HPBC 2015 Keynote Speaker

Big Data: What happens when data actually gets big?
*Timothy G. Mattson, Intel Science and Technology Center for Big Data, Intel, USA*

**Abstract:**

There are big data problems today. Usually when researchers talk about “big data”, however, the data isn’t that big. An inefficient solution stack based on direct products over data-sets (MapReduce/Hadoop), cramming data (the square peg) into whatever data-store is available (the round hole), and moving data around between data-severs and computer-servers … all of this is OK if “big-data” is not that big. But what happens when the data in big data actually gets big?

In this talk we will discuss how Big Data software solution stacks must evolve to address future big data problems.  Our guiding principles are that in a world where big data is really big:

- One size does not fit all; data must match the data store.
- Data Movement is everything; move queries and processing to the data
- Visualization must be a first class citizen of the Big Data Solution Stack

We are developing a reference implementation of the sort of Big Data solution stack we envision will play a key role in the future. We call this the “BigDAWG” solution stack. The research behind BigDAWG is occurring within the Intel Science and Technology Research Center at MIT with collaborators from Brown University, University of Washington, University of Tennessee, and Portland State University.