The Data Vortex: From Interbellum Polish Mathematics to a Novel Topology for Connecting Cores

Coke S. Reed[1], Reed E. Devany[1], Michael R. Ives[2], Santiago I. Betelu[1,3]

[1]Data Vortex Technologies, [2]Plexus Corp, [3]University of North Texas

Slide 1 (Title Slide):

Good afternoon. Or, if you're here in the States, good morning. My name is Reed Devany and I am one of the authors of "The Data Vortex: From Interbellum Polish Mathematics to a Novel Topology for Connecting Cores." My co-authors are here today and will be available at the end of this presentation to answer any of your questions.

Slide 2:

Today's talk will be divided into five categories: the Polish roots of the Data Vortex, a description of the Data Vortex topology, the nature and performance of legacy Data Vortex validation systems, using the Data Vortex for direct core-to-core communication, and novel application areas that can address today's pressing problems. Data Vortex Technologies holds the unique distinction of being a multi-generation family computing company. My grandfather, Dr. Coke Reed, pictured here of the left, is the inventor of the Data Vortex and Chairman, my mother, Carolyn Devany, is the CEO, and I, Reed Devany, am the Communications Director and an intellectual property author and coordinator. Here we are at SC16 in Salt Lake City, receiving an *HPCwire* award from Tom Tabor for best collaboration between industry and government.

Note: This paper was presented at Supercomputing Frontiers Europe (SCFE) 2020 on March 25, 2020. The event was remotely held as a teleconference due to the 2020 COVID-19 outbreak. Thus, while SCFE was hosted in Warsaw, Poland, the talk itself was delivered in Austin, Texas. The paper is broken up to match the presentation. If you would like a copy of the slides, pleas reach out to reed@datavortex.com.

My grandfather's story and the story of the Data Vortex have deep roots in the Polish mathematical tradition, so I am honored to present this paper to our hosts at the University of Warsaw.

Slide 3:

During the Interbellum period between the World Wars, Lwow, Poland (present-day Lviv, Ukraine) was a major seat of topological mathematics. The University of Lwow's mathematics faculty, including John von Neumann, Stefan Banach, Stanislaw Ulam, and Leopold Infeld, met at the Scottish Café to discuss their theorems and problems. Some mathematicians speculate it was the most productive café in the history of Europe. Problems that an individual could not solve were written down in the "Scottish Book" and left as a challenge for their colleagues, oftentimes with prizes offered.

Many of these mathematicians left Poland for the United States prior to the Second World War and joined the Manhattan project, primarily in Los Alamos, New Mexico. The Scottish Book was left behind and buried under Lwow University's football pitch for safekeeping. Following the successful development of the atomic bomb and the end of the War, many of the Poles accepted positions with American universities, bringing with them their unique traditions and research areas, including topological dynamics. Banach's wife, who had remained in Poland with her husband, retrieved the Scottish Book and sent it to Los Alamos.

Slide 4:

"Problem 110" of the Scottish Book is the forefather of the Data Vortex. On October 1st, 1935, Ulam posited a problem regarding the flow of particles in a dynamical system:

LET *M* BE A GIVEN MANIFOLD. Does there exist a numerical constant *K* such that every continuous mapping *f* of the manifold *M* into part of itself which satisfies the shown condition: | $f^{n}x - x$ | < *K* for *n* = 1,2, . . . [f^{n} denotes the n^{th} iteration of the image *f(x)*] posses a fixed point: $f(x_0) = x_0$?

This problem, which von Neuman attempted in 1937, remained unsolved well into the 1970s. Coke Reed and Polish Mathematician Krystyna Kuperberg (both of Auburn University) discovered the solution and provided a counter-example, in which the manifold is Euclidian Three-Space. The associated prize? A bottle of wine. Upon sending Ulam the solution, Reed and Kuperberg received a three-word telegram, "Red or white".

Slide 5:

The solution, shown here, was published in Fundementae Mathematica as, "A rest point free dynamical system on R^3 with uniformly bounded trajectories" in 1981. As you can see in Figure One, the form of a vortex has begun to take shape.

Slide 6:

Coke Reed spent the closing decades of the 20th century working with the United States intelligence community and had accounts on Seymour Cray's first machines. During this time, he began to consider how the mathematical solution to "Problem #110" could be modified to describe and reinvent data movement within a system. As many of those present can attest, great ideas rarely come in the laboratory or at one's desk. One afternoon in the early '90s, Reed went on a walk in Rocky Mountain National Park with his beloved Labrador, Chloe. While staring into the mountains, the solution to his quandary appeared, and thus the "Data Vortex" was born.

What began as a walk in the woods has since grown into a vibrant company, made up of an intelligent team and exciting partnerships.

Slide 7:

In 1995, Reed patented "A Multi-Level, Minimum-Logic Network". The ring structure of the Problem 110 solution, which carried particles, inspired this structure, which carries information, a dynamic system of three-dimension Euclidean space.

Slide 8:

The Data Vortex topology allows for small packet data flow that is high radix, self-routing, congestion free and is enabled by fine-grained parallelism. This is important for networks that need high bandwidth yet low latency and, importantly, are linearly scalable.

Slide 9:

The Data Vortex topology can replace the crossbar within all points of the IT ecosystem. Crossbars like long packets and with long packets come a series of problems. Time is spent using an algebraic algorithm to set the switch, which is "hidden" by data transmission. Since low radix switches require many hops, high radix cross bar switches are chosen to reduce the number of hops. Higher radix crossbars require more time to set, resulting in longer packets.

Slide 10:

The Data Vortex does not have this setting problem, thus the packets can remain short. It consists of a collection of richly connected rings. The rings and the connection between the rings are built using parallel data busses. In a Radix R switch with $R=2^N$, the rings are arranged at $N+1$ levels. A packet on the entry level, level N, the outermost level, can travel to any of the output ports. When a packet travels from level N to level (N-1), the most significant bit of the binary address of the output is fixed so that a packet on level N-1 can reach only half of the output ports; a packet on level N-2 can reach only one fourth of the output ports and so forth. This process continues so that when a packet reaches level 0, the target output is determined. To overcome the problem of algebraic setting (causing the use of long packets) the network must be self-routing. A possible solution is for the network to be a dynamical system DV that is discrete in both time T and space S. DV is a function that is defined on the whole space S. The state of the whole space S at time t must be known in order to know the state of the whole space S at time t+1. By definition of a dynamical system: if x and y are in T and p is a point of S then $DV(0,p) = p$ and $DV[x+y,p] = DV[x,DV(y,p)]$

Slide 11:

The Data Vortex Network is thus a dynamical system that carries data. There is no switch setting for data movement management as in a crossbar, which is set by an algebraic algorithm. Data can be dropped into the flow of the system and transfer can be variable in size and can originate from a variety of inputs. A packet's trajectory through the network is dependent on the packet header and the location and movements of the other packets in the switch.

Slide 12:

What started as a solution to an unsolved problem has become central element of our technology, the "Data Vortex switch." On the right is Kuperberg's work on the Siefert conjecture, another example of derivative work from Problem #110. More than two dozen global patents on the Data Vortex switch have been filed and published over the past twenty years, each of them playing an important role in solidifying the portfolio of this revolutionary technology. This intellectual property effort is also tied to ongoing research and development projects that validate the mathematically proven claims in hardware.

Slide 13:

In legacy validation systems, this was done using FPGAs. Since it is costlier to efficiently distribute a DV network among many FPGAs, an entire radix 64 Data Vortex switch was put into a single FPGA. To increase compute node-to-compute node bandwidth, we interface each compute node with 16 parallel Data Vortex networks. Each Vortex Interface Controller (VIC) on a compute node selects a Data Vortex network at random, thus balancing the traffic load. The VIC contains an Intel Altera FPGA and 16 8 GB/s SerDes and communicates with the Data Vortex switch box. In these legacy validation systems, there are 2 switch boxes within a system and 8 Altera FPGAs per switch box. Each VIC connects to all 16 radix 64 networks in a switch box. Though Altera was chosen in this instance, the Data Vortex topology is not specifically tied to any single technology.

Slide 14:

In this diagram of the legacy system architecture, we see the route taken by long data packets, sent from commodity processors across the PCIe bus and broken up into smaller messages by the VIC. These smaller messages are then sent across 16 Radix *n* Data Vortex networks.

Slide 15:

These systems have been deployed across the United States at government and academic sites and have users from around the world, ranging in applications from few body physics and 3D-FFTs for NAMD and VASP to Quantum Simulation and big data graph analytics.

Slide 16:

Data Vortex users have seen noticeable performance improvement in many key applications that are data movement dependent. This has particularly been the case when (a) large amounts of small data packets are transferred in non-structured chaotic traffic and (b) there is congestion in the network. The former happens in situations where aggregation is otherwise very costly or impossible. In an instance of the Random-Access benchmark (Giga Updates per Second – “GUPs”), the Data Vortex network achieves greater than 100x acceleration compared with industry leading crossbar network topologies. Congestion in other networks happens when the number of compute nodes grows larger, as crossbar topologies are unable to handle the higher bandwidth. Data Vortex performance in this area has been explored with three dimensional FFTs and Breadth First Search. A Data Vortex-enabled system also has the highest performance per core per multi-node machine on the Graph500 Breadth First Search List.

Slide 17:

Here, we see latency as a function of switch loading. As the load increases, latency remains the same. The second graph comes from runs on *Mountain Dao*, a multi-level Data Vortex computer at Pacific Northwest National Laboratory. These runs prove the scalability of the Data Vortex topology as benchmark performance remains unchanged as the network expands to another hop.

A few years ago, Data Vortex users at PNNL identified limitations of the Data Vortex validation systems, particularly the PCIe bottleneck as the source of congestive limitations. The fault, therefore, lies not in the Data Vortex topology, but rather in the legacy protocol in present-day commodity chips and servers. While newer consortiums, such as CXL, Gen-Z, and CCIX, are exploring PCIe alternates, they fail to exploit the entire Data Vortex potential.

Slide 18:

Data Vortex Technology’s most recent global patent application, “Method and Apparatus for Improved Data Transfer Between Processor Cores,” solves this problem, as the Data Vortex switch is used to directly connect cores, bypassing all other busses and allowing for seamless exploitation of all Data Vortex properties. Simply stated, it is an embodiment of an interconnect apparatus that can enable improved signal integrity, even at high clock rates, increased bandwidth, and lower latency. In this instantiation, a sending processing core can send data to a receiving core by forming a packet whose header indicates the location of the receiving core and whose payload is the data to be sent. The packet is sent to the Data Vortex switch. This switch is on the same chip as an array of processing cores and routes the packet to the receiving core first by routing the packet to the processing core array containing the receiving processing core. The Data Vortex switch then routes the packet to the receiving processor core in a processing core array. Since the Data Vortex switches are not crossbar switches, there is no need to globally set and reset the Data Vortex switches as different groups of packets enter the switches.

Mounting the Data Vortex switch on the same chip as the array of processing cores and accessible, on-chip memory, reduces the power required and reduces the latency. The Data Vortex switch on-chip can talk to a larger Data Vortex network, increasing the scale from core-to-core communication across a chip to core-to-core communication across a server and core-to-core communication across a system. The benefit of this idealized system, internally referred to as "Chloe" (named after Dr. Reed's beloved dog, with whom he originally discovered the Data Vortex) is a unique computing environment with direct small packet core-to-core communication across thousands of cores. The user of this system can benefit from tremendous compute opportunity without paying the bandwidth and congestion costs that are constrained by present bottlenecks. If these Data Vortex switch on chips are manufactured on silicon substrates, then the SerDes setting problem is also erased. The resulting system would be self-thinking and unlike any other machine on the market. It fits the parameters of a smarter, more self-sufficient computer, such as the idealized "Future Computing System" outlined in IARPA's RFI-19-01, which was published in 2018.

Slide 19:

This market agnostic solution has the potential for applicability across multiple scales within the IT ecosystem and across the spectra of computing fields. Beyond traditional HPC, there exist possible entry points into the cloud, edge computing/IoT, data center retrieval and elsewhere. Already, efforts in cloud computing are underway using legacy Data Vortex technology and users have begun developing the Data Vortex for high performance middleware, such as RabbitMQ. At this very conference two years ago our partners at Providentia Worldwide presented "Vortex-topology Messaging: Data Vortex for high performance middleware", which compared RabbitMQ runs on Data Vortex and FDR Infiniband. The comparison looked at gradually larger messages, from 8 to 128,000 Bytes. In all instances, the Data Vortex showed at least twice the performance, with, 10x performance at 128,000 Bytes.

We are also playing in the realm of Quantum computing. Dr. Santiago Betelu, our Director of Research Science and an adjunct professor at the University of North Texas developed QuanSimBench with researchers at Los Alamos National Laboratory. The GitHub url is shown here and we encourage you and your colleagues to run the benchmark.

Slide 20:

Importantly, we believe that the Data Vortex can be used to address many of the problems that we are facing today, March 25, 2020. Hundreds of thousands, if not millions, of students and workers around the global are joining the already congested telecommuting infrastructure for the first time. New systems are needed to address the rapid acceleration of COVID-19 and other viruses. And current unrest is highlighting flaws in our financial infrastructure and national security. In this confusing period in human history, the need for a novel topology that can reimagine computing and communication is increasingly evident. As implementation opportunities and use cases for the topology expand, the central Data Vortex idea has remained unchanged and its ties to Stanislaw Ulam and the Polish mathematical tradition remain important.

Slide 21:

Members of our technical team, Mike Ives, Ron Denny, and Santiago Betelu, are here today and able to answer any of your specific questions. Thank you very much for your time.

References:

Gioiosa, R., Tumeo A., Yin, J., Warfel, T., Haglin, D., & S. Betelu (2017) "Exploring DataVortex Systems for Irregular Applications." *IPDPS 2017*.

Kolster, A. & S. R. Quick (2018) "Vortex-topology Messaging: Data Vortex for high performance middleware." *Supercomputing Frontiers 2018*.

Reed, C. S., Murphy, D., Denny, R, R., Ives, M. I. (2019) Patent Application *PCT/US19/65967* "Method and Apparatus for Improved Data Transfer Between Processor Cores".

Reed, C.S., Betelu, S. I., et. al. (2018) The Data Vortex Network: An overview of the architecture, implementation, and performance. *White Paper*. Data Vortex Technologies: Austin, Texas.

Reed, C. S. (1999) *US Patent 5996020A* "Multiple level minimum logic network."