

# The CGIAR Global Cluster Grid of HPCs for Bioinformatics

Anthony McL. Collins,  
International Potato Center, Lima, Peru  
[a.collins@cgiar.org](mailto:a.collins@cgiar.org)

## Abstract

*The Consultative Group on International Agricultural Research (CGIAR) has implemented since 2005 a Grid of HPC Cluster systems at CIP in Lima (Peru), ICRISAT in Hyderabad (India), ILRI in Nairobi (Kenya) and IRRI in Manila (Phillipines) as an advanced global bioinformatics research platform supported by the Generation Challenge Program project (GCP). With remote systems support being coordinated by CIP in Lima, such a widespread global network presents many technical challenges in the area of network, systems and above all user management. Beyond the bioinformatics goals, lessons have also been learnt exploiting the complementary technologies of Internet2 and Access Grid video conferencing. The Grid of HPCs is now being expanded to include systems and users at CIMMYT in Mexico, with CIAT (Colombia) and ICARDA (Syria) pending.*

## 1. Introduction

### 1.1. CGIAR and the GCP project background

The CGIAR is a group of 15 agricultural centers with the shared mission “to achieve sustainable food security and reduce poverty in developing countries through scientific research and research-related activities in the fields of agriculture”. In 2004, the global Generation Challenge Program (GCP) project started “to use plant genetic diversity, advanced genomic science, and comparative biology to develop tools and technologies that help plant breeders in the developing world produce better crop varieties for resource-poor farmers” [B1].

## 1.2. GCP Bioinformatics research goals

One of the five sub-programs of the GCP is dedicated to advances in molecular biology and biotechnology, based on germplasm and crop information systems and bioinformatics resources. Many bioinformatics tasks are highly computing intensive, thus in 2005 a GCP Cluster Grid system was implemented, as an advanced tool at the service of all collaborating scientists and partners within the GCP:

[http://www.generationcp.org/sccv10/sccv10\\_upload/HPC\\_brochure.pdf](http://www.generationcp.org/sccv10/sccv10_upload/HPC_brochure.pdf).

## 2. The CGIAR global Cluster Grid

### 2.1. The Cluster Grid systems

The goal was to establish a Cluster Grid HPC facility for bioinformatics and similar high throughput scientific computing, based on the LINUX operating system, with future scalability to increase bioinformatics throughput through the addition of further nodes and/or compatible specialized accelerator hardware. The initial Cluster Grid system was created by purchasing small Paracel cluster systems for three sites, linked together with a 4th and much larger Paracel cluster at ILRI in Africa:

- Lima, Peru: CIP, 4-nodes
- Hyderabad, India: ICRISAT, 4-nodes
- Los Banos, Phillipines: IRRI, 8-nodes
- Nairobi, Kenya: ILRI, 33-nodes, plus a Genematcher Co-processor

Each center has installed a Paracel Linux cluster HPC based on AMD Opteron 64-bit processors to support both 32- and 64-bit applications. This enables programmers to seamlessly migrate applications to a 64-bit multi-

processing architecture to handle larger problems while preserving the functionality and investment of existing 32-bit applications evolving from PCs. All sites run the Rocks LINUX Cluster operating system: <http://www.rocksclusters.org>, configured with the Platform Computing LSF MultiCluster (Load Sharing Facility) program, to manage the global Cluster Grid system: <http://www.platform.com/Products/Platform.LSF.Family/home.htm>, the Ganglia Cluster toolkit, and the MPICH portable implementation of MPI

Exploiting the cross-platform power of LSF, a non-homogeneous system based on a cluster of similar Sun AMD Opteron nodes running SuSE LINUX has subsequently been incorporated:

- Texcoco, Mexico DF, 10-nodes

## 2.2. Internet2 and Access Grid

Meanwhile working in parallel supported via a separate project, fast Internet2 connectivity is being progressively extended to improve Grid performance. To date sites connected are

- Los Banos, Phillipines: IRRI, 155 mbits
- Lima, Peru: CIP, 2 mbits
- Nairobi, Kenya: ILRI, 512 kbits
- Texcoco, Mexico DF, 2 mbits

## 3. Bioinformatics research goals

### 3.1. Out of the box processing

All systems offer the highly optimized Paracel BLAST routine with the Paracel Bioview Workbench web interface:

<http://coe04.ucalgary.ca/bwb/>

Benchmarking of the free NCBI open source Blast routine shows similar performance on standard Intel Pentium 4 CPU desktop PCs as running (unoptimized) on a single Opteron CPU node of the HPC. In contrast the Paracel optimized Blast version runs on the 4 dual-Opteron CPU nodes at about 11X speed, which is a “bonus” of nearly 40% over the nominal 8X linear scaling speed of the 4 node Cluster (ie. 8 CPUs). This is outstanding performance given that even linear scaling is frequently not attainable on HPCs, refer Figure 1 [B2].



# Application: BLAST



## • Paracel BLAST performance

Dataset: Potato

Type	DNA sequences
Number of Sequences	32553
Total Length	26242629
Max Sequence Length	4968

Arabidopsis

Type	DNA sequences
Number of Sequences	45683
Total Length	55938335
Max Sequence Length	16011

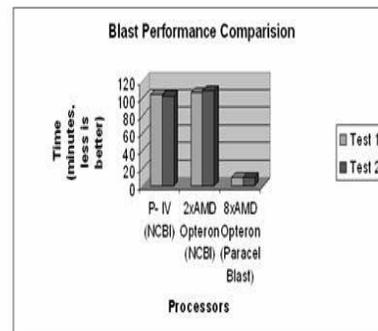


Figure 1. Paracel BLAST performance

## 3.2. Global HPC Community cluster grid programs

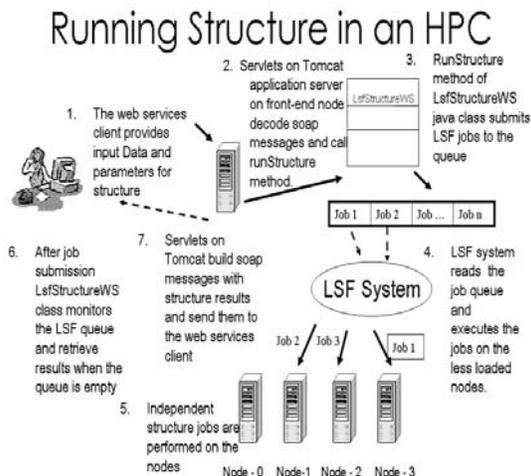
The global HPC community now offers a series of standard cluster-ready programs, including the “Rmpi” statistics package, Structure for molecular marker analysis, and the European Molecular Biology Open Software Suite (EMBOSS: [www.emboss.org](http://www.emboss.org)) for sequence analysis. The web services interface, Soaplab, is also being used to “wrap” a variety of genomics analysis tools, and includes all EMBOSS applications.

Soaplab is a tool that can automatically generate and deploy Web Services on top of existing command-line analysis programs. It can create two types of Web Services; *Analysis Service* and *Derived Analysis Service*. While the former allows users to send input data as weakly typed name-value pairs, the later has strongly typed methods for sending input data and receiving results. The weakly typed services allow integration of many applications within a single programming interface. Axis with Tomcat is also being used to publish Java class programs as web services.

Structure is software usually used as command line utility that must be run several times for population structure analysis. That makes this software an ideal candidate for parallel processing. CIP has achieved improved performance by putting jobs on a LSF Queue, LSF handling the parallel processing and load balance.

A client application has been developed using Webservices to send job requests and receive results. The whole stack advancing from the user to the HPC and back to the user is:

- User logon
- Client application (developed using eclipse IDE and Eclipse RCP framework)
- WSDL Webservices, developed in Java, deployed as Webservices using Axis libraries running on a Tomcat application server
- Java Webservices receive parameter values and data from the client application and put jobs on the LSF Queue
- LSF does load balancing
- Client polls the LSF queue, when it is empty (all jobs finished), it retrieves results
- User receives results, refer Figure 2.



**Figure 2. Running Structure**

### 3.3. CGIAR Custom HPC Community programming

Other programs with user-friendly web services interfaces are being locally developed at each site [B3]. The management node of each system includes the basic Linux LAMP program suite and the MPI compiler, with MPI libraries on all cluster nodes. A user interface has been developed for “Rmpi” using the Eclipse Workbench and Java, which are being used to develop rich clients under Service Oriented Architectures (SOA).

## 4. Technical challenges

### 4.1. Operations management – global user account security

The CGIAR maintains a global staff directory based on Microsoft Active Directory (MS AD) implemented on Windows 2003 servers at Palo Alto, USA. This MS AD manages the database of usernames and passwords for the CGIAR global windows network of over 8000 users. Using an LDAP connector and PAM modules developed at CIP, all HPC system user names and passwords are synchronized with the MS AD so that individual CGIAR users of different systems on the global grid retain the same credentials as on their desktop windows PC. This substantially simplifies management of the HPC systems security, as password use is automatically monitored. An auxiliary program, Password Station, enables registered MS AD users to personally manage the status of their network accounts by direct web access, and so to immediately recover from any routine lockout situations.

Thus, finally, all CGIAR staff using windows desktop PCs, LINUX workstations, or the HPC facility can retain the same user name and password on all systems globally. This relieves the LINUX systems administrators of tedious user account housekeeping, password maintenance responsibility, and simplifies life for users.

A global HPC systems consolidated job statistics processing tool is being developed at CIP to monitor usage and track users of all

systems. This tool has broad systems management objectives including

- Monitoring LSF Multicluster load sharing submitted job requests
- Monitoring actual the weekly processing load of all systems
- Reviewing the job profile on all systems
- Reviewing the user profile on all systems
- Providing operations management indicators for system update
- Monitoring for bottlenecks, failures etc.
- Monitoring LINUX systems log files for critical events

#### 4.2. Internet2 – to connect the CGIAR global Grid

To date Internet2 connectivity has been achieved by four of the five HPC sites, but in fact in Kenya the bandwidth is a nominal 512K via satellite, which is almost unusable in practice. Two of the three remaining sites have only recently attained stability with the IPv6 protocol and multicast functionality is now being tested. In summary, Internet2 has not yet significantly contributed to performance of the global grid, but it is of a great potential benefit, although still essentially experimental.

A background problem to date has been a lack of experienced technical support staff to integrate existing CGIAR campus networks using IPv4 via gateways to the Internet2 using IPv6, to provide transparent user access to the Internet2 network.

The primary goals envisaged by CGIAR for Internet2 connectivity are to

- Enable tight coupling of the LSF Multicluster systems into a global Grid for bioinformatics processing
- Maintain image databases and datasets for Grid processing, automated by replication services where possible
- Implement Access Grid video conferencing for global training seminars

#### 4.3. Access Grid video conferencing

*“Access Grid is a collection of resources and technologies that enables large format audio and video based collaboration between groups of people in different locations. The Access Grid is an ensemble of resources, including multimedia large-format displays, presentation and interactive environments, and interfaces with Grid computing middleware and Visualization environments. In simple terms, it is advanced videoconferencing using big displays and with multiple cameras at each node (site)”*: Wikipedia.

There have been technical problems maintaining Access Grid connectivity from Peru until the recent implementation of the IPv6 and multicast protocols. Nevertheless, we have run collaborative events interconnecting an IRRI multi-media conference room in the Phillipines, a CIP multi-media group workstation in Peru, and with CIMMYT “sitting in” on a PC window in Mexico.

With CIAT now on Internet2, we plan to launch 6-way advanced video conferencing seminars for bioinformatics global workgroups later in 2008, with the two non-Internet2 HPC sites “sitting in” on PC windows.

### 5. Lessons learnt

#### 5.1. Systems and user management

**PLUS:** exploitation of the existing CGIAR global MS AD user database, accessed via LDAP, has simplified systems management

**MINUS:** The initial HPC project budget was essentially limited to installation and local site management, thus it has been extremely difficult to create and maintain a global systems technical support team. There has been only minimal technical support training, limited to the most rudimentary level of support for the Paracel cluster systems hardware and software. On the one hand, this has caused overloading of the central HPC support staff in Lima, and on the other hand, it has impeded generating and

implementing global standards to promote use of the Grid.

## 5.2. Community of practice

**PLUS:** based on common and/or shared scientific goals, several specialized mini-communities of practice for HPC users have arisen.

**PLUS:** using Access Grid conferencing and collaboration tools via Internet2, interactive global seminars are now viable within the growing CGIAR bioinformatics HPC community and they underwrite virtual team building.

**MINUS:** lacking adequate resources for globally distributed technical support, HPC user growth has been slower than initially envisaged.

**MINUS:** insufficient funds were available for the training of scientists in LINUX and HPC cluster user tools, and the parallelizing of program algorithms.

**MINUS:** centers do not in general have adequate budgets to expand or replace their current mini-HPC systems as they start aging after 3 years. In reality, this period represents the basic lifespan of a generation of HPC hardware.

## 5.3. Achievements to date

CGIAR has nurtured a small but active global team of biometricians working at the leading edge of molecular biology and genomics research.

However, CGIAR does not have the budget resources to compete in the “Top 100” of global HPC systems, and must look towards accessing a central HPC service resource for production processing in the near future. Nevertheless, while the actual CGIAR Grid of dispersed and relatively small HPCs is beginning to age, the bioinformatics community of practice is maturing, encouraged by virtual team building tools such as Access Grid video conferencing.

Sometimes “easy achievements” are undervalued: a number of users have run non-optimized “PC” programs on just one HPC node, but as jobs running for extended periods like 2 weeks.

Although apparently trivializing the advanced cluster processing goals of HPCs, in the context of aiming towards advanced research, such background jobs would normally interfere with a user’s desktop PC foreground jobs, and indeed a two week job run may not survive on a PC without crashing. So permitting low-value use of HPC resources may nevertheless support high value results, noting that minimal additional support resources are required, unlike specialized programs such as MUSCLE, which in general are difficult to re-develop as parallelized versions for HPCs [B4].

## 6. Conclusions

### 6.1. The cost/benefit balance sheet

Due to location of the CGIAR HPCs at relatively isolated sites with limited resources to establish a strong and common global technical support team, user growth has been somewhat limited. Nevertheless, a number of specialized science teams have advanced to exploit state of the art bioinformatics tools and processing to form a community of practice, which would not have occurred without the availability of their local HPC systems. Thus, the systems have been instrumental in disseminating HPC technology at the service of scientific research in lesser developed parts of the world.

### 6.2. The future of Cluster Grid computing in CGIAR

As the user applications consolidate, and as the existing HPCs approach the time for renewal of equipment, it is clear that greater cost/benefit for high throughput processing could be achieved by a centralized service resource. However, the current small HPC sites are effectively acting, and will continue to act, as training and feeder sites for any such centralized, large-scale HPC facility in the future. In this way, the small HPC site utility will be prolonged by downplaying any requirement to maintain them as high powered distributed processing sites. The most likely future scenario for CGIAR is to evolve towards a service bureau approach for high throughput processing. This is the way to move beyond the

currently fragmented investment with relatively high support costs at remote sites, where it is difficult to generate a critical mass of activity. This is also in phase with the global movement promoting service oriented architectures for systems (SOA).

## **7. Bibliography**

[B1] Itoh, T., Tanaka, T., et al, "Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*", Genome Research, vol. 17, no. 2, pp. 175-183, Feb. 2007

[B2] Naruse, A., Nishinomiya, N., Kumon, K., Yamaguchi, M., "Hi-per BLAST: High Performance BLAST on PC Cluster System", Genome Informatics, vol. 13, pp. 254-255, Dec. 2002.

[B3] Rosenberg, N. A., "Distruct: a program for the graphical display of population structure", Molecular Ecology Notes, vol. 4, pp. 137-138, 2004.

[B4] Edgar, R. C., "MUSCLE: a multiple sequence alignment method with reduced time and space complexity", BMC Bioinformatics, vol. 5, pp. 113, 2004.