

Offering GridWay users more power by means of Amazon public cloud



Alejandro Lorca

**Distributed Systems Architecture Group
Universidad Complutense de Madrid**



Introduction of framework

Motivation

Cases of Interest

Technical quests and DEMO

Summary and Outlook

GridWay



- A tool for interoperability (grid, clusters)
- Powerful metascheduling
- Based on Globus services
- <http://gridway.org>

Amazon Web Services



- Public Cloud Computing provider
- Compute power (EC2) and Storage (S3, EBS)
- Pay as you go
- Instant delivery
- <http://aws.amazon.com>

We all want more...

Cases of interest:



- **User** decides to speed up his work



- **Group of users** require more resources for a given **experiment (VO)**



- The **organization** extends computability to reduce the **overall “time on queue”**

Use Case I: User's job speed up

User 

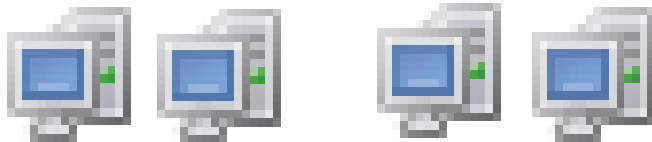


Job templates

GridWay daemon (gwd)

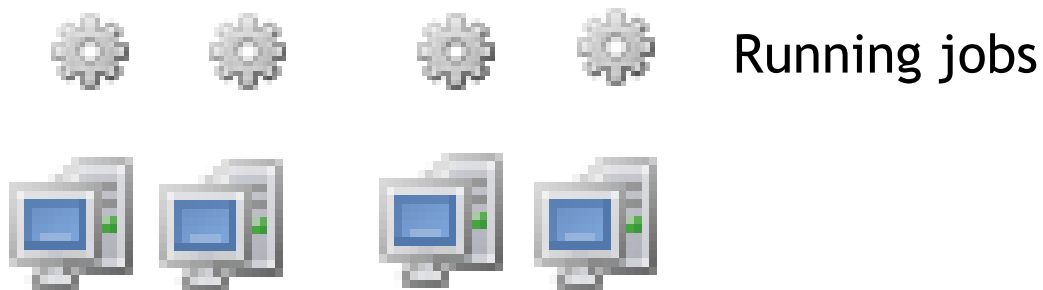
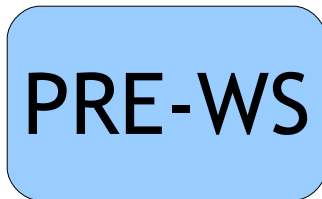
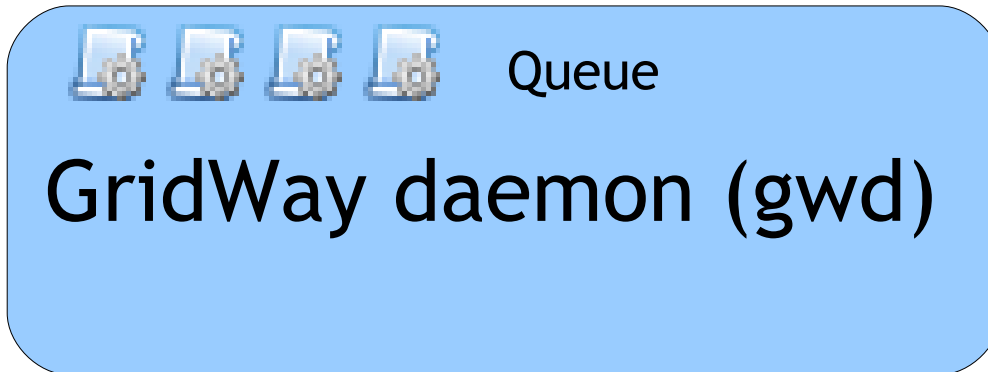
PRE-WS

WS

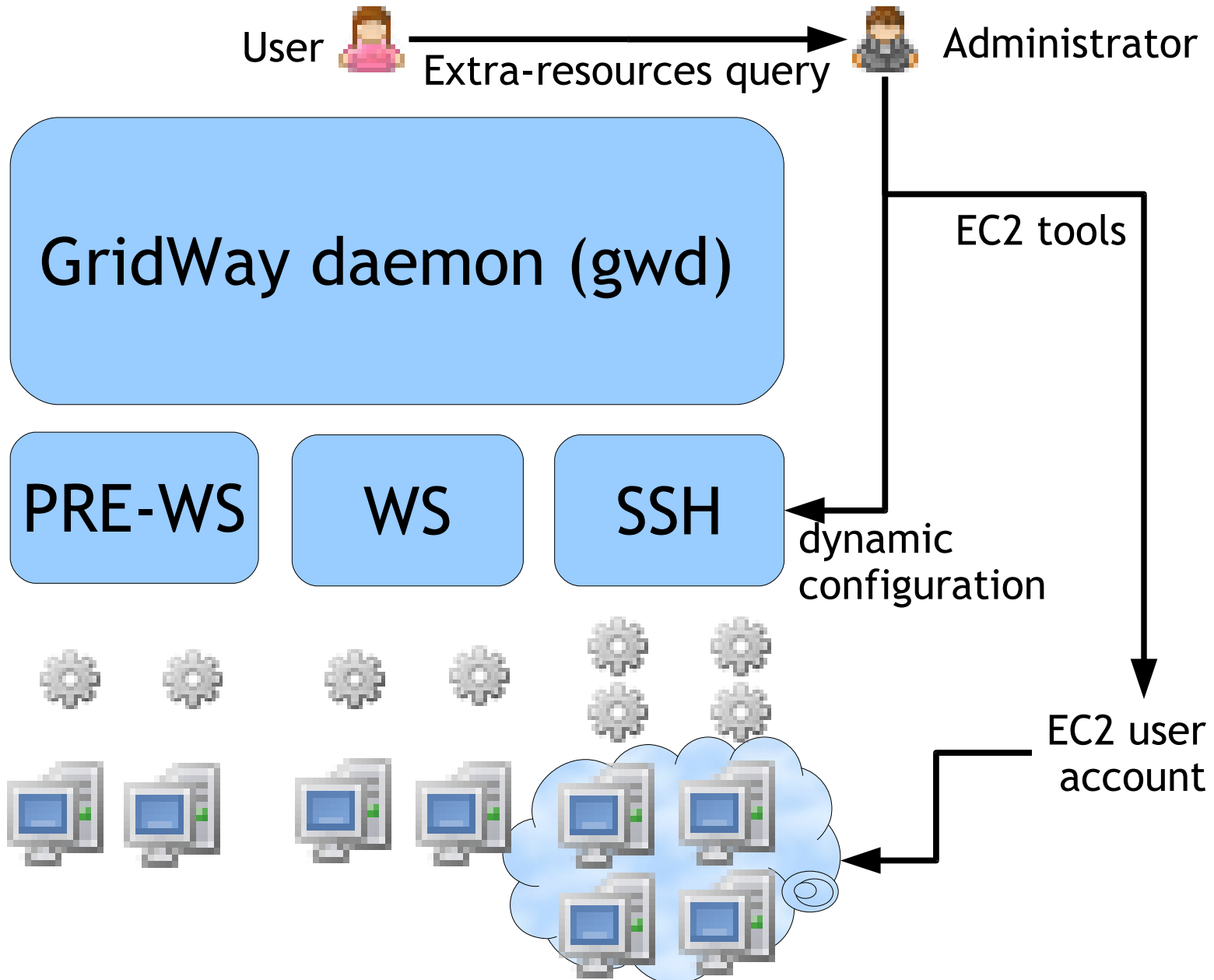


Use Case I: User's job speed up

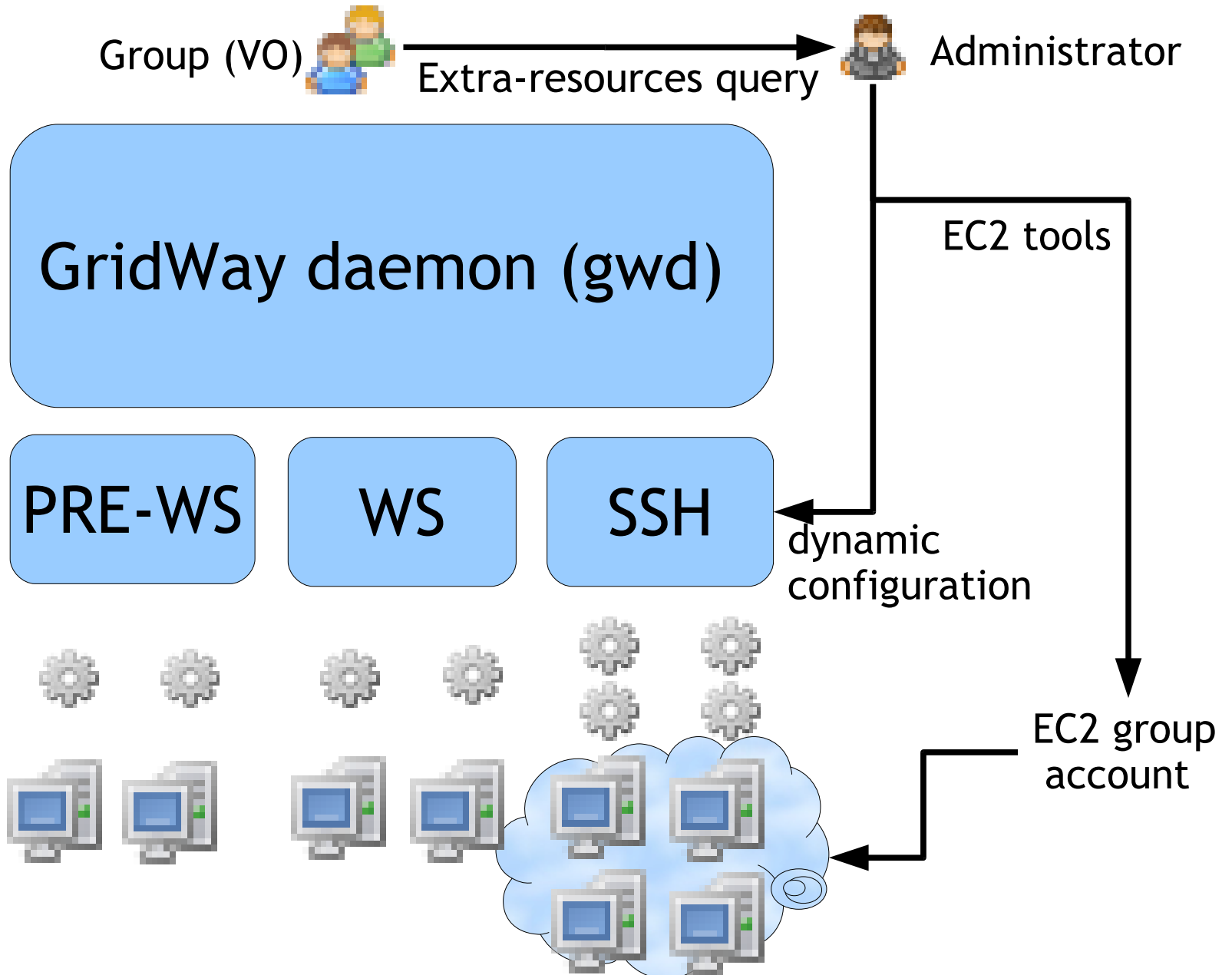
User 



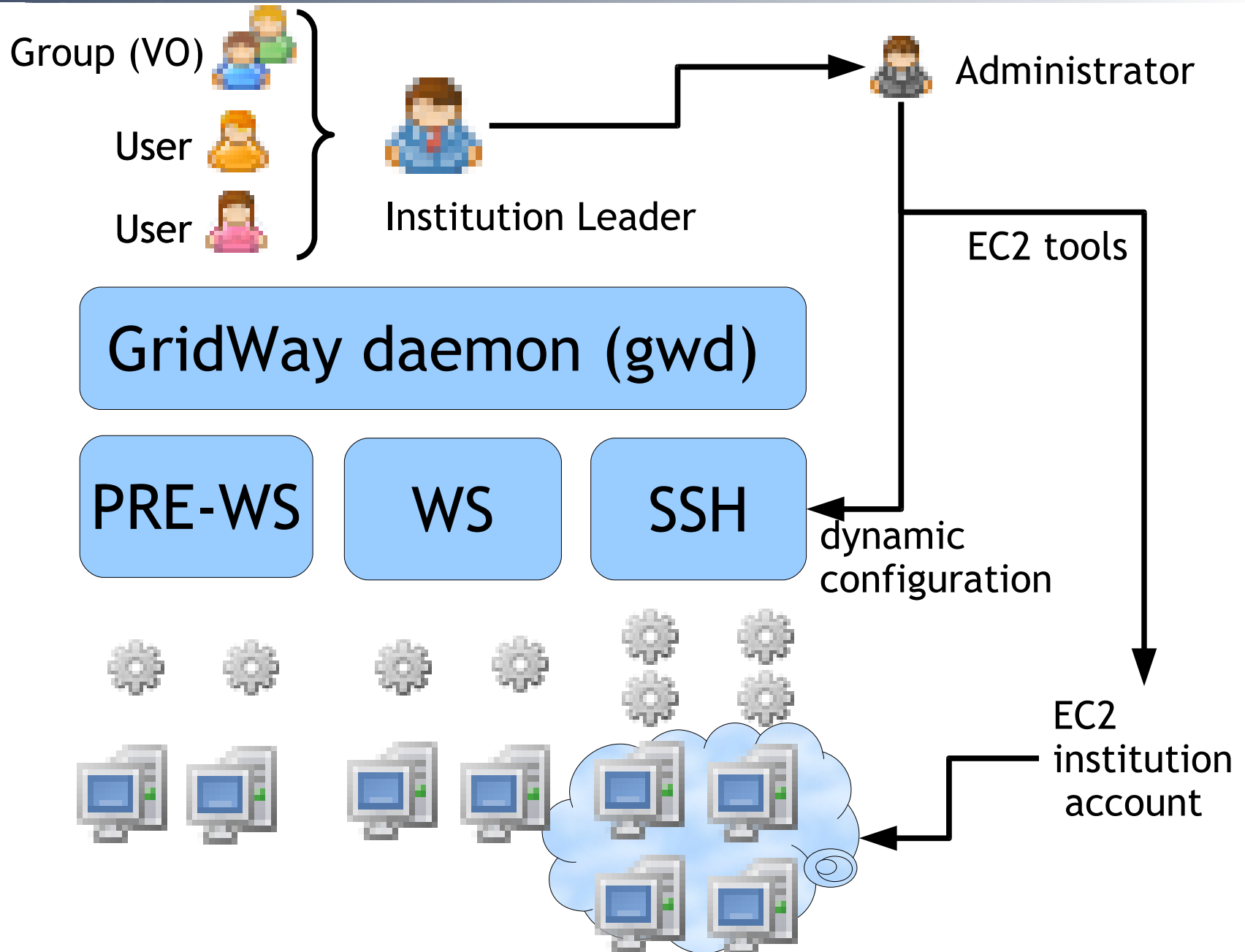
Use Case I: User's job speed up



Use Case II: Group experiment extension



Use Case III: Institutional extension



Making it feasible:



- User, Groups or Institutions
EC2 accounts (certificate and private key)
are **provided to gwadmin** for launching VMs.



- This is like giving a checkbook to the local resource provider (**Trust/Risk**). The gwadmin acquires great responsibility.



- The legitimation of use is left to the standard user-group-other Unix permissions.
- Everyone reading the **machines'**
private keys can access those machines.

Key issues:

- **Security:** gwadmin needs +r on **Public CERT** and **private KEY** from the EC2 account
- **GW-cloud CLI/API:** the gwadmin should apply cli tools quite *similar to ec2-tools* to **enroll machines** into the static config files.
- **VMs contextualization:** Depending on **AMI**, work is still to be done to **become a CE**.
- **Protocol limitations:** ssh-plugin works faster, but **hangs for ~20** simultaneous connections



Don't try it at home, it costs money!

dsa-research.org

DEMO

Summary:

- **Extension:** Users are given more power on a specific situations, but this is not free
- **Fast and simple:** SSH and SFTP are the only protocols for communication and execution
- **Flexibility:** The GridWay administrator might prepare plenty of different images aiming at different needs.
- **It works!:** DEMO part

Outlook:

- **Automation:** The gwadmin would give up the VMs managing to the **service manager daemon** according to some queue/job policies.
- **VMs Infrastructure:** a **service manager** decides which infrastructure combination is **most adequate** following some SLAs and knowing different Resource providers.
- **Simulation/Tests:** more experience needed, ssh-plugin, workload, price policies, service manager.