



Success Brief

Intel® Core™ 2 Duo
processor technology
Quad-Core Intel® Xeon®
processor 3200 series

CLC bio



“Intel has helped us bring super computing power to the desktop while underpinning far greater accuracy in genetic research findings.”

Jan Lomholdt, Vice-President Sales and Marketing, CLC Bio

Life-changing research receives powerful boost

Intel core microarchitecture delivers new levels of analytical capability to bioinformatics industry

CLC bio is the world's leading supplier of full service bioinformatics solutions. Bioinformatics is the use of computer science, mathematics and information theory to model and analyse biological systems, typically at the molecular level and especially systems involving genetic material. Solving bioinformatics problems can for example, accelerate drug discovery research helping find solutions to pressing healthcare needs like HIV or Avian Flu.

Headquartered in Aarhus, Denmark, CLC bio provides a wide range of tools that aid researchers in these and other fields. It is the only bioinformatics company to provide software, hardware and data analysis; as well as custom designed bioinformatics algorithms.

Search for greater accuracy

Much of the current bioinformatics research uses industry-standard algorithms such as Smith-Waterman, BLAST and HMMER to find sequencing matches within vast quantities of data.

However, CLC bio acknowledged that the algorithms were not being utilised to their full potential. For example, Smith-Waterman is widely regarded by researchers to be the more precise algorithm but it is also very time consuming to use. As a result, many prefer BLAST because it is a lot faster. But it is not as accurate, up to 50% of hits can be missed. Consequently, some researchers run the slower Smith-Waterman instead of BLAST.

However, CLC bio has accelerated other known and computationally intensive algorithms, such as ClustalW and the very popular HMMER. The company recognised that by adopting a Predictive Enterprise approach, technology could be used to maximise return on investment (ROI). In turn business growth could be enhanced through the development of technology that is adaptive to its customers' changing needs.

Jan Lomholdt, Vice-President of Sales and Marketing at CLC bio, said: “Bioinformatics research is growing rapidly and we wanted to bring greater accuracy within the reach of all researchers. We were keen to explore the potential of Single Instruction Multiple Data (SIMD) technology embedded into Intel® processor technology, which we believed would significantly streamline parallel data processing.”

Measures of Success

- Provides bioinformatics researchers with Predictive Enterprise capability enabling them to capture and manage the explosion in bioinformatics data while delivering the tools to analyse the information at high speed
 - The speed of the accurate but relatively slow bioinformatics Smith-Waterman algorithm is significantly increased due to Cell* software that has been developed to work with Single Instruction Multiple Data (SIMD) technology embedded into the latest Intel® processor technology
 - Laptop and desktop PCs powered by SIMD-enabled Intel® Core™2 Duo processor technology gain computing capability that previously was only available from high performance clusters
-

Single Instruction Multiple Data technology embedded into the latest Intel® processor technology turns desktop PCs into supercomputers

To achieve its aims the company developed Bioinformatics Cell* software designed to run specifically on SIMD-enabled Intel® Core™ 2 Duo processor technology. To carry out a series of benchmarking tests¹, Cell software was loaded onto Dell desktop hardware powered by Intel Core 2 Duo processor technology and running Linux Fedora Core 6* operating system.

The performance of Smith-Waterman version BLASTn was measured against traditional software running on the same platform. One hundred query sequences were made of a 20 million nucleotides (DNA sequences) database. This test revealed that Intel Core 2 Duo processor technology with Cell* software was up to 130 times faster than traditional software, reducing previous search times from 12 hours to seven minutes.

In speeds measured by Giga Cell Updates per Second (GCUPS), the CLC bio Cell software ran Smith-Waterman at 13.00 GCUPS on the SIMD-enabled Intel Core 2 Duo processor technology, compared to a speed of 0.05 GCUPS on a Single-Core Intel® Pentium® 4 processor not using the accelerated version.

Running an identical benchmark setup on a SIMD-enabled Quad-Core Intel® Xeon® processor 3200 series, revealed that the speed of Smith-Waterman was up to 144 times faster than traditional software.

Return on Investment

- **Sense** – CLC bio Cell* 2.0 software powered by SIMD-enabled Intel® Core™ 2 Duo processor technology is up to 130 times faster than traditional bioinformatics software. Standard bioinformatics Smith-Waterman algorithm searches carried out on typical desktop reduced from 12 hours to approximately seven minutes enabling researchers to gather or sense far more information.
- **Predict** – Server powered by SIMD-enabled Quad-Core Intel® Xeon® processor 3200 series, running second version of Cell software increased the speed of Smith-Waterman algorithm by up to 144 times compared to traditional bioinformatics software, enabling researchers to assimilate and analyse the data to form bioinformatics hypothesis.
- **Act** – A huge number of versions of the Intel SIMD-enabled Cell software were downloaded by researchers as soon as it was made available – and in the following months, several highly-respected and influential scientific journals published a series of articles on the advances it enabled.
- Cluster servers powered by SIMD-enabled next-generation 45nm Intel Core 2 Duo processor technology expected to radically enhance the accuracy and speed of bioinformatics algorithms.



The popular ClustalW alignment algorithm and the commonly used HMMER algorithm, tested on the Cell*, achieved speeds 21 times and 17 times faster respectively, than those gained through the use of traditional software.

Lomholt said: "These performance increases are groundbreaking. But they would have been even higher if we had used the Cell* software on SIMD-enabled Quad-Core Intel Xeon processor 5300 series. The performance scales very well and already supports next-generation 43nm Intel Core 2 Duo processor technology."

"With the Cell* software allied to Intel® technology, we've brought super-computing capability to laptops and desktop PCs equipped with SIMD-enabled Intel® processor technology. Previously you would need high-performance clusters to get this level of power, which means researchers can carry out much faster and far more accurate research – all from a laptop if necessary."

Find a business solution that is right for your company. Contact your Intel representative or visit the Intel® Business/Enterprise Web site at: <http://www.intel.com/go/digitaloffice> or visit the industry solutions-specific sites at: <http://www.intel.com/business/bss/industry>.

Copyright © 2007 Intel Corporation. All rights reserved. Intel, the Intel logo, Intel. Leap Ahead, the Intel. Leap Ahead. logo, Intel Core, Core Inside, Xeon and Xeon Inside are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. This document is for informational purposes only. INTEL MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS DOCUMENT.

¹ Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel® products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing.

* Other names and brands may be claimed as the property of others.

1107/SKD/RU/XX/PDF 318747-001E

