



Polish Roadmap towards  
Domain-Specific Infrastructure  
for Supporting Computational Science  
in European Research Area

# GridSpace2 Towards *Science-as-a-Service* Model

Eryk Ciepiela, Bartosz Wilk, Daniel Haręźlak,  
Marek Kasztelnik, Maciej Pawlik,  
Jan Meizner, Marian Bubak

Academic Computer Center CYFRONET AGH

KUKDM 2013, Zakopane, 28.02.2013



- **Science-as-a-Service**
- **e-Science-as-a-Service**
- From computational experiment to **service**
- From computational experiment to **executable scientific publications**
- **GridSpace2** – web-oriented distributed computing platform
- **Collage Authoring Environment** – executable publication framework
- Science made possible by GridSpace2/Collage Authoring Environment
- Conclusions
- Future work

# Science-as-a-Service

Applying software-as-a-service principles  
and IT best practices to science



## “as-a-service” model

- Proved successful in multiple areas in IT
- Infrastructure-as-a-Service
- Platform-as-a-Service
- Software-as-a-Service

## Software-as-a-service model applied to science

- Core service providers
- Outsourcing
- Service market, competitiveness, ecosystem
- Cost-effectiveness, resource sharing, low entry costs, pay-per-use
- Scientific experiments as a service
- Science as an enterprise
- Crowdsourced science
- Open science



# e-Science-as-a-Service

Applying software-as-a-service principles  
and IT best practices to e-science



## ...and especially e-science

- Data centers as computing power and data storage providers
- Software providers
- IT staff to hire
- Base software platform, tools, problem solving environments
- Computational experiment as a service
- Shared e-infrastructure, reusable and reproducible experiments
- Executable scientific publications



# From computational experiments to services

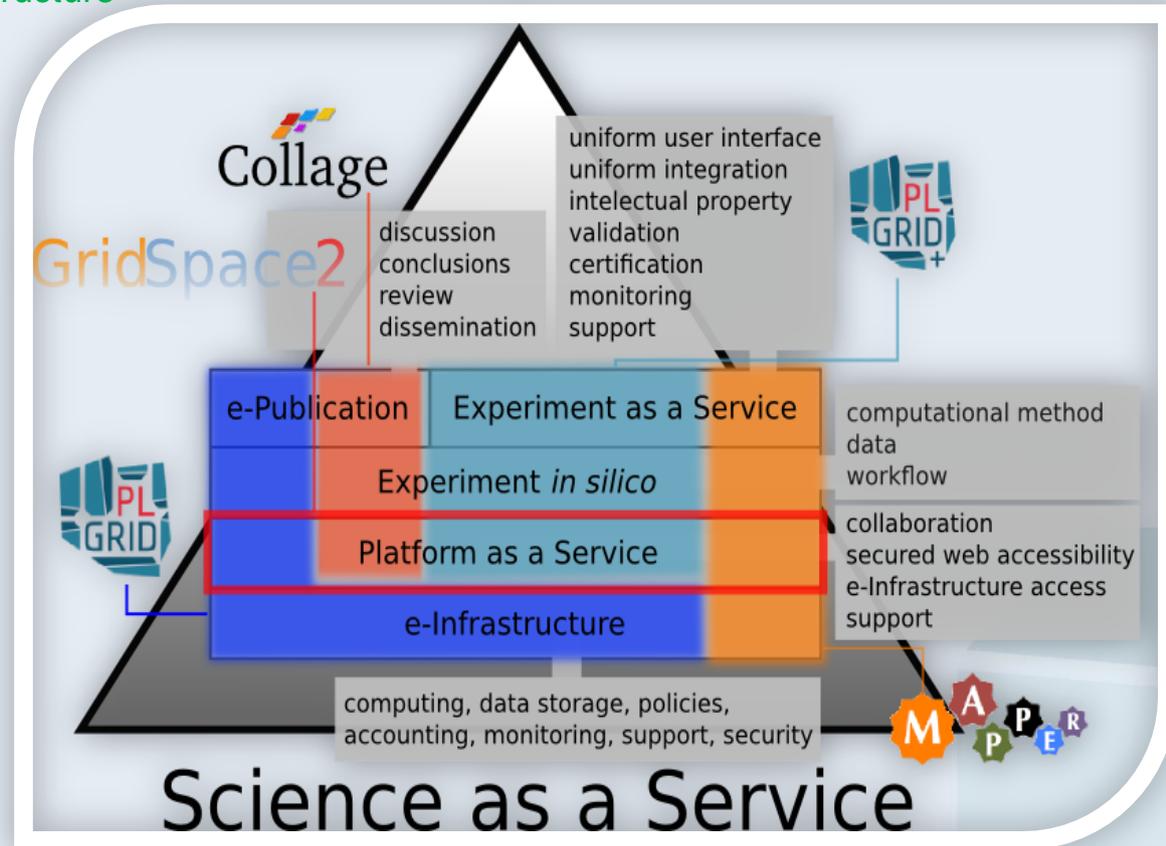
Common aspects **ensured** or **streamlined** by GridSpace2



## From computational experiment to service

(e.g. PL-Grid Plus domain service)

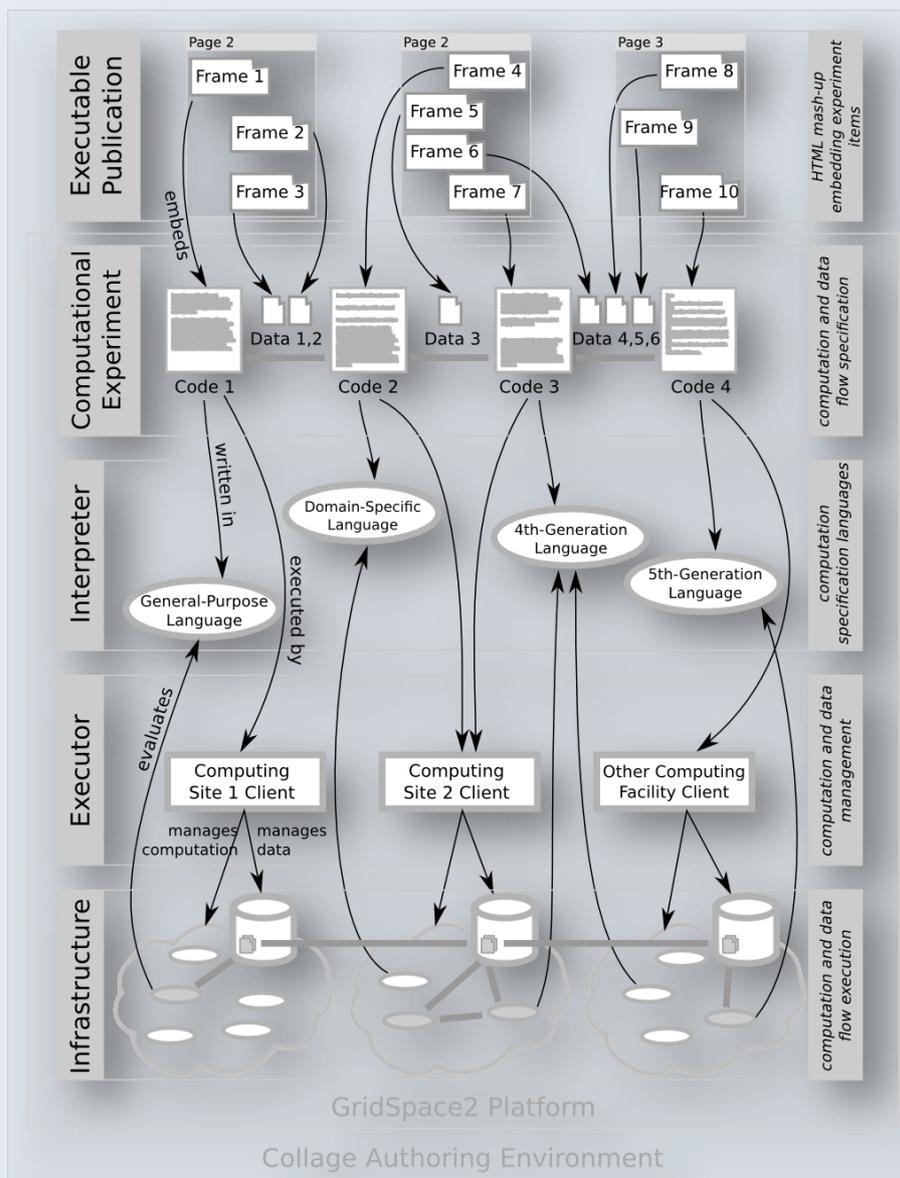
- Integration with common e-infrastructure
- User access management
- Respected intellectual property rights
- Cataloging, indexing
- Accessibility
- Reusability
- Experiment availability
- Documentation availability
- Examples availability
- Monitoring
- Accounting
- Maintenance
- User support
- Quality assurance
- Security assurance



**Provisioning of services at little cost with GridSpace2 in platform-as-a-service model**



# GridSpace2 and Collage Authoring Environment Concept



# GridSpace2 and Collage Authoring Environment

## Managing computational experiments



Collage Authoring Workbench

You are logged in to **collage-exphost.elsevier.com** as **eciepiela**

collage-exphost.elsevier.com: collatz/collatz | collage-exphost.elsevier.com: hello

The Collatz Conjecture

Computing sequences for **Ruby 1.8.7** with collage-exphost.elsevier.com

Generating plots **GnuPlot 4.2.6** with collage-exphost.elsevier.com

```
set output 'collatz/hwm.png'  
plot 'collatz/results_raw.txt' using 1:3 title "The biggest  
number (high water mark) reached when iterating"  
  
set output 'collatz/last.png'  
plot 'collatz/results_raw.txt' using 1:4 title "Value in  
last iteration"
```

1 input data defined for this code. 3 output data defined for this code.

Output

```
gnuplot> set output 'collatz/hwm.png'  
gnuplot> plot 'collatz/results_raw.txt' using 1:3 title "The  
biggest number (high water mark) reached when iterating"
```

© 2012 Spectrum - IT Division. Please report any problems related to this portal to our [Issue Tracker](#). Acknowledgements

# GridSpace2 and Collage Authoring Environment

## Accessing services and interacting with executable publications

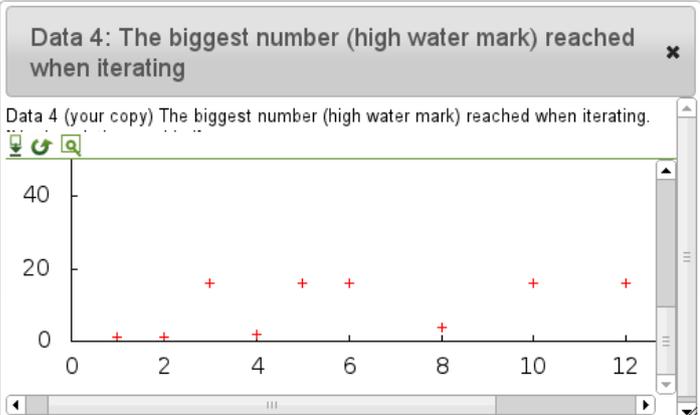


### The Collatz Conjecture

eciepiela, DOI: 10.0000/1358511059290

The experiment was released by **eciepiela** on **Fri Jan 18 13:10:59 CET 2013** in the **private** scope. No, below is not an article. It's only generated text with injected labels that navigate to particular experiment items to show you how in-text links work.

Euismod in pellentesque massa placerat dui ultricies lacus. Penatibus et magnis dis parturient. Ipsum consequat nisl vel pretium lectus. Interdum velit euismod in pellentesque. Nascetur ridiculus mus mauris vitae ultricies leo integer malesuada nunc. Sapien nec sagittis aliquam malesuada. Donec ultrices tincidunt arcu non sodales neque sodales. Sed velit dignissim sodales ut eu. [Data 1](#) Ac tortor vitae purus faucibus ornare. Mattis enim ut tellus elementum sagittis vitae et. Vitae semper quis lectus nulla at. Vitae purus faucibus ornare suspendisse sed nisi lacus sed viverra. Amet tellus cras adipiscing enim eu. [Code 1](#) Tellus integer



non nisi est sit amet. Proin ut massa tunc neque aliquam vestibulum morbi blandit. [Data 5](#) Non consectetur a erat nam at lectus urna duis. Diam volutpat commodo sed egestas egestas fringilla phasellus faucibus scelerisque. Magna eget est lorem ipsum dolor sit.

**Data 1: Arguments**

---

**Code 1: Computing sequences for given arguments**

Code 1 (your copy) Computing sequences for

Source: Ruby 1.8.7 | Output

```
step+=1
end
results_raw.puts("#(argument) \t#"
```

---

**Data 2: Raw results to be visualized afterwards**

Data 2 (original) Raw results to be visualized

1	0	1	1
2	1	1	1
3	7	16	1
4	2	2	1
5	5	16	1
6	8	16	1
7	10	52	10

---

**Code 2: Generating plots**

---

**Data 3: Number of iterations**

**Data 4: The biggest number (high water mark) reached when iterating**

**Data 5: Value in last iteration**

Save experiment | Reset

# Science made possible by GridSpace2/Collage platform



- E. Ciepiela, L. Zaraska, G. D. Sulka: *GridSpace2 Virtual Laboratory Case Study: Implementation of Algorithms for Quantitative Analysis of Grain Morphology in Self-assembled Hexagonal Lattices According to the Hillebrand Method*. In: M. Bubak, T. Szepieniec, K. Wiatr (Eds) Building a National Distributed e-Infrastructure - PL-Grid - Scientific and Technical Achievements (2012)
- T. Jadczyk, M. Malawski, M. Bubak, I. Roterman: *Examining Protein Folding Process Simulation and Searching for Common Structure Motifs in a Protein Family as Experiments in the GridSpace2 Virtual Laboratory*. In: M. Bubak, T. Szepieniec, K. Wiatr (Eds) Building a National Distributed e-Infrastructure - PL-Grid - Scientific and Technical Achievements (2012)
- E. Ciepiela, T. Jadczyk, D. Harężlak, M. Kasztelnik, P. Nowakowski, G. Dyk, M. Malawski, M. Bubak, I. Roterman: *Computations of Protein Hydrophobicity Profile as Virtual Experiment in Gridspace Virtual Laboratory*. In: Bio-Algorithms and Med-Systems (2012)
- L. Zaraska, W. J. Stępniewski, E. Ciepiela, G. D. Sulka: *The effect of anodizing temperature on structural features and hexagonal arrangement of nanopores in alumina synthesized by two-step anodizing in oxalic acid*. In: Thin Solid Films, accepted for publication (2013)
- L. Zaraska, W. J. Stępniewski, G. D. Sulka, E. Ciepiela, M. Jaskuła: *Analysis of nanopore arrangement and structural features of anodic alumina layers formed by two-step anodizing in oxalic acid using the dedicated executable software*. In: Applied Physics A, accepted for publication (2013)
- \*\*\*. In: **Computers & Graphics, Special Issue featuring Executable Papers**, to appear (2013)
- Your publication maybe?

Pilot case studies for GridSpace2 and Collage

Publications made possible by GridSpace2

Executable publications powered by GridSpace2 and Collage



# Publications describing GridSpace2/Collage and its applications



- E. Ciepiela, D. Hareżlak, J. Kocot, T. Bartynski, M. Kasztelnik, P. Nowakowski, T. Gubała, M. Malawski, M. Bubak: **Exploratory Programming in the Virtual Laboratory**. In: Proceedings of the International Multiconference on Computer Science and Information Technology (2010)
- P. Nowakowski, E. Ciepiela, D. Hareżlak, J. Kocot, M. Kasztelnik, T. Bartyński, J. Meizner, G. Dyk, M. Malawski: **The Collage Authoring Environment**. In: Proceedings of the International Conference on Computational Science (2011)
- E. Ciepiela, P. Nowakowski, J. Kocot, D. Hareżlak, T. Gubała, J. Meizner, M. Kasztelnik, T. Bartynski, M. Malawski, M. Bubak: **Managing Entire Lifecycles of e-Science Applications in the GridSpace2 Virtual Laboratory - From Motivation through Idea to Operable Web-Accessible Environment Built on a Top of PL-Grid e-Infrastructure**. In: M. Bubak, T. Szepieniec, K. Wiatr (Eds) Building a National Distributed e-Infrastructure - PL-Grid - Scientific and Technical Achievements, Springer (2012)
- D. Hareżlak, M. Kasztelnik, E. Ciepiela, M. Bubak: **Scripting Language Extensions Offered by the GridSpace Experiment Platform**. In: M. Bubak, T. Szepieniec, K. Wiatr (Eds) Building a National Distributed e-Infrastructure - PL-Grid - Scientific and Technical Achievements, Springer (2012)
- J. Meizner, E. Ciepiela, P. Nowakowski, J. Kocot, M. Malawski, M. Bubak: **Flexible and Extendable Mechanism Enabling Secure Access to e-Infrastructures and Storage of Confidential Data for the GridSpace2 Virtual Laboratory**. In: M. Bubak, T. Szepieniec, K. Wiatr (Eds) Building a National Distributed e-Infrastructure - PL-Grid - Scientific and Technical Achievements, Springer (2012)
- E. Ciepiela, P. Nowakowski, D. Hareżlak, M. Kasztelnik, G. Dyk, J. Meizner, M. Bubak: **The Collage Authoring Environment: A Platform for Executable Publications**, abstract for Cracow Grid Workshop (2012)
- E. Ciepiela, D. Hareżlak, M. Kasztelnik, J. Meizner, G. Dyk, P. Nowakowski, M. Bubak: **The Collage Authoring Environment: from proof-of-concept prototype to an exploitable service. Development and deployment of a web-oriented platform for executable scientific publications**, accepted to International Conference on Computational Science (2013)
- P. Pierzchała: **Multiscale Applications in the GridSpace Virtual Laboratory**, Master of Science Thesis supervised by Katarzyna Rycerz; AGH University of Science and Technology (2012)
- B. Bodziechowski, E. Ciepiela, M. Bubak: **Assessment of Software Quality with Static Source Code Analysis: GridSpace2 Case Study**, abstract for Cracow Grid Workshop (2012)



[www.plgrid.pl](http://www.plgrid.pl)



[www.plgrid.pl/plus](http://www.plgrid.pl/plus)



[www.mapper-project.eu](http://www.mapper-project.eu)



[collage.elsevier.com](http://collage.elsevier.com)

more on [dice.cyfronet.pl/products/gridspace](http://dice.cyfronet.pl/products/gridspace)

# Conclusions



- Inefficient traditional model pushes scientists and research organizations to seek new business models for science
- Software-as-a-service model proved effective in IT market, science-as-a-service can too, especially in e-science
- E-science services are not only software and data but complete and consumable products
- Scientific findings and methods need new ways for communicating and disseminating to embrace reusability, verifiability, reproducibility, transparency, executability
- GridSpace2 is a platform-as-a-service that facilitates provisioning of experiment-as-a-service at little cost
- Collage Authoring Environment overlays GridSpace2 and enables executable scientific publications
- Collage Authoring Environment is integrated with Elsevier ScienceDirect portal

- In the context of PL-Grid Plus project GridSpace2 will be used as a platform for streamlined provisioning of domain services
- Multiscale applications as services will be powered by GridSpace2 platform in the scope of the Mapper project
- Collage Authoring Environment will empower special issue of “Computers and Graphics” journal featuring executable publications
- We're eager to support research teams in provisioning of services and e-publications

