

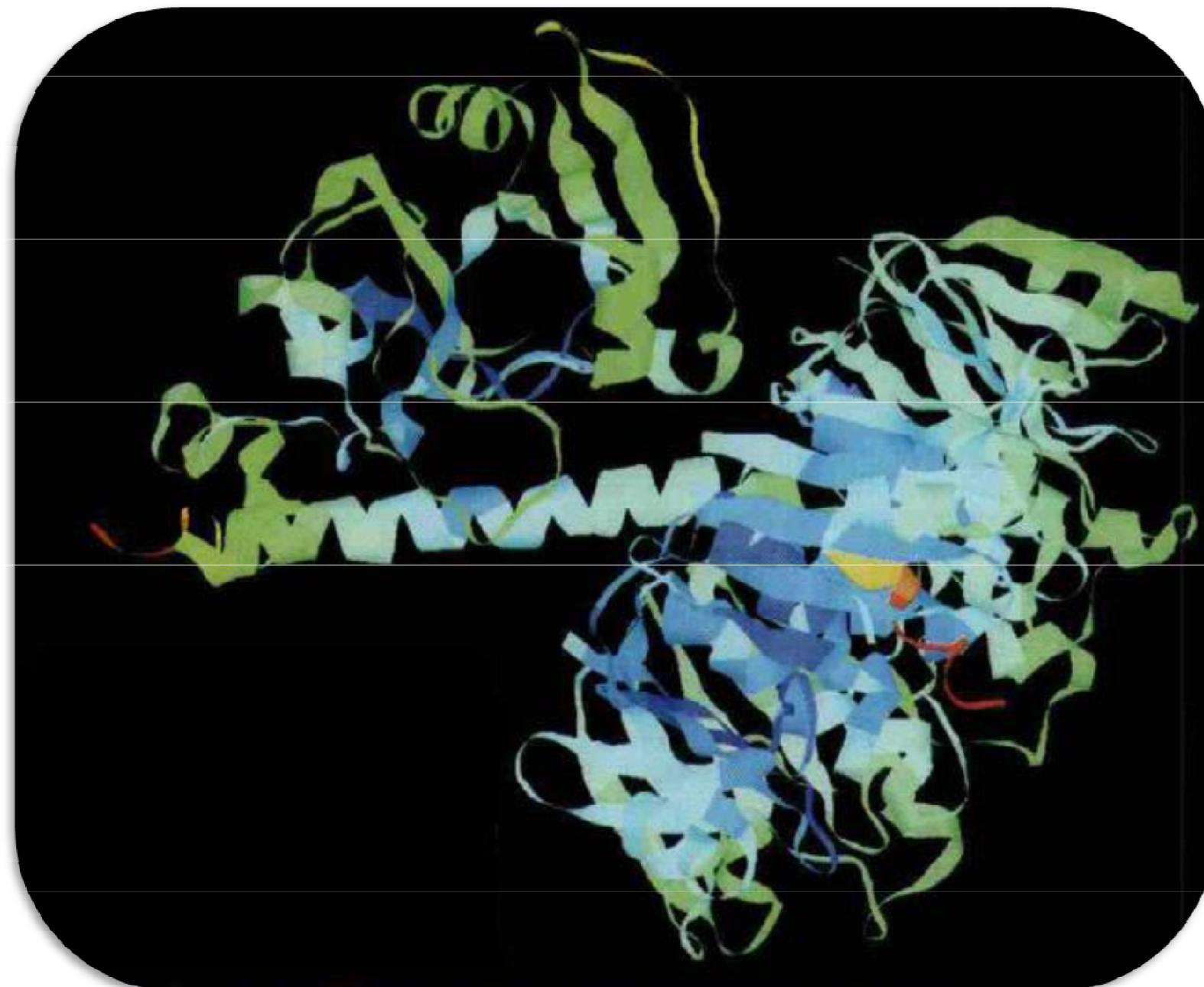
# Analyzing Grid Software: The Condor Project and Globus Toolkit

## The Problem

There is a need today to gather distributed computing resources for use in specific scientific problem solving.

Why? Places like the European Laboratory for Particle Physics are currently producing several petabytes of data per year -- a million times the storage capacity of the average desktop computer.

The challenge comes in harnessing these computing resources from a heterogeneity of networked systems all possessing unique configurations and characteristics.



Determining the structure of the complex Cholera Toxin molecule.

## Globus Toolkit

- Grid projects are being built on protocols and services provided by the Globus Toolkit.
- Developers: Ian Foster, Argonne National Laboratory and Carl Kesselman and team at the University of Southern California's Information Sciences Institute, and other institutions.
- Southern California Earthquake Center simulations cover a very large space with very high resolution and can generate up to 40TB of data per simulation run.

## Clustering Solution

- Provide significant improvements in total computing power.
- However, a cluster remains a dedicated facility, built at a single location.
- Example: 8,000-processor ASCI White system at Lawrence Livermore National Laboratory in California.



Decommissioned July 27, 2006.

## Grid Computing Defined

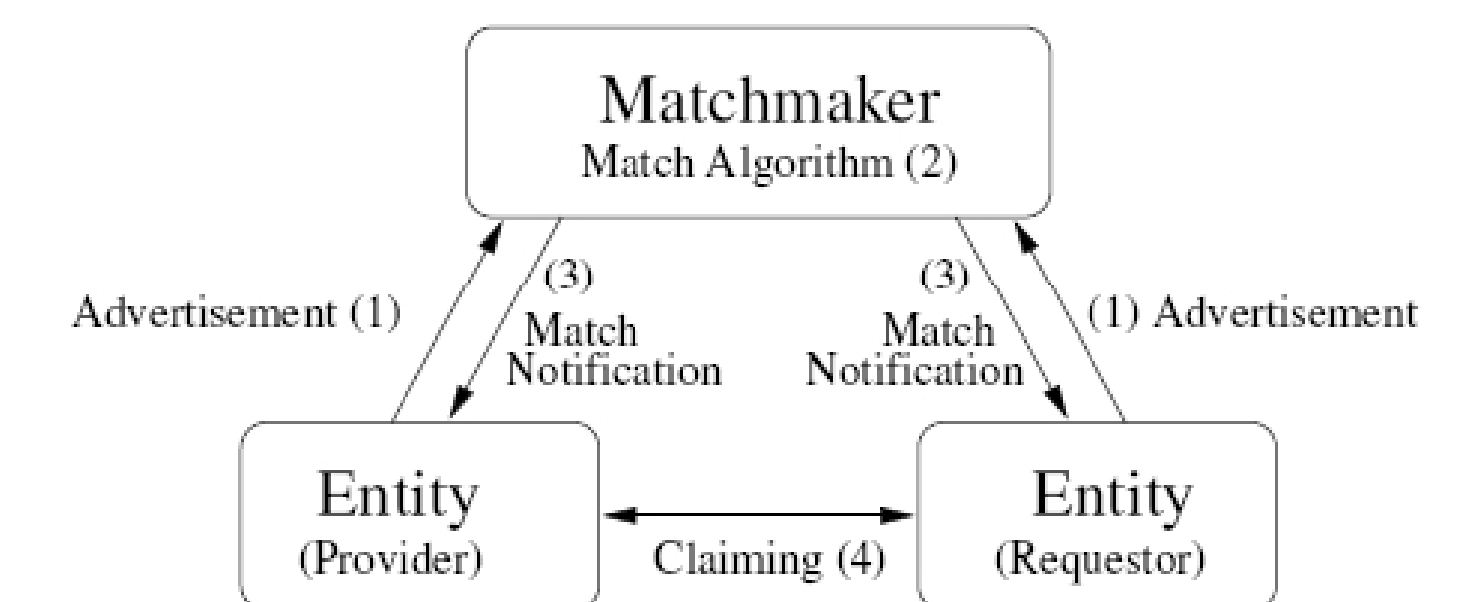
Carl Kesselman wrote that "a computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities."

### Implementation: Condor Project

- Provides a grid of processing resources for High Throughput Computing.
- High Performance Computing (HPC) attempts to process floating point operations per second or minute.
- However, many scientific experiments around the world need to process floating point operations that normally take weeks, months, or even years to complete.

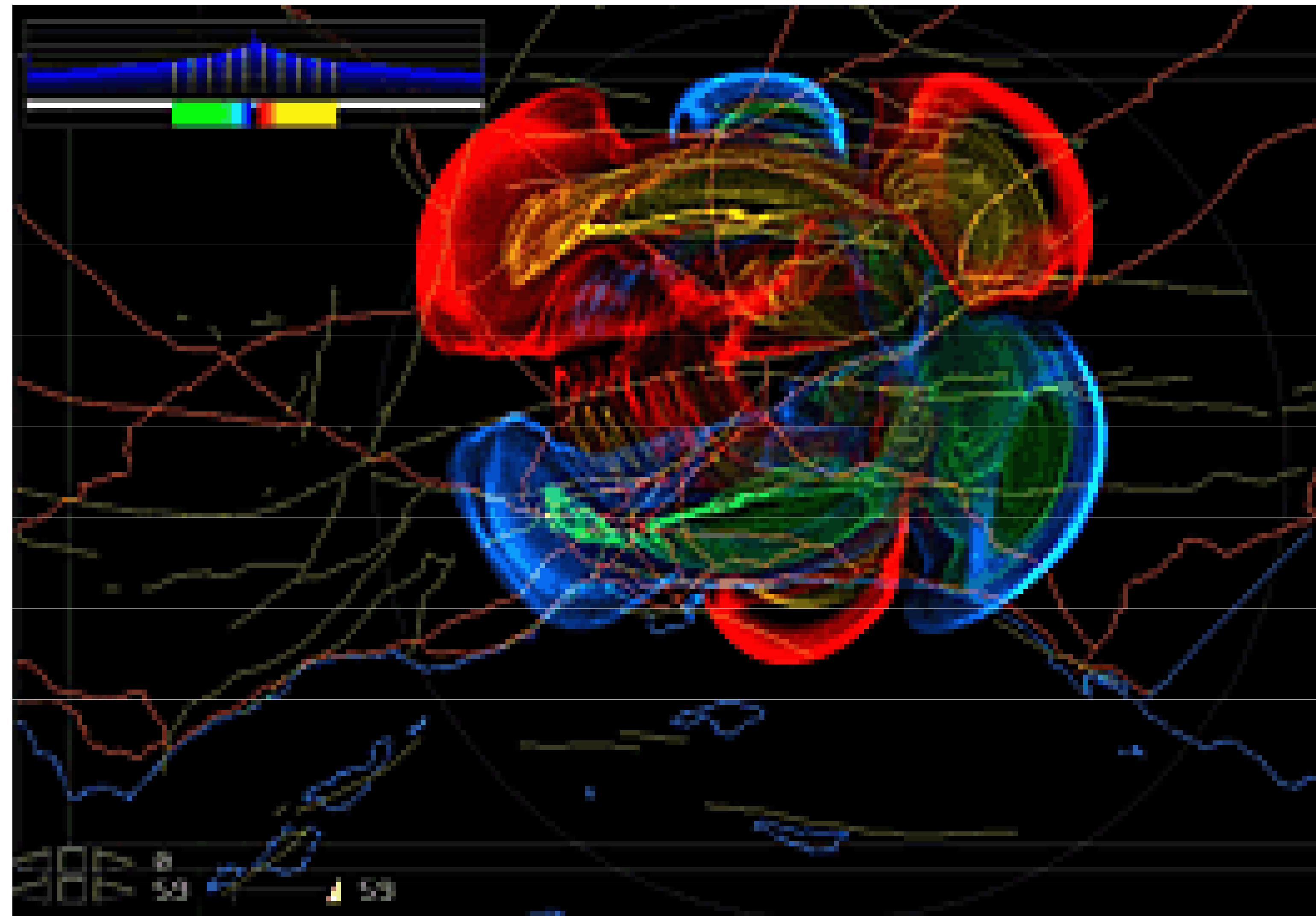
## Matchmaking Framework

- Entities that provide or require a service advertise their resource requirements and characteristics in classified advertisements (ClassAds).
- A matchmaker matches the ClassAds according to the specified constraints.
- The service then notifies each entity of the match without any further intervention.
- It is then the responsibility of the matched entities to establish further communication.



## Conclusion

- The Globus Toolkit has allowed Grid technology to finally begin its standardization process.
- Analyzing the Condor system's functionality and architecture helps in viewing how a Grid system is intended to operate.
- Projects like Condor tackle only part of the problems that Grid computing aims to solve as a whole. Software in Grid computing will continue to be developed in a more standardized way that will allow true resource sharing without centralized control.



Earthquake simulation