Editors: Volodymyr Kindratenko, kindr@ncsa.uiuc.edu
Pedro Trancoso, pedro@cs.ucy.ac.cy

NOVO-G: AT THE FOREFRONT OF SCALABLE RECONFIGURABLE SUPERCOMPUTING

By Alan George, Herman Lam, and Greg Stitt

The Novo-G supercomputer's architecture can adapt to match each application's unique needs and thereby attain more performance with less energy than conventional machines.

Throughout computing's long history and within the many forms of computers existing today, from hand-held smartphones to mammoth supercomputers, the one common denominator is the fixed-logic processor. In conventional computing, each application must be adapted to match the fixed structures, parallelism, functionality, and precision of the target processor—such as the CPU, digital signal processor (DSP), or graphics processing unit (GPU)—as dictated by the device vendor.

Although advantageous for uniformity, this "one size fits all" approach can create dramatic inefficiencies in speed, area, and energy when the application fails to conform to the device's ideal case. By contrast, a relatively new computing paradigm known as *reconfigurable computing* (RC) takes the opposite approach, wherein the architecture adapts to match each application's unique needs, and thus approaches the speed and energy advantages of application-specific integrated circuit (ASIC), while offering the versatility of CPUs.

Many RC systems have emerged in the research community and marketplace, addressing an increasingly broad application range, from sensor processing for space science to proteomics for cancer diagnosis. Most of these systems are relatively modest in scale, featuring from one to several reconfigurable processors, such as field-programmable gate arrays (FPGAs). At the extreme RC scale, however, is the Novo-G machine at the US National Science Foundation's Center for High-Performance Reconfigurable Computing (NSF CHREC Center) at the University of Florida. Initial Novo-G studies show that, for some important applications, a scalable system with 192 reconfigurable processors can rival the speed of the world's largest supercomputers at a tiny fraction of their cost, size, power, and cooling. Here, we offer an overview of this novel system, along with its initial applications and performance breakthroughs.

Reconfigurable Computing

Demands for innovation in computing are growing rapidly. Technological advances are transforming many data-starved science domains into data-rich ones. For example, in genomics research in the health and life sciences, contemporary DNA sequencing instruments can determine 150 to 200 billion nucleotide bases per run, routinely resulting in output files in excess of 1 terabyte per instrument run.

In the near future, DNA sequence output from a single instrument run will easily exceed the size of the human genome by more than 100-fold. Thus, it's increasingly clear that the discordant trajectories growing between data production and the capacity for timely analysis are threatening to impede new scientific discoveries and progress in many scientific domains—not because we can't generate the data, but rather because we can't analyze it.

To address these growing demands with a computing infrastructure that's sustainable in terms of power, cooling, size, weight, and cost, adaptive systems that can be dynamically tailored to the unique needs of each application are coming to the forefront. At the heart of these systems are reconfigurable-logic devices, processors such as FPGAs that under software control can adapt their hardware structures to reflect the unique operations, precision, and parallelism associated with compute-intensive, data-driven applications in fields such as health and life sciences, signal and image processing, and cryptology.

The benefit of using RC with modern FPGAs for such applications lies in their reconfigurable structure. Unlike fixed-logic processors, which require applications to conform to

 Copublished by the IEEE CS and the AIP 1521-9615/11/$26.00 © 2011 IEEE

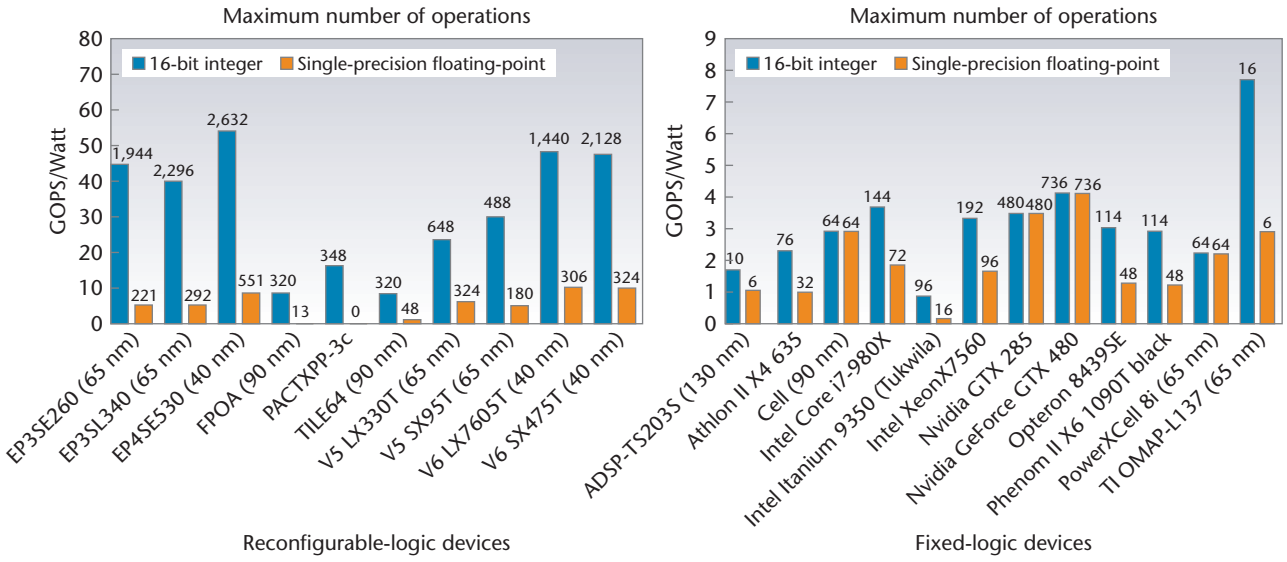


Figure 1. Computational density (in giga operations per second) per watt of modern fixed- and reconfigurable-logic devices.[2] As the figure shows, reconfigurable-logic devices often achieve significantly more operations per watt than fixed-logic devices.

a fixed structure (for better or for worse), with RC the architecture conforms to each application's unique needs. This adaptive nature lets FPGA devices exploit higher degrees of parallelism while running at lower clock rates, thus often achieving better execution speed while consuming less energy.

Figure 1 shows a comparative suite of device metrics,[1,2] which illustrates performance (in terms of computational density) per watt of some of the latest reconfigurable- and fixed-logic processing devices for 16-bit integer (Int16) or single-precision floating-point (SPFP), assuming an equal number of add and multiply operations. The number above each bar indicates the peak number of sustainable parallel operations.

Generally, FPGAs achieve more speed per power unit compared to CPUs, DSPs, and GPUs. For example, the study's leading FPGA for Int16 operations (Altera Stratix-IV EP4SE530) can support more than 50 billion operations per second (GOPS) per watt, while the leading fixed-logic processor (TI OMAP-L137 DSP) attains less than 8 GOPS/watt. With SPFP, the gap is narrower, but FPGAs continue to enable more GOPS/watt. Similarly, although not shown in the figure, the gap widens increasingly for simpler (byte or bit) operations. With RC, the simpler the task, the less chip area required for each processing element (PE), and thus the more PEs that can fit and operate concurrently in hardware.

However encouraging, this promising approach has thus far been limited primarily to small systems, studies, and datasets. To move beyond these limits, some key challenges must be overcome. Chief among these challenges is parallelization, evaluation, and optimization of critical applications in data-intensive fields in a way that's transparent, flexible, portable, and performed at a much larger scale commensurate with the massive needs of emerging, real-world datasets. When successful, however, the impact will be a dramatic speedup in execution time concomitant with savings in energy and cooling.

As we describe later, our initial studies show critical applications executing at scale on Novo-G, achieving speeds rivaling the largest conventional supercomputers in existence—yet at a fraction of their size, energy, and cost. While processing speed and energy efficiency are important, the principal impact of a reconfigurable supercomputer like Novo-G is the freedom that its innovative approach can give to scientists to conduct more types of analysis, examine larger datasets, ask more questions, and find better answers.

Our initial studies show critical applications executing at scale on Novo-G, achieving speeds rivaling the largest conventional supercomputers in existence—yet at a fraction of their size, energy, and cost.

Novo-G Reconfigurable Supercomputer

The Novo-G experimental research testbed has been operating since July 2009 at the NSF CHREC Center, supporting various research projects on scalable RC challenges. Novo-G's primary emphases are *performance*

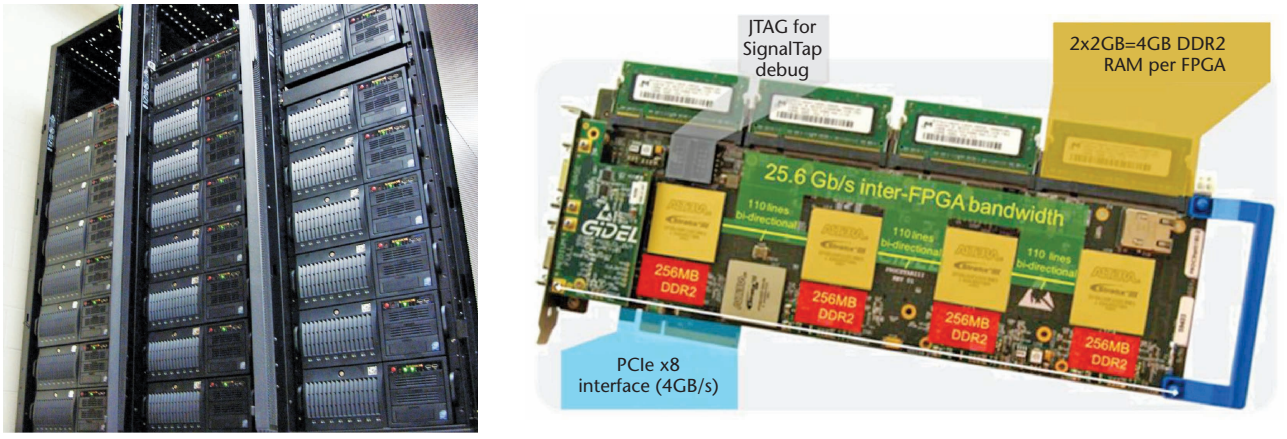


Figure 2. Novo-G and a processor board. (a) The Novo-G supercomputer and (b) one of its quad-FPGA reconfigurable processor boards. The current configuration includes 24 compute nodes housed in three racks. The head node is a single 1U server with twin quad-core Xeons.

(device, subsystem, system), *productivity* (concepts, languages, tools), and *impact* (scalable applications).

Figure 2 shows the Novo-G machine and one of its quad-FPGA boards. The current Novo-G configuration consists of 24 compute nodes, each a standard 4U Linux server with an Intel quad-core Xeon (E5520) processor, memory, disk, and so on, housed in three racks. A single 1U server with twin quad-core Xeons functions as head node. Compute nodes communicate and synchronize via Gigabit Ethernet and a nonblocking fabric of 20 Gbits/s InfiniBand. Each of the compute nodes houses two PROCStar-III boards from GiDEL in its PCIe slots. Novo-G's novel computing power is derived from these boards, each containing four Stratix-III E260 FPGAs from Altera, resulting in a system of 48 boards and 192 FPGAs. (An impending upgrade will soon add 72 Stratix-IV E530 FPGAs to Novo-G, each with twice the reconfigurable logic of a Stratix-III E260—yet roughly the same power consumption—thereby expanding Novo-G's total reconfigurable logic by nearly 80 percent.) Concomitantly, when fully loaded, the entire Novo-G system's power consumption peaks at approximately 8,000 watts.

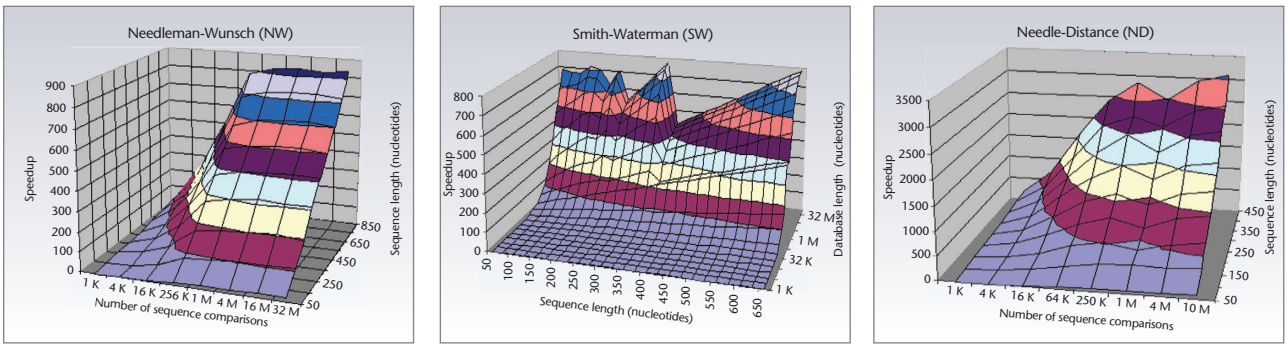
While this set of FPGAs can theoretically provide the system with massive computing power, the memory capacity, throughput, and latency often limit performance if unbalanced. As Figure 2b shows, 4.25 Gbytes of dedicated memory is attached to each FPGA in three banks. Data transfer between adjacent FPGAs can be made directly through a wide, bidirectional bus at rates up to 25.6 Gbps and latencies of a single clock cycle up to 300 MHz, and transfer between FPGAs across two boards in the same server is also supported via a high-speed cable. By supplying each FPGA with large, dedicated memory banks, as well as high bandwidth and low latency for inter-FPGA data transfer, the system strongly supports RC-centric applications. Processing engines on the FPGAs can execute with minimal involvement by the host CPU cores, enabling maximum FPGA utilization.

Alongside the architecture, equally important are the design tools available and upcoming for Novo-G. RC's very nature empowers application developers with far more capability, control, and influence over the architecture. Instead of stipulating all architecture decisions to the device vendors—as with CPUs and GPUs—in RC, the application developer specifies a custom architecture configuration, such as quantity and types of operations, numerical precision, and breadth and depth of parallelism. Consequently, RC is a more challenging environment for application development, and productivity concepts and tools are thus vital.

Novo-G offers a broad and growing range of academic and commercial tools in areas such as

- strategic design and performance prediction tools for parallel algorithm and mapping studies;
- message-passing interface (MPI), Unified Parallel C (UPC), and shared memory (SHMEM) library for system-level programming with C;
- Very High-Speed Integrated Circuit Hardware Description Language (VHDL), Verilog, and an expanding list of high-level synthesis tools for FPGA-level programming;
- an assortment of core libraries;
- middleware and APIs for design abstraction, platform virtualization, and portability of apps and tools; and
- verification and performance-optimization tools.

To help expand the applications and tools available on Novo-G and establish and showcase RC's advantages at scale, the Novo-G Forum was formed in 2010. This forum is an international



Baseline: $192 \cdot 2^{25}$, length 850 sequence comparisons
Software runtime: 11,026 CPU hours on 2.4 GHz Opteron

# FPGAs	Runtime (sec)	Speedup
1	47,616	833
4	12,014	3,304
96	503	78,914
128	391	101,518
192 (est.)	270	147,013

(a)

Baseline: Human X chromosome v 19200, length 650 Seqs
Software runtime: 5,481 CPU hours on 2.4 GHz Opteron

# FPGAs	Runtime (sec)	Speedup
1	23,846	827
4	5,966	3,307
96	250	78,926
128	188	104,955
192 (est.)	127	155,366

(b)

Baseline: $192 \cdot 2^{24}$, length 450 distance calculations
Software runtime: 11,673 CPU hours on 2.4 GHz Opteron

# FPGAs	Runtime (sec)	Speedup
1	13,522	3,108
4	3,429	12,255
96	144	291,825
128	118	356,125
192 (est.)	77	545,751

(c)

Figure 3. Performance results on Novo-G for three bioinformatics applications: (a) Needleman-Wunsch (NW), (b) Smith-Waterman (SW) without traceback, and (c) Needle-Distance (ND). Each chart illustrates the performance of a single FPGA under varying input conditions. Each table shows performance with varying number of FPGAs under optimal input conditions.[3]

group of academic researchers and technology providers working collaboratively with a common goal: to realize the promise of reconfigurable supercomputing by demonstrating unprecedented levels of performance, productivity, and sustainability. Faculty and students in each academic research team are committed to contributing innovative applications and tools research on the Novo-G machine based upon their unique expertise and interests. Among the forum participants are Boston University, Clemson University, University of Florida, George Washington University, University of Glasgow (UK), Imperial College (UK), Northeastern University, Federal University of Pernambuco (Brazil), University of South Carolina, University of Tennessee, and Washington University at St. Louis. Each academic team has one or more Novo-G boards for local experiments and has remote access to the large Novo-G machine at Florida for scalability studies.

Initial Applications Studies

Of Novo-G's three principal emphases, *impact* is undoubtedly the most important. What good is a new and innovative high-performance system if the resulting applications have little impact in science and society? We now offer an overview of Novo-G's initial performance breakthroughs on a set of bioinformatics applications for genomics, which we developed in collaboration with the University of Florida's Interdisciplinary Center for Biotechnology Research (ICBR). Results of such breakthroughs can potentially revolutionize the processing of massive genomics datasets, which in turn might enable revolutionary discoveries for a broad range of challenges in the health, life, and agricultural sciences.

Although more than a dozen challenging Novo-G application designs are underway in several scientific domains, our focus here is on our first case studies. These include two popular genomics applications for optimal sequence alignment based upon wavefront algorithms:

- Needleman-Wunsch (NW) and
- Smith-Waterman (SW) without traceback,

and a metagenomics application—Needle-Distance (ND)—which is an augmentation of NW with distance calculations. We're nearing completion on an extended version of SW with the traceback option—SW+TB—by augmenting our SW hardware design to collect and feed data for traceback to the hosts so that FPGAs can perform SW while CPU cores perform TB. Initial results indicate that, after adding TB, execution times increase less on Novo-G than on the C/Opteron baseline, and thus Novo-G speedups with SW+TB exceed those of SW.

Each of the applications features massive data parallelism with minimal communication and synchronization among FPGAs, and a highly optimized systolic array of processing elements (PEs) within each FPGA (and optionally spanning multiple FPGAs). Using a novel method for in-stream control,[3] we optimized each of the three designs to fit up to 850 PEs per FPGA for NW, 650 for SW, and 450 for ND, all operating at 125 MHz.

Figure 3 shows a contour plot for each application that illustrates relative design performance on one FPGA under varying input conditions. The corresponding tables show how the three designs scale when

executed on multiple FPGAs in Novo-G. In all cases, speedup is defined in terms of an optimized C-code software baseline running on a 2.4 GHz Opteron core in our lab. More details on these algorithms, architectures, experiments, and results are provided elsewhere.[3] All data except the tables' final rows came directly from measurements on Novo-G and include the full execution time, including not just computation time, but data transfers to and from the FPGAs.

Speedup with one FPGA on each of the three applications peaked at approximately 830 for NW and SW and more than 3,100 for ND. When ramping up from a single FPGA to a quad-FPGA board, we measured and observed speedups to grow almost linearly to about 3,300 for NW and SW and more than 12,000 for ND. At the largest scale of our testbed experiments—32 boards, or 128 FPGAs—speedups for NW and SW exceeded 100,000 and ND exceeded 356,000.

Because not all 48 boards in Novo-G were operational during our study, we extrapolated these trends, estimating speedups on all 192 FPGAs of Novo-G of about 150,000 for NW and SW and almost 550,000 for ND. Putting these numbers in context, the latter implies that a conventional supercomputer would require more than a half-million Opteron cores operating optimally to match Novo-G's performance on the ND application. By contrast, none of the world's largest supercomputing machines (as cited, for example, in the www.top500.org's top rankings) has this many cores, and thus none could achieve such performance on this application despite being orders of magnitude larger in cost, size, weight, power, and cooling. Although Novo-G won't provide all applications with the same speedups as these examples, they do highlight RC's potential advantages, especially in solving problems where conventional, fixed-logic computing falls far short of achieving optimal performance.

For a growing list of important applications from a broad range of science domains, underlying computations and data-driven demands are proving to be underserved by conventional "one size fits all" processing devices. By changing the mindset of computing—from processor-centric to application-centric—reconfigurable computing can provide solutions for domain scientists at a fraction of the time and cost of traditional servers or supercomputers. As we describe here, the Novo-G machine, applications, research forum, and preliminary results are helping to pave the way for scalable reconfigurable computing.

References

1. J. Williams et al., "Characterization of Fixed and Reconfigurable Multi-Core Devices for Application Acceleration," *ACM Trans. Reconfigurable Technology and Systems*, vol. 3, no. 4, 2011; to appear.
2. J. Richardson et al., "Comparative Analysis of HPC and Accelerator Devices: Computation, Memory, I/O, and Power," *Proc. High-Performance Reconfigurable Computing Technology and Applications Workshop, ACM/IEEE Supercomputing Conf. (SC10)*, IEEE Press, to appear.
3. C. Pascoe et al., "Reconfigurable Supercomputing with Scalable Systolic Arrays and In-Stream Control for Wavefront Genomics Processing," *Proc. Symp. Application Accelerators in High-Performance Computing*, 2010; www.chrec.org/pubs/SAAHPC10_F1.pdf.

Alan George is director of the US National Science Foundation Center for High-Performance Reconfigurable Computing and a professor of electrical and computer engineering at the University of Florida. His research interests focus upon high-performance architectures, networks, systems, services, and applications for reconfigurable, parallel, distributed, and fault-tolerant computing. George has a PhD in computer science from the Florida State University. He is a member of IEEE Computer Society, the ACM, the Society for Computer Simulation, and the American Institute of Aeronautics and Astronautics. Contact him at ageorge@ufl.edu.

Herman Lam is an associate professor in the Department of Electrical and Computer Engineering at the University of Florida. His research interests include design methods and tools for RC application development, particularly as applied to large-scale reconfigurable supercomputing. Lam has a PhD in electrical and computer engineering from the University of Florida. He is a member of IEEE and the ACM and is a faculty member of the NSF Center for High-Performance Reconfigurable Computing. Contact him at hlam@ufl.edu.

Greg Stitt is an assistant professor in the Department of Electrical and Computer Engineering at the University of Florida and a faculty member of the US National Science Foundation Center for High-Performance Reconfigurable Computing. His research interests include design automation for reconfigurable computing and embedded systems. Stitt has a PhD in computer science from the University of California, Riverside. He is a member of IEEE and the ACM. Contact him at gstitt@ece.ufl.edu.