The VESTEC project has received funding from the European Union’s Horizon 2020 Programme for research, technological development and demonstration under grant agreement n° 800904
The VESTEC Consortium

- DLR German Aerospace Center (Coordinator)
- The University of Edinburgh
- Sorbonne Université
- KTH Stockholm
- Kitware SAS
- Intel Deutschland GmbH
- Fondazione Bruno Kessler
- Université Paul Sabatier Toulouse
- Tecnosylva SL
Visual Exploration and Sampling Toolkit for Extreme Computing (VESTEC)

The Vision

- Support urgent decision-making (UDM) with a generic software architecture incorporating:
  - Real-time data
  - (interactive) visualization
  - (ensemble) simulations
  - (topological) data analysis

**Urgent Decision-Making**

- Disaster Relief Forces
- Crisis Management Center
- High Performance Computing Center

- Sensor Networks
- Social Networks
- Web Search Statistics

**Data Analysis**

- Ensemble simulations
- Topological data analysis

**Visualization**

- Crisis Manager
- High Performance Computing Center
- HPC Cluster

**Simulations**

- Ensemble simulations support urgent decision-making (UDM) with a generic software architecture incorporating real-time data, (interactive) visualization, (ensemble) simulations, and (topological) data analysis.
Three Pilot Applications for Urgent Decision-Making

Wildfire Forecasting

Spread of Mosquito-borne Diseases

Space Weather Forecasting
VESTEC – A Combination of HPC Computing Paradigms

Urgent Computing (UC)

Interactive Computing (IC)

We inherit both the opportunities and the challenges of UC and IC
A Challenge for Both Urgent and Interactive Computing: Batch Schedulers

• Effectively all HPC systems schedule jobs by using a batch queue (e.g., SLURM, PBS, etc.)

→ Jobs might wait in the queue for an **unbounded amount of time**

• Obstacle for Urgent Computing:
  
  *results may not arrive in time* (e.g., before a disastrous event)

• Obstacle for Interactive Computing:
  
  *users do not want to wait for their interaction* (at least, it is impractical)
Three Possible Solutions

A dedicated machine for UDM

- Machine would be idling most of the time
- A large HPC system is expensive to buy and maintain
- One system might not be enough in case of a disaster

Special job priorities for UDM jobs

- Other jobs need to be cancelled
- HPC resource owners are bound by SLAs
- Other users would need to be rewarded if their jobs are cancelled

Federating UDM jobs over multiple machines

- Jobs still wait in queues
- Probability of running in time is much higher
- No policy changes needed → works with current systems

We don't expect that policies change in the short and medium term → We decided for the federator strategy
The VESTEC Approach: Federate over Multiple Machines
Additional Challenges

• Interactivity may require access to ports on the compute nodes, which is often not allowed for security reasons

• We must not DDoS an HPC machine!

• Simulation codes produce a high amount of data → we use in-situ topological data analysis for feature extraction

• The decision maker is not an HPC or even computing expert → the VESTEC system should work automatically in the background

• Different applications need different workflows, different codes have different interfaces → the VESTEC system should be designed as general as possible
VESTEC
Architecture
Topological Data Analysis in VESTEC
Identification of Different Scenarios from Ensemble Simulations

$N$ ensemble simulations $\rightarrow$ simulation outputs $\rightarrow$ clustering $\rightarrow$ representing different outcome scenarios $\rightarrow$ decision maker

cluster 1

cluster 2

\vdots

cluster $K$
Topological Data Analysis in VESTEC
Identification of Different Scenarios from Ensemble Simulations

$N$ ensemble simulations $\rightarrow$ simulation outputs $\rightarrow$ clustering $\rightarrow$ \{cluster 1, cluster 2, ..., cluster $K$\}

very large data sets
Topological Data Analysis in VESTEC
Identification of Different Scenarios from Ensemble Simulations

- $N$ ensemble simulations
- simulation outputs
- clustering
- topological feature extraction
- data reduction, e.g., Persistence Diagrams
Topological Data Analysis in VESTEC
Identification of Different Scenarios from Ensemble Simulations

Conclusions & Outlook

• The VESTEC system will support urgent decision making by using HPC infrastructure, interactive visualization, and real-time sensor data

• We face a lot of challenges – not only technical, but also regarding policies and security:
  • batch queues
  • closed ports
  • a variety of applications
  • large amounts of data

• In the second half of the project, we will continue implementing the developed architecture

The VESTEC vision
All tier 0 and tier 1 HPC machines in Europe subscribe to a disaster response scheme managed by the VESTEC system
Any Questions?

Contact:
Dr. Max Kontak
Institute of Software Technology
DLR German Aerospace Center
Cologne, Germany
max.kontak@dlr.de

http://www.vestec-project.eu/

Twitter: @VESTECproject

The VESTEC project has received funding from the European Union’s Horizon 2020 Programme for research, technological development and demonstration under grant agreement n° 800904