The Computerworld Honors Program

Honoring those who use Information Technology to benefit society

Final Copy of Case Study

Status:
Laureate

Year:
2013

Organization Name:
Novartis Institutes for BioMedical Research

Organization URL:
http://www.nibr.com/

Project Name:
Molecular Pathways Research Using Utility Supercomputing

Please select the category in which you are submitting your entry
Emerging Technology

Please provide an overview of the nominated project. Describe the problem it was intended to solve, the technology or approach used, how it was innovative and any technical or other challenges that had to be overcome for successful implementation and adoption. (In 300 words or less.)

In September 2011, Novartis Institutes for BioMedial Research used Cycle Computing software and provisioned a 30,000-core utility supercomputer in the Amazon Web Services (AWS) cloud to accelerate molecular modeling with the goal of focusing on molecular pathways shared by various diseases. The goal of this compute was to integrate clinical insights with mechanistic understanding of diseases. Novartis' current techniques and lab-based PC systems could not effectively process the high volume of compound data they were creating, and their internal HPCC was already running at capacity. They needed the ability to screen millions of compounds and get results in less than a week, and they needed the ability to screen the compounds on demand. The solution they found was HPC in the AWS cloud, using Cycle Computing's software to optimize data processing. The global 30,000-core cluster was run with CycleCloud, Cycle's flagship HPC in the Cloud software. Automating the process of provisioning resources and replicating data across two continents, CycleCloud performed hundreds of thousands of computational tests with the run time per job averaging 37 minutes and

the total work completed nearing 100,000 hours. Novartis' researchers completed nearly 11 years of molecular dynamics computing in a few hours, at a cost of $1,279 per hour at peak, and it required no upfront capital.

When was this project implemented or last updated? (Please specify month and year.) Has it incorporated new technologies and/or other innovations since its initial deployment? (In 300 words or less.)

Novartis ran their 30,000 cluster run in September of 2011, and has been refining the process during 2012. As discussed at AWS re:Invent in Nov of 2012, there have been additional compounds and tweaks to the runs themselves to improve the science.

Is implementation of the project complete? If no, please describe the project's phases and which phase the project is now in. (In 300 words or less.)

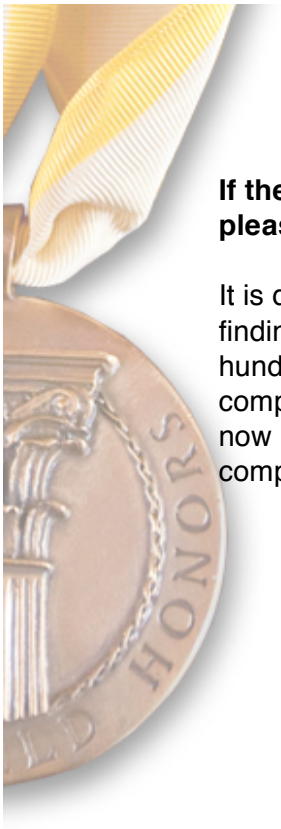
Yes, it is complete and in ongoing use during 2012.

Please provide at least one example of how the technology project has benefited a specific individual or organization. Feel free to include personal quotes from individuals who have directly benefited from the work. (In 300 words or less.)

Novartis was researching the molecular pathways shared by diseases in the hopes of discovering drugs to cure them. The primary benefits to Novartis were the ability to cost-effectively even run this scale of a problem, which was a 150,000% cost improvement over building a 30,000 core environment in-house. Utility supercomputing capacity enabled by Cycle Computing's software and Amazon Web Services (AWS) cloud infrastructure enabled impossible science: a complete scan of the candidate molecular space, allowing Novartis researchers to push the boundaries of drug design research.

Would this project be considered an innovation, a best practice or other notable advancement that could be adopted by or tailored for other organizations and uses? If yes, please describe that here. (In 300 words or less.)

Yes, this project is an innovative best practice, and a model for how pharmaceutical companies can compute massive amounts of information quickly and cost effectively. At the time of this run, the 30,000 core environment was the largest cloud-based utility supercomputer to have been created. This record breaking utility supercomputer yielded tremendous, clear benefits in better science, with better compounds. It was publicly stated in November that this run directly led to 3 drug leads that are in the wet lab undergoing testing, and wouldn't have been identified otherwise. This serves as a best practice model for how rational drug design should be done moving forward.



If there are any other details that the judges should know about this project, please note them here. (In 300 words or less.)

It is difficult to communicate to non-pharmaceutical industry personnel the importance of finding 3 additional drug candidates as a result of this run. The goal of the many hundreds of millions spent by pharmaceutical companies on drug targets is to find viable compounds that might provide therapies. This is tremendous result, and the fact that it is now possible for under $10,000 to run this calculation will radically improve how computer-aided drug design is done.