

BDEC2 Platform Demonstrator Proposal

1 Overview

The BDEC2 effort has identified the need for a demonstrator of a potential platform for programming the computing continuum that is the future of big data and extreme scale computing. A recent paper by BDEC2 organizers and participants¹ describes the notion of the continuum, enumerating classes of computing elements from *nano* (count = 10^9 , size = 10^1) to *facility* (count = 10^1 , size = 10^9).

This proposed contribution to the platform demonstrator is based on the InLocus architecture², which operates on streams of data. Due to its limited execution model and semantics, it is suitable for processing streams of e.g. sensor data on a message by message basis. This primitive execution allows InLocus to execute on devices at the *nano* or *micro* end of the computing continuum.

By processing data a message, or tuple, at a time, it resembles cloud data processing systems like Apache Storm or Twitter Heron, both of which assemble topologies of stream processing functions (“bolts”) and route messages via these processing pipelines. These processing elements can be considered analogous to “microservices” in a serverless computing environment. This type of decomposed application design pattern underlies web-scale applications running at the Cloud end of the computing continuum.

Finally, there is a growing movement in high-performance computing to move to Asynchronous Many-Task (AMT) runtimes to expose parallelism. Applications can be defined in terms of a “dataflow” model, where small

¹P. Beckman, J. Dongarra, N. Ferrier, G. Fox, T. Moore, D. Reed, and M. Beck. Harnessing the Computing Continuum for Programming Our World, https://www.researchgate.net/publication/332246123_Harnessing_the_Computing_Continuum_for_Programming_Our_World

²L. Brasilino, A. Shroyer, N. Marri, S. Agrawal, C. Pilachowski, E. Kissel, and M. Swamy. Data Distillation at the Network’s Edge: Exposing Programmable Logic with InLocus. In IEEE International Conference on Edge Computing, July 2018. <https://doi.org/10.1109/EDGE.2018.00011>

“patches” of computation are performed when the requisite data is available. This needs to be done carefully as over-decomposed structures may not be efficient, but the model is the same as that above – small computational operators await an input of some sort and then execute over it.

2 Proposed Platform Demonstrator

Each of the three above environments achieves scalability by decomposing applications into ensembles of small operations, working together. Given that this data-driven, message-oriented model can be productively implemented across the computing spectrum, we assert that it can enable a *transcontinuum workflow* with explicit *edge computing* support.

A straightforward application would be processing sensor data. This would be feasible to implement using combinations of various BDEC2 stakeholder’s projects. An environmental sensor data application running across various scales of the computing continuum is a powerful vision for the convergence we seek. Further, early demonstrations of this are possible by combining existing components.