

# Reconfigurable Accelerator for Bioinformatics

## ゲノム解析からシミュレーションへ。 計算能力は足りているか？

ここ数年の間に、ゲノム解読技術は急速に進展し、既にヒトやイネなどのゲノムが解読されています。その解析には大規模なPCクラスタが活躍してきました。生物の働きの全貌を理解するにはまだまだ多くのデータが必要ですが、細胞などを数理的なモデルとして表現し、シミュレーションする研究が数多く行われるようになってきています。

さまざまな病気の発生メカニズムを分析し、新たな治療薬を開発すること、悪条件に耐えて生育する農作物を開発することなど、これからのバイオテクノロジーの夢は計算機によるシミュレーション抜きには語れないといっても過言ではありません。

DNA配列の解析はクラスタシステムが得意とする分野ですが、細胞のはたらきはとて複雑で、シミュレーションをクラスタで効率よく行うのは難しい課題です。クラスタでは通信がボトルネックになってしまうような処理も、Reconfigurable Systemであれば、必要な演算器を高密度に並べることで、高速に実行できる可能性を秘めています。現在は実時間よりもシミュレーション時間のほうが長くなりがちですが、これを実時間よりも速くすることができれば、大きなブレイクスルーにつながることを期待されます。

## STRATEGY:

生物学の分野で用いられるアルゴリズムは多種多様であり、そのひとつひとつに対して専用計算機を開発することは現実的には不可能であるといえます。では、ソフトウェア処理をクラスタなどの並列システムを用いて加速するしか方法はないのでしょうか？

FPGAをはじめとするプログラマブルデバイスを用いた Reconfigurable Computing はそのひとつの解決法です。同一の基板上に、ソフトウェア的にさまざまな問題を直接解く回路を構成することが可能であり、さらにデバイスの容量が許す限り多くの演算器を用いて並列処理を行うことができます。これは、演算器資源が限られている汎用のCPUに対して大きなアドバンテージで、比較的低い動作周波数で大きな計算能力を得ることが可能になります。

## ARCHITECTURE:

このボードの中心になっているのは、Xilinx Virtex-II (XC2V6000) で、システムゲート数は600万ゲート相当です。高速動作とともに、大きな回路を載せることができ、計算処理のほかにボード上のすべてのコンポーネントの制御を受け持ちます。

Virtex-IIの周辺には、4MBの同期型SRAMを4セット、64MBのSDRAMを接続しており、メモリへの並列アクセスを可能にすることでFPGA上に構成した演算器を効率よく並列動作させることができます。ホストインターフェイスは64bit/66MHzのPCIバスを備えています。



## APPLICATION 1: Cell Simulation

このボードを用いたアプリケーションとして、最も重点がおかれているのが細胞の代謝系シミュレーションです。現在、代謝系シミュレーションによく用いられるいくつかのアルゴリズムの実装を行っており、ボード上のFPGAとホストCPUで適切な計算負荷の分配を行うことで、PentiumIII 1GHzの10倍から100倍程度の性能が得られる見込みです。

## APPLICATION 2: image processing

顕微鏡から得られる画像をCCDカメラ経由でボードに入力し、細胞の分裂過程を自動認識するシステムを構築中です。従来は長い時間をかけて手作業で行われてきた作業ですが、これを自動化することでさまざまな変異体に対して網羅的に実験を行うことが可能になります。性能的には、PentiumIII 450MHzの100倍以上の処理速度を発揮しており、32台のPCクラスタを1枚のボードで置き換えることが可能になります。

## APPLICATION 3, 4, 5: ...

その他にも、このボードを用いて、さまざまな分野における計算を加速する試みを行っています。

## すべての研究者に、スーパーコンピュータを。

生物学分野の研究者が取り扱うデータは巨大で、高速な計算処理が要求されるのはシミュレーションだけではなく、

自動化された高速な計測器の出力する大量のデータをリアルタイムに分析し、次の実験計画を立案しようとするときに、数千にもものぼるサンプルに対して網羅的に実験を行い、そこから得られるデータを分析しようとするときに、従来考えられてきた計算科学的な分野とは比べものにならないほどの計算資源が必要になります。

しかし、誰もが自分の研究室にクラスタを設置することができるわけではありません。消費電力、サイズ、価格、設置場所、そのどれもがユーザーを選びます。専用計算機というアプローチでは、性能やサイズの問題は解決しますが、特定の問題しか解くことができません。データ分析にもシミュレーションにも使える高速なシステムが出現すれば、大いに歓迎されることは間違いありません。Reconfigurable Systemは、専用計算機の性能と汎用計算機の柔軟性を両立しうるソリューションであるといえます。

Amano Lab. Keio University

bio@am.ics.keio.ac.jp

〒223-8522 横浜市港北区日吉 3-14-1 慶應義塾大学理工学部情報工学科天野研究室

# ReCSiP: a Reconfigurable Cell Simulation Platform

## 目標

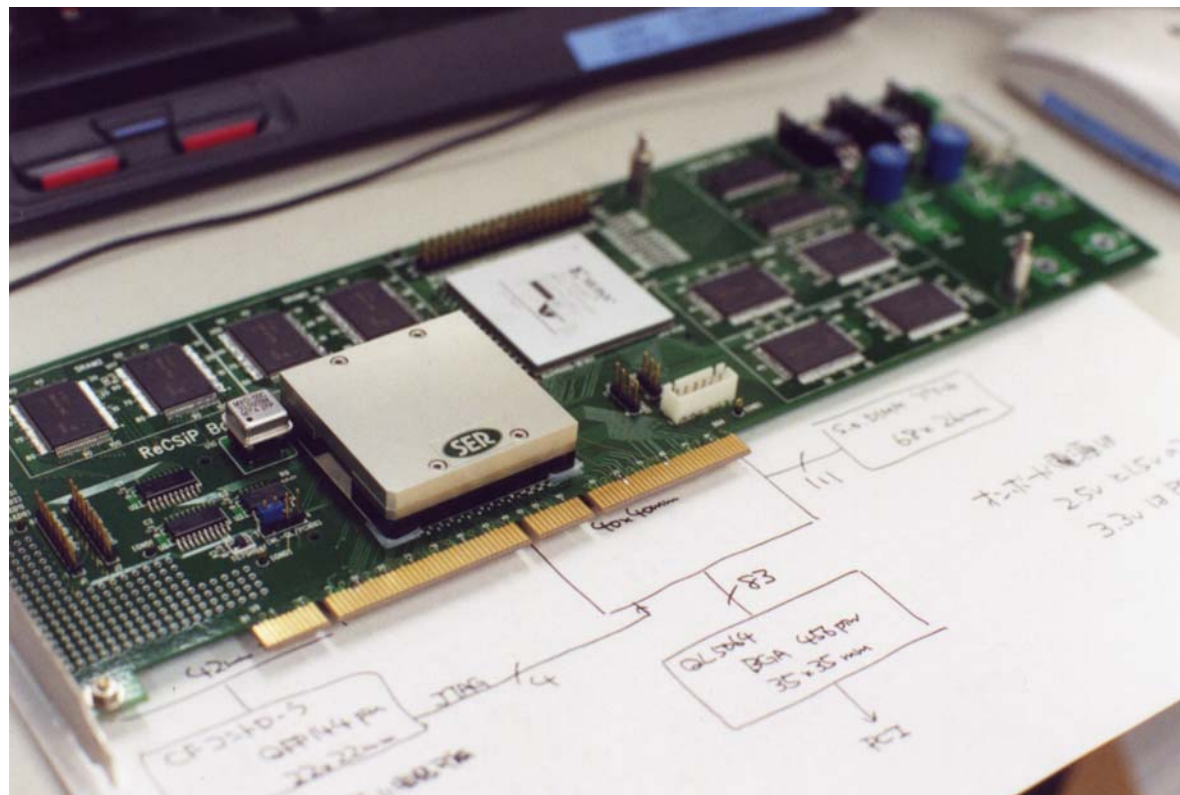
- ◎ FPGA の Bioinformatics への応用可能性の検証
- ◎ 代謝系シミュレーション、画像処理など

## 現状

- ◎ 実機が動作中
- ◎ いくつかのアプリケーションの実装が進行中
  - ✕ 代謝系シミュレーション
  - ✕ 画像処理（細胞系譜）
  - ✕ など

## 構成

- ◎ FPGA
  - ✕ Xilinx Virtex-II (XC2V6000)
- ◎ メモリ
  - ✕ SSRAM (4MB x 4 セット)
  - ✕ SDRAM (64MB x 1 セット)
- ◎ ホストインターフェイス
  - ✕ 64bit/66MHz PCI



# ReCSiP2: The next version of ReCSiP

## 目標

- ◎ さらなる性能向上
- ◎ ソフトウェア・ハードウェア協調シミュレーション
- ◎ FPGAを用いた大規模システムへの布石

## 変わること

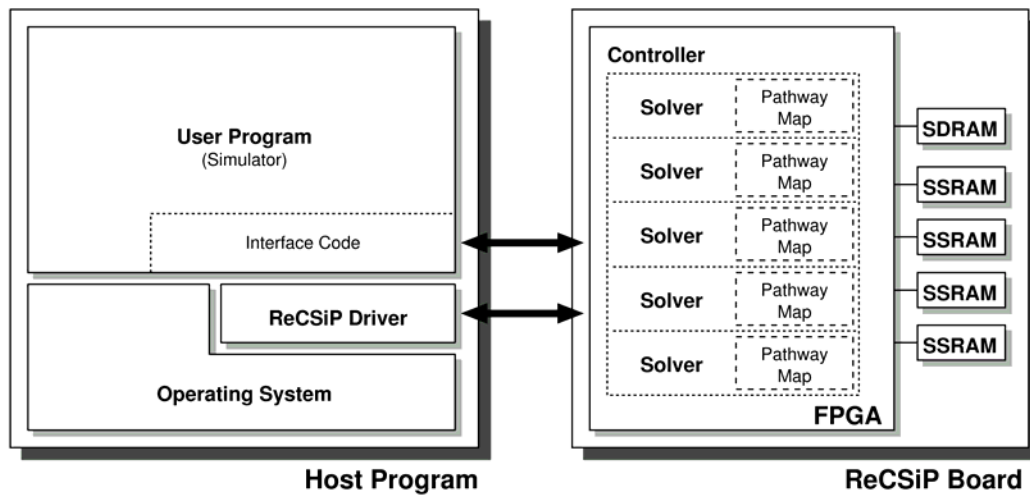
- ◎ FPGA の変更
  - ✕ PowerPC を内蔵 (Virtex-II Pro)
- ◎ PCIに頼らないボード間相互接続
  - ✕ 複数のFPGAを用いたシミュレーションを実現
  - ✕ ボード間接続用の 10Gbps インタフェースを装備（電氣的・機械的規格は Infiniband 4x に準ずる）
- ◎ メモリのモジュール化
  - ✕ 1GB までの SO-DIMM を装備可能にすることにより、大容量のシミュレーションデータに対応

## 進捗状況

- ◎ 現在設計が進行中
  - ✕ 2004 年 2 月実機完成予定



## ReCSiP: a ReConfigurable Cell Simulation Platform



FPGA を用いたシステムの性能を最大限に発揮するには、FPGA とホスト CPU の協調処理環境が必要になります。FPGA が得意とする処理と CPU が得意とする処理の性質は異なるからです。しかし、ReCSiP のユーザは必ずしも計算機の専門家ではなく、生物学を専門とする研究者である可能性も想定しなければなりません。ハードウェアに関する十分な知識がなければ、デバイスドライバをはじめ、FPGA と通信するようなプログラムを書くのは容易ではありません。

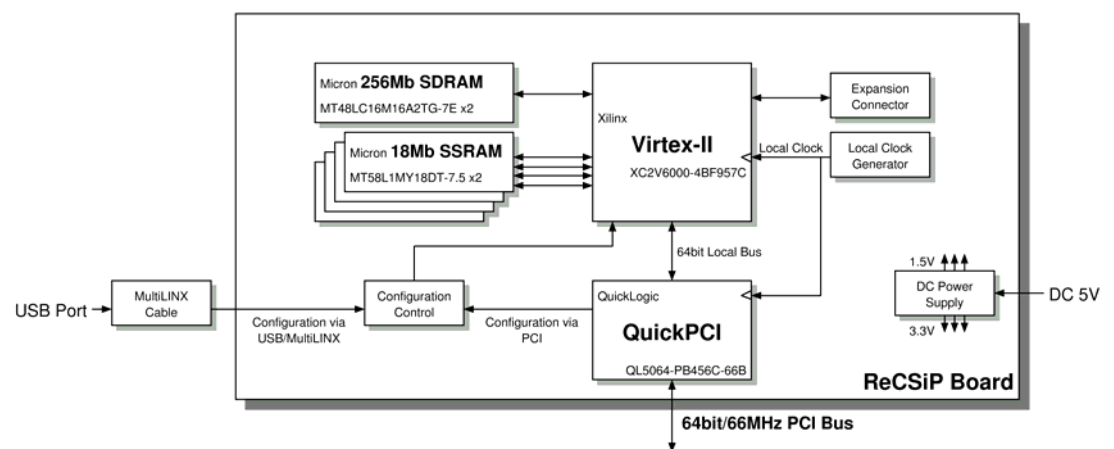
ReCSiP では、ボードやそのドライバだけでなく、

- FPGA に搭載する標準化されたホストインターフェイス
- このインターフェイスと通信するためのプログラム開発を支援するツールを提供することで、FPGA に載せる回路のデザインを容易にするとともに、ソフトウェアの開発も容易に行えるように配慮されています。

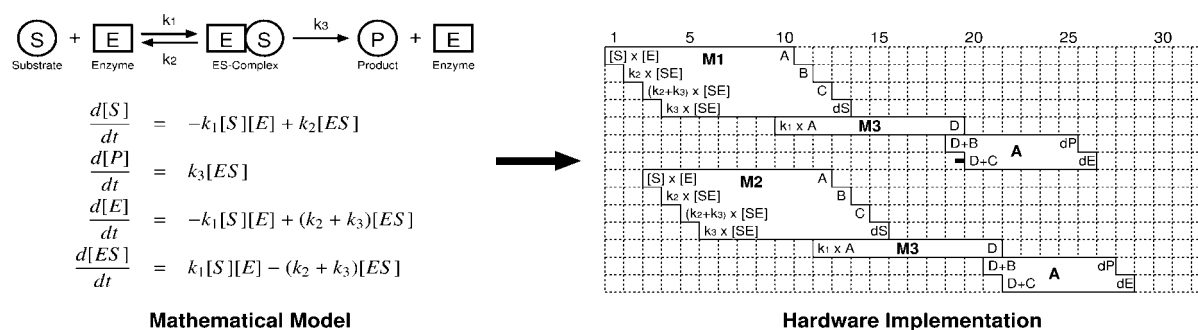
## Overview of the ReCSiP Board

ReCSiP ボードの中心となるのは、Xilinx の FPGA である Virtex-II (XC2V6000)です。その周辺には 1 セット 4MB の SRAM が 4 セット、64MB の SDRAM が 1 セットと、PCI インターフェイスとして動作する QuickLogic QL5064 などが接続されており、Virtex-II 上には演算回路とこれら周辺回路のコントローラが構成されます。

各メモリは最大 133MHz で動作し、複数のメモリをボード上に備えることにより、FPGA 上に構成した演算器が効率よく並列処理を行うことが可能になります。また、ボード上には拡張用のコネクタも備えており、測定器に直接接続してデータ収集を行ったり、あるいは外部出力のインターフェイスとして用いることも可能です。



## Simulating the Metabolic Network



細胞内で起こる各種の化学反応をモデル化し、細胞内に存在する物質ごとに濃度の時間変化を求めるのが代謝系シミュレーションです。

左図に示したように、反応は微分方程式などを用いて反応速度式の集合の形でモデル化されます。これらの一連の反応速度式をタイムステップごとに解いてやればシミュレーションができるわけです。

いくつかの代表的な反応モデルを、いくつかの代表的な数値解法によって解くハードウェアを実装し、必要に応じてそのうちのいくつかを FPGA 上で動作させることで、多くのシミュレータを高速化することが可能になります。