CLOUDFLOW OPEN CALL 1

Short Technical Description of the CloudFlow Infrastructure

Call identifier: CloudFlow-1
Submission Deadline: 30th September 2014, at 17:00 h (Brussels local time)
Expected duration of participation: 1st January 2015 to 31st December 2015

Foreseen budget for CloudFlow-1: Up to 770,000 € funding for new beneficiaries

This amount of funding is planned to be spent on seven experiments. The maximum funding for new beneficiaries per experiment is expected to be: 110,000 €.
TABLE OF CONTENTS

1 Introduction to short technical description ........................................................................... 1
2 CloudFlow: System architecture Overview ........................................................................... 2
   2.1 CloudFlow: Definitions ..................................................................................................... 3
   2.2 The interplay of the CloudFlow components ................................................................... 6
   2.3 Required steps to integrate with the CloudFlow infrastructure ....................................... 8
3 CloudFlow: Hardware layer .................................................................................................... 9
   3.1 Cloud configuration .......................................................................................................... 9
   3.2 Hardware configuration .................................................................................................... 9
   3.3 Pricing information .......................................................................................................... 10
4 Examples of CloudFlow applications and services ............................................................... 11
   4.1 Generic services ............................................................................................................. 11
   4.2 Examples for experiment-specific services and applications ...................................... 12
5 CloudFlow: partners, competences, contributions ............................................................... 14
   5.1 Independent Software Vendors ...................................................................................... 14
      5.1.1 Missler .................................................................................................................... 14
      5.1.2 NUMECA .............................................................................................................. 14
      5.1.3 Jotne ...................................................................................................................... 15
      5.1.4 ITI ......................................................................................................................... 16
   5.2 Research institutes .......................................................................................................... 16
      5.2.1 Fraunhofer ............................................................................................................. 16
      5.2.2 SINTEF ............................................................................................................... 17
      5.2.3 DFKI .................................................................................................................. 18
      5.2.4 UNott ................................................................................................................... 18
   5.3 Innovation and technology consultancy ........................................................................... 19
      5.3.1 CARSA ............................................................................................................... 19
   5.4 End user .......................................................................................................................... 19
      5.4.1 Stellba ................................................................................................................... 19
   5.5 Hardware HPC provider .................................................................................................. 20
      5.5.1 Arctur .................................................................................................................... 20
1 INTRODUCTION TO SHORT TECHNICAL DESCRIPTION

This document gives a short technical description of the currently existing CloudFlow infrastructure and a look ahead to its future development, since all proposed experiments are supposed to be run using the CloudFlow infrastructure. This document is thus intended to give proposers the necessary information how to relate their experiments to the CloudFlow infrastructure and how to describe this relationship convincingly in their proposals.

There is further accompanying material for the first open call CloudFlow-1 that includes the Proposal Template


and the Guide for Applicants


These documents provide detailed information on the structure of the envisaged proposals and the rules for their submission.

In Section 2, a general overview of the CloudFlow system architecture and the interplay of its software components is given, leading to some technical considerations important for newly proposed experiments. This is followed in Section 3 by a description of the hardware layer including the corresponding pricing information. Section 4 illustrates CloudFlow concepts by providing some examples of generic services as well as experiment-specific services for an already running experiment. Since a new experiment will also include work – at varying levels – by the current consortium, the existing CloudFlow partners are presented in Section 5 to give an idea how they can contribute to newly proposed experiments.

For the homepage of the CloudFlow project please see:

www.eu-cloudflow.eu

The ongoing experiments concerning water power plant design and water turbine development (maintenance, repair and overhaul) are described at:

www.eu-cloudflow.eu/experiments

For further questions please first consult our frequently asked questions list (FAQ) under:

www.eu-cloudflow.eu/open-calls/faq

In case you do not find the answer to your question there, you can drop a question to us by contacting:

info@eu-cloudflow.eu

which the CloudFlow Competence Center will try to answer within 3 working days.
In this chapter we first present the abstract system architecture. This description ‘abstracts from’ the individual per-experiment services and applications and focusses on the common components of the CloudFlow Platform as presented in the following figure.

**FIGURE 1: THE LAYER STRUCTURE OF THE CLOUDFLOW PLATFORM**
2.1 CLOUDFLOW: DEFINITIONS

The following definitions are adapted to the CloudFlow project. The used terms may have a slightly different or wider meaning outside of this project.

**CloudFlow Platform**

The CloudFlow Platform encapsulates all of the components illustrated in Figure 1. It comprises everything which is required by the described CloudFlow components to perform their tasks and to communicate with each other.

**CloudFlow Portal**

The CloudFlow Portal is a web site that serves as a user entry point to the CloudFlow infrastructure. It is a web-based client through which a user may log in and interact with the CloudFlow system. As such, the CloudFlow Portal provides a web-based Graphical User Interface (GUI) that allows activating other CloudFlow tools and applications using a standard web browser.

**Desktop Application**

A Desktop Application is a program installed on a user’s local machine that interacts with the CloudFlow infrastructure. It can be used as an alternative to the Portal.

**Workflow**

A workflow is a sequence of connected steps needed to be performed to achieve a certain work task. Every workflow step is either a service or an application. CloudFlow workflows consist of one or more applications and services. Workflows (as well as services and applications) have semantic descriptions and are executed by the Workflow Manager.

The main idea of this concept is to hide the complexity of data flows and internal communication from the user so that she/he can focus on the actual engineering task. To further ease the use of workflows, pre-defined workflows will be created by system administrators in collaboration with experienced engineers and will be ready to use for end users. In the future, users will also be able to adapt pre-defined workflows depending on their individual needs using the Workflow Editor or to rely on an automatic assembly using semantic concepts.
**Semantic Service / Workflow Descriptions**

The machine-readable semantic descriptions of services, applications and workflows comprise the respective location, functionality, inputs and outputs. This kind of service description guarantees the compatibility between software from various vendors. The automated checking of semantic consistency and the inference of implicit knowledge are only two of many advantages offered by semantic service descriptions. In the future development, these will be further exploited, e.g., in the Workflow Editor.

---

**Workflow Manager**

The Workflow Manager is a CloudFlow component that automates the execution of workflows consisting of several services and applications and monitors their execution status. It hides the complexity of executed chains of services and applications by taking care of the data flow between the inputs and outputs of the single services and applications.

---

**Workflow Editor**

The Workflow Editor is a CloudFlow component that will allow users with the appropriate rights to specify new and/or edit/change already existing workflow definitions. The Workflow Editor backend will make use of the semantic service and workflow descriptions to enable a (semi-)automatic workflow creation/adaptation. It will provide a SOAP\(^1\) interface, to which a web-based GUI will be connected that enables comfortable access to the Workflow Editor’s capabilities via the CloudFlow Portal. The Workflow Editor is not yet included in the current state of development, but will be available for future experiments.

---

**Service**

A service is a piece of software / program that runs remotely on the Cloud performing computations. A service provides the computation results back to its caller. During its execution, it can, e.g., interact with the CloudFlow Storage Space or a dedicated data base. In CloudFlow, two types of services are distinguished:

- a **synchronous service** performs its computation before returning;
- an **asynchronous service** returns immediately and

---

\(^1\) Simple Object Access Protocol
notifies the Workflow Manager using SOAP when it is done.

Services do not rely on user input during their execution, but may display their status to the user, e.g., through a small web page. Services will be made available as SOAP web services with a standardized interface and must provide a corresponding interface description in form of a WSDL².

### Application

Applications provide a graphical interface for the CloudFlow user to interact with a running workflow. The purpose of applications may range from simple web pages to let the user set parameters for the following services or receive (intermediate) output values to complex interactive visualization and even remote access to commercial software packages. Since the execution time of an application is user-dependent and will certainly exceed standard SOAP communication timeouts, applications use the same mechanisms as asynchronous services to interact with the Workflow Manager.

### Storage Space

Storage Space is disc space (physical or virtual) provided by the CloudFlow Platform that enables to store input, intermediate, and output data. Data from this storage space can be downloaded to a local machine by a CloudFlow user for archiving the data in company data bases or additional use in other standard applications.

### VM – Virtual Machine

Virtual Machines are used to run services on the Cloud. They contain one or more services. To start a virtual machine, its image is loaded from the image repository located in the Cloud layer and executed using dedicated virtualization software.

In order to transparently allocate on-demand compute nodes for the execution of compute-intensive workflows, compute-resource-provisioning will be implemented for future experiments. This will provide abstractions independent from Cloud vendors for the CloudFlow Portal and Workflow Manager in order to bring up and shut down on-demand instances with the correct virtual machine for a given workflow execution.

---

² Web Services Description Language
**Image Repository**

The (virtual) Image Repository is a repository that holds images of all VMs that are available in the CloudFlow context. To deploy a service in the Cloud, a corresponding virtual image has to be created by the service provider, bundling the service executable with the needed environment. Usually, one starts with an image provided by the Cloud provider and installs and configure additional software (if any). The new image is stored in a database in the Cloud infrastructure.

**Hardware**

Hardware is the actual compute resource on which services are executed. We distinguish between Cloud hardware resources and HPC resources (for more details see Chapter 3).

### 2.2 THE INTERPLAY OF THE CLOUDFLOW COMPONENTS

Figure 2 shows the interplay between the main software components of the CloudFlow Platform during the execution of a workflow in a high-level schematic way.

**FIGURE 2: THE INTERACTION OVERVIEW OF THE MAIN CLOUDFLOW SYSTEM COMPONENTS**

The Workflow Manager plays the central role of a broker to hide the complexity of executed chains of services and applications. It allows to invoke and to monitor the execution of services and applications, automating complex workflows that consist of many services / applications.

The Workflow Manager is able to start a selected workflow and to pass the needed input parameters. Then semantic descriptions of relationships between services and applications which
take part in the ordered workflow are used to pass output values from finished services and applications to the later stages in the workflow. Asynchronous services have to provide a current state of computation progress on request during their execution. The progress status of the workflow can then be requested from the Workflow Manager and displayed to the end user. With the same mechanism, it is also possible for the Workflow Manager to pass the whole graphical user interface from an asynchronous service (in this case called an application) to the end user. At the very end of the workflow, its final results are stored by the Workflow Manager and can later be accessed by the user upon request.

All functionality of the Workflow Manager is available through a SOAP web service interface. It can then be accessed by user-friendly clients, as, e.g., the Portal. Other clients, such as Desktop Applications, are also possible.

Distinct services and applications run in separate Cloud environments and thus need the Workflow Manager as the centralized system to initiate and control the workflows in which they are integrated. Consequently services and applications have to implement a common interface that allows full compatibility with the Workflow Manager and with other services and applications from various vendors. The semantic descriptions of services, applications and workflows chosen by CloudFlow guarantee this compatibility and let the user select workflows without the need of specific expert knowledge for each of the services contained in a chosen workflow.

The CloudFlow software components aim at being neutral to Cloud providers and Cloud APIs. Commonly used functionality is implemented as separate services, hiding vendor specific APIs. The current implementation uses a mix of Cloud software components. For example, the commercial software vCloud is used for controlling virtual machines, while solutions from the open source project OpenStack are used for authentication.

With the CloudFlow Platform, engineers as the intended main end users are given a universal tool to configure, track and interact with expert software without any knowledge about its implementation. To complete a certain engineering task end users can select a suitable pre-defined workflow using the CloudFlow Portal and start its execution. In the future, using a Workflow Editor component currently under development, they will also have the possibility to create new workflow definitions or adapt existing ones.

The end user is only supposed to interact with the workflow at certain points, when its execution reaches an application. In this case, the user may for example be asked to provide additional input of domain specific configuration information which is not available as output from previously executed services in the workflow and cannot be specified in advance. Applications may also display intermediate output values or can even be interactive visualizations. Applications are implemented to run in a browser and can thus be displayed inside the CloudFlow Portal or in Desktop Applications.

Note that the current version of the Workflow Manager only supports services using a SOAP API, where all parameters are embedded in an XML document compliant with the respective SOAP schema. WSDL (Web Service Description Language) interface descriptions are used to specify the interfaces of services and applications. The machine-readable semantic descriptions also contain references to the information stored in the WSDL files of the respective services and applications, as for example their exact invocation endpoint.
2.3 REQUIRED STEPS TO INTEGRATE WITH THE CLOUDFLOW INFRASTRUCTURE

New project partners joining the CloudFlow project need to cast the functionalities of their software into services and applications which are compatible with the CloudFlow infrastructure.

The following important criteria and steps are required for taking advantage of the CloudFlow infrastructure:

- The encapsulation of certain software functionality into independent services. Independent means that their execution should solely be dependent on their declared input parameters.
- Services must be realized as SOAP web services and must provide an interface description in the form of a WSDL file.
- The implemented web services must be deployed in a server which can be installed on a Virtual Machine in the Cloud.
- Long lasting computations must be implemented as asynchronous services as defined by our infrastructure. This means that the SOAP request spawns the computation and returns immediately, but the computation continues its execution in the background. The server provides a SOAP method at the same web service endpoint, which can be used to get the status of the long lasting computation upon request. When the computation has finished, the Workflow Manager has to be given the final output values. This must happen by calling a certain SOAP method offered by the Workflow Manager’s web service interface.
- Applications are also implemented as SOAP web services and need to provide a WSDL interface description. In addition to the behavior of services, they must provide a web-based Graphical User Interface (GUI) as one of their return values. Either they return their complete web-based user interface (which may, e.g., contain Java Script code) directly, or deliver a new, updated version of the user interface (depending on the actual state of the computation) whenever they receive a status request by the Workflow Manager. In the latter case, the applications need to provide a certain SOAP method at the same web service endpoint that returns the actual user interface.
- For GUI applications running in the Cloud, it is possible to view and interact with the user interface using a remote desktop-like connection, e.g. VNC, through the web browser.
- Services that produce large amounts of data can use the generic storage services offered by the CloudFlow infrastructure to store and share data between services. The link to the stored results can then be returned as output of the service.
3 CLOUDFLOW: HARDWARE LAYER

In this section we describe the current CloudFlow hardware infrastructure as provided by our partner Arctur. Please note that it is possible to suggest running an experiment on an alternative hardware infrastructure including a new hardware infrastructure provider as an experiment consortium partner, provided that the CloudFlow architecture is used.

3.1 CLOUD CONFIGURATION

Arctur uses CentOS 6 distributions of Linux in its datacenters wherever possible. If Linux is not a possibility, then Windows Server 2008R2 or Server 2012 can be used. Arctur’s goal is to reduce manual work to a minimum; therefore most of the server infrastructure is operated by means of the Puppet configuration management system.

VMware vCloud is used for Cloud services. On top of the actual hardware ESX 5.1 provides a hypervisor layer. In the application layer different operating systems can be used by the consumer. Amongst the most used ones are the Ubuntu and CentOS distributions of Linux followed by Microsoft Windows Server solutions.

3.2 HARDWARE CONFIGURATION

Arctur’s systems to run production tasks in the Cloud can be categorized as:

- hosts for virtual machines,
- hosts to provide storage over iSCSI and
- hosts covering infrastructure support roles.

They are generally equipped with dual core Xeon 51xx and quad core 54xx CPUs. iSCSI runs over 1Gbit Ethernet, provided by 3com 4500G switches.

The Arctur HPC system is based on an IBM iDataPlex solution. Each of the 84 nodes is equipped with two Xeon X5650 CPUs and 32GB of memory. They are connected with a high speed low latency InfiniBand network, used mainly for MPI communications between nodes. This provides for 10TFlops of computing power. They are also connected with Gbit Ethernet for management and another separate Gbit Ethernet for the parallel file system Lustre. The Lustre system is aggregating over 100TB capacity and 4.5GB/s throughput.

Head node to the HPC cluster is an IBM machine with the same CPUs as the compute nodes but 48GB of memory and more local storage.

One node in the system is equipped with an NVidia GPU. Due to the configuration of the InfiniBand network this GPU node can be reached by every node in the HPC system, thus enabling a boost in performance if the application is able of taking advantage of CUDA capabilities.

The HPC system/cluster is running CentOS6 with usual HPC add-ons (OFED, MPIs, etc.). Queuing and resource management is implemented with Torque PBS + Maui.

User applications and software are installed and configured on a pre-experiment basis during the testing and debugging period. Therefore the software stack is constantly being adapted to current demands. The software configuration is primarily left to the actual users to configure as they need. If the configuration is needed on an administrative level, a team of HPC system administrators supports or independently deploys the software solutions that are needed.
3.3 PRICING INFORMATION

In the following we present a general pricelist for usage of the above described infrastructure to indicate which running costs an experiment may imply. As stated above, experiment consortia can contain other independent HPC / Cloud providers.

HPC rental is done on a per usage basis. The calculation basis for the rental is one CPU-hour (CPUh). This unit is represented by one CPU working for one hour of wall clock time. One hour of actual usage is defined as every period of time when a single CPU core is running above its idle speed for more than 5 minutes in total.

The user always gets access to minimum one complete server, meaning 12 compute cores and 32GB of RAM.

Each user (account) on the system is given 100GB of storage space for free. If the user requires more storage, additional space can be rented. The minimum rental of additional storage is 1TB per month.

The rental prices of the HPC infrastructure are as follows:

- 0.05 € per CPUh
- 50 € per 1TB of storage per month

In addition to the use of the physical HPC infrastructure the end user use also the Cloud infrastructure that supports the services and workflows. The rental is done on top of the HPC rental. The rental of the Cloud infrastructure is also done on a monthly basis, but here the resources can be dynamically allocated as defined by the user. The changes of configuration can be done on an arbitrary basis by the end-user. The monthly fee is in this case calculated by calculating full single days in which a configuration has been used.

The general pricelist of the available resources is:

- 1,000 Mhz CPU = 5 €/month
- 1,024 MB RAM = 9 €/month
- 100 GB storage tier-2 = 16 €/month (tier-2 = SATA based storage)
- 100 GB storage tier-1 = 19 €/month (tier-1 = SAS based storage)
- 100 GB storage tier-0 = 27 €/month (tier-0 = SSD based storage)
- Windows Server 2012 license = 38 €/month

To provide a general idea of the costs that arise from the usage of the Cloud infrastructure we present here some indicative cases of infrastructure usage:

<table>
<thead>
<tr>
<th>Size of VM</th>
<th>Specs</th>
<th>Price / month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiny</td>
<td>1 CPU, 1 GB RAM, 16 GB HDD</td>
<td>28 €</td>
</tr>
<tr>
<td>Small</td>
<td>2 CPU, 4 GB RAM, 20 GB HDD</td>
<td>66 €</td>
</tr>
<tr>
<td>Medium</td>
<td>8 CPU, 16 GB RAM, 32 GB HDD</td>
<td>208 €</td>
</tr>
<tr>
<td>Large</td>
<td>12CPU, 32 GB RAM, 100 GB HDD</td>
<td>388 €</td>
</tr>
</tbody>
</table>

The prices indicated are just for information and can change on a daily basis, depending on the specifics of the VM. Additionally if Microsoft licenses are needed this also has to be accounted for in the monthly rental price. We suggest Windows Server 2012 at a price of 38 € per month.

All potential proposers willing to use Arctur's infrastructure are encouraged to contact Arctur during the proposal preparation phase for a more detailed discussion of the Cloud computing needs for their envisaged experiment at hpc@arctur.si. Arctur also has a web page giving indications about pricing at https://www.arctur.si/solutions_and_services/data_center.virtual_server_hosting.
4 EXAMPLES OF CLOUDFLOW APPLICATIONS AND SERVICES

The CloudFlow Platform, with its current infrastructure and six internal experiments, consists of dozens of services and applications. In this section we want to further illustrate the CloudFlow concepts, applications and services by relating them to one of the ongoing experiments in the project. First we present some examples of selected generic services that all experiments are supposed to use for basic operations. Then we exemplify experiment-specific applications and services using one experiment of the current first wave of CloudFlow experiments.

4.1 GENERIC SERVICES

We first describe some generic services that provide functionality that is likely to be required by many CloudFlow workflows. These services provide functionality to interact with the Cloud infrastructure. It is intended that the different software vendors rely on these services instead of using their own implementations. Since some of the services are handling large amounts of data, they have a REST\(^3\) instead of a SOAP interface.

**TABLE 1: GENERIC STORAGE SERVICES**

<table>
<thead>
<tr>
<th>Identification</th>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service STORAGE-1</td>
<td>getResourceInformation</td>
<td>This service provides a common interface for different file storage solutions. It takes a fileId as input and returns information of how to manipulate the file via a REST API.</td>
</tr>
<tr>
<td>Service STORAGE-2</td>
<td>listFiles</td>
<td>This service provides a common interface for different file storage solutions.</td>
</tr>
</tbody>
</table>

Another important set of services are the Cloud Management Services which are mainly used by the CloudFlow infrastructure internally. Of course, services or applications from vendors may use it if appropriate. The Cloud Management Services enable to gather information about the CloudFlow vCloud environment such as which virtual applications (V Apps) and virtual machines (VMs) are available, what their status is and what properties they have. Furthermore, they allow starting and stopping VMs. The Cloud Management Services are available as web services as well as a Java library (Jar) to be directly integrated in services requiring additional computer resources.

**TABLE 2: LIST OF CLOUD MANAGEMENT SERVICES**

<table>
<thead>
<tr>
<th>Identification</th>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service vCloud 1</td>
<td>getAvailableVMs</td>
<td>This service delivers functionality for listing available virtual machines. Return value is the IP address as well as the port where the webserver is accessible.</td>
</tr>
<tr>
<td>Service vCloud 2</td>
<td>startVM</td>
<td>This service enables to start v Apps and VMs containing services given a virtual image.</td>
</tr>
<tr>
<td>Service vCloud 3</td>
<td>stopVM</td>
<td>This service enables to stop VMs in order to save computational resources.</td>
</tr>
</tbody>
</table>

---

\(^3\) Representational State Transfer
4.2 Examples for Experiment-Specific Services and Applications

As an example, we consider one of the six experiments currently being carried out by the CloudFlow consortium, see

www.eu-cloudflow.eu/experiments/experiment_6

Use case to be examined: The experiment is intended to compare the actual shape of water turbine blades (of Kaplan or Francis type) with their nominal CAD models, addressing

- the comparison of the manufactured turbine blades with their nominal CAD shapes;
- the checking of older turbines and their blades for wear and
- the decision on the need for repairs and upgrades.

The experiment is a first step for facilitating the efficient use of Cloud resources for updating and reverse engineering of turbine geometries from measured point clouds.

The partners for this experiment make up a valid consortium as defined in the GfA, which comprises the end user Stellba providing this use case, the software vendor Jotne providing pieces of its commercial software for Product Life-cycle Management, Arctur as the HPC provider and SINTEF as the R&D institution. In addition, the experiment will be evaluated by partner UNott and possible business models examined by partner CARSA. See Section 5 for more information on all these partners and their roles in CloudFlow.

The main implementation tasks of this experiment are to:

- read a CAD model of a turbine and read a point cloud data set measured from a corresponding physical turbine,
- register the point cloud and the CAD model directed by the user providing information on approximate corresponding points in the point set and on the CAD model,
- calculate the distance between the points in the registered point cloud and the surfaces of the CAD model and
- visually compare the registered point cloud and original CAD model.

The experiment is built upon existing software tools from project partners, including SINTEF’s GoTools (Mathematics and data structures for design, analysis and manufacturing) and partner Jotne’s commercial software EPM Data Manager™ (for reading STEP files). Interactive visualization is realized using SINTEF’s Tinia framework with an added adaption layer for the CloudFlow Platform, exposing the functionality through a CloudFlow compatible SOAP API.

The services are implemented using Java EE, which makes it easy to deploy function calls as SOAP services. Simple web applications, e.g., file chooser, are written using JavaScript and HTML5.

Note that this particular example should in no way be considered as binding for the scope of an experiment and the definition of its corresponding applications and services.

The workflow is performed in the following steps.

1. If either the file containing the point cloud or the CAD model are not already uploaded to the Cloud infrastructure, the user starts the application "Upload files to Cloud infrastructure" to upload local files (CAD, point cloud) to the Cloud. The application accesses the corresponding service "Upload file to the Cloud infrastructure". The engineer uses the application "Selection of CAD model and point cloud" to choose which two files to compare. The Workflow Manager creates graphical representations for the CAD model and the point cloud by triggering "Generate representation of CAD model for visualization" and "Generate representation of point cloud for visualization".
2. The Workflow Manager starts the application "Identify corresponding points in CAD model and point cloud", which visualizes the two files to the user for input of the set of corresponding points to be used for registering.

3. The service "Guided registration of CAD model and point cloud" is called to calculate the registration matrix (4x4 transformation matrix).

4. Then the Workflow Manager starts the service "Calculate difference between CAD model and point cloud" to perform the actual difference calculation of the CAD model and the point cloud after registration.

5. Afterwards the service "Prepare visualization of difference between CAD model and point cloud" is started by the Workflow Manager to perform the necessary preparation for visualization of the differences between CAD model and point cloud.

6. Finally the application "Visualize difference between CAD model and point cloud" is started by the Workflow Manager, which allows the user to examine the differences between the CAD model and the measured point cloud.
5 CLOUDFLOW: PARTNERS, COMPETENCES, CONTRIBUTIONS

In this section, the existing CloudFlow partners are presented to give proposers an idea how these can contribute to newly proposed experiments. More detailed descriptions of the six existing experiments can be found at

www.eu-cloudflow.eu/experiments/experiment_1
www.eu-cloudflow.eu/experiments/experiment_2
www.eu-cloudflow.eu/experiments/experiment_3
www.eu-cloudflow.eu/experiments/experiment_4
www.eu-cloudflow.eu/experiments/experiment_5
www.eu-cloudflow.eu/experiments/experiment_6

First, the independent software vendors are presented, then the R&D institutions, the partner for innovation and business consultancy, the end user partner and finally the partner providing HPC and Cloud infrastructure.

5.1 INDEPENDENT SOFTWARE VENDORS

5.1.1 MISSLER

General company info: Missler Software is a leading global developer and supplier of CAD and CAM solutions (see www.topsolid.com/). The company provides integrated software solutions for product design, manufacturing and data management. Missler Software offers uniquely integrated solutions for the mechanical engineering, sheet metal and wood industries.

Experiment involvement: Experiments 1 & 2

Specific contributions to CloudFlow and its experiments: The CloudFlow project will allow Missler to explore new business models and to acquire new types of technology.

Missler provides CAD & CAM applications for the experiments 1&2. It is developing a dedicated CAD application for specialized design, which is hosted on the Cloud and can be loaded on demand in a local machine (Software as a Service - SaaS). The CAM application is being developed to work in an asynchronous and remote mode for the second experiment.

Future possibilities: Missler is ready to supply CAD and CAM functionality to new Application Experiments to “cloudify” additional parts of its offering of integrated software solutions.

5.1.2 NUMECA

General company info: NUMECA Int. is a Belgian SME, with headquarters located in Brussels, active in the development and marketing of advanced Computational Fluid Dynamics (CFD), grid generation, multi-physics and optimization software (see www.numeca.com). NUMECA’s software allows simulation, design and optimization of fluid flow and heat transfer and is recognized worldwide for its superior performance, reliability and ease-of-use.

Experiment involvement: Experiment 3

Specific contributions to CloudFlow and its experiments: NUMECA provides its CFD Solver and its Mesh Generator that are going to be used during CFD experiments. For the needs of this project NUMECA adapted its software and developed the Web Services that allow a desktop application to
be able to operate through the Web. The Web Services are currently deployed at the CloudFlow HPC provider.

The Mesh Generator is a fully automatic hexahedral grid generator for all types of rotating machinery: complex axial, radial, and mixed-flow configurations. The Grid Generator (AutoGrid5™) produces grids in a few minutes, with a few clicks through a wizard based interface. The Grid Generator receives an input in CAD format, along with the NUMECA mesh configuration file and produces a grid that is exported in CGNS or AP209.

The CFD Solver is 3D Navier-Stokes flow solver, recognized as the fastest and most accurate for turbomachinery configurations. Acceleration of several orders of magnitude is achieved thanks to the CPUBooster™ technique, NLH approach and HPC scalability. The CFD Solver receives as input the files that are exported by the Grid Generator along with the CFD Solver configuration file and produces the results that are then used for post-processing.

Future possibilities: Potential future NUMECA contributions to CloudFlow could include a continuation of the “CFD on the Cloud” objective, but also shape design optimization using numerical CFD simulations coupled with evolutionary algorithms, to predict optimal shapes with respect to user-defined objective functions (i.e., efficiency maximization, drag reduction,...).

5.1.3 JOTNE

General company info: Jotne’s software products (see www.jotne.com/epmtech) have successfully reduced development and product lifecycle costs through the use of intelligent and standards-based data management in the areas of Defense, Aeronautics, Oil & Gas, Built Environment and Aerospace. EXPRESS Data Manager™ (EDM) offers interoperability functionality for data exchange, data sharing, data integration and data archival by fully implementing the methodology of ISO 10303 (STEP).

Experiment involvement: Experiments 1, 3, 4 & 6

Specific contributions to CloudFlow and its experiments:

- EDMOpenSimDM™ - adopts Product Data Management (PDM) paradigms to engineering analysis needs and configures and manages design and analysis data jointly
  o Is accessed via SOAP web-services with data and files that are of proprietary formats or compliant to ISO 10303-209 (AP209).
- EDMinterface(C++/AP209)™ - Application Programmers Interface (API) to ease access to AP209 data, for example, to write data translators to and from proprietary formats
  o Reads and writes data compliant to ISO 10303-209.
- EDMdeveloper™ - Development platform for data modelers, application developers and method developers with powerful functions for data validation and data conversion
  o Can be programmed to read and write any format, but has special provisions for ISO 10303-21, 26 and 28 and for all data dictionaries written in EXPRESS (ISO 10303-11), including AP209.

Future possibilities: Jotne’s main focus for future experiments is to complement the current solution for Simulation Data Management with Test Data Management and to extend this further into through-life support, that is, true Product Life-cycle Management (PLM), to make industry benefit from the collaborative potential of Cloud-hosted engineering services throughout the product life.
5.1.4 ITI

General company info: ITI GmbH is an internationally active CAE specialist for virtual system engineering. The company’s core business activities are the development of standard software tools for engineers and scientists as well as engineering services at a high level of know-how, reliability and transparency. ITI solutions are successfully used in the fields of automotive technology, power transmission, fluid power systems, thermal and energy systems and for plant construction. ITI’s SimulationX tool package is based on leading software technologies like Modelica, intuitive GUI, TypeDesigner, database links, efficient code generation and solvers.

Experiment involvement: Experiment 5

Specific contributions to CloudFlow and its experiments: ITI focuses in CloudFlow on technology platforms to implement Cloud-based simulation management. An important objective is the interactive user interface for defining, executing and analyzing system simulations from any computer without software requirements:

- Systems simulation web services for
  - Management of simulation data (models, parameters, results)
- HTML based GUI providing access to:
  - Parameterization
  - Running simulations
  - Result visualization
  - Result analysis

Future possibilities: New experiments could reuse the systems simulation services and the user interface as long as models provide a FMI for co-simulation.

5.2 RESEARCH INSTITUTES

5.2.1 FRAUNHOFER

General info: Fraunhofer is Europe’s largest application-oriented research organization. The research efforts are geared entirely to people’s needs: health, security, communication, energy and the environment.

Fraunhofer is represented by two institutes: Fraunhofer IGD ([www.igd.fraunhofer.de/iet](http://www.igd.fraunhofer.de/iet)) concentrates on technology research dedicated to the application of science and computing in different tasks and workflows, including design, analysis, visualization, and interpretation. The main focus is on coupling real time visualization and interactive simulation through computer graphics technology. Fraunhofer EAS ([www.eas.iis.fraunhofer.de](http://www.eas.iis.fraunhofer.de)) focuses on deploying research results in the field of simulation, co-simulation, Grid and Cloud computing to industry and research partners.

Experiment involvement: Experiments 4 & 5

Specific contributions to CloudFlow and its experiments: Fraunhofer contributes the following competencies and software components to the CloudFlow Infrastructure:

- GridWorker Service for parallel simulation and optimization as well as variant simulations using the FMU (functional mockup unit) or other simulation tools
• Interactive fluid dynamics and structural mechanics simulations in 2D and 3D with efficient simulation data transfer
• HTML converter service where selected 3D simulation data can be converted into an HTML page for 3D visualization using X3DOM and WebGL in a Web Browser
• Visualization client as a post-processing functionality for 3D simulation data
• Cloud Management Service to start and stop virtual applications and machines

**Future possibilities:** For new experiments Fraunhofer can provide the following know-how and software components:

- massively parallel programming, GPU programming
- simulation data transfer
- simulation data visualization (scientific visualization - SciVis) on desktop machines and in Web browsers

The interactive simulations in the Cloud focus on high update rates of time-dependent simulation and therefore offer new opportunities for coupling simulations. Furthermore, for selected simulation results the HTML converter service or the visualization client may be beneficial for future partners to visualize 3D simulation results – also in Web Browsers. Future experiments could also benefit from ongoing developments regarding new approaches for interactive simulation and visualization.

GridWorker allows parallel simulations and optimizations as well as variant analyses supporting a multitude of simulation tools. Further simulators or even programs or algorithms may be supported on demand for further experiments.

### 5.2.2 SINTEF

**General info:** SINTEF is an independent research foundation in Norway predominantly funded through contract research. In CloudFlow SINTEF is represented by the Department of Applied Mathematics (see [www.sintef.no/math](http://www.sintef.no/math)). The department has long traditions within simulation, geometry, optimization, GPUs, modern parallel architectures, and visualization. The deployment of research results to industry through easy-to-use program libraries is represented in CloudFlow by GoTools and Tinia.

**Experiment involvement:** Experiment 6

**Specific contributions to CloudFlow and its experiments:** SINTEF is the technical coordinator of CloudFlow, heads the system design group and leads Experiment 6.

SINTEF contributes the following software components to the CloudFlow Infrastructure:


**Future possibilities:** For new experiments SINTEF can provide tailored geometry processing and visualization by using/extending the functionalities of GoTools and TINIA.
5.2.3 DFKI

General info: DFKI is one of the largest non-profit contract research institutes in the field of innovative software technology based on Artificial Intelligence methods (see www.dfki.de). The department of Innovative Factory Systems (IFS) investigates how developments in ICT can be transferred into the manufacturing and engineering domain. One main focus is the process control using semantic knowledge representation and service-oriented devices. A significant amount of experience in this research area has been attained and forms the basis for DFKI’s activities in CloudFlow.

Experiment involvement: All experiments

Specific contributions to CloudFlow and its experiments: Contributed software components to the CloudFlow Infrastructure are

- Workflow Manager - The central component of the internal CloudFlow infrastructure, which is responsible for orchestrating workflows, composed of several distinct applications and services using their semantic descriptions.
- Semantic service and workflow descriptions - The inputs and outputs of single services and applications are described semantically in a vendor-independent fashion. Based on this, semantic workflow descriptions define the sequence of services and applications to be executed and model the individual connections between their inputs and outputs. The descriptions are based on the OWL-S standard and guarantee a maximal degree of interoperability and flexibility.

Future possibilities: Future experiments including new applications and services may introduce additional requirements for the Workflow Manager, which needs to be continuously maintained and extended. While the semantic workflow descriptions are static for the first wave of experiments, they build the basis for a guided flexible workflow adaptation and creation by end users which will be implemented for the next waves of experiments.

5.2.4 UNOTT

General info: The University of Nottingham is in the Russell Group as one of the leading UK teaching and research institutions. It is represented in CloudFlow by the Human Factors Research Group (HFRG), which has a long-standing international reputation in user-centred design and human factors methodologies (www.nottingham.ac.uk/engineering-rg/manufacturing/humanfactors/index.aspx). In CloudFlow, we are conducting user requirements gathering with end users, providing recommendations towards the development and integration of technologies, and evaluating the impact of technologies in the contexts of design, engineering and manufacturing.

Experiment involvement: All experiments

Specific contributions to CloudFlow and its experiments: UNott contributes to tasks which concern research and application of human factors, specifically requirement analysis, validation methodology, as well as assessment and validation of all experiments in the first wave.

Future possibilities: For the new experiments, UNott will assist new partners to identify experiments requirements analysis and methods to evaluate experiments. UNott will also provide expertise in supporting the evaluation for aspects that are related to human factors, e.g., usability.
5.3 INNOVATION AND TECHNOLOGY CONSULTANCY

5.3.1 CARSA

General company info: CARSA operates in three main areas, namely innovation, technology and internationalization (see www.carsa.es). These three areas are the mainstays for a number of specific services targeted at private firms such as the design and definition of competitive business models depending on the context and domain of each business. CARSA provides expertise to generate revenue and make a profit from our customers' operations, and employs different strategy tools and processes in order to define the best possible business model for each company.

Experiment involvement: All experiments

Specific contributions to CloudFlow and its experiments: CARSA contributes to designing, studying and testing suitable Cloud business models for each software vendor currently involved in CloudFlow. CARSA analyses the seven key factors of the Osterwalder Canvas, both individually (selling particular software tools) and collectively (exploiting a common commercialization platform).

In all experiments of the first wave CARSA is addressing the future uptake and sustainability of CloudFlow outcomes by analyzing the current market outlook, providing requirements according to existing needs and developing a strategy and business model for better exploiting those results in the market.

Future possibilities: For the new experiments, CARSA can support the new partners by examining different Cloud-based business options and their impact on the company's market and revenues. Thus CARSA can help each partner in the selection and fine tuning of the best mechanisms to succeed in the marketplace.

5.4 END USER

5.4.1 STELLBA

General company info: Stellba Hydro is a German SME. Stellba’s main competence is the service, repair and modernization of small and medium sized hydro power plants (see www.stellba-hydro.de/). Several tasks such as mechanical design, hydraulic and structural optimization and reverse engineering form the daily engineering workflow.

Experiment involvement: All experiments

Specific contributions to CloudFlow and its experiments: Stellba is the only end user in the first wave of ‘internal’ ongoing experiments and consequently involved in all experiments. CAD and CAM systems, simulation software, data handling and reverse engineering are part of the daily work. Therefore Stellba can define different use cases for the experiments, delivers experiment descriptions in collaboration with the experiment leaders (software vendors) and contributes to the CloudFlow infrastructure by means of practical issues from the end-user view. In addition, Stellba’s engineering tasks allow defining a specific workflow per experiment and even linking several experiments by means of a comprehensive workflow.

The execution and evaluation of the experiments is the main part of Stellba’s contribution to CloudFlow. A benchmark of the “traditional” way for performing the engineering tasks sets the baseline for the later evaluation of the CloudFlow experiments.
Future possibilities: The main contributions of Stellba are in the first wave of internal experiments. Further contributions in terms of more complex workflows and a higher degree of automation are possible but have to be coupled with another end user as a new project partner.

### 5.5 HARDWARE HPC PROVIDER

#### 5.5.1 ARCTUR

**General company info:** Established in 1992, Arctur has progressed to become the main Slovenian commercial supplier of HPC (High Performance Computing) services and solutions (see [http://hpc.arctur.net/](http://hpc.arctur.net/)). In 2009/2010 Arctur has established its own High Performance Computing infrastructure (in the 10 TFlops range) to be used as the technological foundation for advanced HPC and Cloud solutions and innovative web services in a distributed, high-redundancy environment. The company has extensive experience in server virtualization and deployment, integration of disparate IT-systems, IT support of project management and server farm leverage for the deployment of Software as a Service (SaaS).

**Experiment involvement:** All experiments

**Specific contributions to CloudFlow and its experiments:** Arctur is the provider of hardware and Cloud resources to the project (see Section 3) and thus supporting all of the current CloudFlow experiments. In addition to the ICT support that Arctur brings to the project, the company is also contributing strongly to the development of all of the layers of the CloudFlow infrastructure, effectively acting as the consultant for the development of the platform from the infrastructure provider’s point of view.

**Future possibilities:** Arctur will continue to support the project with the infrastructure and the specific know-how that is crucial to the intended future development of the CloudFlow Platform. Furthermore, Arctur is currently deploying the OpenStack Cloud environment to be used in the project. In addition, Arctur will support the project and its experiments with any new additions to the infrastructure equipment that might be put into production use in its datacenter during the duration of the project. As HPC and Cloud computing experts, Arctur’s staff members can also support consultancy demands that come from the end users that are connected to and using their infrastructure.