Floreon+: a web-based platform for flood prediction, hydrologic modelling and dynamic data analysis

Vaclav SVATOŇ[1], Michal PODHORANYI[1], Radim VAVŘÍK[1], Patrik VETEŠKA[1], Daniela SZTURCOVÁ[1,2], David VOJTEK[1,2], Jan MARTINOVIČ[1], Vít VONDRÁK[1]

[1] IT4Innovations, National Supercomputing Center, VSB-Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava, Czech Republic

[2] Institute of Geoinformatics, Faculty of Mining and Geology, VSB-Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava, Czech Republic

vaclav.svaton@vsb.cz, michal.podhoranyi@vsb.cz, radim.vavrik@vsb.cz, patrik.veteska@vsb.cz, daniela.szturcova@vsb.cz, david.vojtek@vsb.cz, jan.martinovic@vsb.cz, vit.vondrak@vsb.cz

Abstract

The main goal of this article is to describe the overview of Floreon+ system, an online flood monitoring and prediction system, which was primarily developed for the Moravian-Silesian region in the Czech Republic. Moreover, the article specifies the basic processes, which are implemented for running automatic and on-demand simulations that utilize the High Performance Computing (HPC) infrastructure. The main purpose of hydrodynamic models in the disaster management context is to provide an accurate overview of hydrologic situation in a given river catchment. In the event of extreme weather conditions, such as unusually heavy rainfall, these models could provide valuable information about imminent flood risk endangering a particular area. In the disaster management context, time plays a very significant role. Up to date and accurate results obtained in a short time can be very helpful. The availability of such results can be significantly improved by utilization of HPC resources and tools. The article describes the individual parts of the system in terms of data types, dynamic data processing, visualization, and the overall architecture.

Keywords: Floreon+, hydromodelling, geovisualization, high performance computing, dynamic data.

INTRODUCTION

Almost all large rivers in Central and Eastern Europe have experienced catastrophic flood events, e.g. the 1993 and 1995 flooding of the Rhine River, 1999 and 2002 Danube/Theiss Rivers, 1997 Oder River, 2001 Vistula River and 2002 Labe River. Each of these flood catastrophes brought damage worth several billions of euros; in some of the cases, as many as hundreds of people lost their lives (Halmo 2006). Floods, however, affect not only Central and Eastern Europe but also they represent a major problem in many regions all around the world (Knebl et al. 2005).

One of the directions contemporary world's science is taking is towards the modelling of flood catastrophes by means of available hydraulic/hydrologic modelling tools. Flood modelling, which has acquired crucial importance over the past years, is constantly enhancing its quality in connection with the development of information technologies (software, hardware). It has also considerably improved with the advent of geographic information systems, radar-based estimates using the Next Generation Radar (NEXRAD), high-resolution digital elevation models (DEMs), distributed hydrologic models and delivery systems on the Internet (Garrote and Bras 1995).

In practice hydrological models are usually coupled into platforms or systems, which are then used as a tool for flood prevention/flood control. These platforms provide similar functionality on different levels of complexity, therefore none of them carry out the process of predicting and modelling in the same way. According to provided information, flood modelling and monitoring platforms may be divided into two groups: (i) Platforms providing modelling information about flood events. Representative of the first group of platforms is eWaterCycle (Hut et al. 2016), which provides global modelling data about floods. (ii) Platforms offering information about monitoring of the current state. Such platforms mainly dispatch systems for coordinating different kinds of resources, preparing emergency plans and deploying experts and equipment. These kinds of platforms are based on exchanging information and data between them and national services (ERCC - Emergency Coordination Centre; EENA - European Emergency Number Association).

Nevertheless, the introduced Floreon⁺ system (Kuchar et al. 2016), which is a combination of both of the above-mentioned platform groups, is slightly different. The main objective of the Floreon⁺ project is to create a platform for integration and operation of monitoring, modelling, prediction and decision support for disaster management mainly in the Moravian-Silesian region. The central thematic area of the project is hydrologic modelling and prediction. The system focuses on acquisition and analysis of relevant data in real time and application of prediction algorithms with this data. The results are then used for decision support in disaster management processes by providing predicted discharge volumes on measuring stations and prediction and visualization of flood lakes in the landscape (Podhoranyi et al. 2016). Modularity of the Floreon⁺ system, which is developed for this science and research platform, allows for simple integration of different thematic areas, regions, and data. In addition to the flood modelling and simulation area, the Floreon⁺ system tackles other thematic areas - air and water pollution and environmental hazards and traffic situation. The results of these areas still focus on supporting the decision and planning disaster management processes.

Floreon⁺ system uses various types of static and dynamic data for operational running and therefore suitable database structure must be applied to ensure optimal and fast access to them. The database has to be prepared for storing big amount of continual data that are obtained from different geo-servers. The part of the Floreon⁺ system related to data and DB structures is very important for correct functionality of the entire system and thus this article tries to explain basic principles and processes inside the system database.

The major objective of the article is to present basic structure and data processing within the Floreon⁺ infrastructure, which covers many types of different data and DB structures. Main goal aims at heterogeneous spatial-temporal data (stored in PostgreSQL database) with a focus on describing the origin of data and the process that is optimal for storing data efficiently in a suitable form for the following spatial analysis and fast visualizations. A detailed description of the ways of publishing spatial-temporal database content in the Geoserver is the partial goal of the article that we have focused on as well.

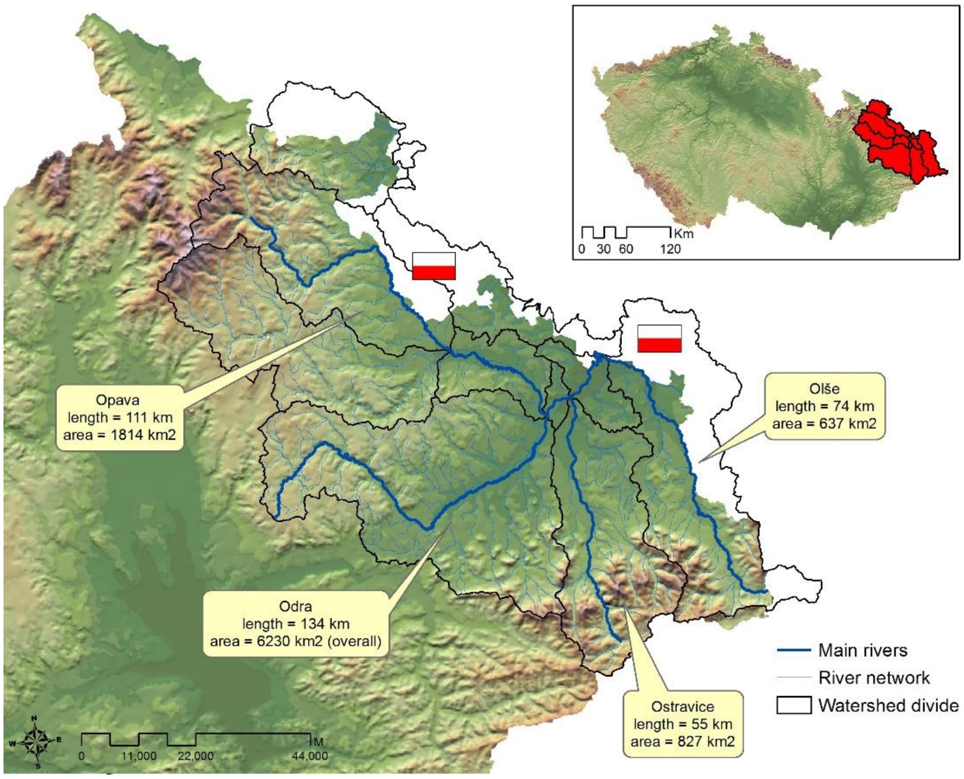

Fig. 1. Area of interest

DATA AND STUDY AREA

This section provides a brief description of hydrological models used in the Floreon$^+$ system and the structured data layer serving as a source for particular simulations.

Characteristics of Moravian-Silesian region

The Moravian-Silesian region (see Figure 1) is situated in the northeastern part of the Czech Republic with the Czech third largest City of Ostrava and the area of 5,427 km^2. Approximately half of the area consists of farmland and another 35 % is covered with forests. With its more than 1.25 million people and the average density of 240 people/km^2, the region is also home to a large number of companies.

From the hydrological point of view, the most of its area falls within the Oder River catchment. The area of the Czech part of the catchment is 6,252 km^2 and is further divided into subbasins belonging to the Oder River itself (length of 134.32 km) and its largest tributaries – the Opava River (110.66 km) on the left side and Ostravice (54.79 km) and Olza rivers (73.10 km) on the right side.

The largest floods occurred on the 5th - 16th July, 1997, when more than half of average annual precipitation was in the form of rain. 536 cities and villages in the Czech Republic were affected, 50 people died and 80 thousand people were evacuated. The material damage amounted to 63 billion Czech crowns.

Floreon$^+$ system hydrologic models

There are two main categories of numerical hydrologic models currently used in the Floreon$^+$ system. The first category consists of rainfall-runoff (R-R) models which allow simulating transformation of predicted precipitations from given areas into discharges in corresponding river outlets. The outputs provided by R-R models form inputs for the second category of models – hydrodynamic (HD) models. HD models enable hydrodynamic simulations of river surface and thus visualization of inundated areas in case of flood events.

HEC-HMS (Hydrologic Engineering Center – Hydrologic Modeling System)

This R-R model was originally developed by the U.S. Army in the 60s. At the beginning of 90s an independent division took over the development (HEC-HMS, 2010).

The modelled catchment is represented by basin model consisting of particular elements – subbasins, channels, junctions, and sinks. Each element typically has several parameters for detailed specification of its properties and behaviour.

Several methods for each of the following categories are available: infiltration loss simulation, transformation of excess precipitation into surface runoff, representation of baseflow contributions to subbasin outflow, simulation of flow in open channels and water impoundments representation.

The schematizations and parameters are stored in text files, meteorological data and resulting hydrographs are stored in a proprietary binary format. For further Floreon$^+$ system processing, hydrographs are exported to a CSV file using the HEC-DSSVue utility.

Math1D

Another R-R model integrated into the Floreon$^+$ system is our in-house model internally called Math1D (Kubíček and Kozubek, 2008).

It is a numerical semi-distributed R-R model that implements the unit hydrograph and SCS curve number methods and accumulates water contribution of interflow modelled by convolution integral and simplified differential equation and contribution of surface runoff using non-stationary linear and non-linear isochrones. This model uses modified schematizations exported from HEC-HMS model into XML files. The simulated hydrographs are directly stored as output CSV files.

HEC-RAS 4.1 (Hydrologic Engineering Center – River Analysis System)

HEC-RAS 4.1 is an HD model for 1D steady and unsteady flow simulations representing water movements (HEC-RAS 2010). It is based on 1D Saint-Venant equations derived from the Navier-Stokes equations. Along with the precise topographic and hydraulic data, discharge volumes for river profiles form inputs to the model. The model offers an evolution of discharge, water surface elevation, current velocity, and other statistical data.

FLOREON$^+$ SYSTEM DATA LAYER

Digital elevation models and results from HD model are stored in PostgreSQL geographical database running on the Floreon$^+$ geo-server machine. Besides this, most of the parameters of the models, their inputs and outputs, and the system management data are stored in the complex relational database structure based on Microsoft SQL.

The stored data can be divided into these two logical groups - the static and dynamic ones. The former includes all the topographic data and catchments characteristics, the latter consists of the data for monitoring, and prediction, which are continuously changing.

Topographic data

In the Floreon$^+$ system, two types of digital elevation models differing by the data source technology are used.

The first type of DEM is created using 2007 LIDAR data from a filtered point field (cloud points) of high density (19 points per m^2). The scanning was performed from the elevation of 300 m above ground level with the resolution of 16 Mpx (4080 x 4076 px). The point field gave rise to two DEM structures, namely a raster and its triangulated irregular network (TIN) equivalent. The resolution of the raster and TIN is 0.2 m, and therefore it can be considered as very accurate.

The second type of DEM was created using a photogrammetry method in 2010. The resolution of the obtained raster is 10 m which is sufficient to capture the detail of a surface.

Catchment parameters

Each catchment is characterised by many parameters describing particular components of the schematization. These parameters might be classified as geometric and hydrologic.

From the entered measured and interpolated cross sections, the width, depth, bank slope, channel slope, and other geometric parameters are derived.

The group of the hydrologic parameters describing attributes of flow paths includes, for example Manning’s friction coefficients, base flow, curve number, land use, landcover, slope, and others.

Data for monitoring

The data for simulation and monitoring purposes are currently provided by the Povodi Odry (the Oder River basin board) state enterprise in the 10-minute interval. Until the May 2014, the data were obtained from the Czech Hydrometeorological Institute (CHMI). The data package contains measured precipitation rates, measured water levels, measured discharge volumes, and temperatures obtained from particular hydrologic gauges.

Data for prediction

Weather forecast data for the Floreon$^+$ system is currently obtained from the Medard project service (ICS 2016), which is based on the Weather Research & Forecasting model. The Floreon$^+$ system is designed to process data from multiple sources. Earlier the forecast data was also obtained from the Aladin project service (Degrauwe et al. 2016). The forecasts are composed of the predicted precipitation rates, temperatures, wind velocities, and directions. Every 6 hours, the new data package containing an updated forecast for the next 72 hours is available. In the first phase, the new package is downloaded by the Floreon$^+$ system using the weather forecast downloader. This routine then launches the Floreon$^+$ web service designed for processing the downloaded data. The process is finished by storing the new values to the Floreon⁺ database structure. Since our simulations provide predictions for 48 hours, we do not use the second 24-hour interval of the weather forecast, which is the most inaccurate by nature.

SIMULATIONS AND DYNAMIC DATA PROCESSING

The Floreon⁺ system can be roughly divided into two main parts in terms of dynamic data processing (see Figure 2). The first system part deals with automatic simulation execution. Its main task is to execute rainfall-runoff models and a hydrodynamic model for the current time and current data on hourly basis. The second part consists of user initiated simulations or computations also called on-demand simulations. These on-demand simulations are represented by modules designed specifically for HPC cluster execution (Buyya 1999). Due to the user-defined parameters, the user initiated simulations can be more computationally demanding than automated simulations, therefore they are executed directly on the HPC cluster. Thanks to the modularity of the system and properly designed architecture, which will be discussed in the following chapter, the system allows for easy integration of additional HPC modules. The Floreon⁺ system already contains several modules, however, because the focus of this paper deals primarily with hydrology, only the module for user-defined hydrological What-if-Analysis (WIA) will be mentioned.

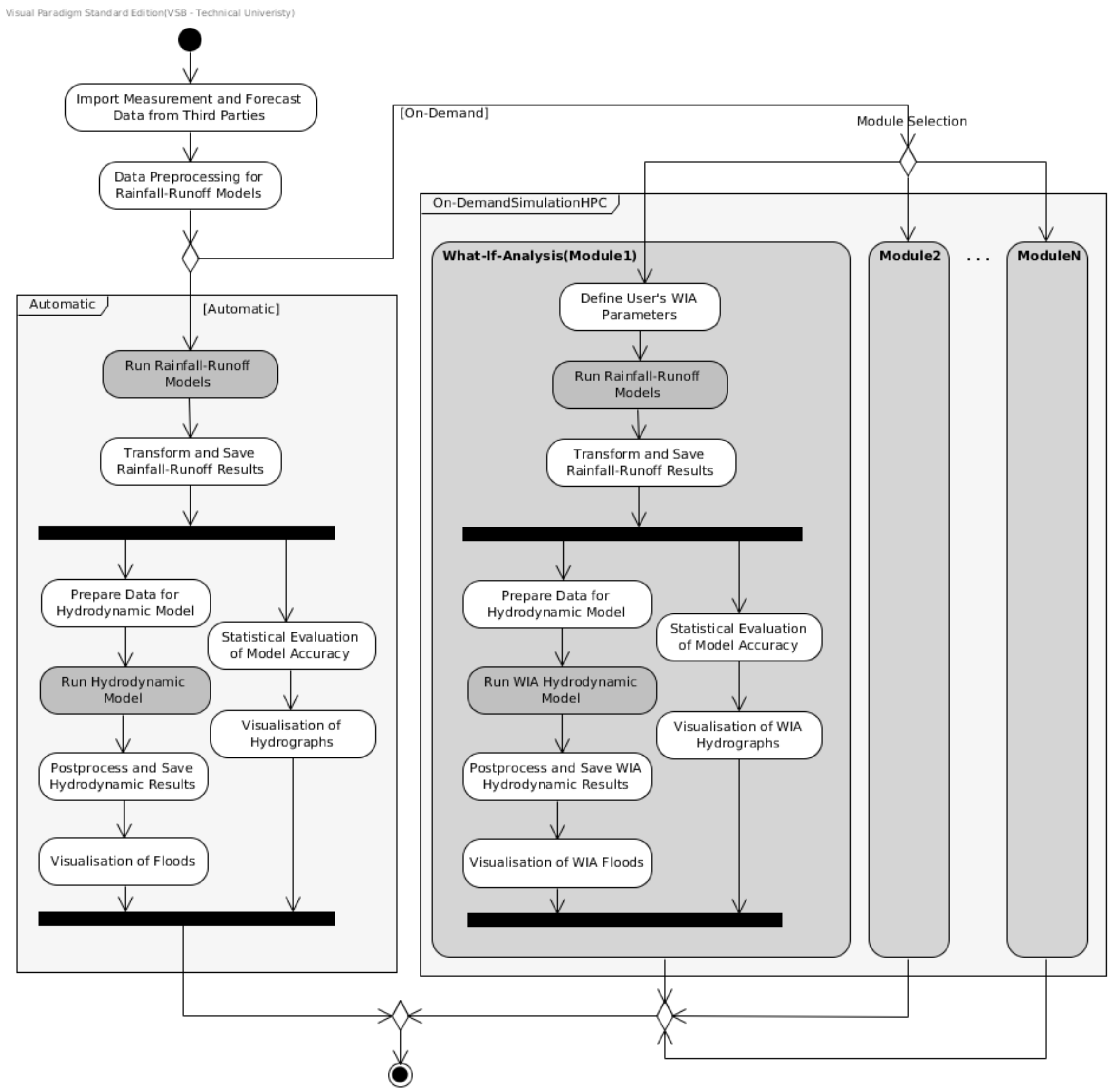

Fig. 2. Dynamic data processing

Automatic Simulations

Every hour, an automatic process, which includes running of rainfall-runoff simulations with hydrodynamics, is triggered on the HPC infrastructure. In this process, the data from the relational database is loaded and preprocessed according to the selected model and subsequently the simulation is run. When the calculation execution is finished, the obtained results are processed and stored in the relational database. In the next step, the process is divided into two parallel runs. Within one run, on the occasion of the user needing to visualize results of rainfall-runoff simulation, statistics evaluations of results accuracy are primarily calculated and visualized in the form of hydrographs. At the same time, hydrodynamic simulation, which pre-processes values obtained for the hydrodynamic model designed for calculation, runs and run it. Model outputs are processed and stored in the spatial-temporal database. Process finishing the possibility of visualizing the results of hydrodynamic simulation and displays it in the form of flood area.

On-Demand Simulations

Under the on-demand hydrologic simulations, the framework for running on-demand What-If Analysis (WIA) is created to simulate crisis situations. This also includes What-if hydrologic simulations. Through the web interface, users can create their own hydrologic What-if simulation running on this framework. They must choose the basic settings from the menu having the option to select river basin, schematization, and rainfall-runoff model for which the rainfall-runoff simulation and hydrodynamic model will be calculated. The framework allows the user to specify precipitation in selected precipitation stations and selected time. This type of simulation also allows the user to edit the default parameters of subbasins and channels. The next step is to run the simulation execution. Execution of What-if simulation is processed on an HPC cluster. As the result from the rainfall-runoff model is used as an input for the hydrodynamic model, the system allows the users to view the hydrographs as soon as they are available, independently of the hydrodynamic computation. As soon as the hydrodynamic model computation is completed and the result values are stored within the spatial-temporal database, the users can view the simulated flood layer within the map interface together with the hydrographs.

Postprocessing

The goal of postprocessing is to transform the results of simulation of hydrodynamic modelling into a form suitable for visualization provided by Geoserver. The result of the hydrodynamic model consists of data in SDF format and metadata. These data are considered as an input for postprocessing. The polygons for floods are calculated upon this data and stored within the PostgreSQL/PostGIS database. An important part of the postprocessing is the archiving functionality for the hydrodynamic results which vary according to the type of simulation execution.

In the case of automatic simulations, the flood prediction is calculated for the next 48 hours with the last 5 hours used to calibrate the model. Thus in the next hour, when another simulation with 48-hour prediction is executed, we will obtain new result values for the same 47 hours (for the same time) as in the last simulation run. To archive the prediction data more efficiently considering the available storage space, only the latest result values for a specific time are archived. However, this does not apply in the case of some significant event. When the flood stage level on some measuring station exceeds the standard value, the system stores the predicted values from every simulation to be used for subsequent analysis or model calibration.

In the case of on-demand simulation, the prediction interval is part of the user's input parameters, therefore the system stores the predicted values for every simulation run.

Visualization

Publication of results of hydrological modelling via web interface (see Figure 3) was guided mainly by an effort to visualize the current state and calculated predictions as fast as possible.

The methods used for the visualization were derived from the need to:

- show the actual state for the current time (measured data and 48-hour prediction),

- visualize the different types of data (hydrographs for the rainfall-runoff models, flood lakes for the hydrodynamic models),
- be able to view the data from the past.

The above mentioned requirements were combined into the time axis widget, a part of Floreon+ GUI (see Figure 3), which is the key element of the application. The users can easily change the current time within the map interface to browse the information for the selected area, no matter if it is (i) prediction of the future, (ii) the current situation based on the real data, or (iii) simulation of a situation in the past.

The time selected on the time axis by the users is passed as an input parameter for the selector of formerly prepared simulation outputs. Geoserver publishes layers corresponding to the selected parameters using the time-WMS service. PostgreSQL/PostGIS is used as a storage for the precomputed layers later to be displayed within the map interface of the Floreon+ system.

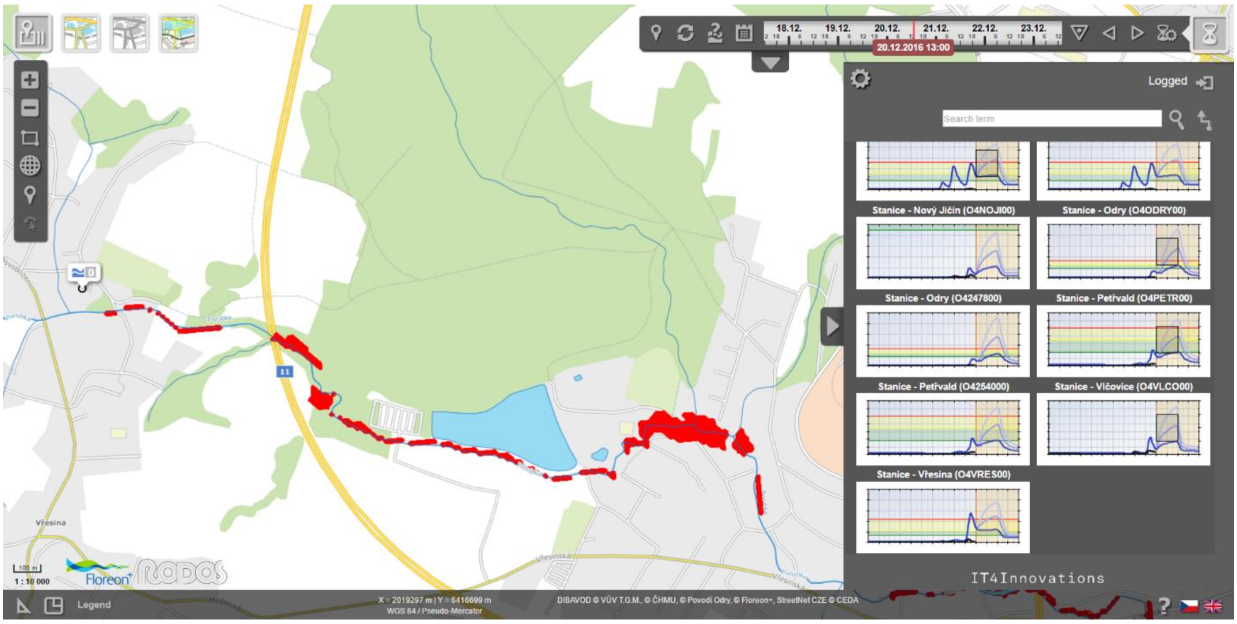

Fig. 3. Floreon+ system web interface

ARCHITECTURE OF THE FLOREON+ SYSTEM

The Floreon+ system architecture uses the high-performance computing infrastructure at IT4Innovations National Supercomputing Center in Ostrava, Czech Republic. High-performance computing (HPC), in other words, means the use of parallel processing for running advanced application programs efficiently, reliably, and quickly. The most common users of HPC systems (Younge, Andrew, 2011) are scientific researchers, engineers, and academic institutions. IT4Innovations operates 2 supercomputers, namely Salomon (24192 cores, 129 TB RAM, 2 Pflop/s) and Anselm (3344 cores, 15 TB RAM, 94 Tflop/s). The supercomputers are available to the academic community within the Czech Republic and Europe and industrial community worldwide. These two HPC clusters are currently available to use within the Floreon+ system to run on-demand simulations. The architecture of the Floreon+ system can be seen in Figure 4.

There was need for an architecture, which would allow users to run complex and computationally demanding calculations on a supercomputer directly from the user interface of a client application without the necessity to connect directly to the HPC cluster and manage the jobs from the command line interface of the HPC scheduler. Therefore, we have developed our in-house application framework called HPC as a Service Middleware. This mid-layer, in software terminology also known as middleware, manages and provides information about submitted and running jobs and their data between a client application and the HPC infrastructure. This middleware is able to submit the required computation or simulation on the HPC infrastructure, monitor the progress, and notify the user in case of need.

Middleware features

- Providing HPC capabilities as a service to client applications and their users
- Unified interface for different operating systems and schedulers
- Authentication and authorization to provided functions
- Monitoring and reporting of executed jobs and their progress
- Current information about the state of the cluster

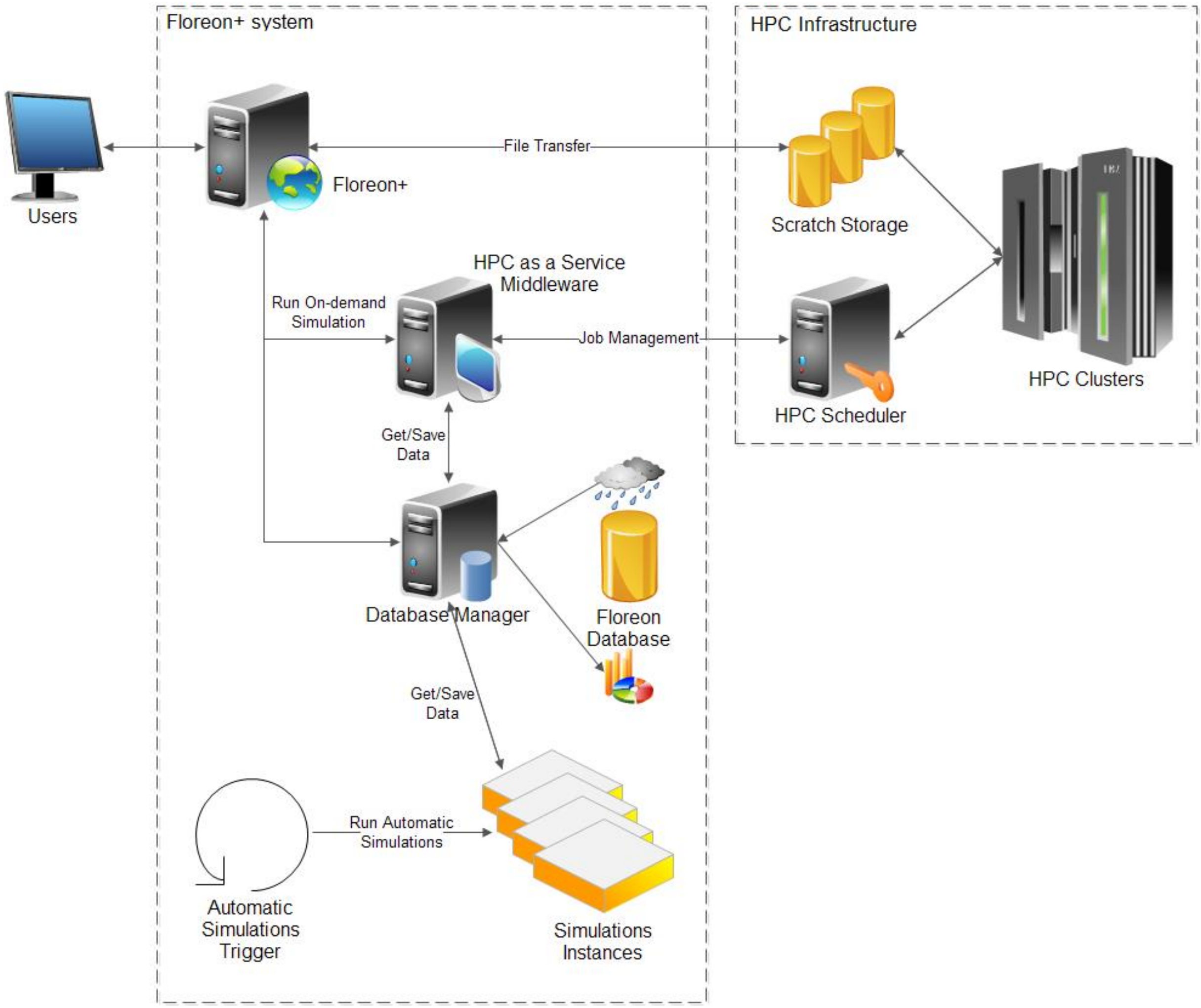

Fig. 4. Floreon+ system architecture

Before running the job (What-If simulation), the users usually need to upload their input data to the application disk storage. After the job setting configuration has been chosen, specified by the cluster configuration and required computational resources, the newly created job is sent to the cluster for processing. Before the job's actual start, the input data are copied from the application storage to the HPC cluster disk storage (Scratch storage). When the copied input data are completely available on the cluster storage, the job is submitted to the HPC scheduler queue to be executed. The middleware also calls the HPC scheduler at regular intervals to acquire the status of all submitted jobs and feeds this information to the client application. The above mentioned job execution workflow is part of the internal logic and therefore fully automated. From the user's perspective, the required steps consist of input parameter selection for the What-If simulation and simulation start only.

CONCLUSION

In this article, we have presented Floreon+ system, an online flood monitoring and prediction system. The main aim of the Floreon+ project is to provide detailed information about the possibility of incoming flood and its extent. In the future, Floreon+ system should estimate the possible property damage before it happens, provide recommendations about flood measures, and online decision support for emergency committees to minimize the number of casualties and property damage in case of extreme weather events.

As the Floreon+ system is very extensive, this article focuses mainly on the area of hydrology and describes the processes necessary for dynamic data processing, all the way from the initial raw data import to the visualization of floods via the map interface. The system utilizes the HPC infrastructure at IT4Innovations and enables the users to execute user-defined hydrologic What-If analyses directly on the HPC cluster. Due to the system's universally designed architecture, both in terms of infrastructure and different methods of visualization, the Floreon+ system enables an easy integration of additional modules that do not necessarily need to address the area of hydrologic modelling.

FUTURE WORK

Given that the Floreon+ system already consolidates the data from several areas of interest (hydrology, traffic modelling, mobility) the future work is mainly focused on integration of new and more accurate models and on the interconnection of these seemingly different areas of interest. We have already experimented with several use cases that are currently in the process of integration into the Floreon+ system. For example, one of these use cases is focused on traffic modelling during floods and finding alternate routes due to the traffic bottlenecks or flooded roads. Another use case is, for example, focused on analysing the arrival times of firefighting vehicles during floods, thus taking the road closures into account.

Acknowledgement

This work was supported by The Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPU II) project „IT4Innovations excellence in science - LQ1602“ and from the Large Infrastructures for Research, Experimental Development and Innovations project „IT4Innovations National Supercomputing Center – LM2015070“ and from the SGS 2016 project „Flood and pollution modelling II - SP2016/150".

REFERENCES

Garrote, L., Bras, R.L. (1995) A distributed model for real-time flood forecasting using digital elevation models. Journal of Hydrology 167 (1-4), 279–306.

Halmo, N. (2006) Flood protection program of Slovak republik In: International conference of flood protection, 4 – 7 December, High Tatras: Slovakia.

Knebl, M.R., Yang, Z.-L., Hutchinson, K., Maidment, D.R. (2005) Regional scale flood modeling using NEXRAD rainfall, GIS, and HEC-HMS/RAS: a case study for the San Antonio River Basin Summer 2002 storm event. Journal of Environmental Management 75, 325-336.

Podhoranyi, M., Kuchar, S., Portero, A.(2016) Flood evolution assessment and monitoring using hydrological modelling techniques: Analysis of the inundation areas at a regional scale.(2016) IOP Conference Series: Earth and Environmental Science, 39 (1)

Kuchar, S., Podhoranyi, M., Vavrik, R., Portero, A.(2016) Dynamic computing resource allocation in online flood monitoring and prediction.(2016) IOP Conference Series: Earth and Environmental Science, 39 (1)

Hut, R., Drost, N., van Meersbergen, M., Sutanudjaja, E., Bierkens, M., and van de Giesen, N.(2016) eWaterCycle: a hyper-resolution global hydrological model for river discharge forecasts made from open source pre-existing components, Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-225, in review, 2016.

Buyya, R.(1999) High performance cluster computing: Architectures and systems (volume 1). Prentice Hall, PTR Upper Saddle River, NJ, USA, 0130137847

Degrauwe D., Y. Seity Y., F. Bouyssel F., and P. Termonia. (2016) Generalization and application of the flux-conservative thermodynamic equations in the AROME model of the ALADIN system. Geosci. Model Dev. Discuss.

Younge, A. J., et al.(2011) Analysis of virtualization technologies for high performance computing environments. Cloud Computing (CLOUD), 2011 IEEE International Conference on. IEEE, 2011.

CPD-74A (2010) Hydrologic Modeling System HEC-HMS User's Manual. U.S. Army Corps of Engineers Hydrologic Engineering Center, Davis, CA.

Kubíček, P. and Kozubek, T. (2008) Mathematic-analytical Solutions of the Flood Wave and its Use in Practice (in Czech). VŠB-TU Ostrava, Ostrava, 150 p.

CPD-68 (2010) HEC-RAS River Analysis System User's Manual Version 4.1. US Army Corps of Engineers Institute for Water Resources Hydrologic Engineering Center, Davis, CA.

Institute of Computer Science (ICS), The Czech Academy of Sciences. http://www.medard-online.cz, December 2016.