**Transcript**

***Steve Begley – High performance solutions to streaming data problems***

We live in a world today where there’s vast amounts of information that’s readily available to people for their day-to-day lives or for business centres, research etc. Typically we use computers to access this information but we find that today hardware is evolving at a faster rate than what traditional methods of accessing this information can allow for. So often the hardware is under-utilised for the available processing power in a computer. So our research is focussing on ways of finding more efficient methods of accessing vast volumes of data with modern hardware.

Our research is focussed on streaming data, that is, accessing data from remote sources where the data is in too vast a volume to consume its entirety on any given computer system. So traditional, state-of-the-art methods to do this would be to take snapshots of data, look at smaller subsets, to look at specific time windows of data to analyse and process this data. What we hope to achieve with our research is to find more expeditious manners of processing this data so there is really two key goals for what we wish to achieve: 1. Get more data into the available resources on any given computer system compared to existing state-of-the-art methods and also to process that data at a higher rate. So we need to look at some unorthodox methods of doing so.

Data is everywhere in our day-to-day lives but looking at streaming data in particular I can give some examples where it is of use right now and how processing it at a faster rate would be beneficial. For example, people use eBay, so we have on one side our auctions and on the other side people bidding on auctions. We need to correlate those bids as fast as possible against any given auction to resolve it. So, that is one common example. But we can look at things where the data is more detached. For example, we may have weather systems around the globe, information on that, and aircraft movements around the world. We need to look at where aircraft are entering bad weather. The volume of data for that is too great to grab in its entirety but we can also look at embedded systems such as the engine management computer in a car. It collects and logs data from various sensors in the engine. It might be looking for, you know, upcoming problems or something in the engine, but it can’t keep all that data at once so it’s got to look at specific time windows as well. So, the faster we can process such data I think it’s a generic solution to a generic problem. So the better we can do it, it applies to many situations.