

# Condor: A Concept, A Tool and A Model

Miron Livny  
Computer Sciences Department  
University of Wisconsin-Madison  
[miron@cs.wisc.edu](mailto:miron@cs.wisc.edu)  
<http://www.cs.wisc.edu/~miron>



# The Condor Project

(Established '85)

Distributed systems **research** performed by a team that faces

- **software engineering** challenges in a UNIX/Linux/NT environment,
- active interaction with **users** and collaborators,
- and the daily maintenance and support challenges of a real-life distributed **production** environment.

**Funding** - NSF, NASA, DoE, DoD, IBM, INTEL, Microsoft and the UW Graduate School



# National Grid Efforts

- National Technology Grid - NCSA Alliance (NSF-PACI)
- Information Power Grid (NASA)
- Particle Physics Data Grid (DoE)



# Applications

- Optimization - UW, ANL, NW
- High Energy Physics - INFN, UNM, UW, Caltech
- Biological Sciences - UW, UMN
- Animation - UW, C.O.R.E
- Software Engineering - Oracle
- JAVA - NASA



# CS Collaborations

- Argonne National Lab (Globus) - Grid middleware
- Universitat Autònoma de Barcelona - Scheduling of Master-Worker Applications
- Clark Atlanta University - User Interfaces



# Funding Distribution

Source	FTEs	50% RAs	Comments
Federal	3.5	12.0	
Industry		3.0	Unrestricted grants
Grad School	2.0	2.0	
128 account	0.5		"For fee" support
Total	6.0	17.0	



# Concept(s)



“ ... Since the early days of mankind the primary motivation for the establishment of *communities* has been the idea that by being part of an organized group the capabilities of an individual are improved. The great progress in the area of inter-computer communication led to the development of means by which stand-alone processing sub-systems can be integrated into multi-computer '*communities*'. ... ”

Miron Livny, “ *Study of Load Balancing Algorithms for Decentralized Distributed Processing Systems.*”,  
Ph.D thesis, July 1983.

Every Community  
needs a  
Matchmaker!



---

[www.cs.wisc.edu/condor](http://www.cs.wisc.edu/condor)

# Why? Because ...

.. someone has to bring together members who have **requests** for goods and services with members who **offer** them.

- **Both** sides are looking for each other
- **Both** sides have constraints
- **Both** sides have preferences



# High Throughput Computing

For many experimental scientists, scientific progress and quality of research are strongly linked to computing **throughput**. In other words, they are less concerned about **instantaneous** computing power. Instead, what matters to them is the amount of computing they can harness over a month or a year --- they measure computing power in units of scenarios per **day**, wind patterns per **week**, instructions sets per **month**, or crystal configurations per **year**.



# Computing is a Commodity

Raw computing power is everywhere - on desk-tops, shelves, and racks. It is

- cheap
- dynamic,
- distributively owned,
- heterogeneous and
- evolving.

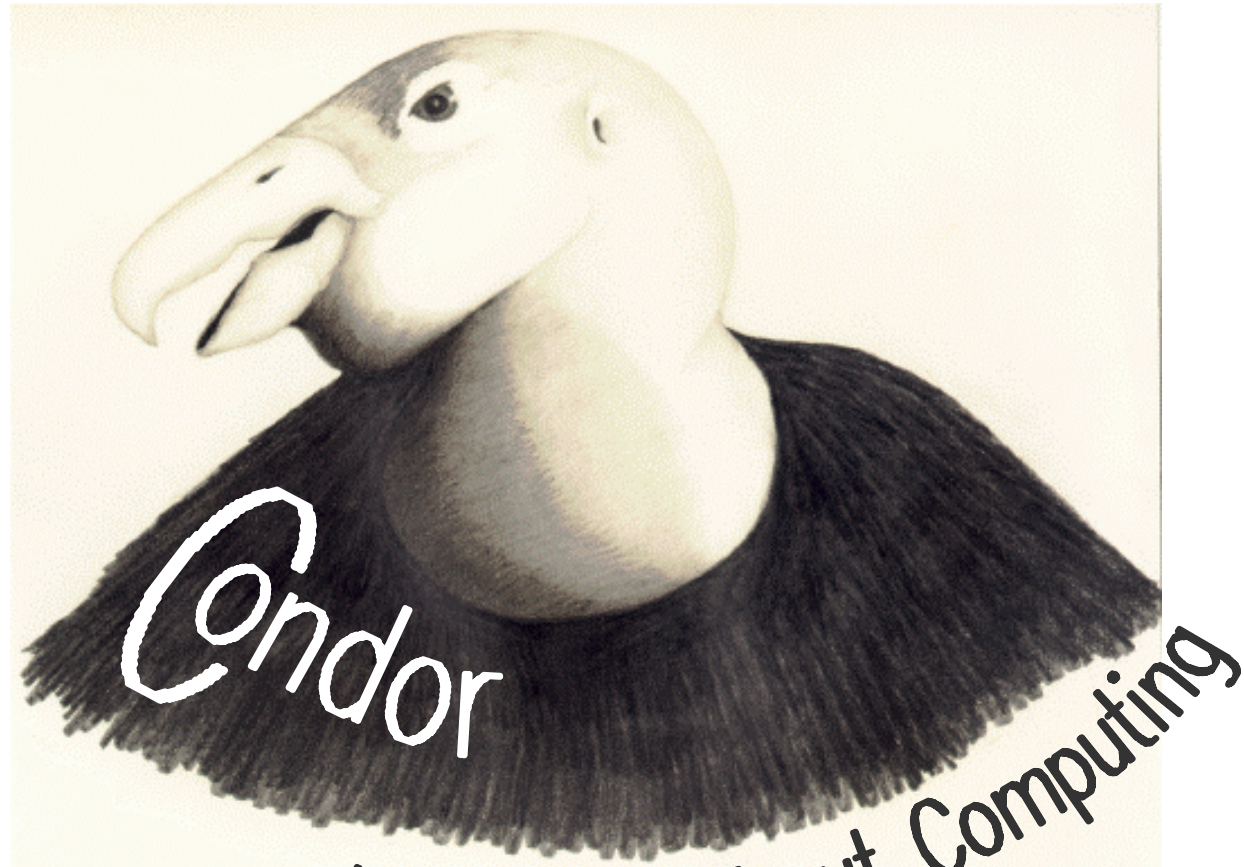


Master-Worker (MW)  
computing is common and  
**Naturally Parallel.**

It is by no means  
**Embarrassingly Parallel.**  
Doing it right is by no means  
trivial.

# A Tool





*Our  
Answer to  
High Throughput MW Computing  
on commodity resources*

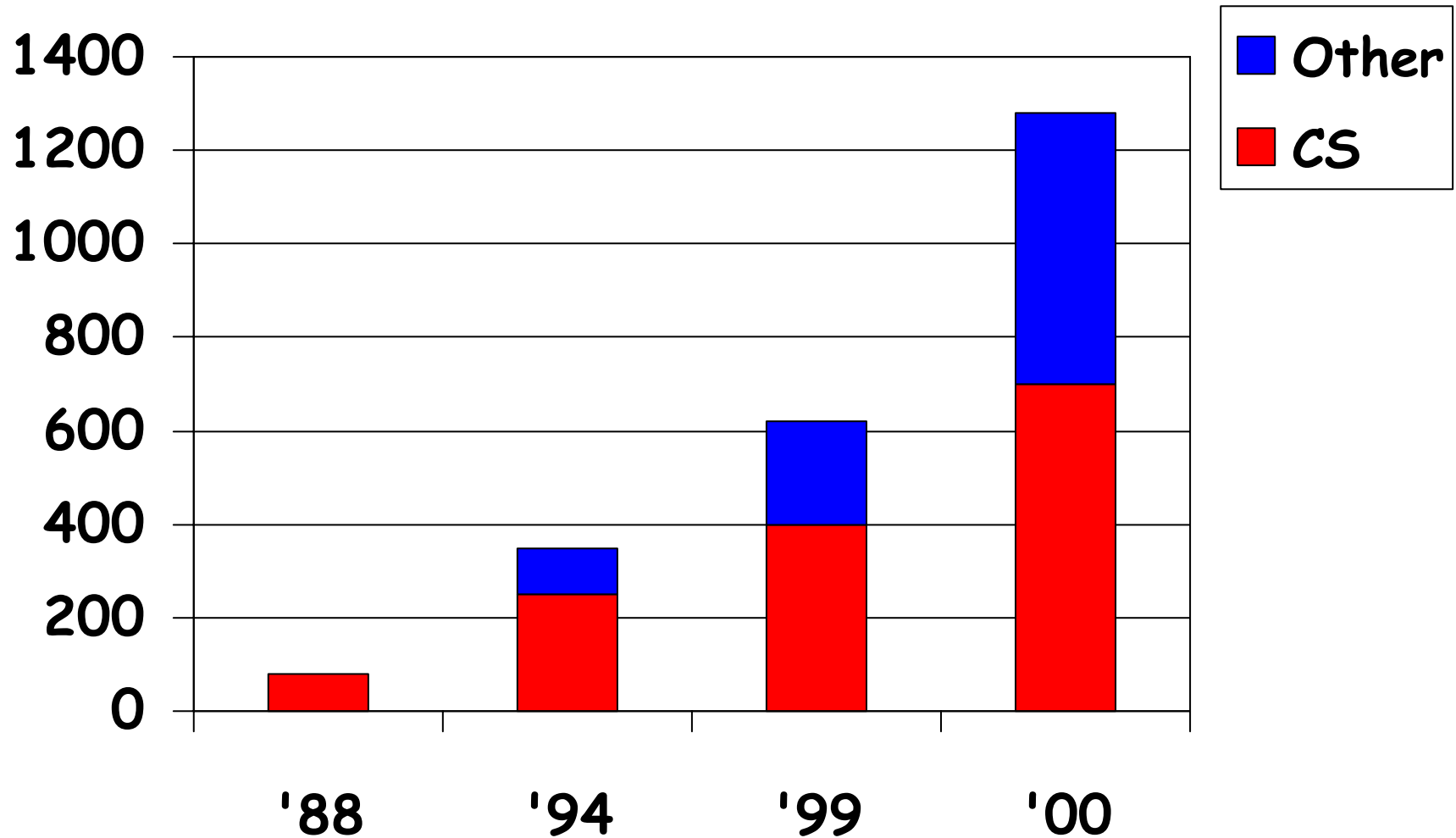
# The Condor System

A High Throughput Computing system that supports large dynamic MW applications on large collections of distributively owned resources developed, maintained and supported by the Condor Team at the University of Wisconsin - Madison since '86.

- Originally developed for UNIX workstations.
- Fully integrated NT version is available.
- Deployed world-wide by academia and industry.
- More than 1300 CPUs at the U of Wisconsin.
- Available at [www.cs.wisc.edu/condor](http://www.cs.wisc.edu/condor).



# Condor CPUs on the UW Campus



# Some Numbers-CS Pool

Total since 6/98	4,000,000	hours	~450 years
"Real" Users	1,700,000	hours	~260 years
CS-Optimization	610,000	hours	
CS-Architecture	350,000	hours	
Physics	245,000	hours	
Statistics	80,000	hours	
Engine Research Center	38,000	hours	
Math	90,000	hours	
Civil Engineering	27,000	hours	
Business	970	hours	
"External" Users	165,000	hours	~19 years
MIT	76,000	hours	
Cornell	38,000	hours	
UCSD	38,000	hours	
CalTech	18,000	hours	



# A Model for ...



# CS-Domain Collaborations

Multi disciplinary research that advances the state-of-the-art in CS and a domain science.

- Based on mutual respect and understanding of objectives, needs, constraints and culture
- Leverage expertise, resources and funding
- Enables experimental Computer Science
- Enables speculative science



# Campus Scientific Computing

Support the increasing demand from domain scientists for advanced computing, storage and networking services

- Computing power
- State-of-the-art middle-ware and libraries
- Access to experts who understand the nature and dynamics of scientific computing
- Cycles for class/research projects



# Software Distribution and Support

Making software developed in academia available to academic and commercial users.

- Legal and technical support for software distribution
- Infrastructure for “for-fee” support
- Blueprint for dealing with IP rights



