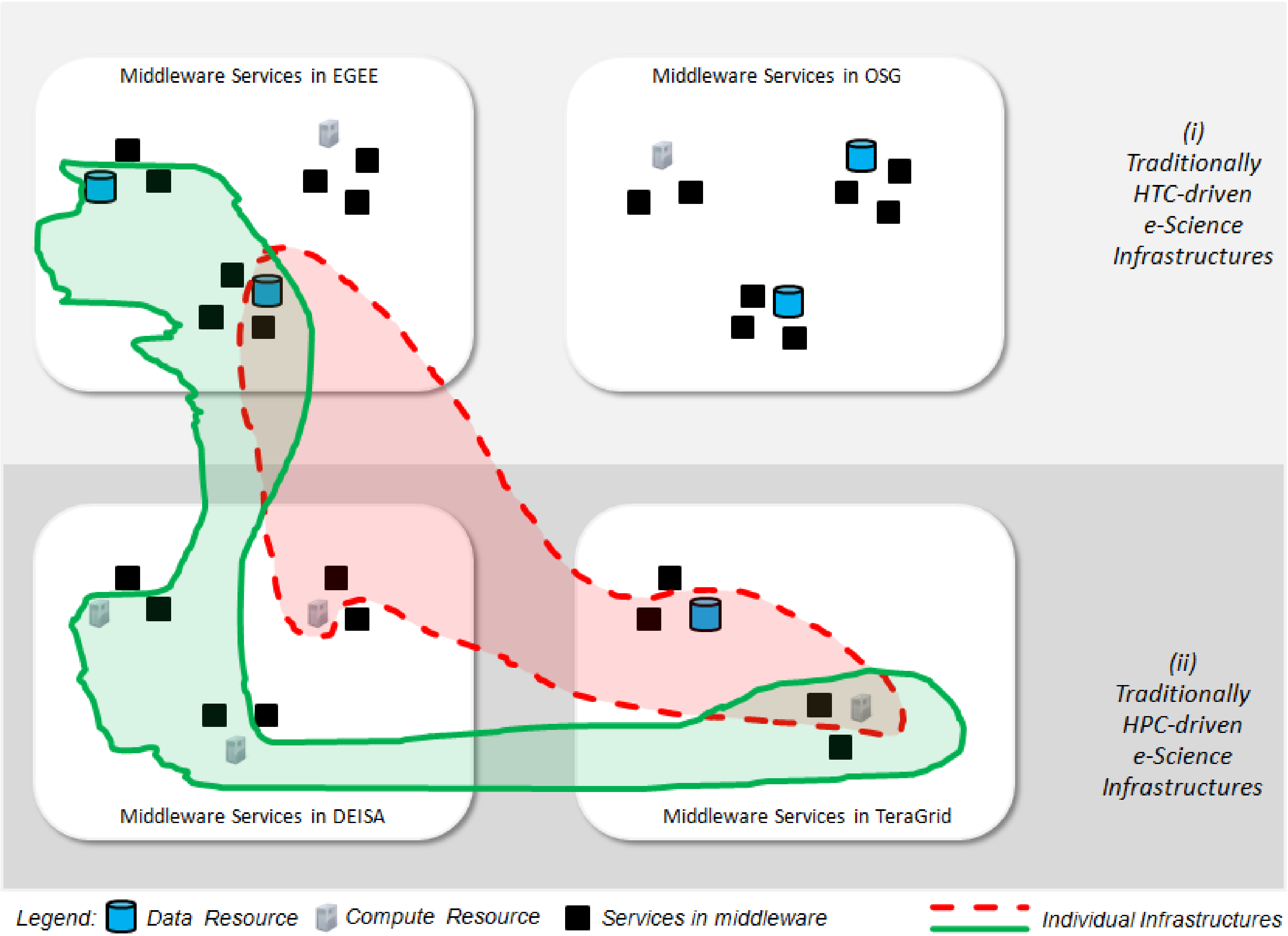
# Step 1: Reference Model



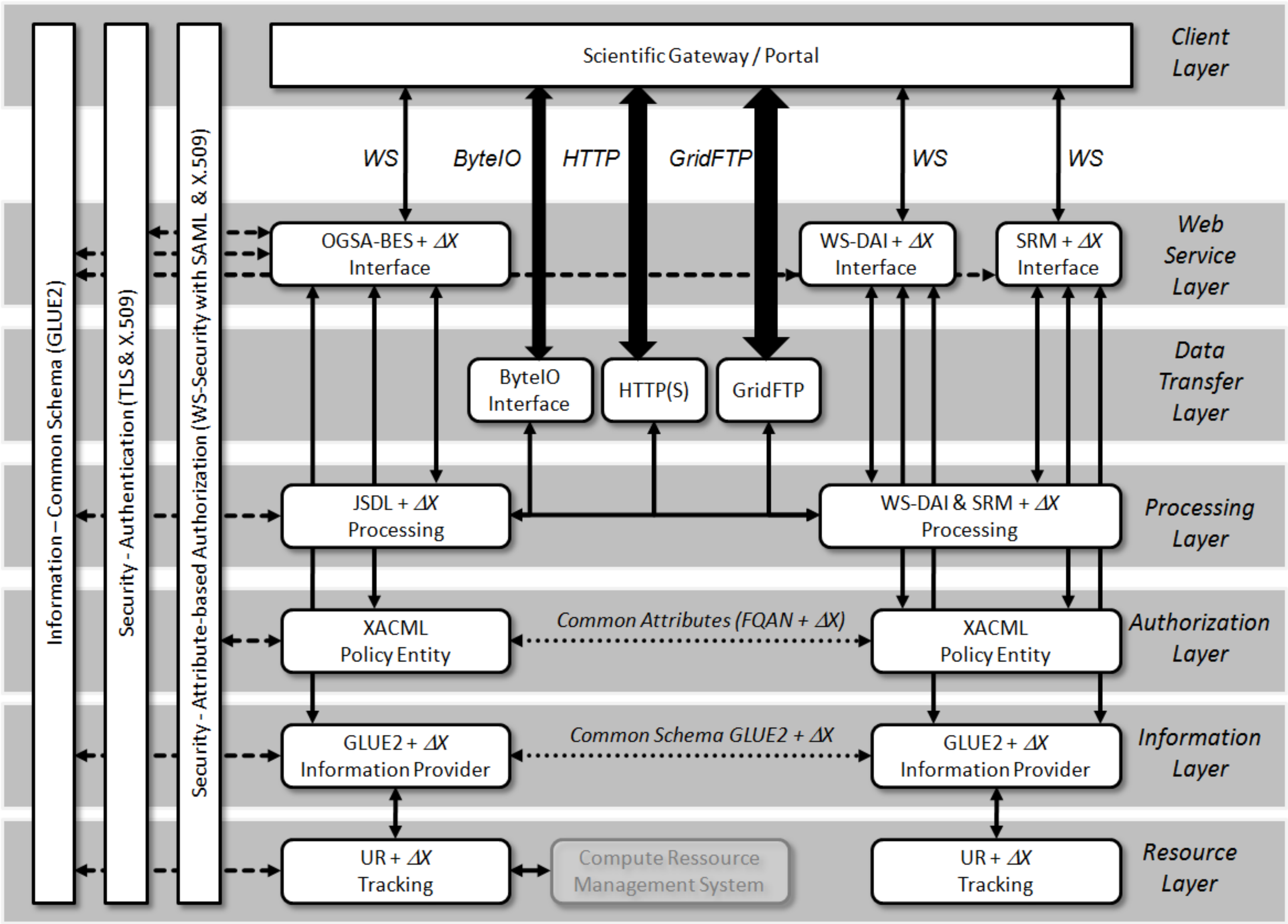




- Open standards are key to success – no transformation logic!
- Many standards defined for special purposes only (security, data, information, compute)
- Standard-based reference models (or profiles) can bring a set of those standards into context

# Step 1: Standards from many areas

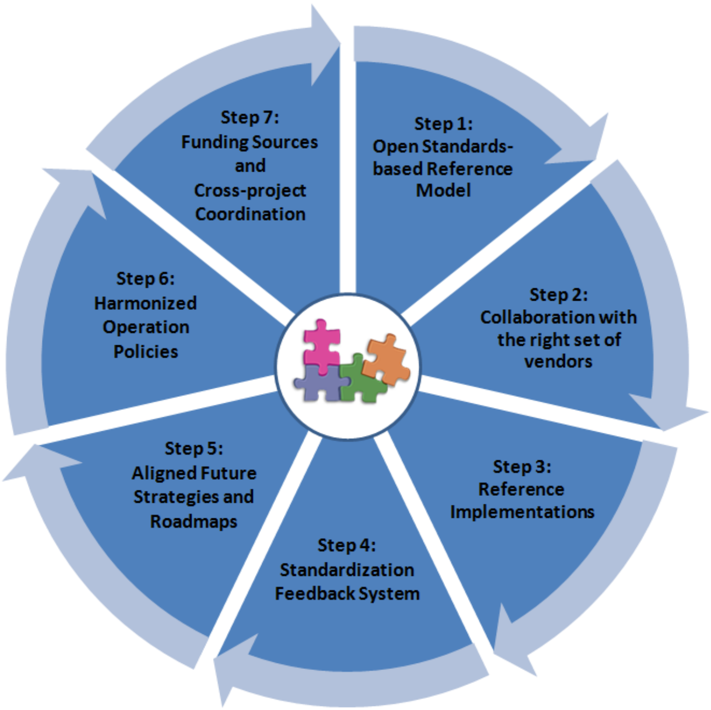


# Step 1: Reference Model Example




[2] Riedel et al., '***Research Advances by using Interoperable e-Science Infrastructures – The Infrastructure Interoperability Reference Model applied in e-Science*** '

# Step 2: Right Set of Vendors




- Closest possible collaboration among vendors required – why?
- Different interests from vendors, collaboration leads to sociology: 'communication among individuals'
- Collaboration as early as possible to get the buy-in from vendors!
- When a sub-fraction of known vendors in the field create a reference model & others joining later: this leads to numerous discussions and higher efforts

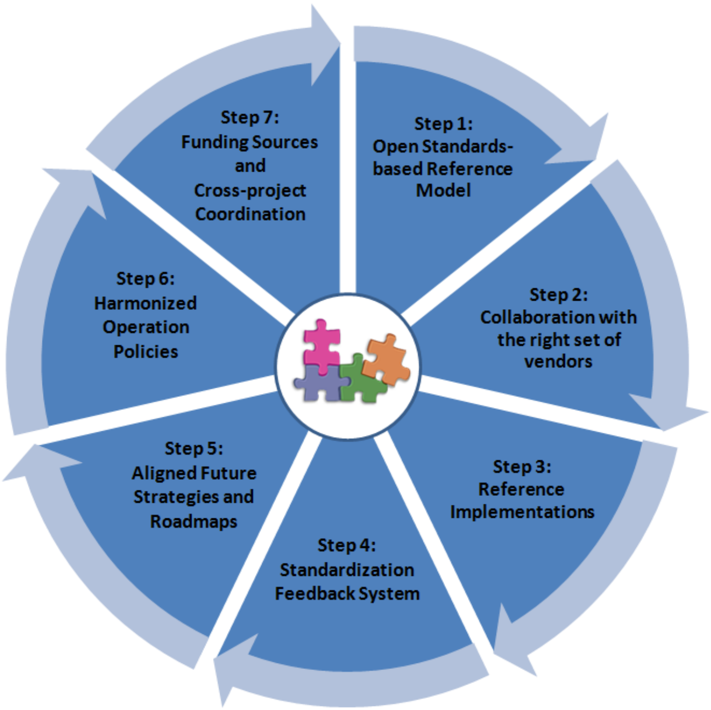
# Step 2: Example within OGF

- Initial OGSA-Basic Execution Service (BES) specification
  - Commercial and academic vendors have been involved (Microsoft, Platform, UNICORE, initially also Globus etc.)
  - Several others in the e-science community have been out of the process (initially ARC not involved, gLite later, etc.)
  - Production deploymens of OGSA-BES still rare (e.g. not in EGEE)

- Production Grid Infrastructure (PGI) Working Group
  - More than a standardization group: collaboration between the ‚right set of vendors‘ in the e-science community

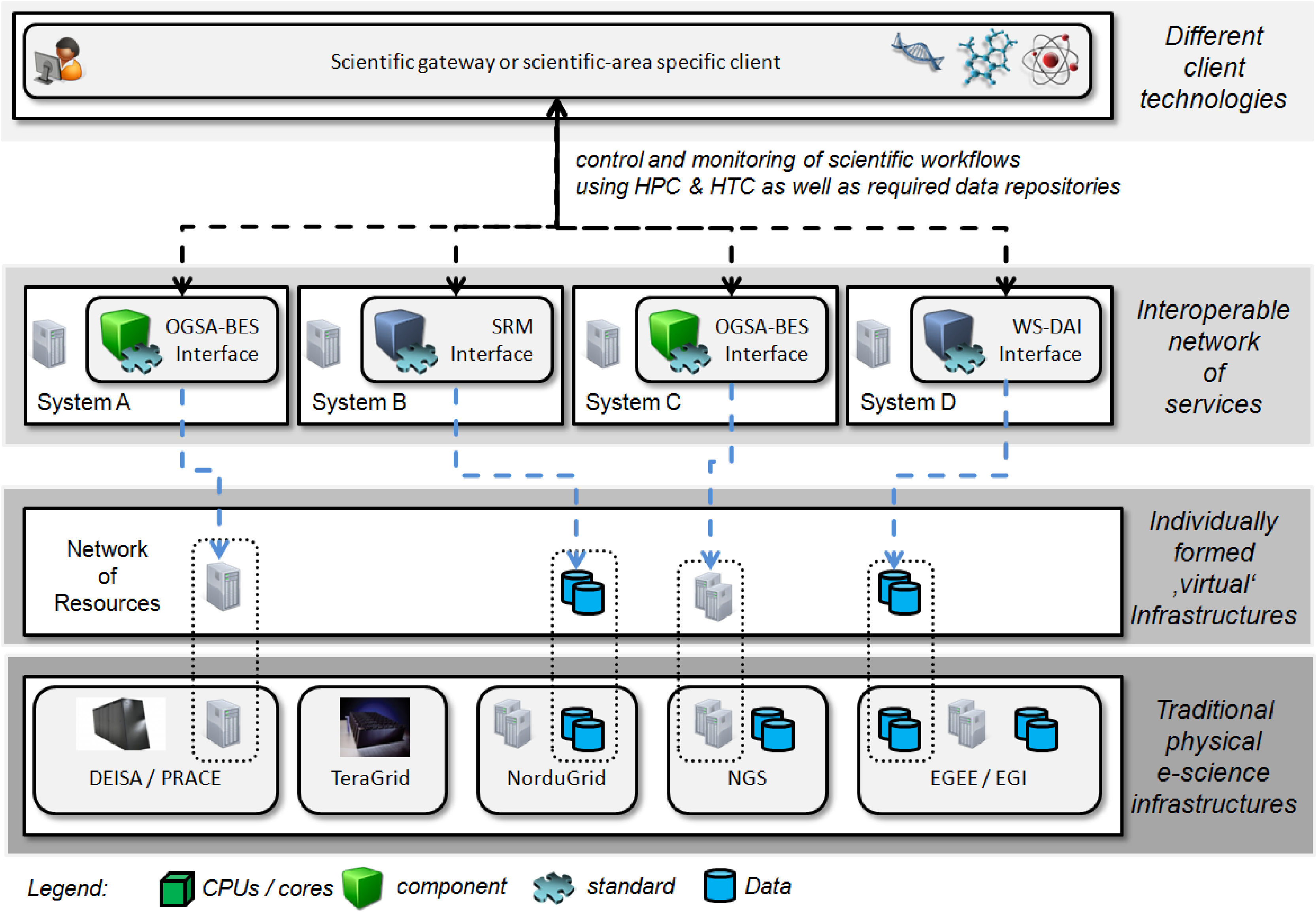
| Grid Technology Vendor/Project | Production e-science Infrastructure |
|---|---|
| ARC | NDGF |
| gLite | EGEE / EGI / OSG (as part of VDT) |
| UNICORE | DEISA / PRACE |
| Globus (IGE project) | TeraGrid |
| OMII-UK Software Stacks | NGS |
| NAREGI | NAREGI Infrastructure |
| EDGES/EDGI | BOINC-based infrastructures (i.e. Desktop Grids) |
| GENESIS | US Campus Grids |

# Step 3: Reference Implementations

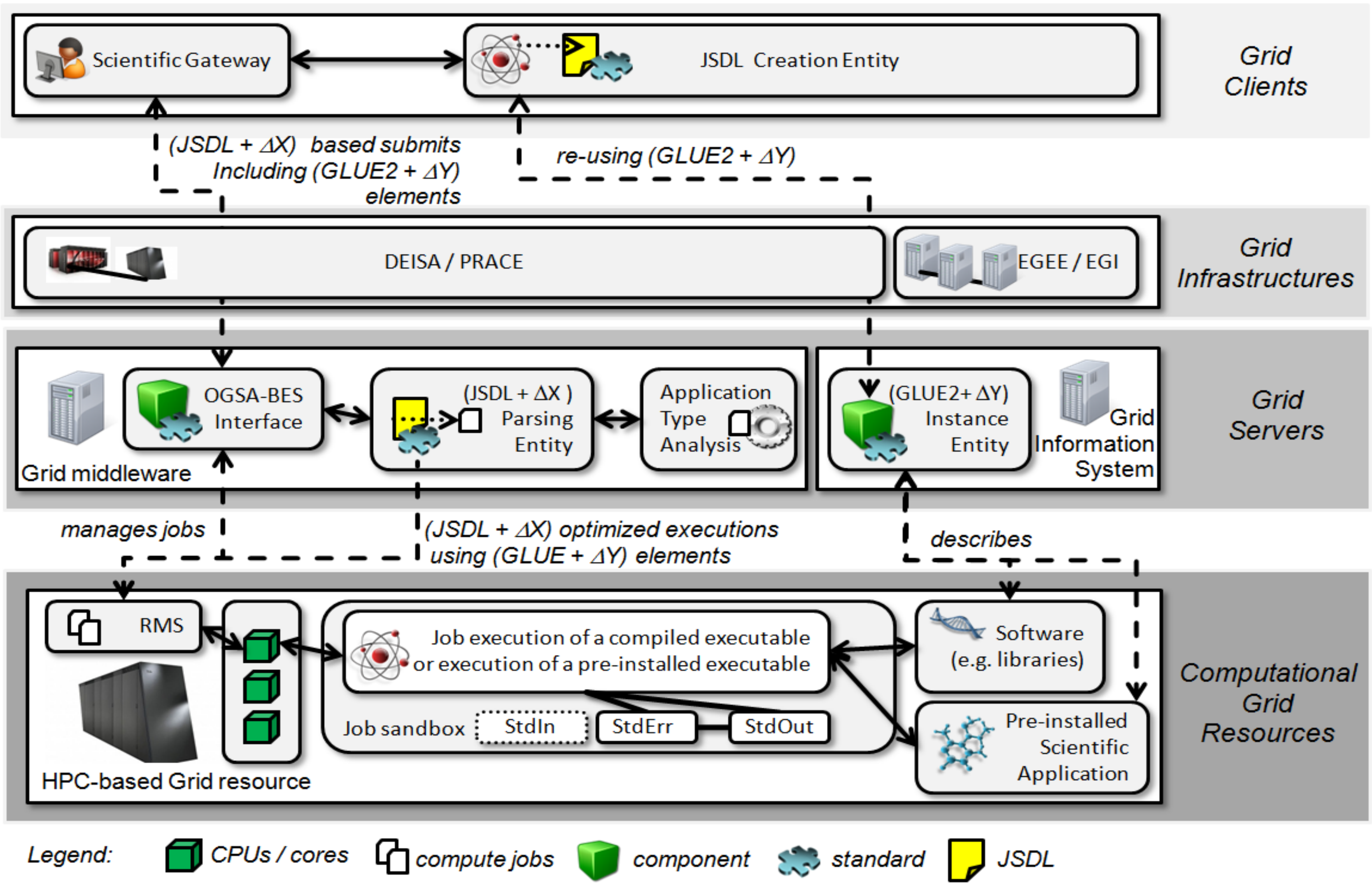



- Standards are the cornerstone
- But many missing links between standard specifications exist
- Reference implementations need to implement together standards (e.g. compute with security)
- Numerous lessons learned how standards work together in order to fill missing links
- Enables consistent standard use

# Step 3: Using Reference Impl.



# Step 3: Using GLUE2 and/in JSDL




[6] Riedel et al., ‚, Improvements of Common Open Grid Standards to Increase HTC and HPC Computing Effectiveness‘

# Step 3: Using GLUE2 in JSDL!

OpenGridForum

```
<jsdl>
  <jsdl:JobSpecification>
    <jsdl:JobIdentification> ...
    ...
    <jsdl:Application>
      <jsdldx:Software>
        <glue2:AppName> PEPC </glue2:AppName>
        <glue2:AppVersion> 2.4 </glue2:AppVersion>
        <glue2dy:ApplicationFamily> LINUX </glue2dy:ApplicationFamily>
      </jsdldx:Software>
      <jsdldx:Executable>
        <jsdldx:ExecutableName> pepc_power6 </jsdldx:ExecutableName>
        <jsdldx:ExecutablePath> /bin </jsdldx:ExecutablePath>
        <jsdldx:ExecutableArgument> particles </jsdldx:ExecutableArgument>
        <jsdldx:ExecutableArgument> 4000 </jsdldx:ExecutableArgument>
      </jsdldx:Executable>
      ...
      <jsdldx:ApplicationType>
         <glue2dy:ComputingActivityType>
            workflownode
         </glue2dy:ComputingActivityType>
     </jsdldx:ApplicationType>
      ...
      <jsdldx:JoinStdOutErr> true </jsdldx:JoinStdOutErr>
    </jsdl:Application>
    <jsdl:Resources>
      <jsdldx:Software>
        <glue2:AppName> VISIT </ApplicationName>
        <ApplicationVersion> 1.3 </ApplicationVersion>
        <ApplicationFamily> LIBRARY </ApplicationFamily>
      </jsdldx:Software>
      ...
    <jsdl:Resources>
  </jsdl:JobSpecification>
  ...
</jsdl>
```

Application Software Statement (can be used without an executable definition for pre-installed software) that re-uses parts of the GLUE2 specification

Revised Application Executable Definition

(must be used with above software statement)

Application Type classification re-using refined elements of the GLUE2 specification

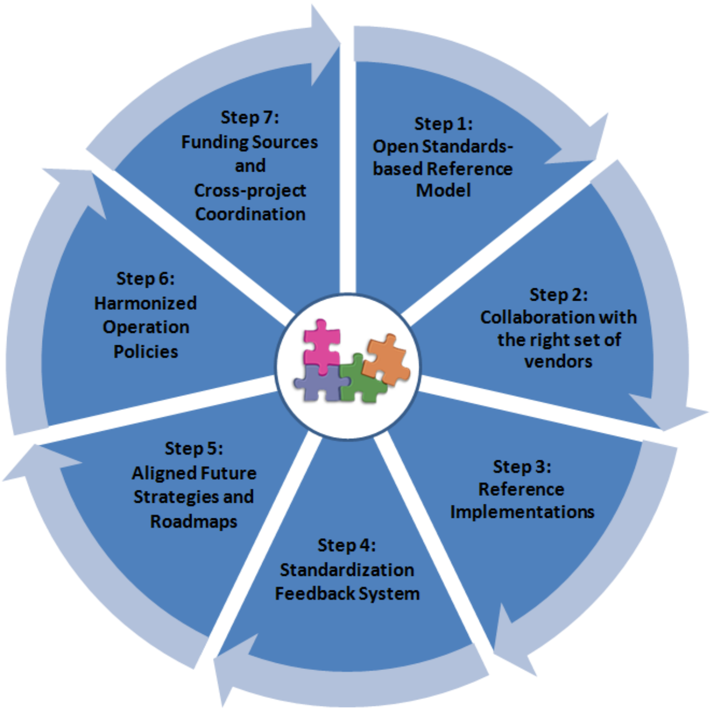
Boolean value that indicates standard out & error joins

Application Software Requirements (re-use of the same construct used above for describing the main application)

*(i) Listing: Example of JSDL + ΔX instance with more meaningful Grid job descriptions, also based on GLUE2 + ΔY.*

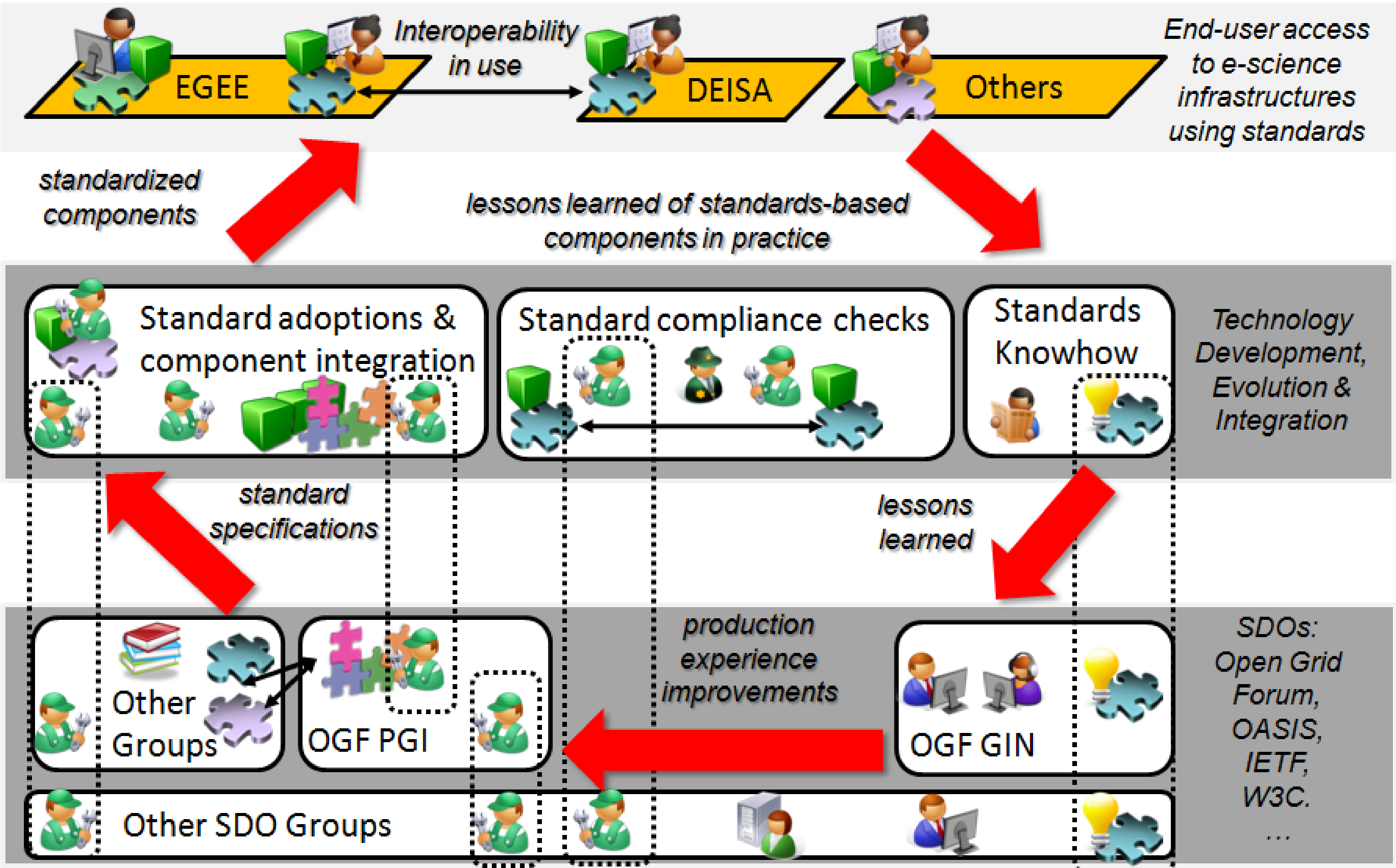
***[6] Riedel et al., ., Improvements of Common Open Grid Standards to Increase HTC and HPC Computing Effectiveness'***

# Step 4: Standardization Feedback




- Experience tells us: Missing links and many missing functionalities in early open standard versions
- Important: Give experience back to the standardization groups!
- Work required to analyse the production lessons learned in order to understand which standards need to be improved
- Goal: Improve the standards! Already 2nd iteration makes a major difference!

# Step 4: Example OGF GIN & PGI



# Step 4: Numerous lessons learned

| | |
|---|---|
| Production Grid-driven realistic reference model based on open standard | Although we used several standards in this drug discovery use case (OGSA-BES, JSDL, GLUE2, GridFTP, security profiles, etc.) their usage in conjunction together as a whole ecosystem so to say was rather unclear. This mainly includes computing, data, security standards as well as information flow aspects and standards. A reference model or greater realistic architecture would be important. |
| Grid Application Improvements | Grid application job descriptions satisfied basic needs in this use case but were not satisfactory enough to describe an application in this multi-Grid setup. Some improvements covering but are not limited to application types classification (e.g. parallel, etc.), application type refinements (e.g. pre-installed, submitted, etc.), revised application executable definition, application software statements, application family extension (e.g. LIBRARY), application software requirements, application output joins, etc. |
| Application Execution Adjacencies | In this workflow, we had several challenges in the execution environment itself. Thus we need better support scientific application executions with standard-based information aspects on the lowest possible level (i.e. resource management system level) covering but are not limited to common environment variables, common execution modules, execution module characteristics |

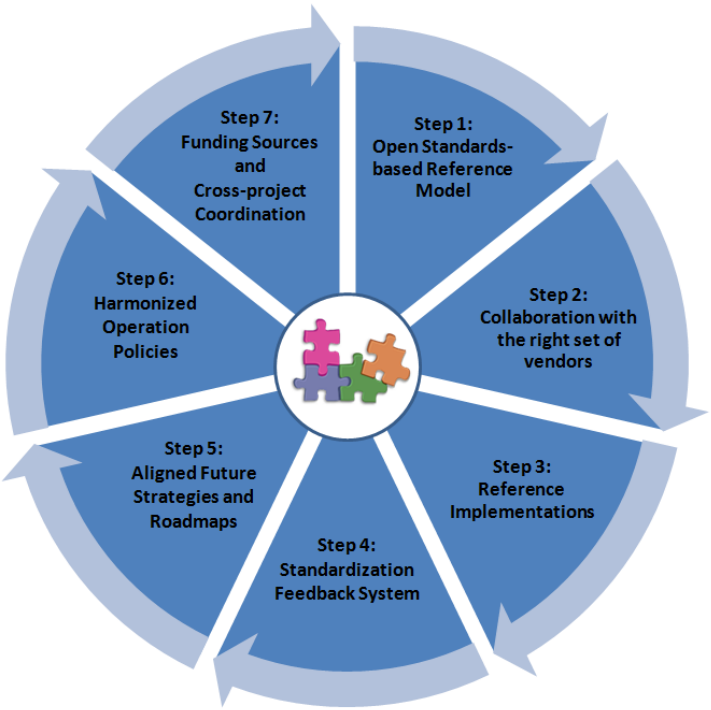
# Step 4: Numerous lessons learned

| | |
|---|---|
| High Performance Computing Extensions to open standards | While executions using CREAM-BES on EGEE had been relatively ok, submission with UNICORE-BES to DEISA lacked important HPC specific information. Therefore, we seek to submit and execute applications more efficiently than possible currently with GLUE2, JSDL, or OGSA-BES covering aspects but not limited to network topology (torus, global tree, Ethernet, etc.), shape reservation (x X y X z ), network information enhancements, available shape characteristics, high message support, task/core mapping definition, available task/core mappings, etc. |
| Sequence Support for Computational Jobs | An analysis of lessons learned obtained from the WISDOM use case leads to specific missing features encountered during production Grid interoperability with respect to the support of automatically started pre- and post-processing functionalities within JSDL using different application execution modes. AMBER, for instance, consists of a set of applications (~80 executables) and some of them are used to transform input data in a suitable format for production runs and/or transform outputs in several other formats necessary for further analysis. Of course, these transformation and short running pre-processing steps should be executed in a serial mode, while the actual corresponding molecular dynamic simulation is executed in a parallel mode. Pre-job sequences (pre-processing, compilation), Post-job sequences (post-processing). |

# Step 4: Numerous lessons learned

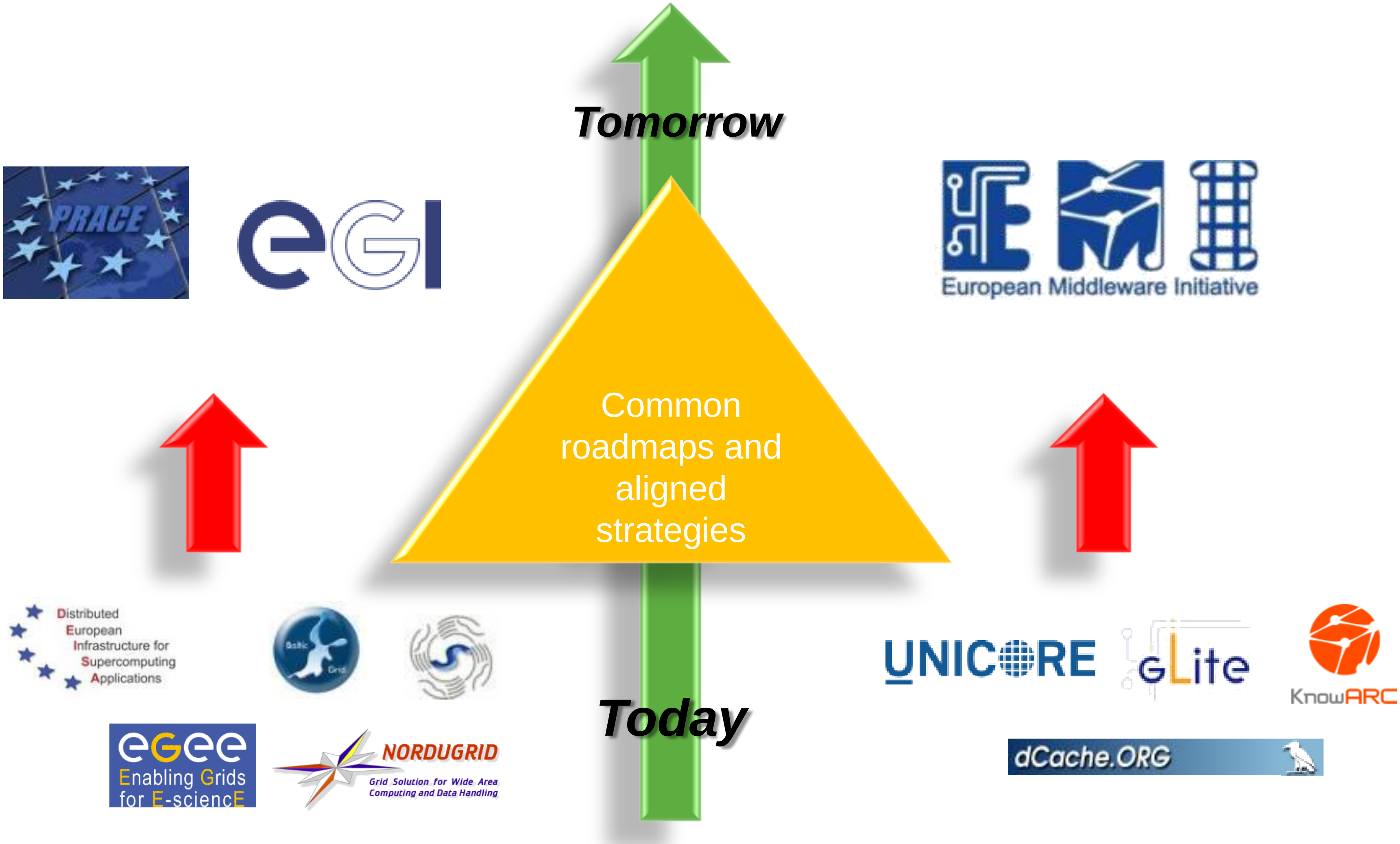
| | |
|---|---|
| Manual Data-staging support | Use cases revealed that in many cases the scientists require a more flexible mechanism during data-staging processes in order to better coordinate distributed data and computation. This is true, irrespective of whether data is transported to where the computational resource resides, or if computation is decomposed and job submissions are performed towards the location of data or even a hybrid of both methods is adopted. One example was the careful manual data input selection (aka manual data-staging) from the outcome of the EGEE workflow step in order to use only good results for the time-constrained workflow step in DEISA. |

# Step 5: Aligned Future Roadmaps

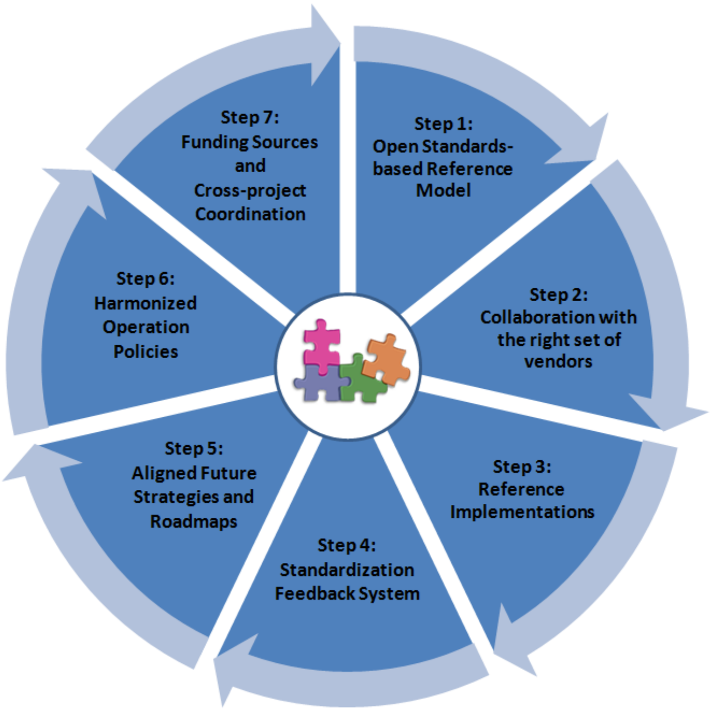



- Future strategies alignment between technology provider is essential
- Not only for standardization endeavors, also general technology development plans
- Better harmonization – reduce costs and service duplication
- Enable the use of components from one provider by different components from another

# Step 5: Example EMI

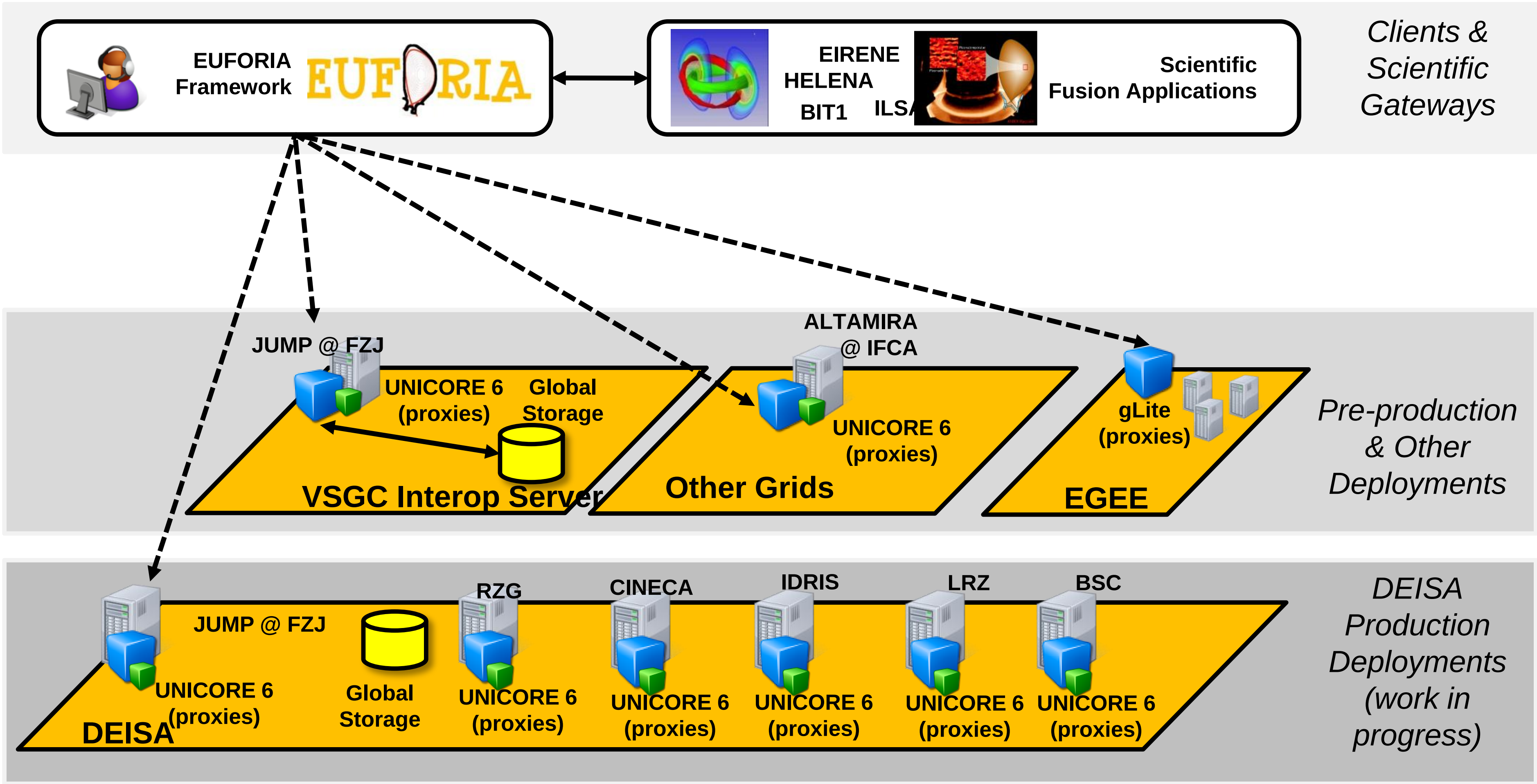


# Step 6: Harmonized Operations



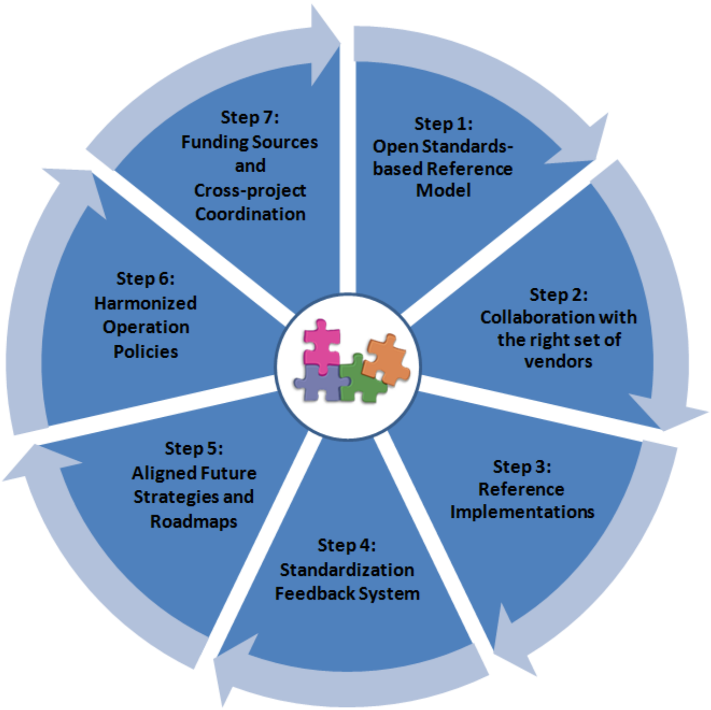

- Important aspect are policies
- Difference between what technology-wise possible … …and what policies allow
- Negotiated and developed on case-by-case basis make sense
- Fine-granular policies for dedicated projects/groups based on technology improvements

# Step 6: EUFORIA Example




[4] Memon and Riedel et al. *'Lessons learned from jointly using HTC- and HPC-driven e-science infrastructures in Fusion Science'*

# Step 7: Cross-project Coordination

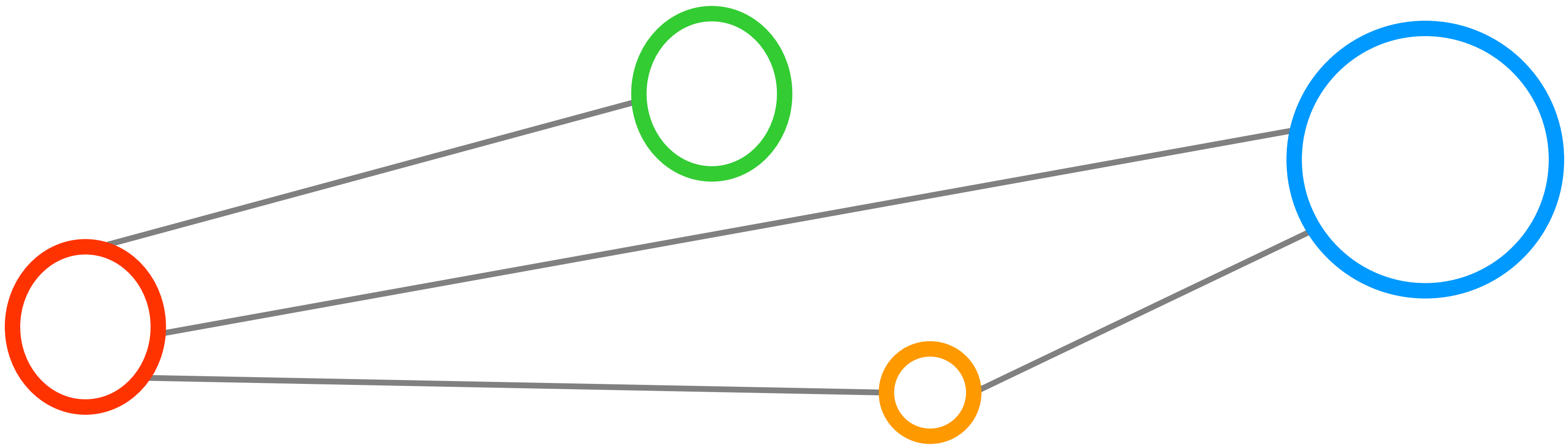



- Key for sustained interoperability
- Funding sources and cross-project coordination…
- Ideal would be one joint funding source and/or non-project-based fundings
  - Many short-term solutions because of 2 to 3 years running projects
  - Standards definition take this time alone
- Funding sources collaboration, e.g. NSF and EC

# Step 7: Example coordination

- SIENA Initiative coordinates the creation of an roadmap between currently funded DCI projects by creating a standards roadmap
- EU requires a common deliverable of DCI projects towards an aligned vision

# Summary



# Summary

- Traditional use: one infrastructure
- Design pattern using HTC and HPC
- Interoperability is desired: How?

a new 'toolset' is given…

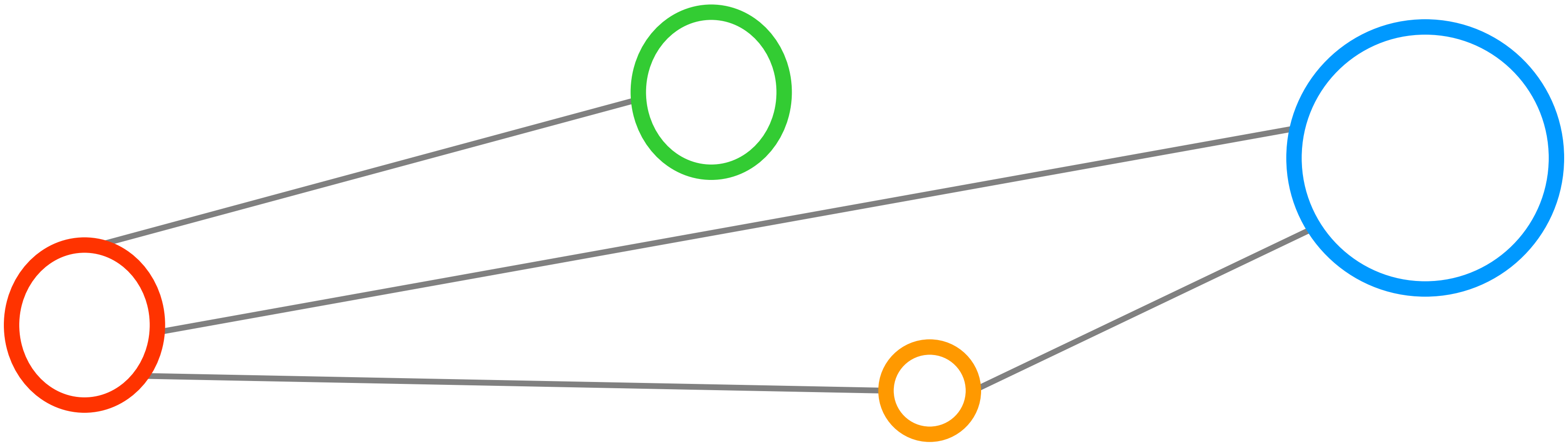
- Standards not enough but important
- Think differently: think process!
- The seven steps provides ‚ways' achieving sustainable interoperability

a new 'mindset' is required…

- Conclusion: We need a new skillset
- Capabilites for social processes
- Broader understanding in technologies

a new 'skillset' is required…

# References



# References


[1] M. Riedel, „E-Science Infrastructure Interoperability Guide – The Seven Steps towards Interoperability for e-Science", in book „Guide to e-Science: Next Generation Scientific Research and Discovery", Editors: X. Yang and L. Wang Springer, to be published in 2010

[2] M. Riedel et al.„ Research Advances by using Interoperable e-Science Infrastructures - The Infrastructure Interoperability Reference Model applied in e-Science", Journal of Cluster Computing, Special Issue Recent Advances in e-Science, Cluster Computing, Vol. 12, No. 4, pp. 357-372, DOI 10.1007/s10586-009-0102-2, 2009

[3] M. Riedel and D. Kranzlmüller et al. „Classification of Different Approaches for e-Science Applications in Next Generation Computing Infrastructures", Proceedings of the 4th IEEE e-Science Conference, Indianapolis,USA, pp. 198 - 205

[4] M. S. Memon, M. Riedel, A. S. Memon, F. Wolf, A. Streit, Th. Lippert, Marcin Plociennik, Michal Owsiak, David Tskhakaya, Christian Konz, Lessons learned from jointly using HTC- and HPC-driven e-science infrastructures in Fusion Science, proceedings of the IEEE ICIET 2010 Conference, Pakistan

[5] M. Riedel et al. "Improving e-Science with Interoperability of the e-Infrastructures EGEE and DEISA"; Proceedings of the 31st International Convention MIPRO, Conference on Grid and Visualization Systems (GVS), May 2008, Opatija, Croatia, Croatian Society for Information and Communication Technology, Electronics and Microelectronics, ISBN 978-953-233-036-6, pages 225 – 231

[6] M. Riedel et al. „*Improvements of Common Open Grid Standards to Increase High Throughput and High Performance Computing Effectiveness on Large-scale Grid and e-Science Infrastructures „* Seventh High-Performance Grid Computing (HPGC) Workshop at International Parallel and Distributed Processing Symposium (IPDPS) 2010, April 19-23, 2010, Atlanta, USA