



Scalable Software Services for Life Science

WP2 – Dissemination and Training

Mihai Duta, OeRC



Objectives:

- **disseminate** widely the results of the ScalaLife project, maximising the impact on future LS activities and developments;
- provide effective **training** to promote the scalable software developed the project.

Task	Description	Coordinator
T2.1	Analysis of Needs and Detailed Planning	Anne Trefethen (OeRC)
T2.2	Web Presence	Helmut Satzger (LRZ)
T2.3	Outreach	Lilit Axner (KTH)
T2.4	Training	Mihai Duta (OeRC)
T2.5	WP Management	Mihai Duta (OeRC)

	Deliverable	Due
D2.1	Dissemination Plan	6
D2.2	Training Plan	12
D2.3	Dissemination Report and Updated Plan	18
D2.4	Training Material	18
D2.5	Training Report and Updated Plan	24
D2.6	Training Material Updated	24
D2.7	Final Report	36

Dissemination ensures good visibility of the project at international level and distribute the results of the project widely and efficiently within the European LS community.

The project results to disseminate are:

- competitive **speed and scalability** demonstrated by the project pilot codebases (Gromacs and Dalton) on small-to-medium systems;
- **libraries and APIs** to facilitate the easy exploitation of acceleration hardware (*e.g.* GPGPUs) or of low-level message-passing communication to reduce the MPI latency;
- new **standards** for LS data storage and exchange;
- the creation of the **Competence Center**, which will provide an expert service for the LS community and be the main vehicle for knowledge transfer.

The initial **Dissemination Plan** (D2.1) was submitted at the end of the 1st half year of the project

- discusses the dissemination vehicles
- sets performance targets

The plan is reviewed by the following **Dissemination Report and Updated Plan**, due in a year.

Dissemination vehicles:

- the creation of a **Web Portal**
- the organisation of **International Events**
- the issue of regular **Newsletters**
- the distribution of **Outreach Material**
- the organisation of **Training**



At the moment the portal

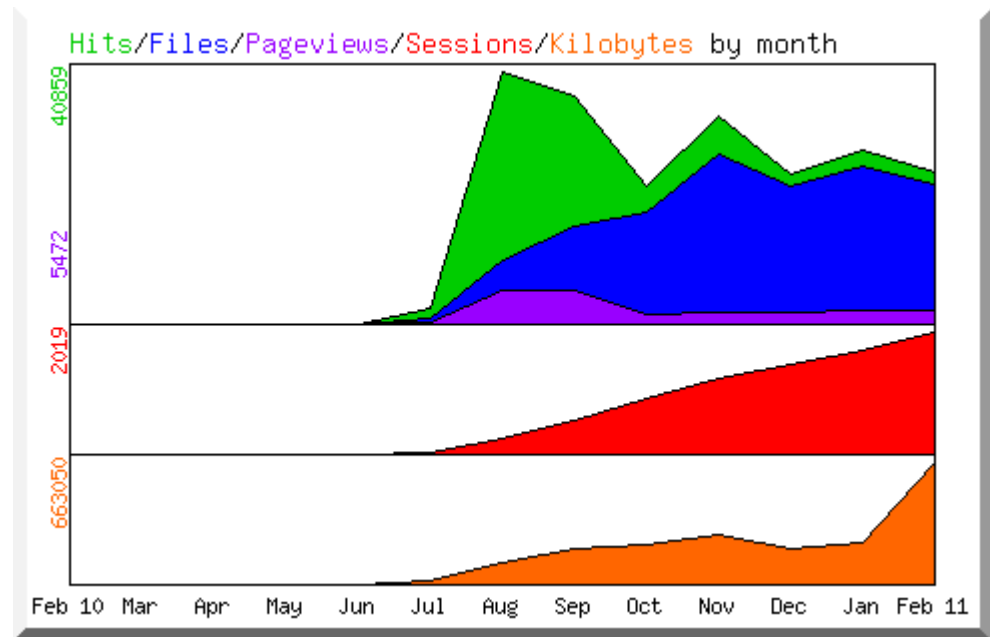
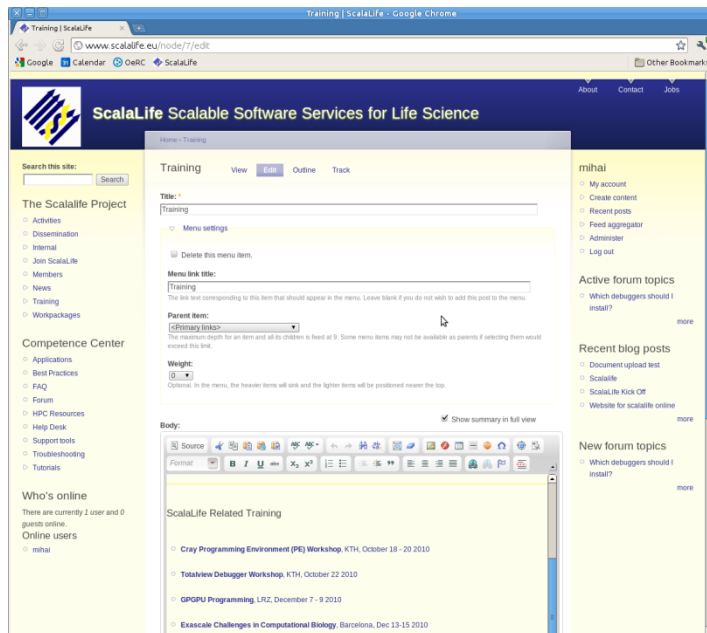
- represents the main project hub, thus enabling the collaboration between project members
- acts as the central information centre for the project, announcing upcoming training events, meetings, project news, achievements, etc.
- acts as the gateway to the Competence Center

In future, it will also

- provide access to software tools and related training material and provides access to best-practice guides
- provide general information on HPC and on existing and emerging technologies relevant to the Life sciences
- support a community collaborative environment, exchange of information and expert advice

The initial design of the Web Portal, finalised very early in the project, already provides a lot of essential functionality.

The variation of web traffic during the 1st six months of the project.



ScalaLife has already been present at:

- The 8th e-Infrastructure Concertation Meeting (Geneva, Nov 2010), one of the FP7 projects present
- Supercomputing Conference 2010 (New Orleans, Nov 2010)
- “Exascale Challenges in Computational Biology” conference (Barcelona, Dec 2010)

ScalaLife will continue to be present at future events:

Dates	Place	Event	Website	ScalaLife
11-12/04/2011	Vilnius, Lithuania	European Grid Infrastructure User Forum 2011	uf2011.egi.eu	Dissemination material
13-14/04/2011	Helsinki, Finland	DEISA PRACE Symposium 2011	www.deisa.eu	Dissemination material
01-03/06/2011	Tsukuba, Japan	ICCS 2011 International Conference on Computational Science	www.iccs-meeting.org	Dissemination material
19-23/06/2011	Hamburg, Germany	ISC International Supercomputing Conference 2011	www.supercomp.de/isc11	Booth (BSC, LRZ)
10-13/09/2011	Vienna, Austria	The EMBO Meeting	www.the-embo-meeting.org	Poster
12-18/11/2011	Seattle, USA	SC Supercomputing 2011	sc11.supercomputing.org	Booth (BSC, LRZ)
21-23/11/2011	Barcelona, Spain	ICREA Network Medicine Meeting		Poster + link
05-08/12/2011	Stockholm, Sweden	7th IEEE International Conference on e-Science	www.escience2011.org	Dissemination material

Intended collaborations:

- CECAM – an application is submitted later this year (April?); if successful, this grant will be used to organise the 1st ScalaLife International Workshop
- MAPPER – this is yet to be organised

Considered:

- INCF neuroinformatics project (www.incf.org)
- DigiGirly project (Microsoft)

Planned and confirmed:

- The Erina+ project (www.erinaplus.eu)

Newsletter:

- bi-annual, reports on the WPs and project in general
- first newsletter already released at the end of Feb 2011

Outreach Material:

- standard material** – the basic ScalaLife informative and advertising material used at events:
 - brochures, project presentations, press releases, posters and flyers
- branding:** logos, colour schemes
- non-standard:**
 - online video presentations
 - Erik Lindahl “Introduction to Molecular Modelling”, Gromacs workshop, Helsinki
 - ScalaLife presentation, MGMS meeting, Oxford
 - collaborative sites

- main channel – the ScalaLife Web Portal
- the ScalaLife information mailing list
- the web sites and communities of external projects with which ScalaLife collaborates
- the communities of Gromacs and Dalton users (which are planned to become integrated within the Competence Centre)
- conferences at which ScalaLife participates
- training events
- announcement(s) on **www.thedigitalscientist.org** and **www.hpcwire.org**

Feedback is crucial for the development and organisation of training and dissemination:

- participants at all training events are asked to provide ratings and comments
- users rate the quality of online material (tutorials, guides, video) and offer suggestions

The project has to have an acknowledgement policy, which should be adhered to by all users of its products:

- standard acknowledgement text to be used in publications
- users should report publications back to project, so ScalaLife can have an extra measure of the effectiveness of dissemination.

To performance of dissemination during the next year will be assessed using the following targets:

- the creation of initial online tutorials and training material;
- the creation of initial best practice guides;
- the creation and organisation of the help desk;
- the formulation of the acknowledgement policy;
- participation at all the events and conferences proposed.

Tentative timeline proposed for some key targets:

Target	Completion
Initial online tutorials and training material	31/03/2011
Initial best practice guides	31/03/2011
Help desk	31/08/2011
Acknowledgement policy	31/08/2011

Quantitative criteria to measure effectiveness of dissemination:

- The ScalaLife grant application to CECAM is successful;
- The first ScalaLife International Workshop is organised successfully as part of the CECAM grant;
- The number of participants at the First ScalaLife International Workshop;
- The uptake of the Competence Centre;
- The number of LS software users who open user trial accounts as part of the Competence Centre;
- The number of LS software users who try the online tutorials, online training material and best practice guides;
- The number of LS software users who use the help desk;
- The number of downloads of optimised (base-code) binaries;
- The number of acknowledgements in scientific publications.

Tightly linked to dissemination, the its role of training is twofold

- general**: train LS scientists to include HPC practices into their routine
- specific**: train LS software users and developers to use the results of the project (base codes, libraries)

Training will take place in two forms:

- training events, both specific to ScalaLife and general training
- online material (tutorials, best practice guides)

The Training Plan (D2.2) is due at the end of the 1st project year

Training is backed up by

- user support (help desk, competence centres)
- hardware provision (provided through the competence centre)
- software provision (optimised binaries)

ScalaLife Specific Training:

Date	Institute	Event
24-26/01/2011	LRZ	Introduction to Molecular Modelling on Supercomputers
21-23/02/2011	KTH	PRACE Scientific Seminar
06-08/04/2011	MGMS, OeRC	“Membrane Proteins:Structure and Function”, ScalaLife Presentation
27/02/2012-02/03/2012	LRZ	ScalaLife Winter School

ScalaLife-related training:

Date	Institute	Event
18-20/10/ 2010	KTH	Cray Programming Environment Workshop
22/10/2010	KTH	Totalview Debugger Workshop
07-09/12/2010	LRZ	GPGPU Programming
13-15/12/2010	IRB, BSC, INB	Exascale Challenges in Computational Biology
13-14/01/2011	LRZ	Scientific 3D-Animation with Blender
09-10/06/2011	LRZ	NWChem Training tutorial
.../07/2011	OeRC	CUDA GPU Programming Course

Creation of good-quality training material and best practice guides is key to the success of the Competence Centre. But ScalaLife needs to offer something on top of manuals and peer advice.


Some general aspects to cover:

- understanding HPC (modern hardware, options available, etc.)
 - MPI processes and threads
 - inter-node communication
 - impact of accelerators
- formulation of LS problem, choice of models and algorithms (Gromacs – mostly PME, Dalton – ?)

What is the best format? (Very) short user guides?

Technical aspects to cover:

- cluster execution
 - fat node clusters, threaded MPI
 - using GPGPUs
- problem tuning
 - experimenting with scaling
 - load balancing: Gromacs processes for PME and PP (automatic), Dalton (?)
 - other parameters influencing message size



What is the best format? Detailed step-by-step examples, which users could try on hardware provided through the Competence Centre.

- CECAM application is not successful. Solution: the 1st International Workshop should be organised under this assumption.
- Low interest in online tutorials and best-practice guides. Solution: material should be diverse and continually updated.
- The production of software libraries and APIs is yet uncertain and depends on several factors.
- The effort going into producing good (online) training material is large:
 - video tutorials should ideally be like the TotalView videos
 - who is going to write new material?
- Online training via the Competence Centre must be backed by a minimal user support, which requires manpower.