Questions and Answers about Inquiry Cloud

This document provides information about Bioteam's new "Inquiry Cloud" product. For further information, to schedule a demonstration, or to purchase a license for this software, please contact us at info@bioteam.net

What Is Amazon's Web Service?

Amazon's web services are one facet of an emerging technology known as "cloud" computing. Amazon, the online merchant, (http://amazon.com) found that they needed a large pool of computers to meet their periods of maximum demand. Rather than let those systems sit idle the rest of the year, Amazon developed a simple and elegant system to rent out the CPU cycles, network bandwidth, and storage.

Over time the service accumulated enough demand to stand on its own as one of Amazon's offerings. It is no longer comprised of spare and excess systems, but is supported by dedicated hardware and staff at Amazon.

Amazon does not actually provide a one-to-one computer rental. Instead, they provide *virtual* servers much like consumer programs such as Parallels and VMWare Fusion allow *virtual* workstations. A customer runs a simple program to request an *instance* in the cloud. Once the instance starts the customer receives its internet address and logs in. It is just like having a real server out on the internet. Amazon also provides storage through their S3 system.

Even if you are not considering inquiry on EC2, it is well worth the time to become familiar with this technology. Amazon's main page for web services is a good place to start: http://aws.amazon.com

What Amazon Web Services does Inquiry Cloud Use?

The services used by Bioteam's inquiry cloud are:

- EC2: The "Elastic Compute Cloud"
- S3: The "Simple Storage Service"

What is Bioteam's Inquiry?

In 2003, Bioteam created the first installation of what is now called "inquiry." It met three critical needs:

- Cluster computing: Inquiry has traditionally included the configuration and setup of a compute cluster, harnessing many individual systems for easy use in either batch or parallel computing.
- Open source bioinformatics tools: A coherent and functional installation of BLAST, EMBOSS, HMMER, R, and other command line tools for bioinformatics.

- Web and Web Services Portal: The public face of inquiry is our web interface: It provides an intuitive and powerful interface to allow use of the cluster without any knowledge of the underlying unix command line.

Since then, inquiry has been installed on many operating systems, including OS X, many variants of Linux, Microsoft's Compute Cluster Server, Irix, and others. We have also created our Data Service to provide automatic updates from the public sequence repositories. Inquiry is now installed and in use on more than 200 clusters worldwide.

What is Inquiry Cloud?

"Inquiry Cloud" is Bioteam's implementation of Inquiry on a cloud computing resource. At present, Amazon's EC2 is the only cloud computing platform that we support. As other platforms become available, we plan to support them.

Historically, customers have always had to own real, physical computers in order to use inquiry. Owning servers can be a lot of work, and it's the sort of work that has almost nothing to do with getting scientific results. With EC2, Bioteam has made it possible to rent those CPU cycles from Amazon rather than insisting that you keep silicon and plastic in a basement or closet. The goal of this release of the software is to maintain as much as possible of the previous experience of purchasing an inquiry license while removing the requirement for owning hardware.

All of the benefits of using Amazon's web services apply directly to inquiry on EC2. Specifically:

- Pay for what you use: When demand is low, switch off the instances and pay nothing for them.
- Scalable computing: When a large task needs to be done, switch on more compute nodes and use them.
- No physical plant: There is no need to provide power, to cool, to install, to physically secure, or (eventually) to dispose of systems that are rented from the cloud rather than owned.

What does Inquiry Cloud Cost?
- $5,995 for academic and government institutions (includes one year of Bronze support)
- $15,995 for commercial entities (includes one year of Bronze support)

What am I buying when I buy Inquiry Cloud?

You are buying a software license to use Inquiry on EC2. That license is perpetual. Once you buy inquiry cloud, you do not have to buy inquiry again. The software we provide does not contain any mechanism to disable itself in the future. Our software license is designed to be as flexible as possible: We charge one price up front ($4,995 for academic and government institutions, $14,995 for commercial entities). That price does not vary with number of users, nor with the size of the cluster, nor with the amount that you use the software. You are permitted to customize the

system and to re-use our code in your in-house software. We forbid is re-distribution (for free or sale) of any of our code.

When you purchase Inquiry Cloud, you will also be required to purchase at least a one-year support contract with the Bioteam. All our support contracts entitle you to software updates, new versions of the inquiry software, and access to our support team via email at support@bioteam.net.   The levels of support are described here: http://bioteam.net/inquiry/support.html.

Once you have purchased inquiry on EC2, Bioteam will authorize your Amazon account to use our Amazon Images. We will also provide a tool to let you start and stop your cluster.

How do I pay for the CPU cycles in the Amazon Cloud?

You will pay Amazon directly. In order to activate your Amazon account for EC2, you will need to provide them with a credit card. Amazon will bill that credit card each month for only the CPU hours and storage that you actually use.

This is identical to the old model, in which you had to buy a computer from some company like Apple, Dell, or HP.

What happens if Amazon stops offering their services?

Amazon recently went on record saying that usage of the EC2 cloud is now well above the "extra capacity" level, and that the service is paying for itself. This means that it is profitable and should be available for the foreseeable future. However, Bioteam recognizes that with a service like Amazon EC2, if they pull the plug then our software becomes unusable.

Therefore, Bioteam commits to the following:

If the Amazon EC2 service changes in such a way that it is permanently unusable for Inquiry, inquiry cloud licenses will become transferrable, at no charge, to any of the other platforms that we support. At present, this includes Apple's OS X and most Linux / Unix based operating systems. Naturally, customers will need to obtain hardware on which to run the software.

For customers with a current support contract, Bioteam will assist in setting up an installation of inquiry on the hardware of their choosing. Though we do not presently support any other cloud platforms, Bioteam does intend to port Inquiry to them as they become available. If we have completed such a port, the unusable license would also be transferrable to any of those platforms.

If EC2 becomes unusable, and we cannot provide a working inquiry system in any other way, we will refund the purchase price (license and support) pro-rated over one year from the date that the customer first launched their licensed instance of inquiry.

Does Bioteam make money off the CPU cycles as well as the license?

No, Bioteam does not impose any additional charge for using our Amazon image. We believe that we can either sell you a license for the software, or charge by the use, but to do both would be wrong.

In the future, we may offer a per-use model as well. That would be a very different piece of software, since we would not be able to allow customization.

How much does an Amazon cluster cost to run?

Bioteam's Amazon instance requires a 64 bit image. Currently the minimum cost for those images is $0.40 per hour. That is just under $10 per day for the portal, and an identical amount for each compute node.

It is important to note that you will be charged when the system is *on*. Amazon doesn't have a way to distinguish between systems that are in active use and those that are sitting around waiting for a user to do something. Therefore, it is up to the customer to manage turning the system on and off as appropriate.

How long does it take to start the system?

The time required to starting cluster depends on the load that Amazon is experiencing, the network speed between the nodes, and the number of compute nodes that you request. The fastest startup we have experienced was less than a minute for a single portal with no compute nodes. The slowest was nearly half an hour for an 8 node system.

In addition, once the system has started, some time is required to put the BLAST targets in place. This usually takes less than two hours. BLAST targets are transferred from our S3 storage space, so there is no additional charge for the bandwidth. We provide a current set of the standard targets available from NCBI.

Can I customize the system?

Yes. In fact, we encourage customization.

When you purchase Inquiry Cloud, you will provide your Amazon identifier to Bioteam. We add that identifier to the list of allowed customers. This gives you access to our base system image. You can always go back to that basic inquiry image if you need to roll back any changes that you make. This is very similar to keeping an installation CD on the shelf in case you need to get back to where you started.

Once you boot up the system and use it, it will often make sense to install additional tools, to customize the web interface, or to upload your own data files. In order to keep the changes you make, you will need to create your own Amazon image and boot from that one instead of from ours. This is a simple procedure, and we've documented exactly how to do it.

Our license terms allow you to keep and use these derived images, but you are not permitted to further re-distribute them to other Amazon customers.

How are my results stored?

Because the system boots "clean" from the original image (or your customized one) each time the cluster restarts, we decided to use the S3 storage system to maintain result files and user accounts.

Every two hours your user information and results are saved to an S3 "bucket" that we define based on your Amazon identifier. This is your space within S3, and you will be charged for the storage.

How can I tell how much money I've spent with Amazon?

Amazon provides comprehensive tools for analyzing usage. Once you have an Amazon account, select "your web services account" and "AWS Account Activity" from the main page: http://aws.amazon.com. For example, here is Bioteam's billing report for a month of development on EC2:

Rate	Usage	Totals
Summary of This Month's Activity as of April 24, 2008		
Billing Cycle for this Report: April 1 - April 30, 2008		
		Expand All | Collapse All
Amazon Elastic Compute Cloud View/Edit Service		
$0.40 per Large Instance (m1.large) instance-hour (or partial hour)	318 Hrs	127.20
$0.10 per GB Internet Data Transfer - all data transfer into Amazon EC2	0.006 GB	0.01
$0.18 per GB Internet Data Transfer - first 10 TB / month data transfer out of Amazon EC2	0.012 GB	0.01
	View Usage Report	**127.22**
Amazon Simple Storage Service View/Edit Service		
	View Usage Report	**0.11**
Taxes Estimated Taxes (Due May 1, 2008)		**0.00**
Charges due on May 1, 2008†		**127.33**

What security measures are in place?

Amazon has given a lot of thought to system level security. They provide a system that can be as secure as any other remotely hosted server that is connected to the Internet. We build on that infrastructure to meet the security needs of most organizations.