

If you can do it on
the desktop,
you can do it
everywhere!

Miron Livny
Computer Sciences Department
University of Wisconsin-Madison
miron@cs.wisc.edu



Claims for “benefits” provided by Distributed Processing Systems

- High Availability and Reliability
- High System Performance
- Ease of Modular and Incremental Growth
- Automatic Load and Resource Sharing
- Good Response to Temporary Overloads
- Easy Expansion in Capacity and/or Function

“What is a Distributed Data Processing System?” , P.H. Enslow, Computer, January 1978

Desktop resources are the
"true" Grid resources.
Why, because they are ...

Execution Site

- > Distributively owned - each resource may have a different owner with a personal usage policy
- > Opportunistic - you can never count on them being there when you need them.
- > Unknown environment - you never know what kind of software your application will find.
- > Heterogeneous - no control over hardware or OS.
- > Can be malicious - you can not trust the information sent to you by the execution site

Client Site

- > Master - in charge of managing resources and work
- > Home - defines home environment for application
- > Identity - source of authentication information
- > Storage - holds workload and data

**What Did We Learn From
Serving
a Quarter of a Million
Batch Jobs on a
Cluster of Privately Owned
Workstations**

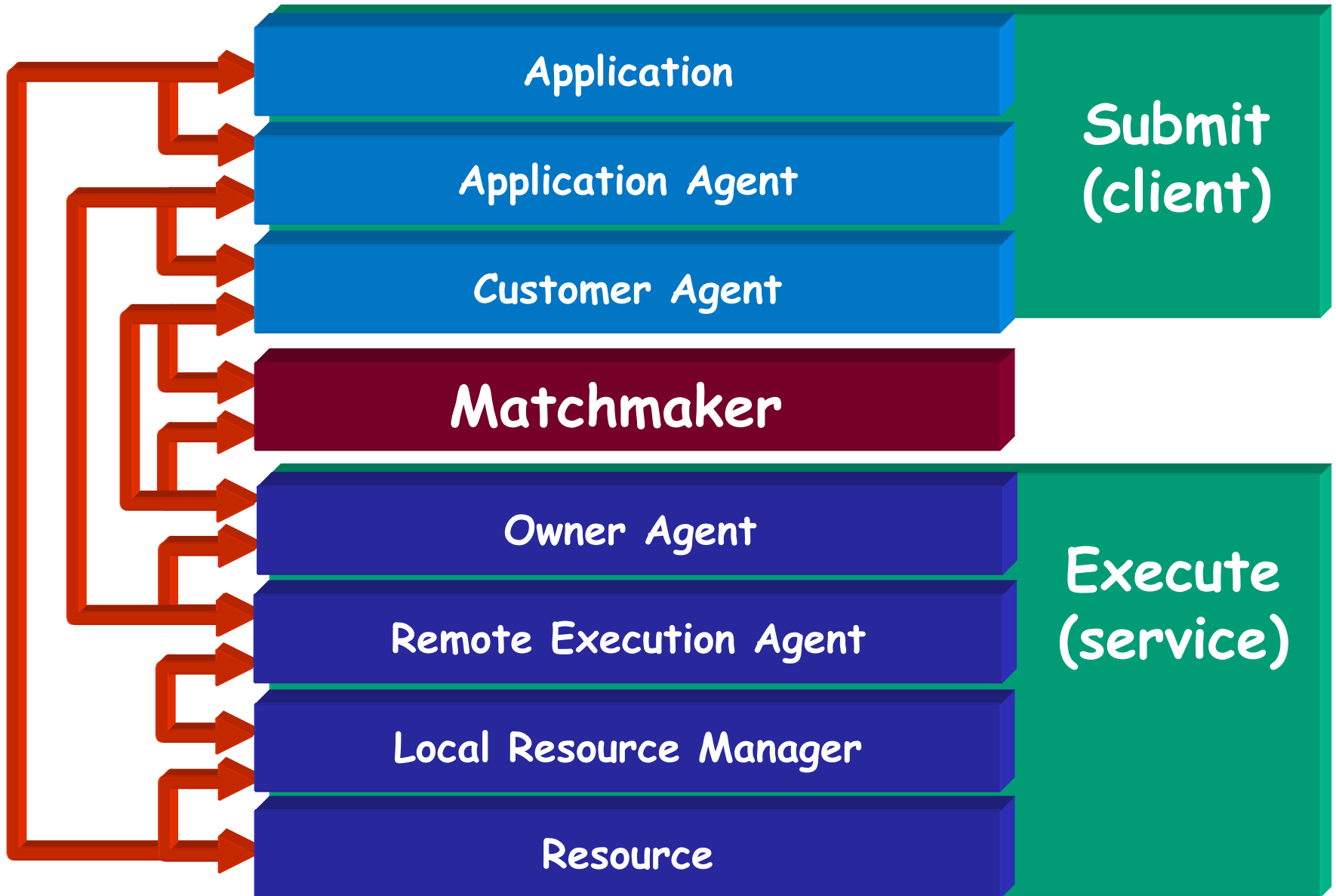
1992

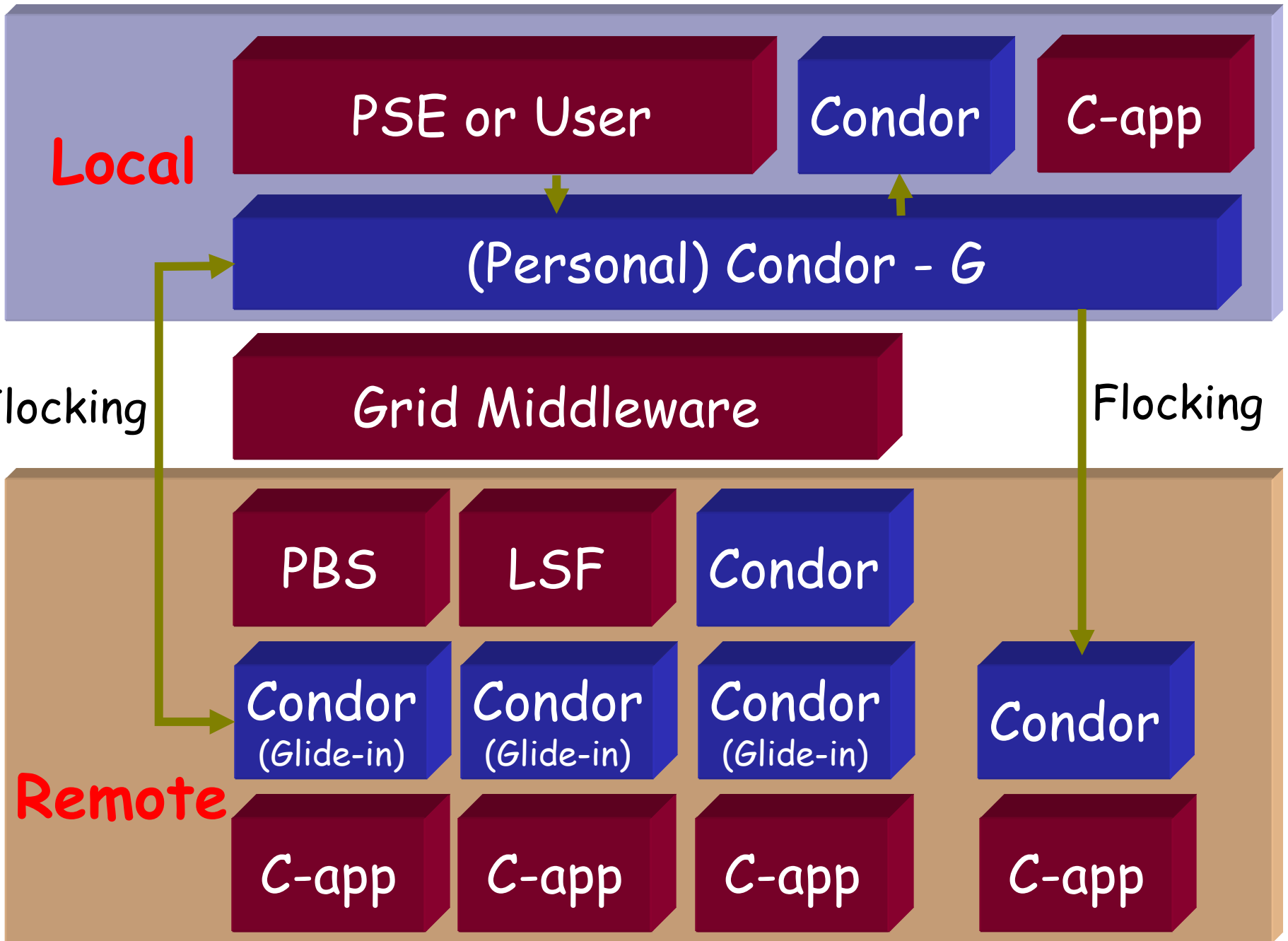
Computer Sciences Department
University of Wisconsin — Madison
Madison, Wisconsin
{mlron@cs.wisc.edu}

**User
Prospective**

- Maximize the capacity of resources accessible via a single interface
- Minimize overhead of accessing remote capacity
- Preserve local computation environment

The Layers of Condor





Terminology

What is "special" about desktop Grids that we need a "special" terminology?

They consists of CPUs, memory, disks and networks and run applications that read and write streams of characters.

Who should I trust?

- > User - sometimes
- > Application - sometimes
- > Administrator - sometimes
- > Owner - sometimes
- > Operating system - sometimes

Applications

- Master worker - YES!!!
- MPI - NO!!!
- Peer to peer - Why? (if you can do master worker?)
- Data intensive - Local disk access is a challenge

There is nothing embarrassing in pleasantly parallel applications