Convey Computer™ GraphConstructor™

The Convey GraphConstructor is a hardware and software solution that accelerates next-generation discovery by speeding up the construction and manipulation of de Bruijn graphs commonly used in short-read genome assembly applications. With increased performance, scientists are able to better manage and analyze escalating amounts of research data.

The flood of life science data gushing from new experimental technologies has grown so large and diverse that bioinformatics software and commodity computing are struggling to keep up. To help researchers better manage and analyze escalating amounts of research data, Convey Computer developed the Convey GraphConstructor (CGC), a software and hardware solution that speeds up some of the world's most popular bioinformatics algorithms.

The Convey GraphConstructor accelerates construction and manipulation of de Bruijn graphs commonly used in short-read, *de novo* genome assembly applications such as Velvet[1] and Abyss[2].

POWER OF HYBRID-CORE COMPUTING

The Convey solution starts with its innovative hybrid-core platforms – the HC-1 and the HC-1^{ex} – that use reconfigurable technology and a parallel memory subsystem to accelerate applications and lower the cost of ownership. Convey's hybrid-core architecture pairs classic Intel® x86 microprocessors with a coprocessor comprised of FPGAs. Particular algorithms—DNA sequence alignment, for example—are optimized and translated into code that's loadable onto the FPGAs at runtime to accelerate the applications that use them. Convey calls these accelerated algorithms "personalities."

While raw processor performance increases are important, improved memory management is often just as important to increasing research throughput. Bioinformatics applications that depend upon random access patterns to large memory spaces, such as graph-based algorithms, experience severe memory performance limitations on cache-based x86 servers and greatly benefit from Convey's memory architecture. Convey's highly parallel memory subsystem allows application-specific logic to concurrently access 8,192 individual words in memory, significantly increasing effective memory bandwidth over cache-based memory systems.

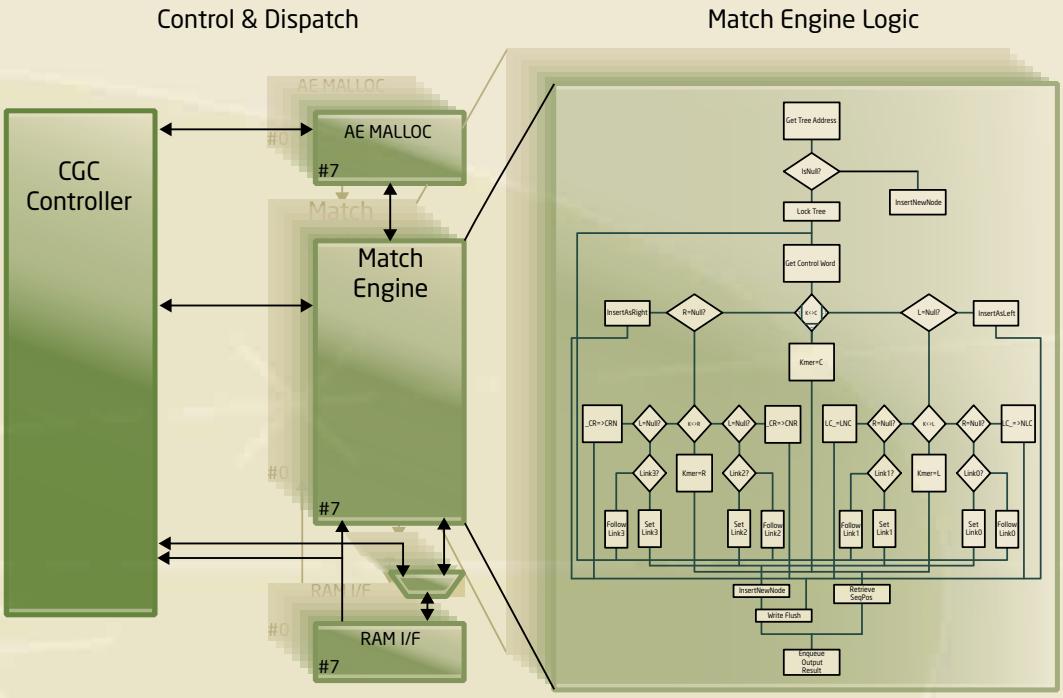
CONVEY GRAPHCONSTRUCTOR

A popular method for *de novo* assembly of short-read sequence data generated by modern sequencers is the de Bruijn graph approach, typified by Velvet[1]. The method is popular because it effectively handles the high redundancy of short-read data, it represents all possible ways to assemble the reads, and it is able to address sequencer errors through graph analysis. The accuracy of resulting assemblies is well understood. However, the long runtimes and extensive memory requirements, particularly for larger genomes, can make its use prohibitive.

The World's First Hybrid-Core Computer.

Convey Computer™ GraphConstructor™

THE CONVEY HYBRID-CORE ARCHITECTURE ALLOWS PERFORMANCE-CRITICAL CODE TO BE IMPLEMENTED DIRECTLY IN HARDWARE



Optimized specifically for the Convey hybrid-core architecture, the Convey GraphConstructor efficiently generates de Bruijn graphs. It reduces the execution time and memory required for graph construction by implementing the read hashing step in hardware on a Convey coprocessor. The CGC also takes advantage of the large randomly accessible memory and parallelism available from multiple large FPGAs. The highly parallel memory of Convey's cache-less approach allows application-specific logic to directly address individual words in memory, increasing effective memory bandwidth as much as ten-fold.

The Convey GraphConstructor approach produces many important advantages, enabling researchers to:

- Tackle previously impractical genome assemblies
- Significantly accelerate throughput
- Reduce memory requirements
- Generate higher quality assemblies by running multiple times with differing parameters
- Lower capital and power/cooling costs

Since roadmap generation requires large amounts of memory, the CGC supports automatic partitioning of the roadmap to be processed in multiple passes. The generated partitions can then fit in available memory, permitting runs on larger data sets that otherwise would be impossible.

Convey's innovative hybrid-core computer system marries the low cost and simple programming model of a commodity system with the performance of a customized hardware architecture. For more information, please see www.conveycomputer.com.

[1]http://www.ebi.ac.uk/~zerbino/velvet/; Velvet is the most widely used program for *de novo* assembly of short-read sequences.
[2]http://www.ncbi.nlm.nih.gov/pubmed/19251739

CONV-11-019.1 © 2011 Convey Computer Corporation. Convey Computer, the Convey logo, GraphConstructor, HC-1 and Convey HC-1ex are trademarks of Convey Computer Corporation. Printed in the U.S.A.

CONVEY COMPUTER CORPORATION • 1302 E. COLLINS BOULEVARD, RICHARDSON, TEXAS 75081
PHONE: 214.666.6024 • TOLL FREE: 866.338.1768 • WWW.CONVEYCOMPUTER.COM