Apple in Research

Rajiv Pillai

Power of UNIX. Simplicity of Macintosh.

Mac OS X: The easy way to be open

Comand Line Interface

FreeBSD 5

Commands and Utilities

Editors

```
Terminal — emacs — 80x(16+8)
My-Computer:Scripts apple$ ls -v
About Scripts.txt       get-data.rb             restore.pl
Debug                   main.c                  restore.pl~
Public                  matchup.py              setup.sh
Test                    profile                 status.csh
counter.php             query                   trials.csv
debug.tcl               query_all               儷黑 Pro.ttf
My-Computer:Scripts apple$ emacs restore.pl
```

Scripting languages

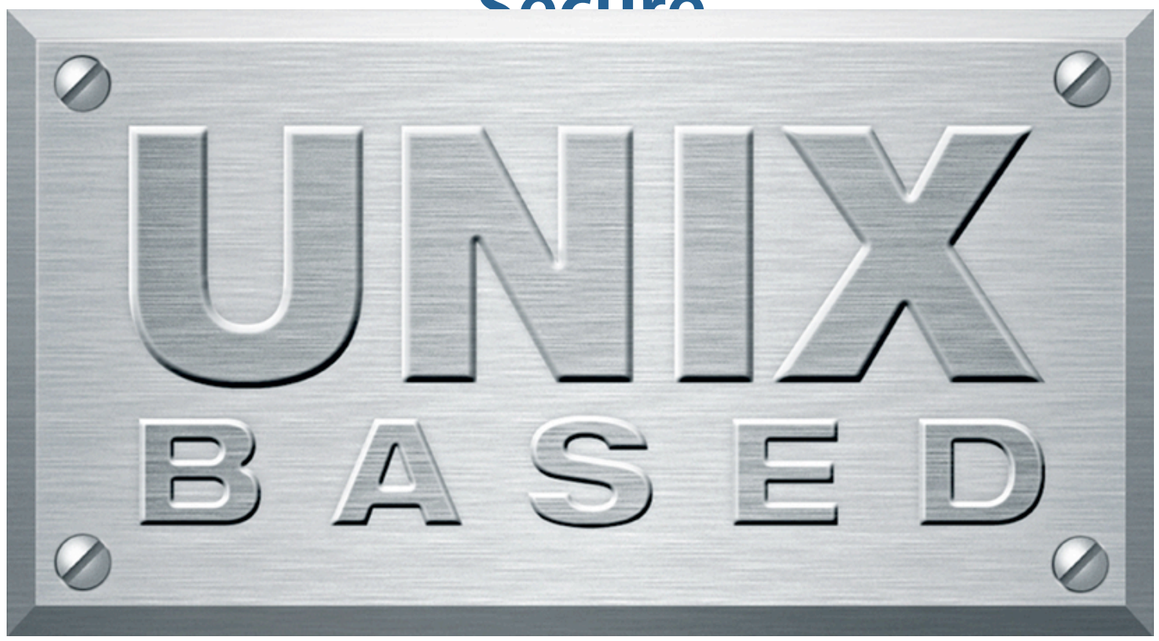
Shells

```
File Edit Options Buffers Tools Help
   my $valuesString = &buildValueString($ipAddress);

   my $sql = qq{ INSERT into condorcet_records
                 VALUES ( $valuesString ) };

   my $sth = $dbh->do($sql);

   #check whether the row was INSERTED
   if ( !$sth)
   {
      $retval = 'false';
   }

-----:**-F1  restore.pl        (Perl)--L346--82%----------------------------------
```

The Best Foundation

The Best Foundation

Secure



Advanced networking

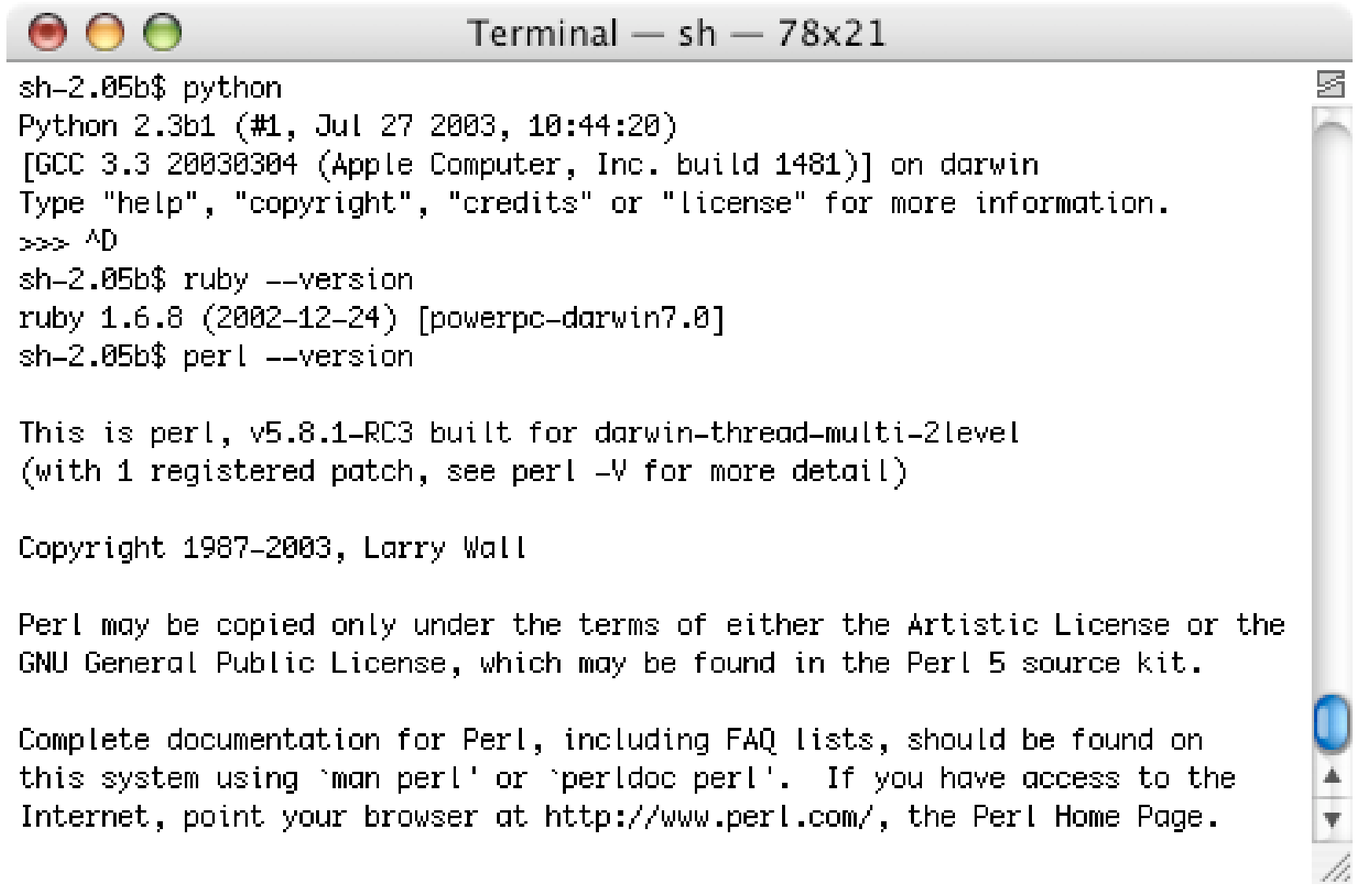
Built on Open Source

Over 100 Open Source Projects

FreeBSD	perl	python	php
Apache	KDE	MySQL	JBoss
samba	UNIX PRINTING SYSTEM	SquirrelMail WEBMAIL FOR NUTS	DELIVERED BY: MAILMAN MAIL LISTS
OpenLDAP PROJECT	Berkeley DB	Kerberos	OpenSSL

Modern Languages

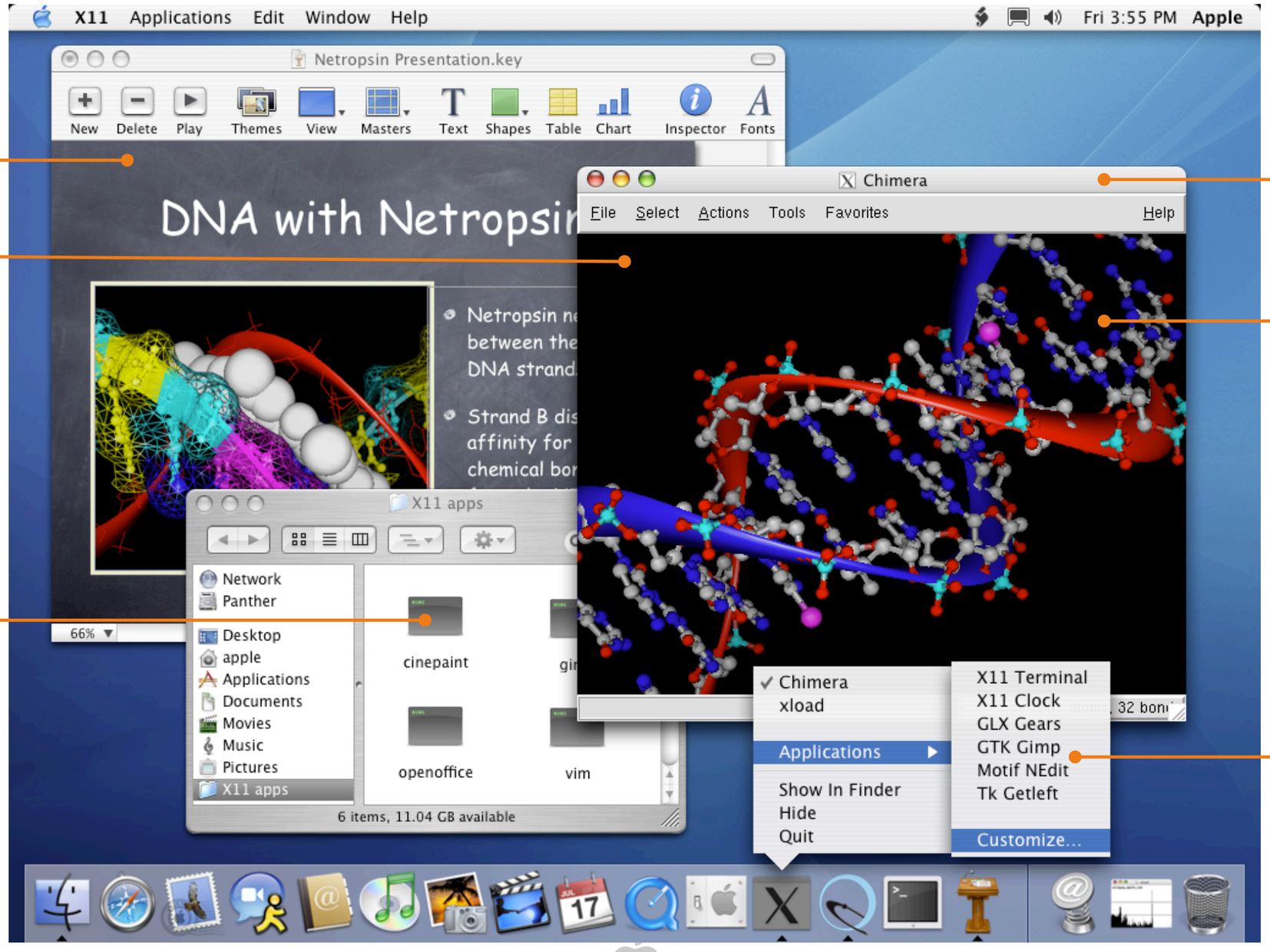
- GCC 3.3
- Perl 5.8.1
- Python 2.3
- PHP 4.3.2
- TCL 8.4.2
- Ruby 1.6.8
- Bash 2.05



Integrated X11

Runs side-by-side with native applications (or full screen)

Launch from Finder



Quartz window manager

Accelerated graphics

Dock menu

A Whole New World of Solutions

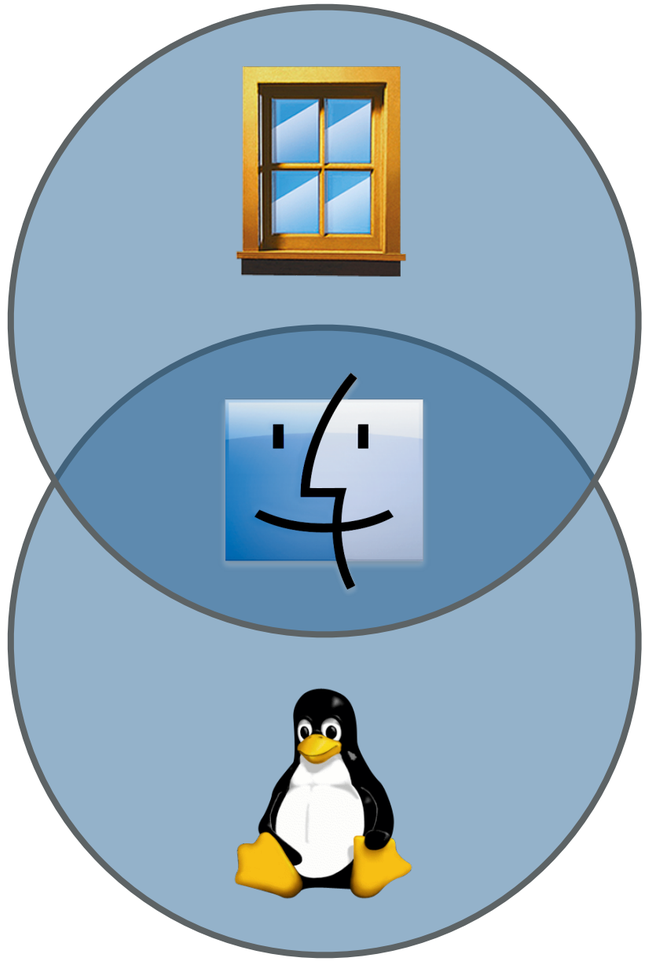
Bringing Mac OS X into new markets

Scientific	Distributed	Enterprise
Mathematica	Platform LSF	Oracle 10g
MATLAB	Globus	Sybase
BLAST	Sun Grid Engine	HP OpenView
HMMER	MPI	SAP client
GROMACS	PBS	SAS
GeneSpring	Myrinet	JBoss (J2EE)
PyMol	Infiniband	Tomcat (JSP)
IBM XL Fortran	iNquiry	Axis (SOAP)

Mac OS X

The Best of Both Worlds

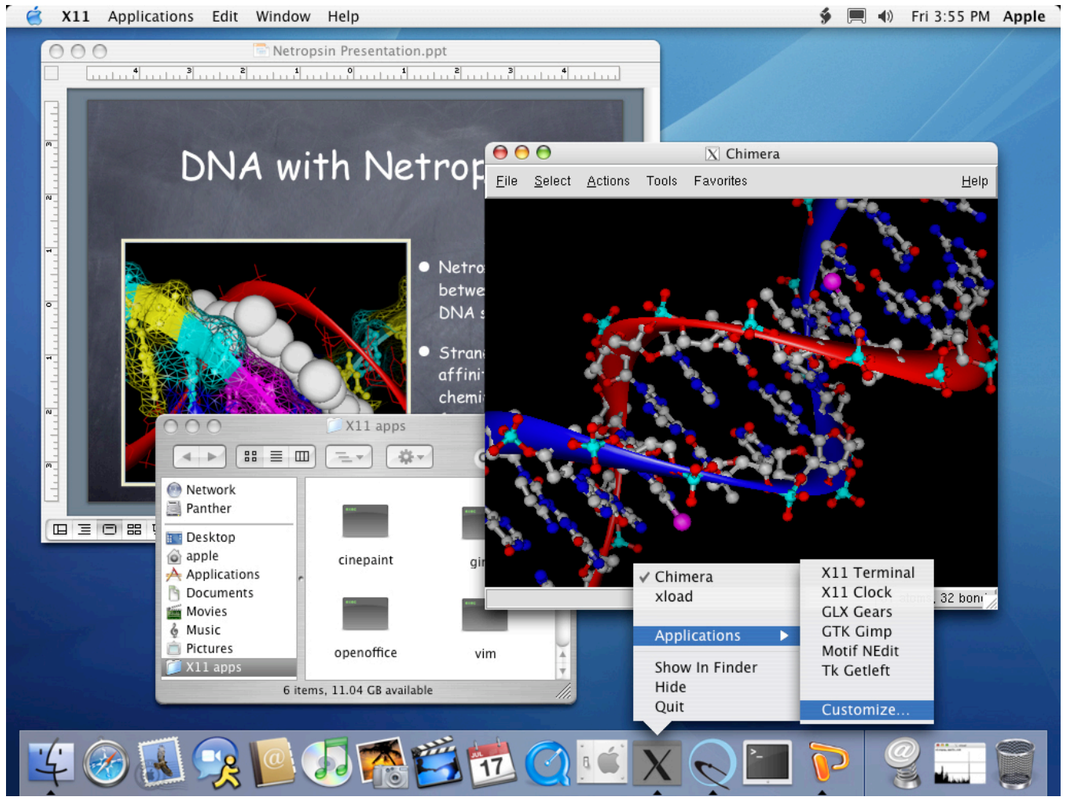
Open Like Linux	Convenient Like Windows
Open Source	Shrink-wrap solutions
Open Standards	Fits in to existing networks
Open APIs & Applications	Single point of support



Runs all the Apps a Scientist Needs

A single system on their desk

- Their favorite GUI applications (e.g. Mathematica, Gaussian, Vector NTI, TurboWorx, more)
- *And* their favorite UNIX applications (e.g. Phred/Phrap, HMMer, BLAST, Smith-Waterman, more)
- *And* their favorite productivity tools (e.g. Photoshop, Microsoft Office, Outlook email client)
- All run simultaneously on Mac OS X (Yes! Side by side)



Over 100 installations since March

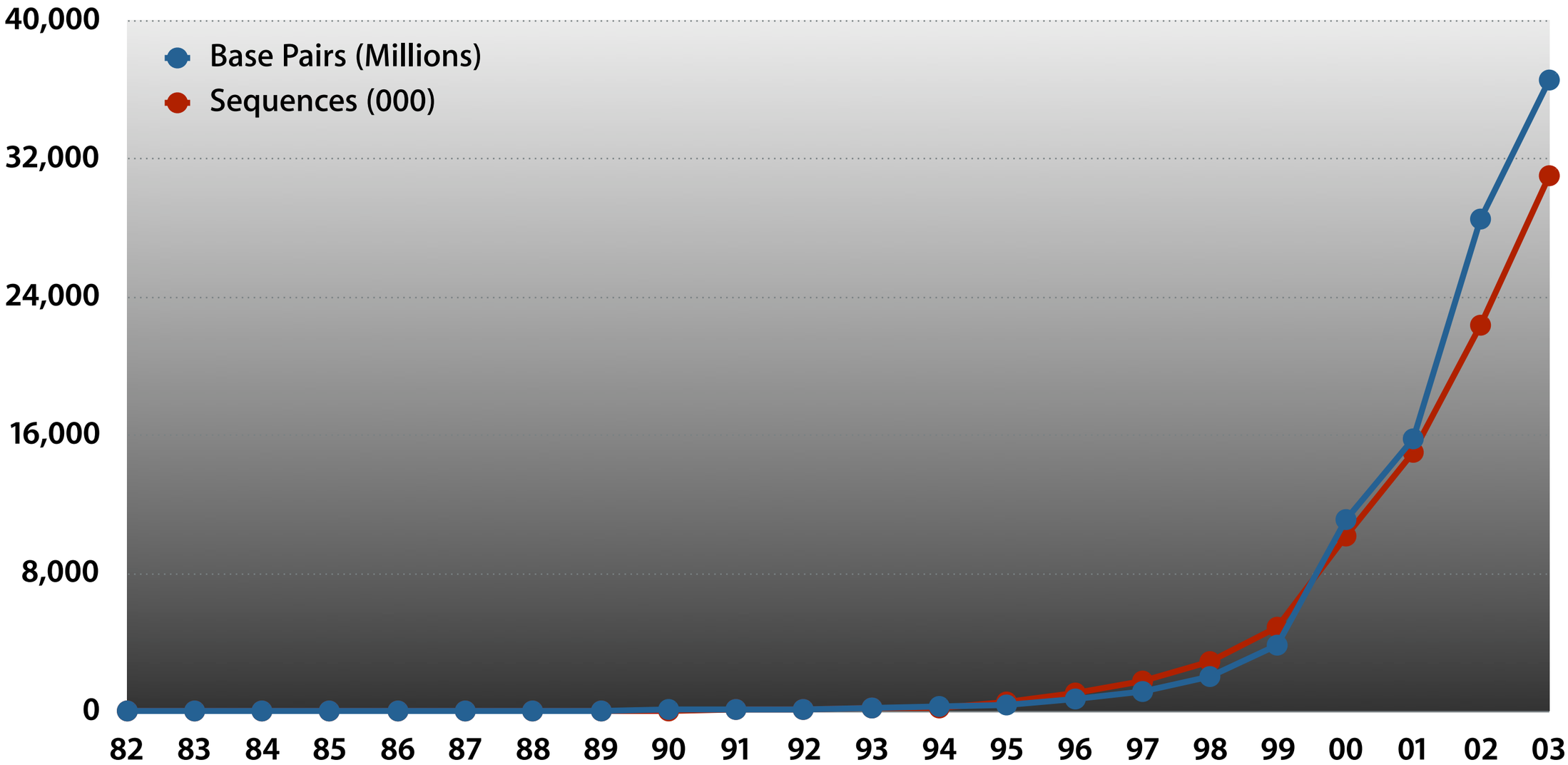
Academic	Government and Commercial	
Harvard University	Naval Medical Research Center	Isis Pharmaceuticals
Stanford University	Scripps Research Institute	Cincinnati Children's Hospital
Cornell University	Children's Mercy Hospital	U.S. Dept. of Agriculture
Yale University	Dana Farber Cancer Institute	Oakland Children's Hospital
McGill University	Beth Israel Cancer Center	Many private companies

Clustering made Easy

Why discuss Clusters ??

Exponential Growth of GenBank

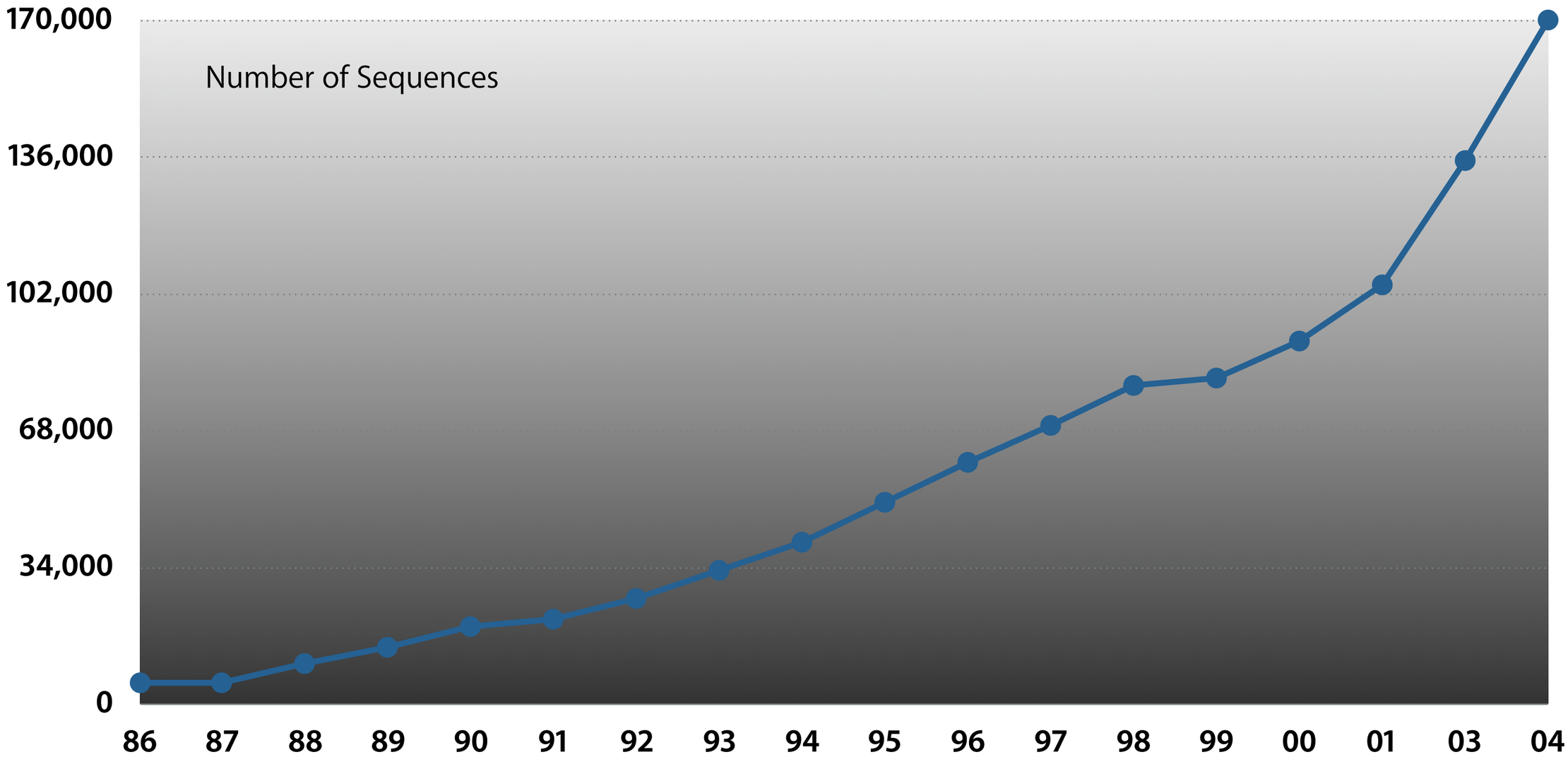
Researchers face a data explosion



Source: National Center for Biotechnology Information

Protein Database Growth

Swiss-Prot Database

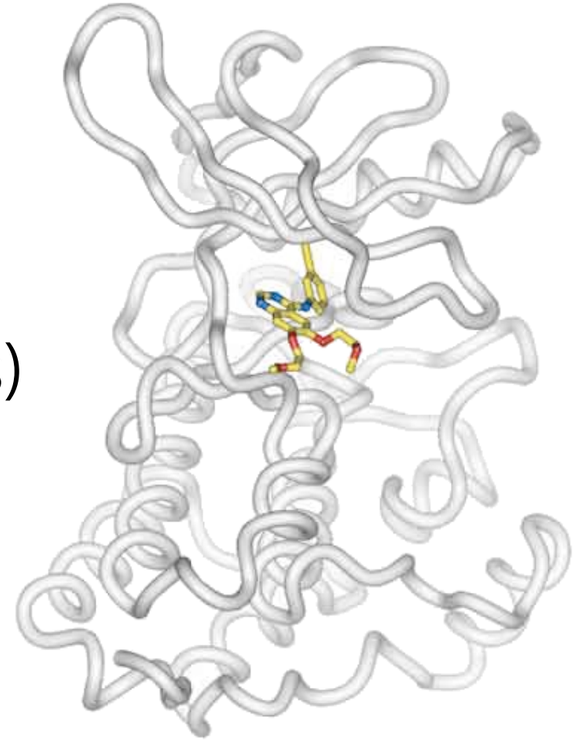


Source: SIB SWISSPROT EBI

Benefits of Clustering in Life Sciences

Immense data processing power

- Accelerates data processing
  - Maintain pace with quickly-growing genome and protein databases
- Faster turnaround time on large computations or searches
  - Reduced time to insight and discovery
  - Better depth of exploration for higher-quality results
- Enables difficult/time consuming computations
  - Large word size BLAST searches
  - Parametric studies (same problem, different parameters)
- Sharing resources allows a lower cost per gigaflop



Cluster Applications

Classes of applications common on clusters

- Bioinformatics
- Business intelligence/data warehousing
- Computational biology
- Computational chemistry
- Computational fluid dynamics
- Data analysis and visualization
- Numerics and statistics
- Simulation

The HPC Landscape is Changing

Smarter, more efficient approaches

- Use off-the-shelf systems whenever possible
  - Cheapest is not always best
  - Price-performance cannot be the only criterion
- Methods are becoming mainstream
  - Clusters are becoming critical components of the research workflow
  - Dependence yields higher standards
- Scientists are realizing there's no "free lunch"

Barriers to Clustering

Why aren't more researchers using their own clusters?

- Complex installation and management
  - Difficult to choose appropriate hardware
  - Complex software presents substantial technical barriers
  - Management resource requirements offset research benefits
- Barriers within the IT organization
  - Funding or logistical barriers to collocating the hardware
  - Lack of human resources to manage additional infrastructure
- Cost
  - Not just hardware, but also operating and maintenance costs
  - Indirectly associated cost - TIME !!

What makes clustering complex ??

Hardware Selection Considerations

- Computational requirements
  - RAM
  - Storage
  - Number of nodes
- Power
- Network
- Cooling
- Noise
- Setup and maintenance
- Security

“We are scientists, not IT professionals. And we wish to remain so.”

What we need

Ready to use clusters

Open source based solution

Low or Zero Maintenance

How does Apple help ?

Apple Workgroup Cluster for Bioinformatics

Apple Workgroup Cluster for Bioinformatics

The freedom to focus on your research

- Easy, do-it-yourself cluster setup
- All-inclusive turnkey solution
- Blazing speed of the G5 processor
- Powerful UNIX-based platform
- Ultraquiet enclosure
- More than 200 preconfigured bioinformatics applications
- Affordable, all-inclusive pricing
- Power- and space-efficient design

Web-Based Tool Access

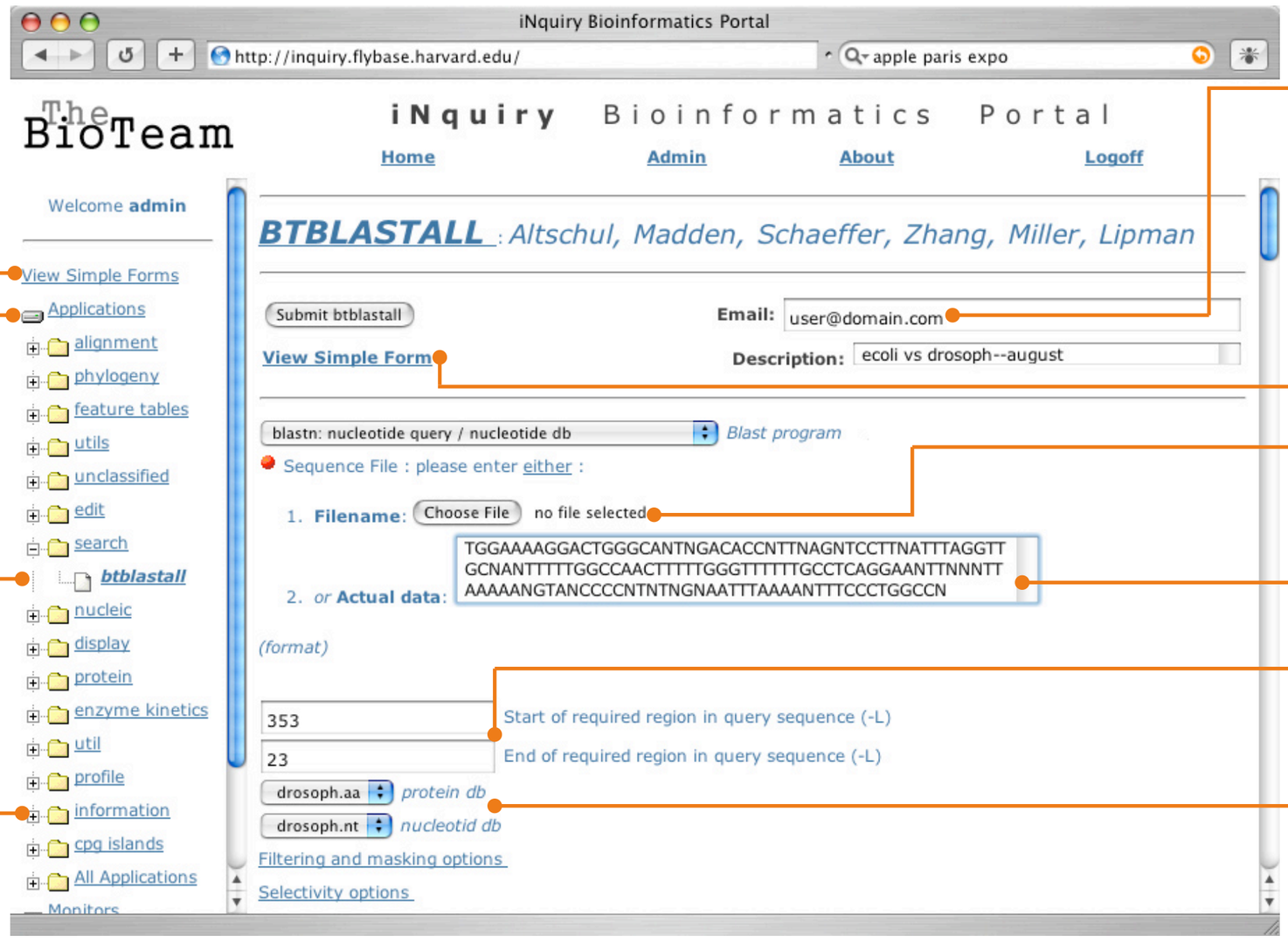
Extensible and modifiable

All tools have an advanced form and a simple form

Tools are organized in folders by function

The active tool is boldfaced

Applications can be viewed in a flat list



Results are emailed to user (and displayed in the web browser)

Quick access to simple forms for standard searches

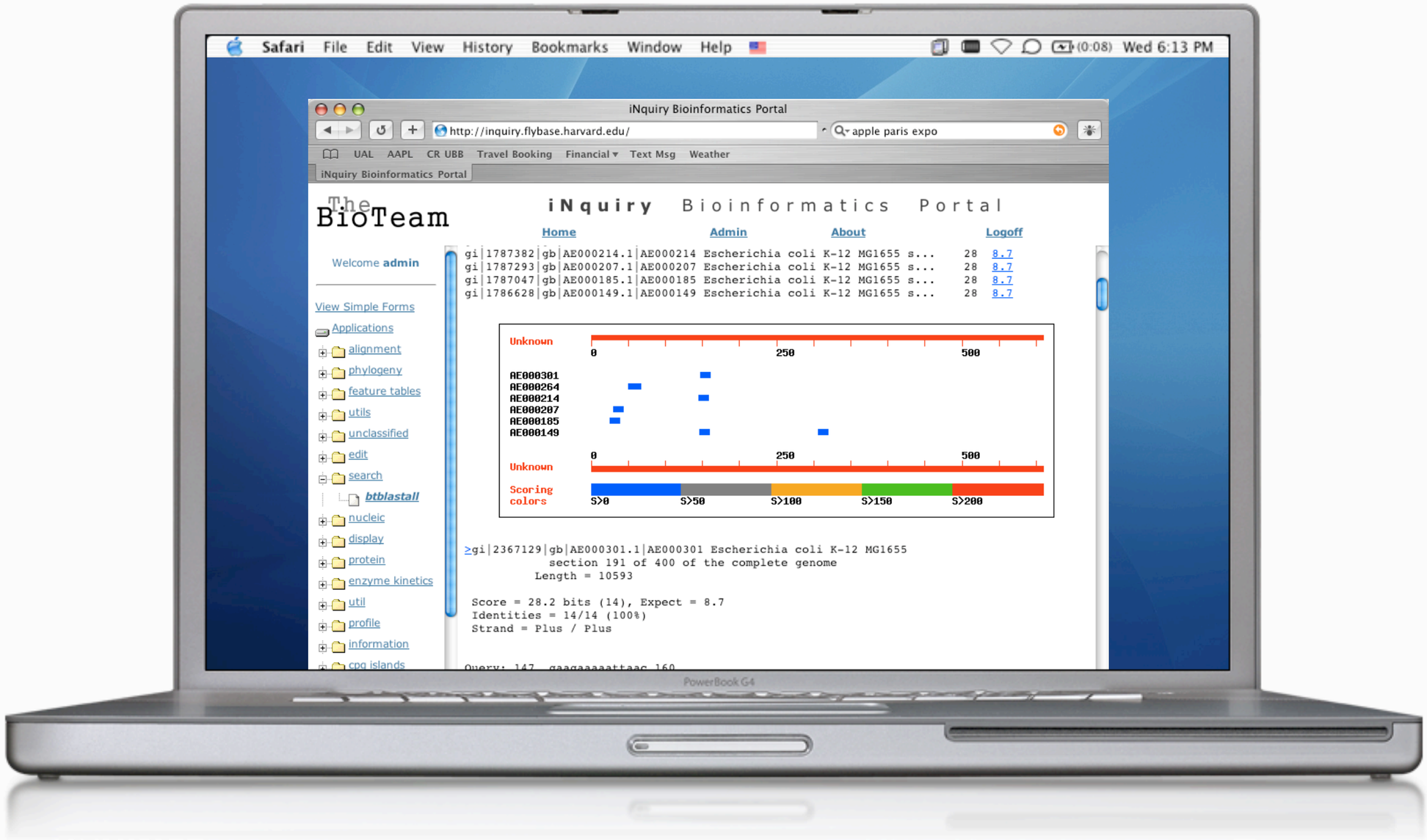
Data entry via file upload

Direct data entry

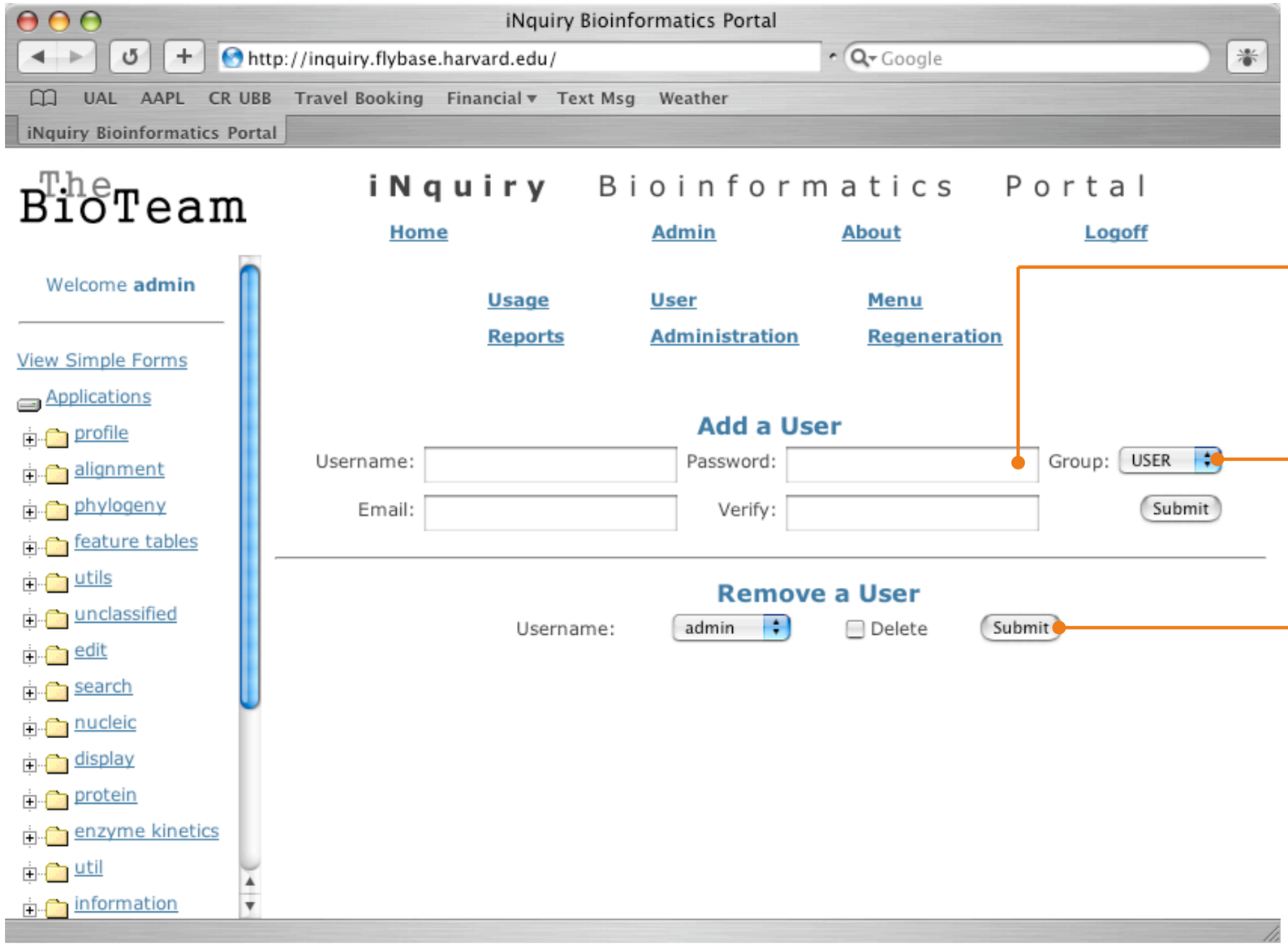
Quick access to installed databases

All command-line options are available in the web interface

Web-Based Results



Simple User Management



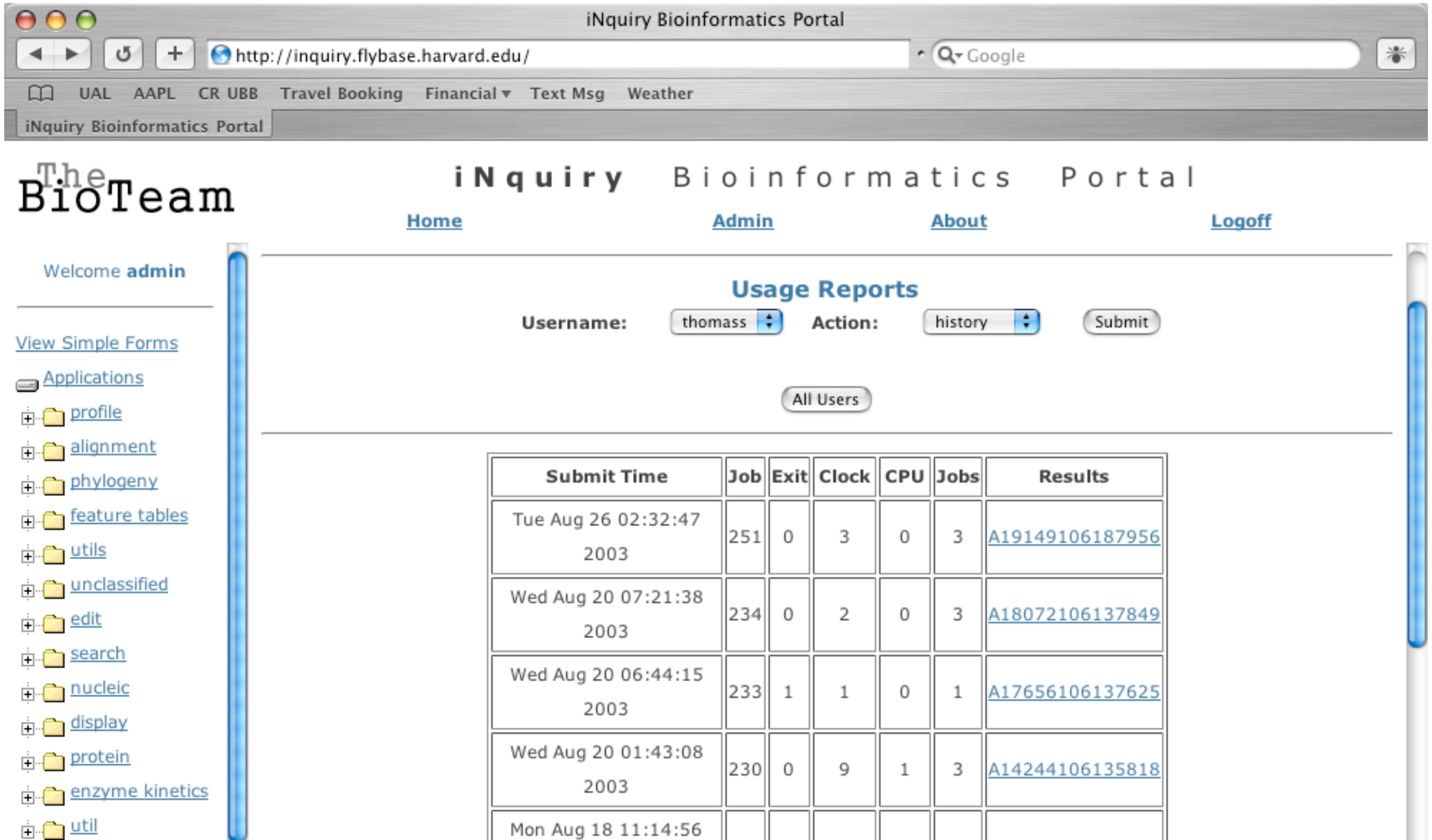
Minimal effort to create a new user

Choose between regular user and administrator

Single-click user disabling or deletion

Usage Reporting

Quickly see results and user statistics



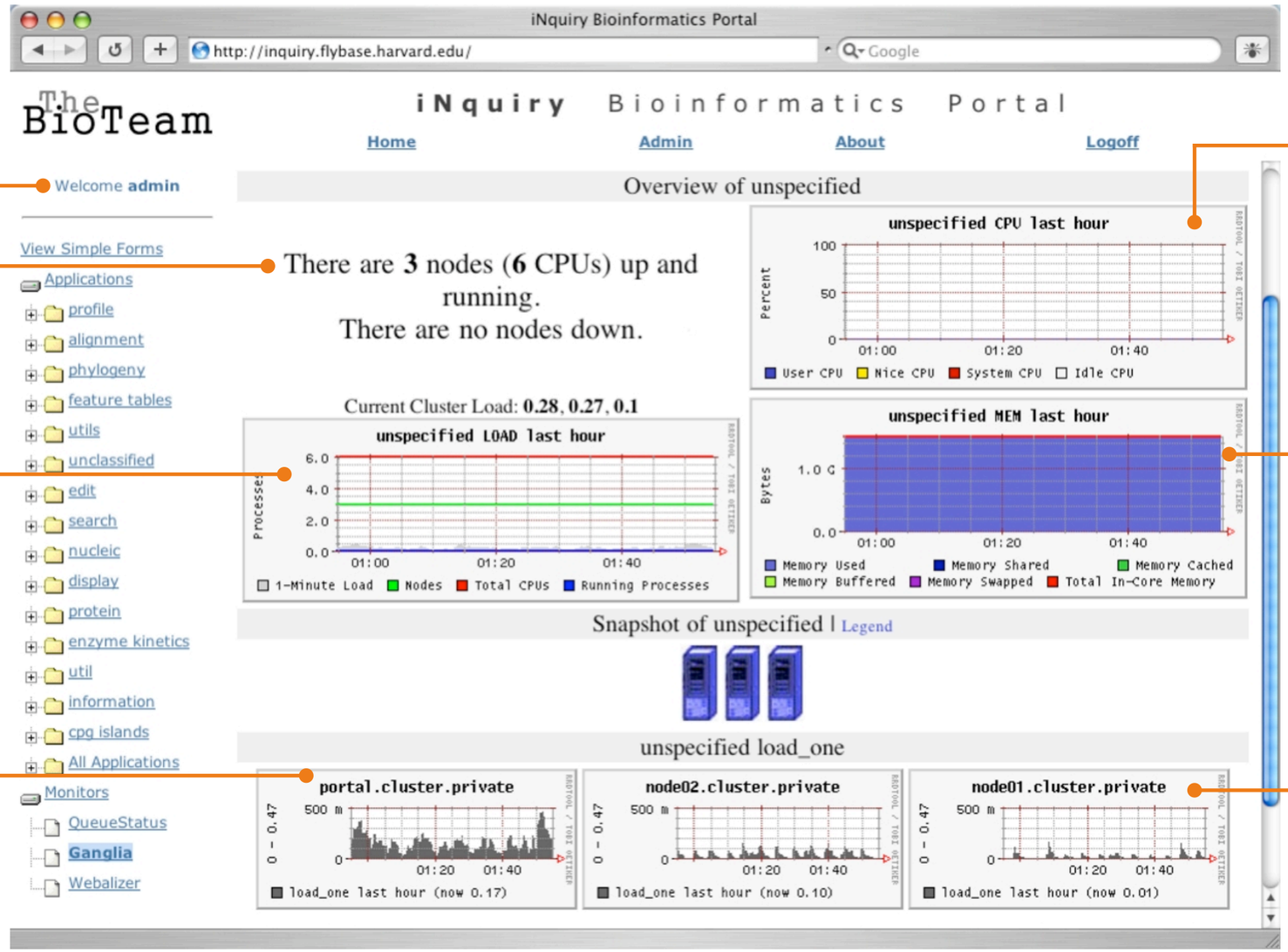
Cluster Monitoring

Only admin users have access to management tools

Instantly see if any cluster nodes are offline

See load statistics for the entire cluster

See load statistics for individual cluster nodes, including the portal (head) node



Histograms of CPU utilization by percentage

Histograms of memory usage by gigabyte

Load histograms for individual cluster nodes

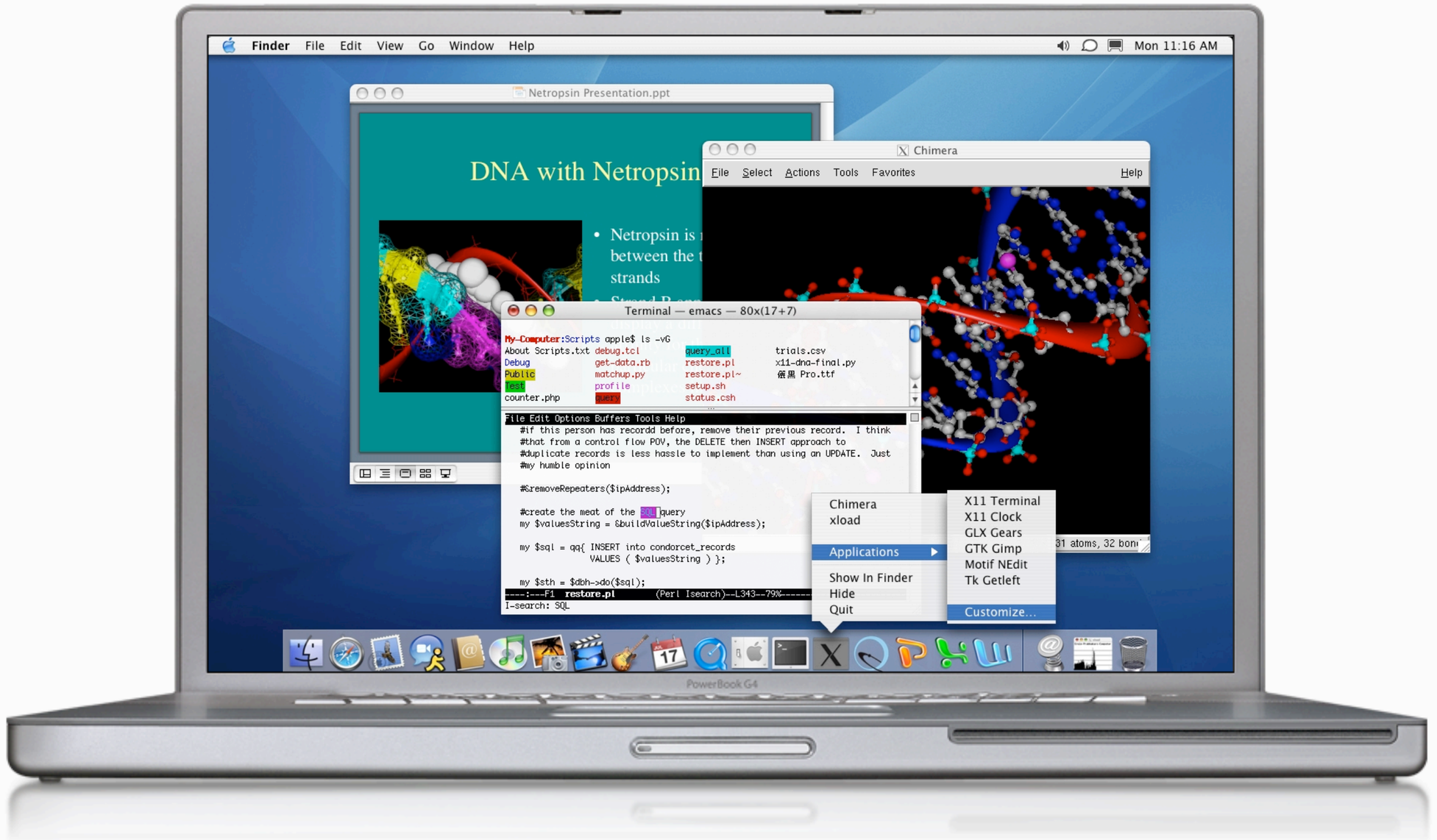
Key Points

- Simple setup process – up and running in 30 minutes
  - approximately two man-weeks of labor and/or $40k professional services of consulting
- Self-sufficiency – minimum IT dependency
- Access to UNIX applications via rich web GUI
- No compiling of tools – optimized versions already installed
- Keeps research data in-house – no outside resources required
- Reduced TCO by simplifying setup and minimizing administration required

Key Advantages

- Everything included – ***complete cluster solution***
- Easy to setup, easy to administer, easy to use – ***produce results!***
- Over 200 bioinformatics applications – ***optimized, ready to go***
- Xserve G5 performance – ***results fast!***
- Support included – ***no worries!***
- UPS, noise deadening rack – ***put it anywhere***

Mobile Super-Computing !!!



Xserve G5 in the Workgroup Cluster

Portal

Two G5 processors at 2.3GHz
Three drive bays for up to 1200GB storage
Up to 8GB PC3200 ECC DDR SDRAM

CD-RW/DVD drive
Mac OS X Server unlimited-client license

Compute Nodes

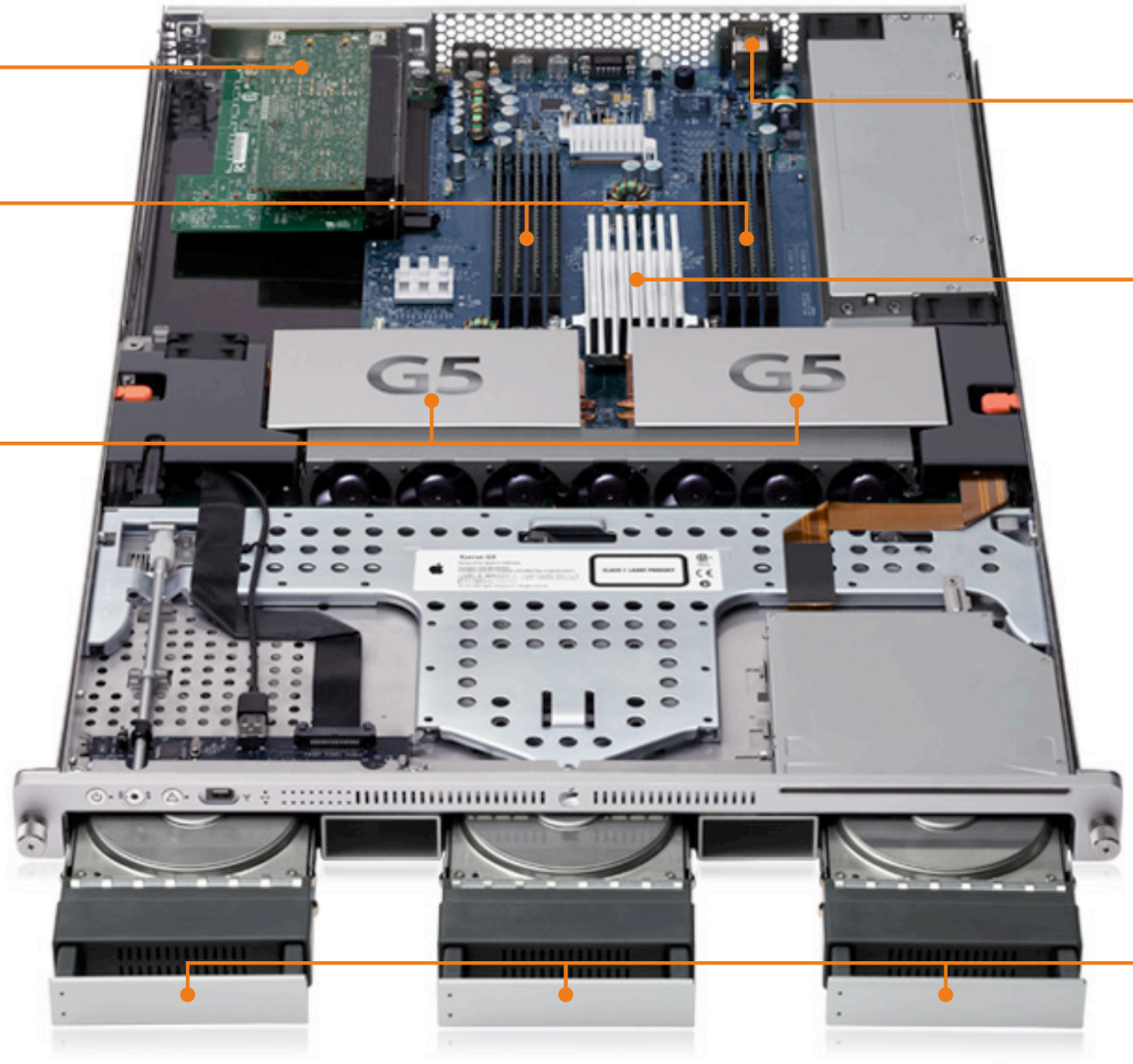
Two G5 processors at 2.3GHz
One drive bay with 80GB module
Up to 8GB PC3200 ECC DDR SDRAM

Breakthrough Performance in a 1U Server

Dual PCI-X slots at up to 133MHz

Up to 8GB ECC DDR main memory

Dual 2.3GHz PowerPC G5 processors



Dual onboard Gigabit Ethernet ports

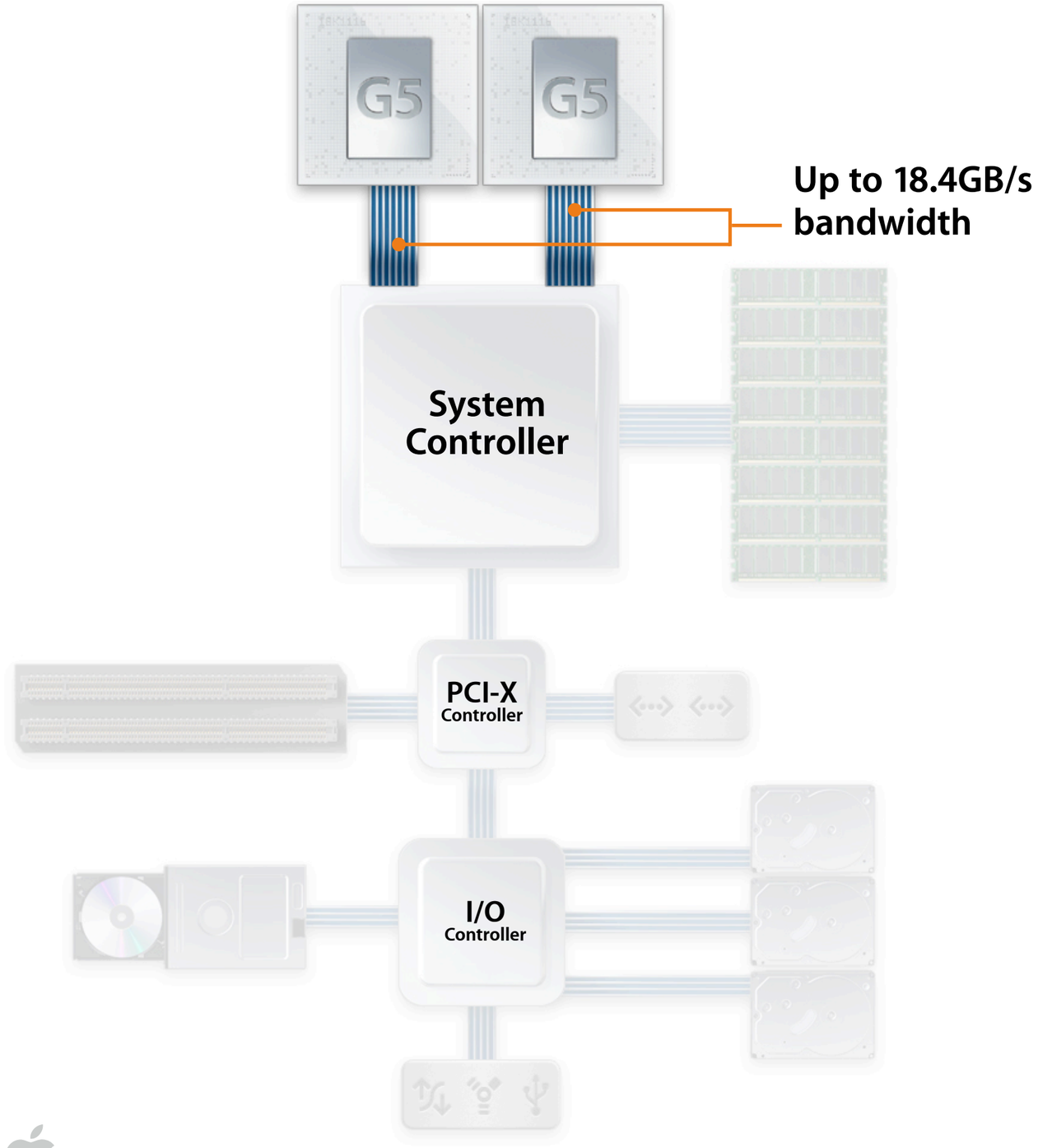
Advanced system controller with dual 1.15GHz system buses

Up to three 400GB drive modules—1.2TB in 1U

1.15GHz Frontside Bus

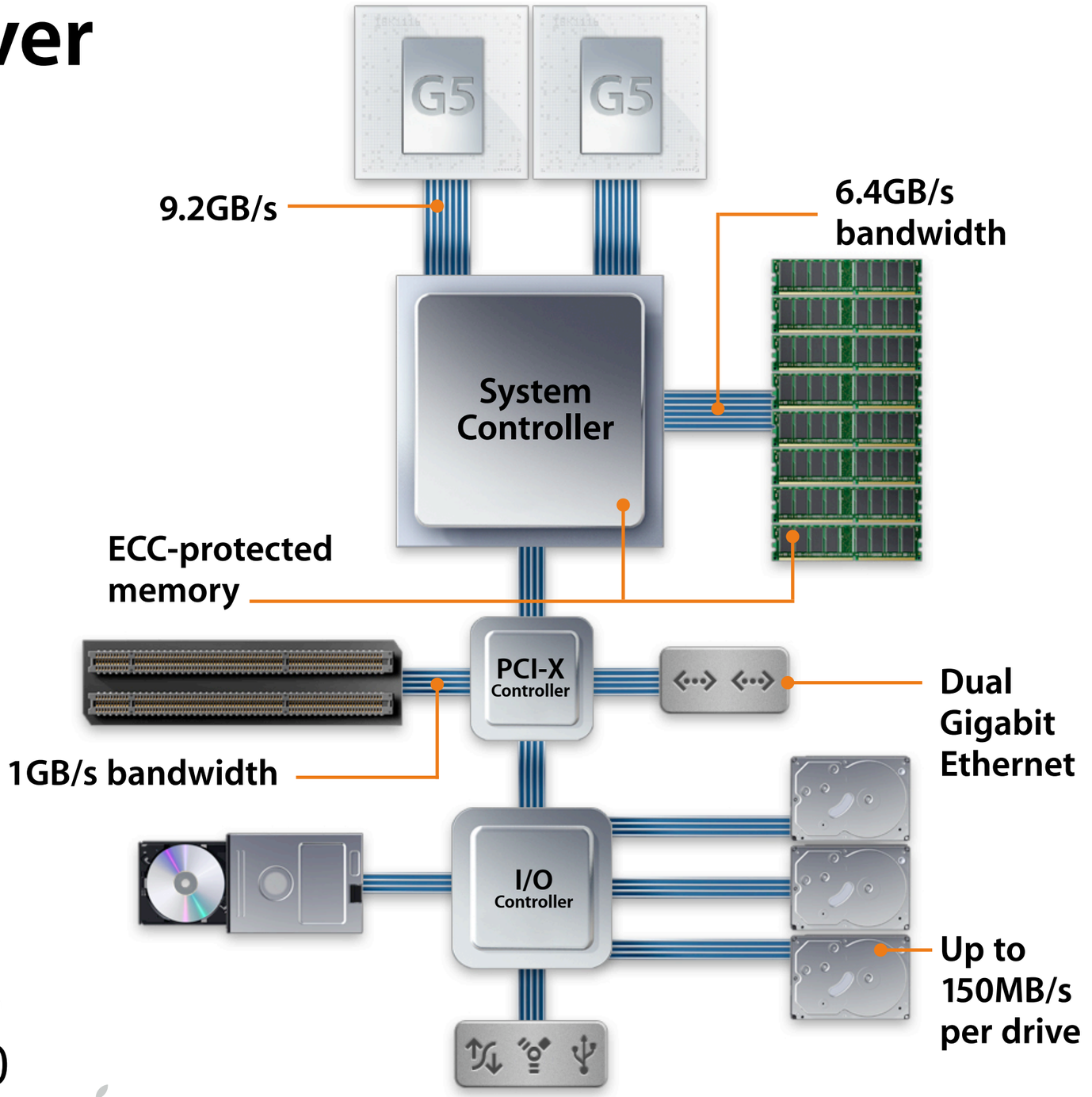
Fastest frontside bus in the industry

- 64-bit DDR bus for massive bandwidth up to 8GB/s
- Dual-channel bus for continuous flow of data
- Superior to standard bus, which negotiates direction
- Elastic interface self-tunes during startup for optimal signal quality



High-Bandwidth Server Architecture

- Apple-designed system controller
  - Dual 1.15GHz frontside buses with 18.4GB/s combined bandwidth
  - 128-bit PC3200 ECC DDR SDRAM with 6.4GB/s bandwidth
- Two PCI-X expansion slots
- Dual Gigabit Ethernet on dedicated internal PCI-X bus
- Three 150MB/s Serial ATA drive bays, FireWire 800, and USB 2.0



Dual 2.3GHz G5 Processing Power

Dual scalar floating-point units plus Velocity Engine deliver tremendous computational capability

Double-Precision Floating Point

10.35

Gigaflops

Single-Precision Floating Point

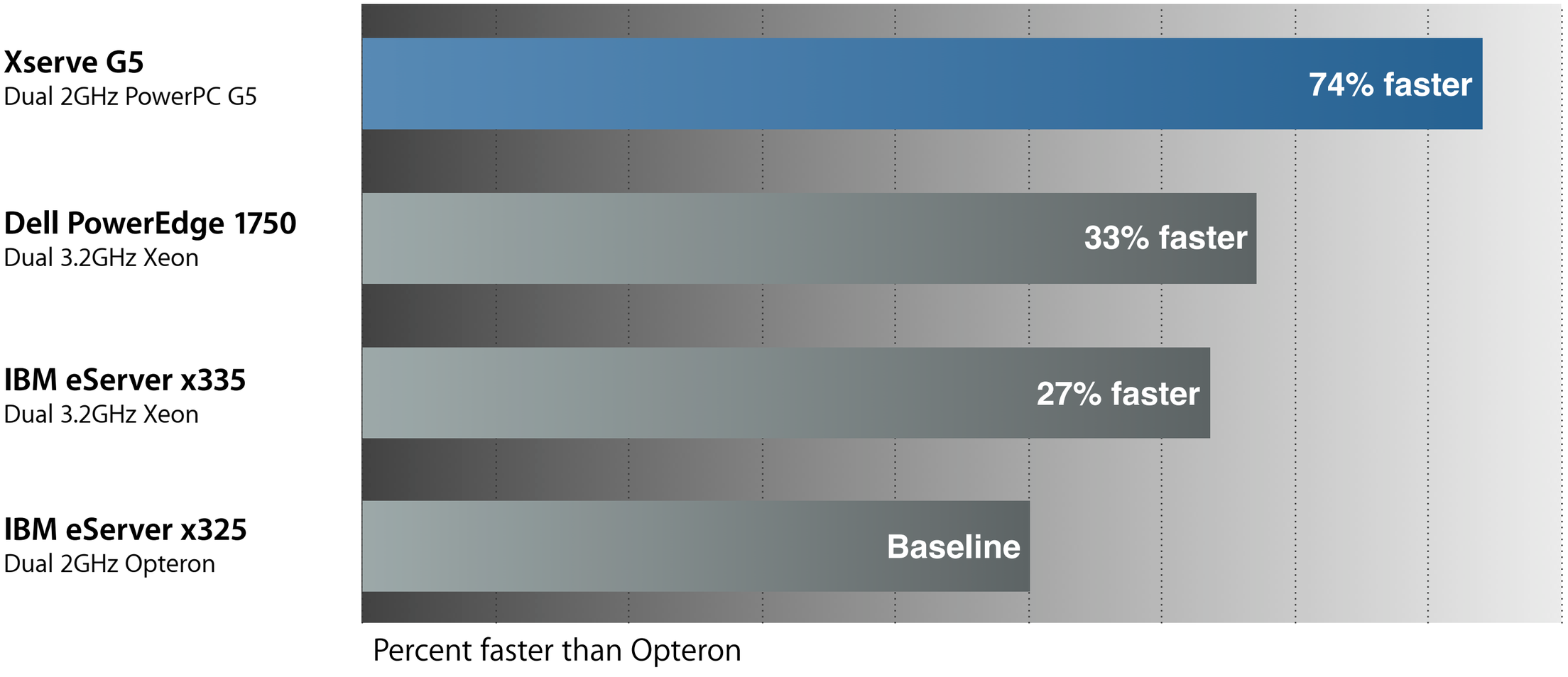
35.22

Gigaflops

*Testing conducted by Apple in December 2004 using preproduction Xserve G5 systems. Double-precision floating-point numbers based on the LINPACK benchmark (www.netlib.org/benchmark/1000d). Single-precision floating-point numbers based on FFT benchmark from Apple.

HMMer Results

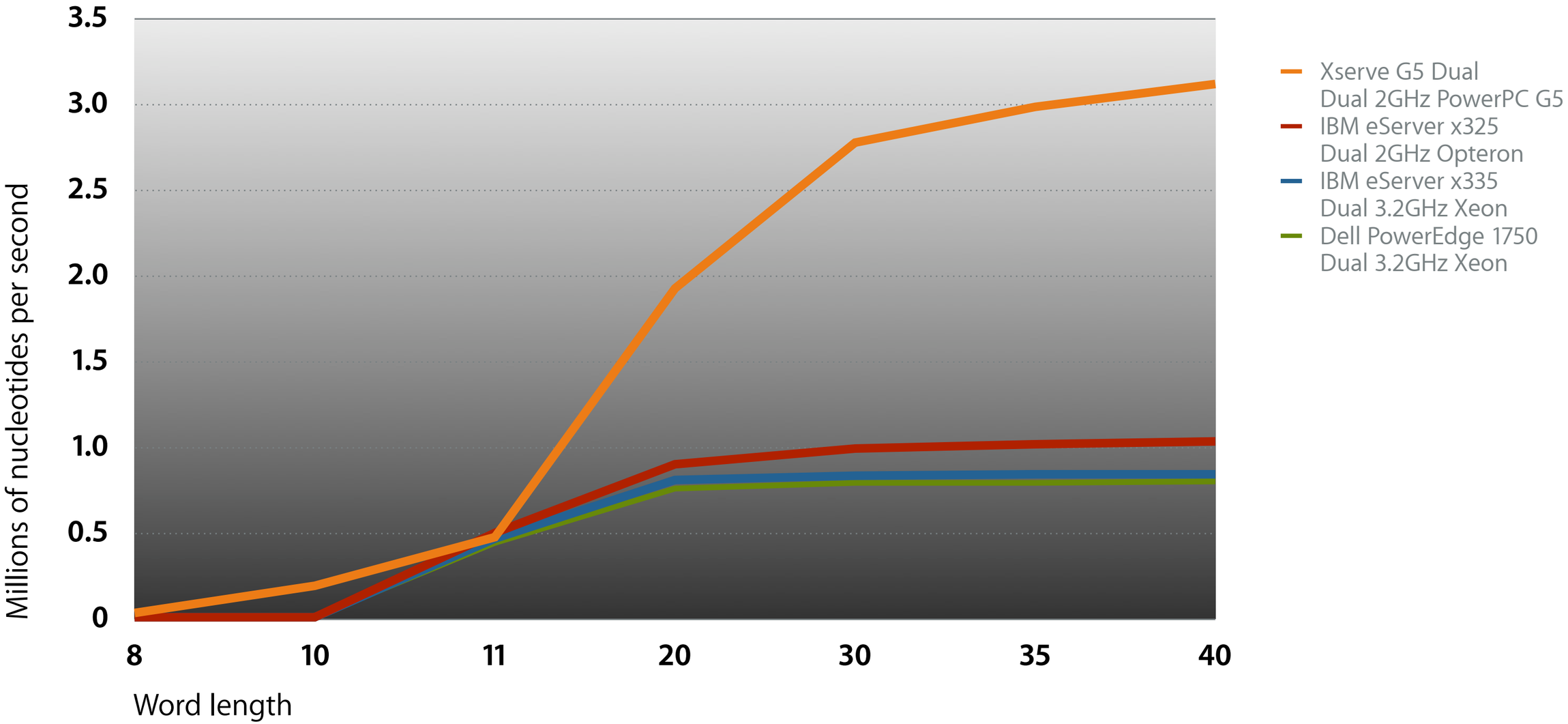
Genome sequence matching



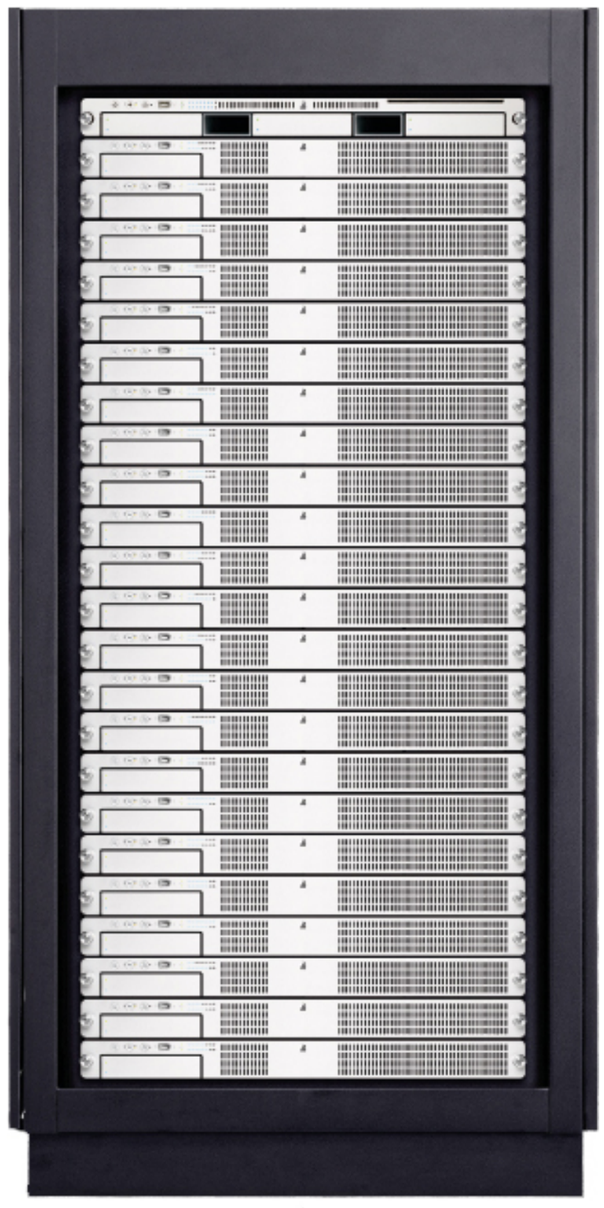
Testing conducted by Apple in January 2004 using preproduction dual 2GHz Xserve G5 units; all other systems tested were shipping units. All Xserve G5 systems were tested using Mac OS X Server v10.3.2. The IBM eServer x335 and Dell PowerEdge 1750 tested using Red Hat Linux 9.0; IBM eServer x325 used Red Hat Enterprise Linux 3.

BLAST Results

DNA sequence matching



Testing conducted by Apple in January 2004 using preproduction dual 2GHz Xserve G5 units; all other systems tested were shipping units.
All Xserve G5 systems were tested using Mac OS X Server v10.3.2 and A/G BLAST. The IBM eServer x335 and Dell PowerEdge 1750 ran NCBI BLAST on Red Hat Linux 9.0.
IBM eServer x325 ran NCBI BLAST on Red Hat Enterprise Linux 3.

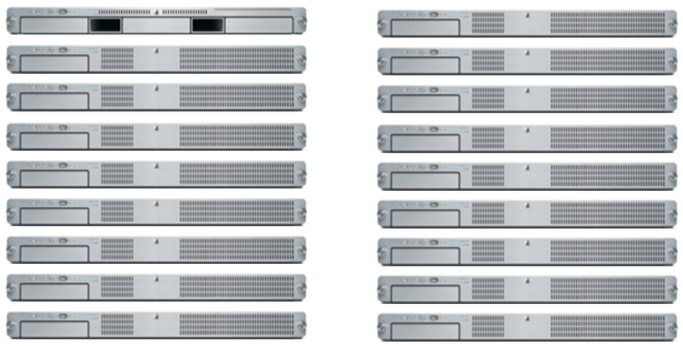


Making your customized cluster

What is a Cluster?

Varied components with a common goal

1. Some number of independently operable computers



Xserve G5 in the Workgroup Cluster

Portal

Two G5 processors at 2.3GHz
Three drive bays for up to 1200GB storage
Up to 8GB PC3200 ECC DDR SDRAM

CD-RW/DVD drive
Mac OS X Server unlimited-client license

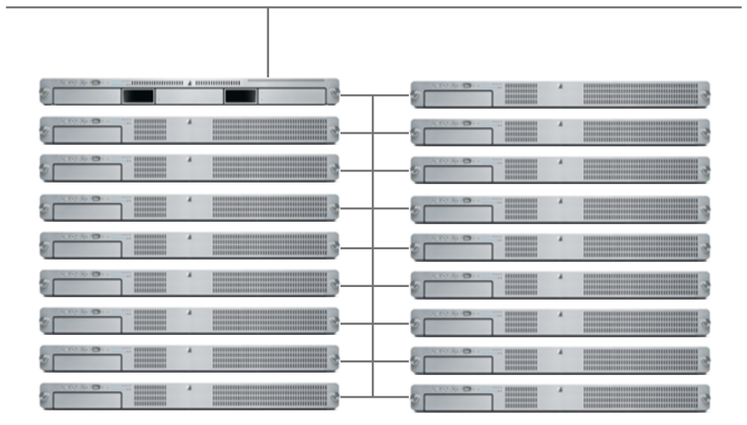
Compute Nodes

Two G5 processors at 2.3GHz
One drive bay with 80GB module
Up to 8GB PC3200 ECC DDR SDRAM

What is a Cluster?

Varied components with a common goal

1. Some number of independently operable computers
2. Interconnection network



Interconnects

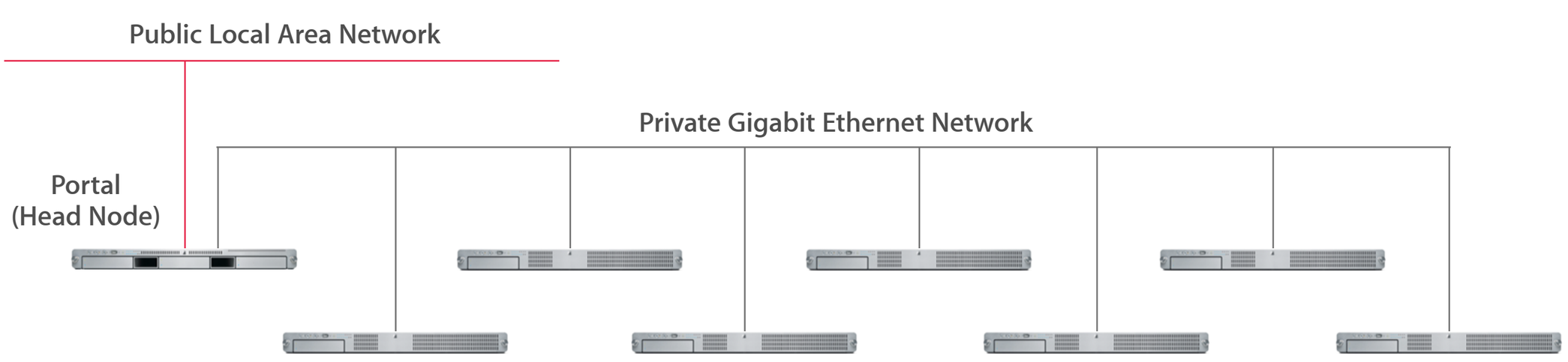
Choosing the Right Interconnect

Let the application drive the selection process

- Type of problem
  - Embarrassingly parallel or tightly coupled
- Type of cluster
  - High performance or high throughput
  - Faster interconnect or more nodes
  - Evaluation of the algorithm should determine the choice
- Interaction of cluster software stack with interconnect
  - OS/DRM/MPI/compilers/libraries/developer tools
- Don’t hesitate to ask for help or look for consultative expertise

Loosely Coupled Cluster

- Some number of computers on a private Ethernet network
- TCP/IP routes communications between nodes
- A ‘portal’ acts as the primary point of contact for people using the cluster
- Example: The Apple Workgroup Cluster for Bioinformatics

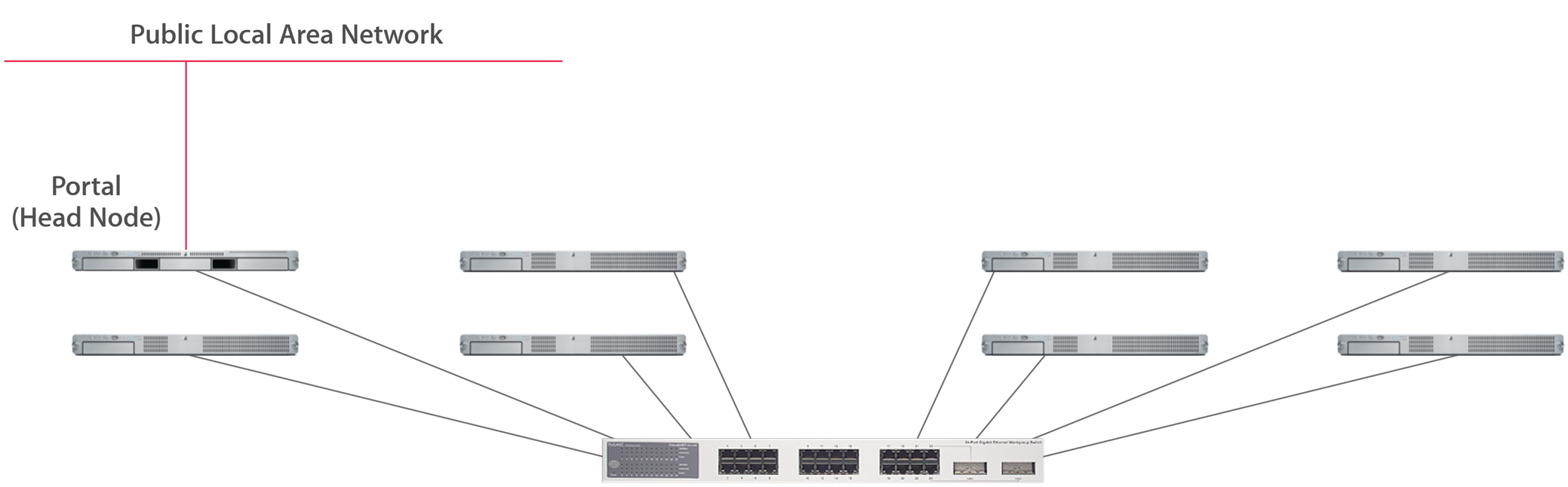


Gigabit Ethernet

Tightly Coupled Cluster

Uses low latency, high bandwidth interconnect

- Some number of computers connected by Myrinet or InfiniBand
- Proprietary protocol handles message passing, may implement MPI directly
- Required by 'tightly coupled' applications such as AMBER
- A 'portal' acts as the primary point of contact for people using the cluster



Myrinet (Myricom)

InfiniBand

High-performance Interconnects

- Myrinet
  - Very well tested and highly reliable
  - Mature driver stack
- Infiniband
  - Open standard
  - Prices expected to drop precipitously

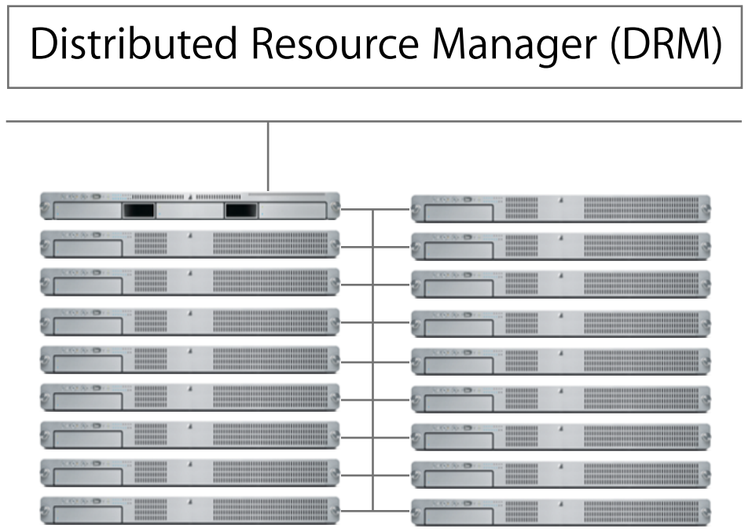
Comparing Interconnects

	GbE	Myrinet	Voltaire InfiniBand
Bandwidth	1 Gb	2.5 Gb	10 Gb
Latency	50 uSec	6 uSec	5 uSec
Linpack Network Efficiency	50%	60%	80%
Storage Connectivity	✗	✗	✓

What is a Cluster?

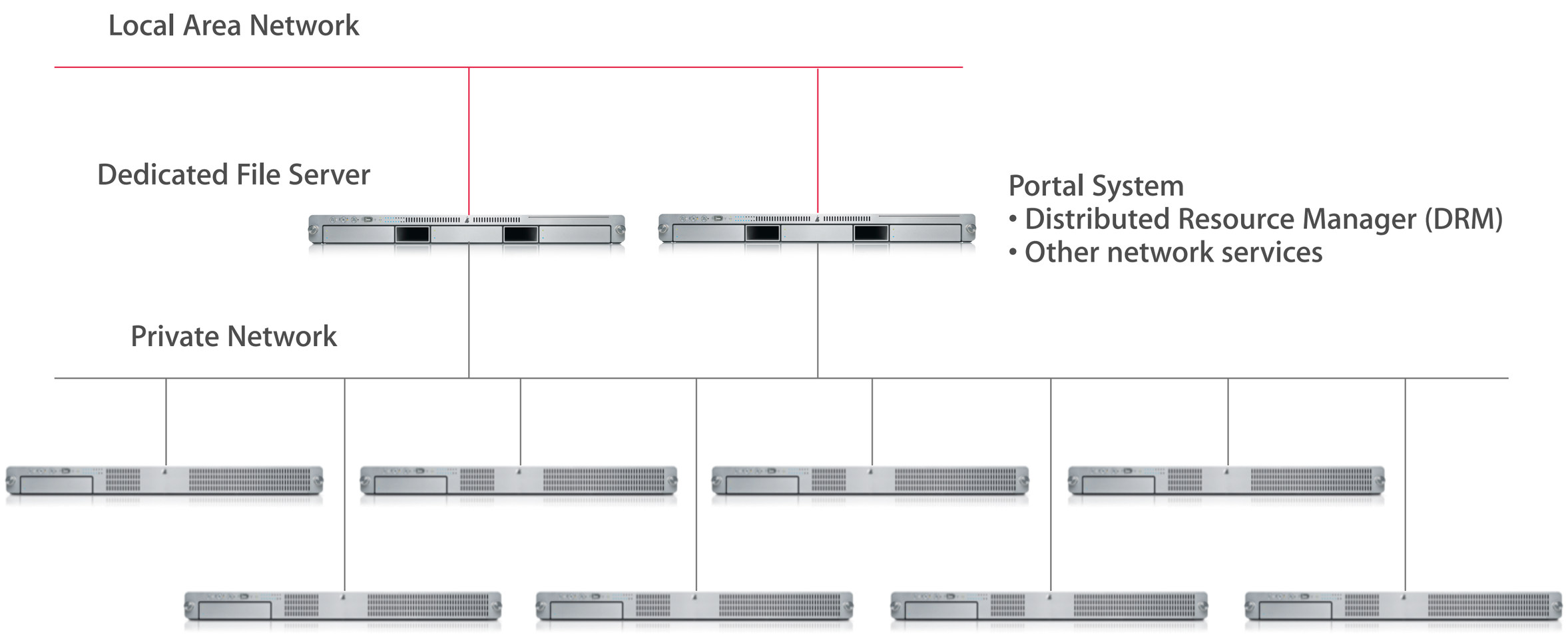
Varied components with a common goal

1. Some number of independently operable computers
2. High performance interconnection network
3. Software for controlling concurrent tasks (DRM)



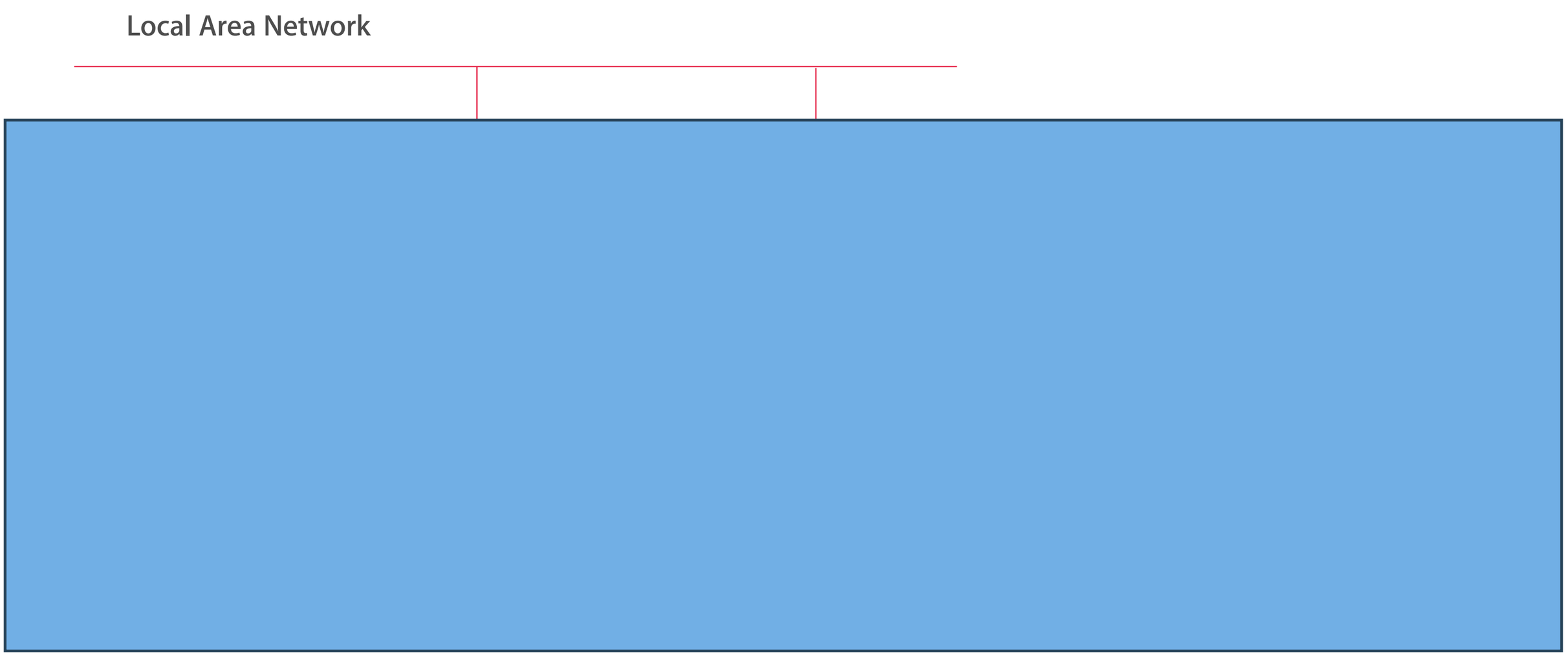
Distributed Resource Management Software

Turns a collection of cluster components...



Distributed Resource Management Software

Into a single virtual resource...



Distributed Resource Management Software

Products available for Mac OS X

Platform	**Platform LSF**	Powerful and flexible workload management software for deploying high-performance computing solutions.
GRID ENGINE	**Sun Grid Engine**	Sun's open source DRM software for deploying powerful compute farm solutions.
	**OpenPBS**	The Portable Batch System (PBS) is a flexible open source batch queueing and workload management system originally developed for NASA.
	**PBS Pro by Altair**	Commercial version of PBS goes beyond OpenPBS with additional features, such as enhanced fault tolerance, reliability and superior cluster support.
INQUIRY	**iNquiry**	Created by The BioTeam, this fully provisioned informatics cluster solution is installed on Xserve G5 systems in minutes from an Apple iPod.
	**Xgrid**	Apple's grid tool that supports both dedicated clusters and idle cycle utilization deployment. Supports basic batch scheduling.

Distributed Resource Management

Queueing and Scheduling

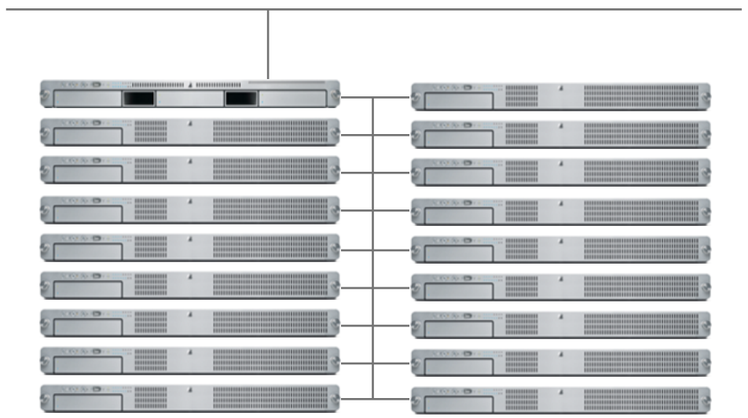
- Xgrid
- Sun Grid Engine
- OpenPBS and PBS Pro
- Pooch
- Platform LSF
- MAUI
- Condor

What is a Cluster?

Varied components with a common goal

Message Passing Interface (MPI)

Distributed Resource Manager (DRM)



1. Some number of independently operable computers
2. High performance interconnection network
3. Software for controlling concurrent tasks (DRM)
4. API for parallel application development (MPI)

Message Passing Interface (MPI)

MPI Basics

Enables parallel computing by allowing an app running across a cluster to pass results between nodes

- Has become the de-facto standard cluster savvy API
- Generally implemented as a library for use by developers
- Supports simple or complex application development schemes
- Provides access to parallel (cluster, grid, other) hardware for:
  - End users
  - Library writers
  - Tool developers

MPI and Other Parallel Computing Tools

Products available for Mac OS X

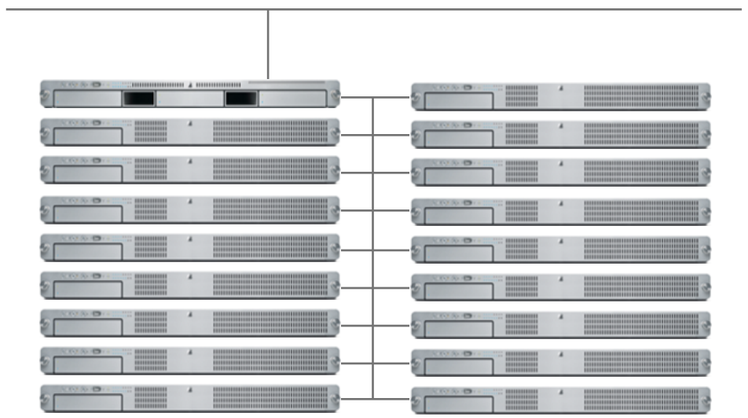
	**MPICH**	Open source, portable implementation of MPI, the standard for message-passing libraries.
MPI/Pro	**MPI Pro by MPI Software Technology**	Commercial MPI solution with performance and feature enhancements.
	**LAM/MPI**	Local Area Multicomputer (LAM) MPI programming environment and development system for heterogeneous computers on a network.
Linda Paradise	**Linda and Paradise by Scientific Computing Associates**	Proven and flexible tools for parallel and distributed computing.
	**Pooch by Dauger Research**	Cluster management software for creating easy-to-use, high-performance parallel computers.

What is a Cluster?

Varied components with a common goal

Message Passing Interface (MPI)

Distributed Resource Manager (DRM)



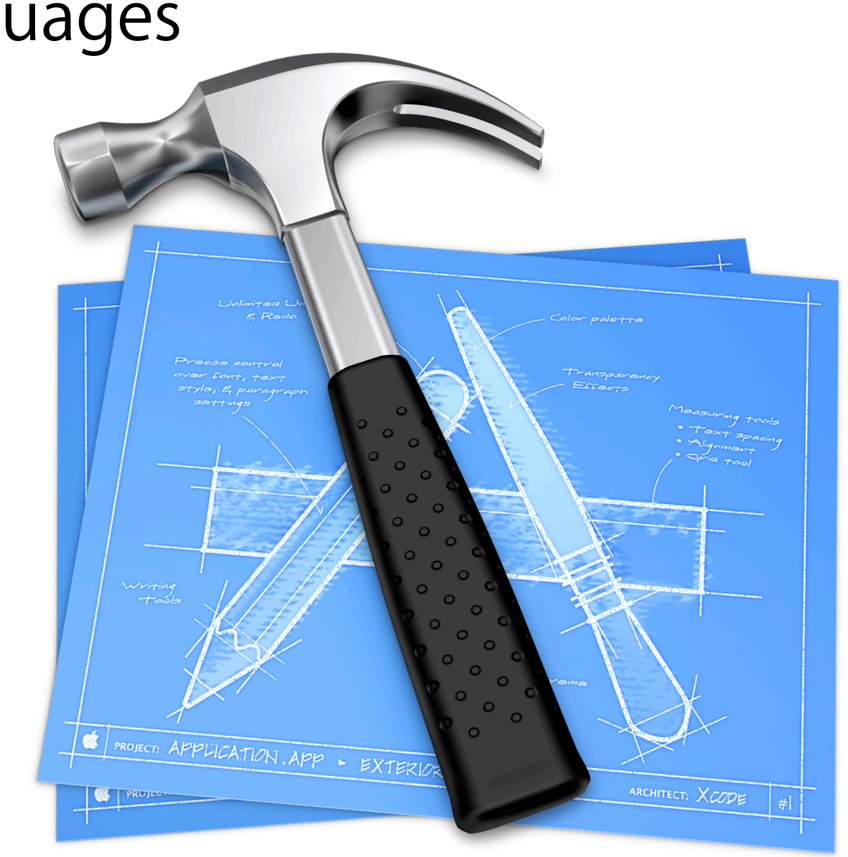
1. Some number of independently operable computers
2. High performance interconnection network
3. Software for controlling concurrent tasks (DRM)
4. API for parallel application development (MPI)
5. Development tools for cluster-aware applications

Xcode

The fastest way to create Mac OS X applications

- GCC 3.3
- Excellent support for C99, ANSI C++, and the C++ STL
- Includes GDB source-level debugger
- Works with C, C++, Objective-C, and other languages

FREE !!!



Compilers and Languages Courses

- Java
  - Standards-compliant optimized implementation
  - Included virtual machine, compiler, and debugger
- Legacy languages
  - Multiple implementations of Fortran F77 and F95
  - ANSI Common LISP and other LISP implementations
  - PowerPC assembly, plus multiple MIPS, x86, and other assembly virtual machines

Compilers

- Numerical Algorithms Group (NAG) f95
- Absoft F77, F90, and F95 compilers
- Veridian
  - VAST-F/AltiVec
- GNU Software Project
  - G77 for GCC 3.1
- IBM
  - XL Fortran
  - XL C/C++
- Auto-vectorizing GCC 4 in Mac OS X v10.4 !!

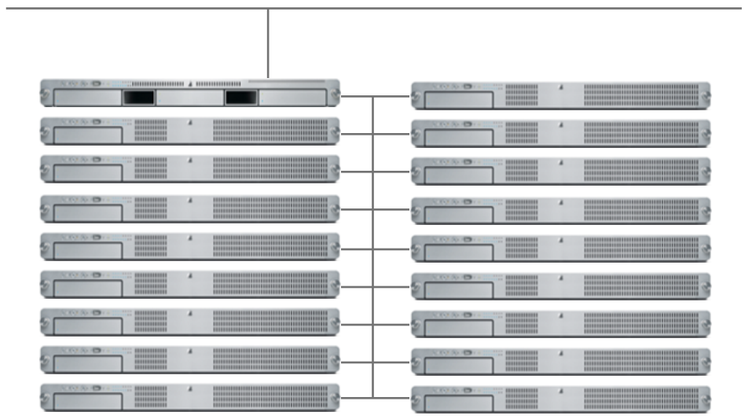
NAGWare f95 Compiler

What is a Cluster?

Varied components with a common goal

Message Passing Interface (MPI)

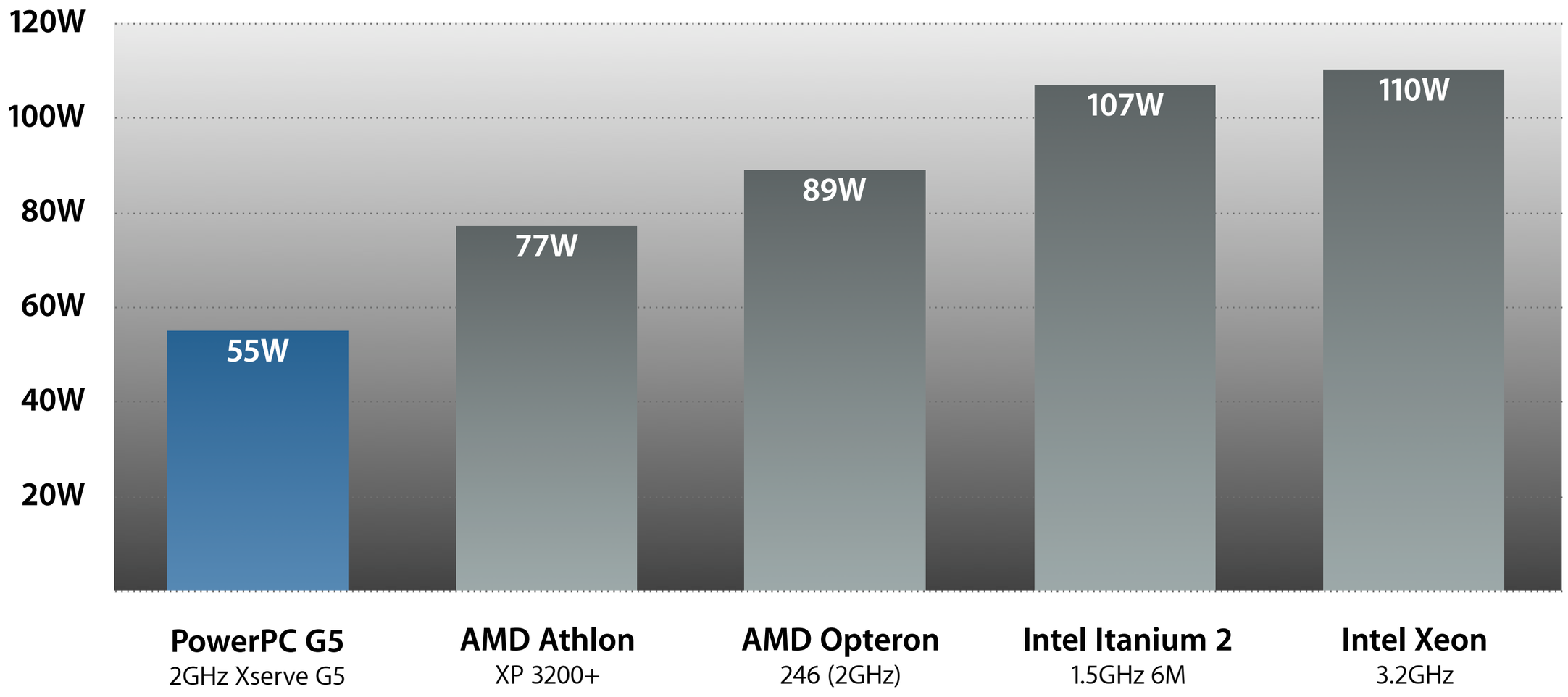
Distributed Resource Manager (DRM)



1. Some number of independently operable computers
2. High performance interconnection network
3. Software for controlling concurrent tasks (DRM)
4. API for parallel application development (MPI)
5. Development tools for cluster-aware applications
6. Physical infrastructure (space, power, HVAC)

Superior Power Efficiency

PowerPC G5 provides outstanding performance with low power consumption



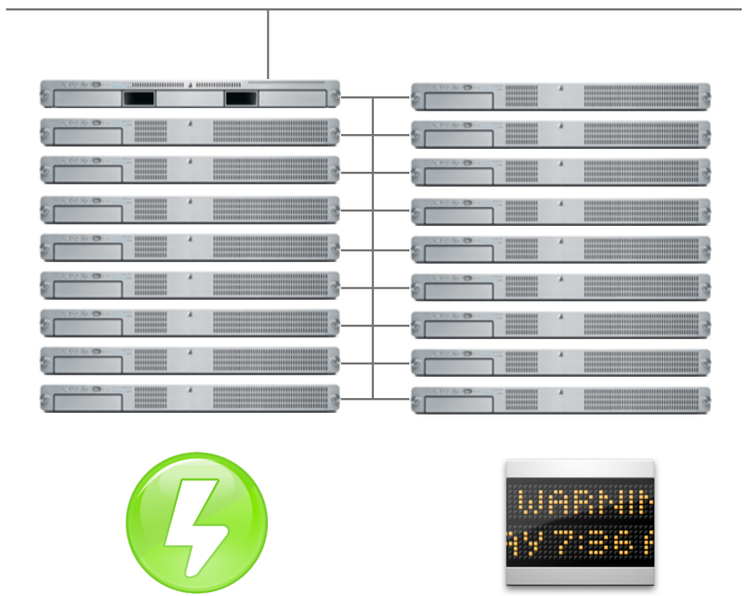
*Maximum power draw. Information from manufacturers' data sheets as of January 3, 2004.

What is a Cluster?

Varied components with a common goal

Message Passing Interface (MPI)

Distributed Resource Manager (DRM)



1. Some number of independently operable computers
2. High performance interconnection network
3. Software for controlling concurrent tasks (DRM)
4. API for parallel application development (MPI)
5. Development tools for cluster-aware applications
6. Physical infrastructure (space, power, HVAC)
7. Administrative and diagnostic tools

Third-party Scientific Software

Open Source Applications

Emacs
OpenOffice
Tcl/Tk
Fink
xTools
LaTeX
SPICE
Staden Package
GRASS
CUViewer
XEphem
Glimmer2
Gammes
Perl
SeqIO
Artemis
SEMPHY
EMBOSS
NCBI BLAST
Ghostscript
Phred-Phrad/Swat/CrossMatch
MUMmer
ClustalW
DOE 2000 Elect Notebook
Pine
MOLMOL
WU BLAST
PyMOL
MPI
FASTA
HMMer
Wise2.2.0
Apollo Genome Browser
OmniGene/OmniView
"R"
"O"
NOMAD
Apache Jakarta
NCBI Toolkit
BioPerl/Java/Python/DAS
MySQL
Primer 3.09
RasMOL
PAUP
MolPro
ACT
Phylip 3.5c
SIDB
XML/Ruby/TeX/CORBA
xFree86
XCluster
ARB
Sun GRID Engine
GIMP
Sequence Analysis
Java TreeView

Productivity and Scientific Applications

Dev Tools

NAGWare f95 Compiler

IBM XL Fortran

absoft development tools and languages

Crescent Bay Software

Math/Stat

MATLAB

SPSS

R

Visualization

Volocity

IDL

LabVIEW 7 Express

JMP The Statistical Discovery Software™

Instrumentation

BD

Molecular Devices

NATIONAL INSTRUMENTS

StereoGraphics

Productivity and Scientific Applications

Informatics

Vector NTI® Suite

Silicon Genetics

Daylight Chemical Information Systems, Inc.

accelrys

HPC

TurboWorx

Sun Grid Engine

Platform

Publishing

EndNote

Office v. X for Mac

CANVAS

OpenOffice.org Source Project

Databases

ORACLE

SYBASE

MySQL

UNIX Sequence Analysis

BLAST, A/G-BLAST
HMMER
Smith-Waterman
EMBOSS
FASTA
ClustalW
Phylip, PAUP

UNIX 3D Structural Analysis

Chimera
AutoDock
FADE, PADRE
PyMOL
VMD
GROMACS

Commercial: Sequence Analysis

Accelrys: MacVector
InforMax®: VectorNTI
GeneCodes: Sequencher
Textco: GeneConstruction Kit

Gene Expression

Silicon Genetics: GeneSpring

Invitrogen: Pathways

BioDiscovery: Imagene

Image Analysis

- NIH: ImageJ
- Improvision: Volocity
- Scanalytics: IPLab
- QImaging
- QED Imaging

Some Clusters

Virginia Tech

“I’d never used a Mac before in my life. Within three days I’d made up my mind that Mac OS X and the new Macintosh G5 were the tools I wanted to use.”

Dr. Srinidhi Varadarajan, Associate Professor of Computer Science
and Terascale Computing Center Director, Virginia Tech

Top 3 Supercomputers*

Rank	#1	#2	#3
Name	Earth Simulator	ASCI Q	System X
Location	Yokohama, Japan	Los Alamos, USA	Blacksburg, USA
Rmax/Rpeak	35.86/40.96 TF	13.88/20.48 TF	10.28/17.60 TF
# Processors	5,120	8,192	2,200
CPU Type	500MHz NEC (NEC)	1.25GHz Alpha (Compaq)	2GHz PowerPC G5 (Apple)
Cost	$300M	$215M	$7.5M
Manufacturer	NEC	HP	Self-made
Cost/GigaFlop	$ 8366	$ 15490	$ 730

* TOP500 List for November, 2003. Please see www.top500.org for additional details.

COLSA, Corp. (U.S. Army)

1,566 nodes

Node Details

- Head Nodes—4 dual 2.0 GHz Xserve G5 units
  - Dual mirrored 80 GB HDD
  - 8 GB RAM
  - 1 CD-RW
  - 1 video card
- Compute Nodes—1562 dual 2.0 GHz Xserve G5 units
  - Single 80 GB HDD
  - 3.5 GB RAM
  - No CD-ROM or video card
- Total of 3132 CPUs, 8 Gflop/s per CPU = 25.056 Tflop/s

Contact Info :

Jatinder Bharadwaj - jatinder.b@asia.apple.com

Rajiv Pillai - rajiv@asia.apple.com