HPC as a Service: Lessons Learned

News

Press Releases Event Announcements CCT Weekly Grants and Funding Student News Archived News (Source: HPCwire)

After a fast-paced three months, round 1 of the HPC Experiment (also known as the Uber-Cloud Experiment) concluded last month, with more than 160 participating organizations and individuals from 25 countries, working together in 25 international teams. In this article we present their main findings, challenges, and their lessons learned.

The aim of the Uber-Cloud Experiment is to explore the end-to-end process of accessing remote computing resources in HPC centers and in HPC clouds as well as to study and overcome the potential roadblocks.

The experiment kicked off in July 2012 and brought together four categories of participants: the industry end-users with their applications, the software providers, the computing and storage resource providers, and the experts. We set up an end-user project by first selecting an end-user and his software provider, assigning an HPC/CAE expert, and matching a suitable resource provider to complete the team. Each team's goal was to complete the project, and to chart the way around the hurdles they identified.

End users can achieve many benefits by gaining access to additional compute resources beyond their current internal resources, such as workstations. Arguably the most important two are the benefits of agility gained by speeding up product design cycles through shorter simulation run times, and those gained by the superior quality achieved by simulating more sophisticated geometries or physics, or by running many more iterations to look for the best product design.

Tangible benefits like these make HPC and more specifically HPC-as-a-Service (HPCaaS) very attractive. But how far are we from an ideal HPCaaS or HPC in the cloud model?

Honestly, at this point, we don't know. However, in the course of this experiment, following each team and monitoring its challenges and progress, we've collected some excellent insight into these roadblocks and how our 25 teams have tackled them.

The main approach for this experiment is to look at the end-users' project and select the appropriate resources, software and expertise that match those requirements.

During the three months of the experiment, we were able to build 25 teams each with a project proposed by an end user. These teams were: Team Anchor Bolt, Team Resonance, Team Radiofrequency, Team Supersonic, Team Liquid-Gas, Team Wing-Flow, Team Ship-Hull, Team Cement-Flows, Team Sprinkler, Team Space Capsule, Team Car Acoustics, Team Dosimetry, Team Weathermen, Team Wind Turbine, Team Combustion, Team Blood Flow, Team Turbo-Machinery, Team Gas Bubbles, Team Side impact, Team ColombiaBio, and Team Cellphone.

The final report, available to all of our registered participants, contains the use cases of many of the teams offering valuable insight through their own words. We look forward to future rounds of the experiment where this accumulating knowledge will yield ever more successful projects.

We recognize that every end-user project requires a slightly different approach, a variety of software and compute resources, a certain expertise to lead the end-to-end process, and a tailored schedule. To be able to keep the entire experiment consistent we asked each team to follow a common roadmap. The expert assigned to each team is the guide in following this roadmap. It calls for communication with the organizers at certain points, although generally the teams are autonomous and make their own decisions.

Based on the roadmap we defined going into round 1 of the experiment, the teams followed six steps to reach their goal:

- **Step 1. Define the end-user project.** The end-user together with the expert and software provider jointly defined the project. Based on this information, as organizers we assigned the appropriate resources to the project.
- Step 2. Contact the resource provider and set up the project environment. The expert contacted the computing resource and performed activities such as assisting in software and license installations, creation of user accounts, and configuration of the project environment.
- Step 3. Initiate the end-user project execution. The expert assisted the end-user with uploading the necessary data, code and configuration files to the remote resource(s). The expert then worked with the resource provider to queue the project up on the HPC system.
- **Step 4. Monitor the project.** The expert remained engaged with the resource providers and at any time had up to date information about the status of the project.
- Step 5. Review results with the end-user. The expert assisted the end-user in downloading the results from the resource provider's environment and discussed the results with the end-user. If any rework or rerun was required it was completed by executing steps 2-5 again.
- **Step 6. Document findings.** During the entire lifecycle of the project, there occurred hurdles, friction and failure points and the expert documented these findings.

Intentionally, we performed the first round of this experiment manually, that is, not via an automated service, because we believe the technology is not the challenge anymore; rather it's the people and their processes, and that's what we wanted to explore. We are continuously improving the roadmap to successful completion of our projects.

The teams reported the following main roadblocks and provided information on how they resolved them (or not):

- Security and privacy, guarding the raw data, processing models and the results
- Unpredictable costs can be a major problem in securing a budget for a given project
- Lack of easy, intuitive self-service registration and administration
- Incompatible software licensing models hinder adoption of Computing-as-a-Service
- High expectations can lead to disappointing results
- Eack of reliability and availability of resources can lead to long delays

Just like all other participants, we as the organizers, treated the experiment as a learning opportunity. In our report we have also summarized what we've found to be shortcomings of the experiment as we put it together in round 1. Learning from these shortcomings we have improved the experiment for round 2. To be specific, we discussed and provided solutions for the following shortcomings:

All participants are professionals with busy schedules and the experiment is not their primary job, so they could only dedicate a few hours per week to the experiment

- >> Vacations delayed most of the teams' progress, especially in the beginning (August) of the Experiment
- Some resource providers ran into resource crunches which delayed team projects
- Some of our projects ran into long delays since the project and the resource provider weren't the best match possible
- Some resource providers struggled with the installation of an application
- Other resource providers had difficulties with providing network access through complex network connections
- Resource providers differ in their service philosophies
- Simply getting started was a challenge
- ▶ A few teams struggled with figuring out which team member needs to do what and when
- Team forming was one of the steps, which took the longest amount of time, each team member needed to exchange significant amounts of information about their background, capabilities, expectations, availability, and commitment levels with one another before the project could even kick off
- Finally, manual processes are just slow; they consumed days, sometimes weeks especially because the various technology and people resources were inherently remote, each with different priorities

We hope that our participants will extract value out of the experiment and the final report. They certainly deserve to do so in return for their generous contributions, support and participation. We now look forward to round 2 of the experiment with its already over 250 participants and the learning that it will result in.

If you are interested in participating in round 2 or just want to monitor its progress, you can register at http://hpcexperiment.com. You can also go there to get the final report for round 1, which details the results and recommendations.

About the Authors

Wolfgang Gentzsch and Burak Yenier are the creators and facilitators of the Uber-Cloud Experiment. Wolfgang is an HPC veteran. Having worked in leading positions in research, academia and industry for some 30 years, Wolfgang is now an HPC consultant and the chairman of the ISC Cloud conference series for HPC and Big Data in the Cloud. Burak is the vice president of operations at CashEdge, a software-as-a-service company in Silicon Valley, which provides innovative payments and aggregation solutions to financial institutions.

Publish Date:

12-12-2012

Home | About | Research | Programs | News | Events | Resources | Contact Us | Log In | LSU | Feedback | Accessibility

Center for Computation & Technology 2003 Digital Media Center • Telephone: +1 225/578-5890 • Fax: +1 225/578-8957 © 2001–2025 Center for Computation & Technology • Official Web Page of Louisiana State University.