

Usage Policies at the Site Level in Grid

Catalin L. Dumitrescu
The University of Chicago

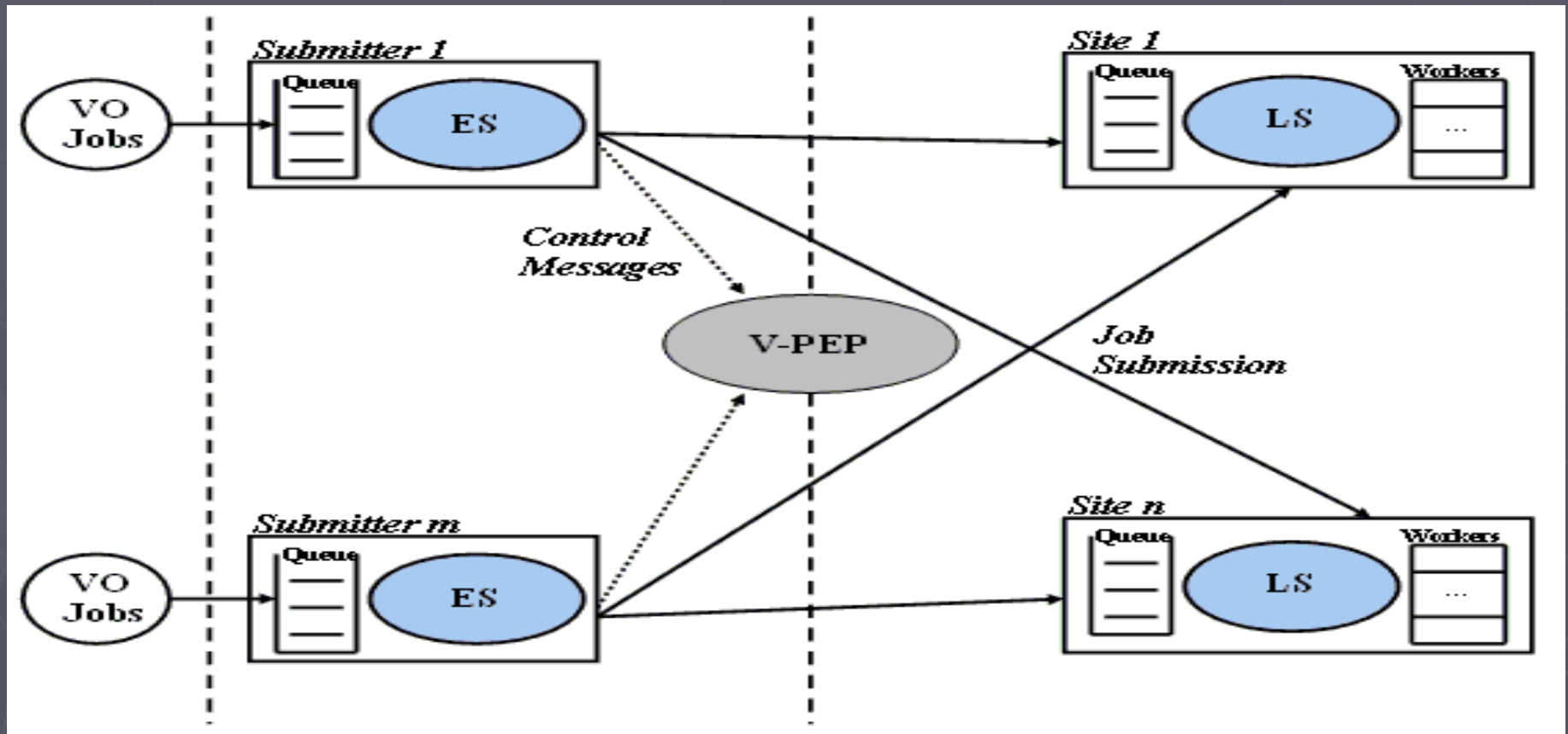
Outline

- Introduction / Scenario
- What is Usage Policy (or Usage SLA)?
- Example & Details
- S-PEP Solutions
- Evaluation Methodology
- S-PEP Applications
- Open Questions

Targeted Scenario (Grid3 Model)

Users/VO/Workloads

Sites/Resources



What is Usage Policy?

- **Usage Policies:** resource owners' (local policy makers) statements about how their resources must be allocated – *high level GOALS*
- **RM Priorities/Rules:** resource administrators mappings of resource owners' statements to different software RMs' syntaxes – *the local POLICY that is actually implemented*

Dummy Example

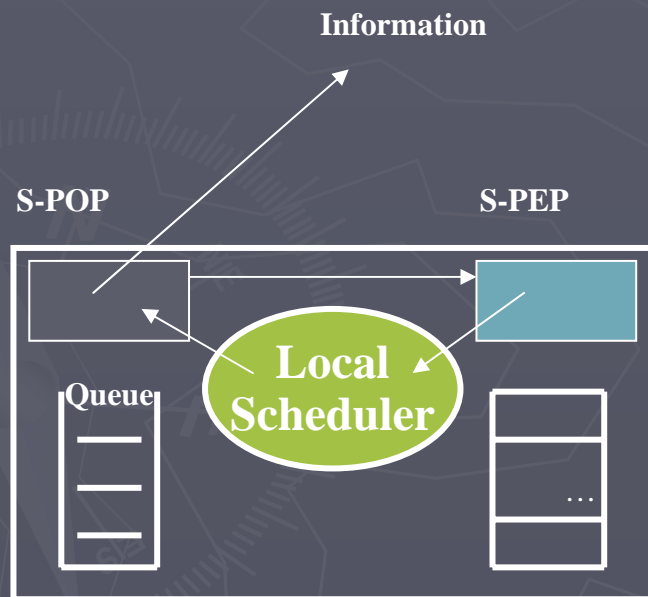
- **Site owner's statements for its site are:**
 - We have a cluster with 380 CPUs
 - At any time: ATLAS has a 30% ; the other VOs together have just 10%
 - When additional resources are available, Grid03's VOs can grab these resources
- **The Condor priorities become:**
 - % condor_userprio -setfactor atlas 2
 - % condor_userprio -setfactor others 9

PEPs

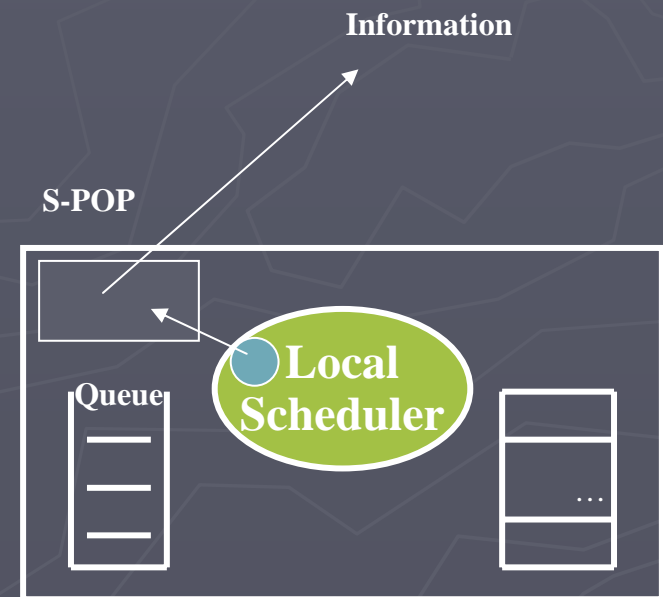
- ▶ conceived as tiny modules coupled to RMs
- ▶ **POPs components:**
 - gather monitoring metrics and other information relevant to their operations
- ▶ **PEPs:**
 - execute policies by different means
 - use this information to steer resource allocations as specified by the usage policies

S-PEP Approaches

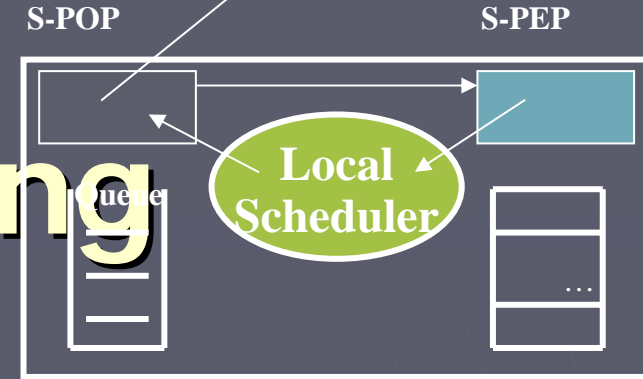
► Continuous Monitoring Approach



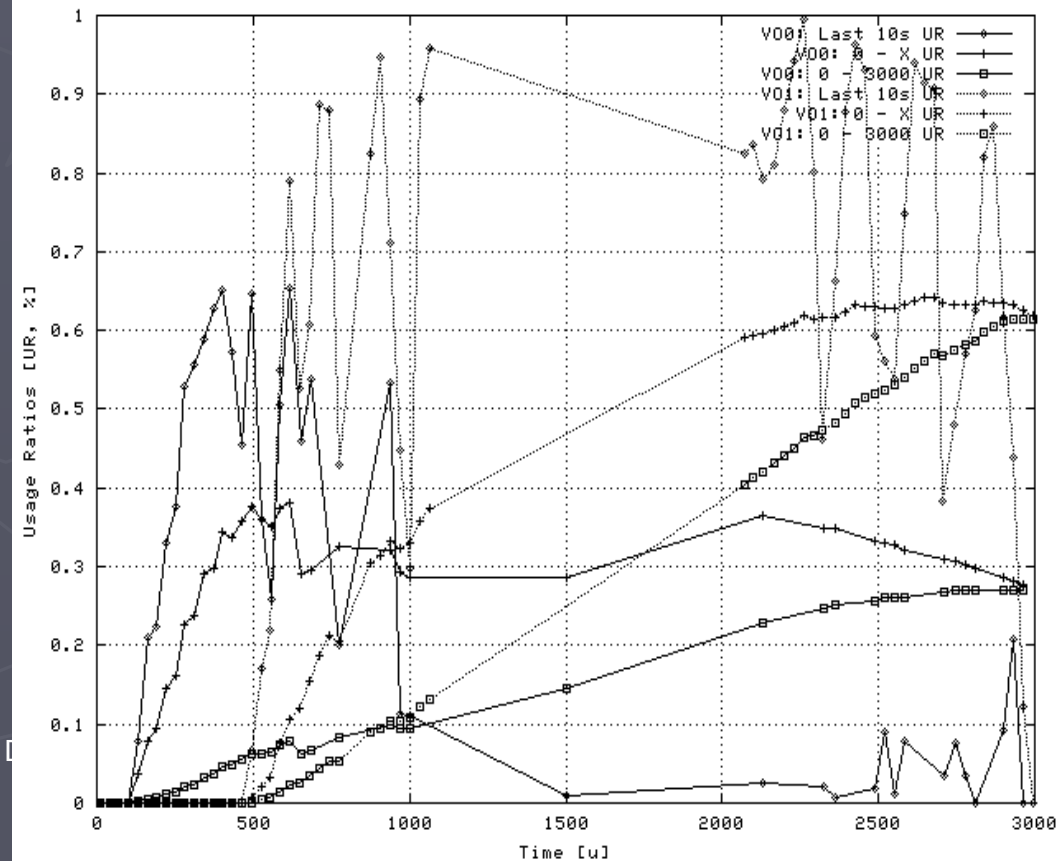
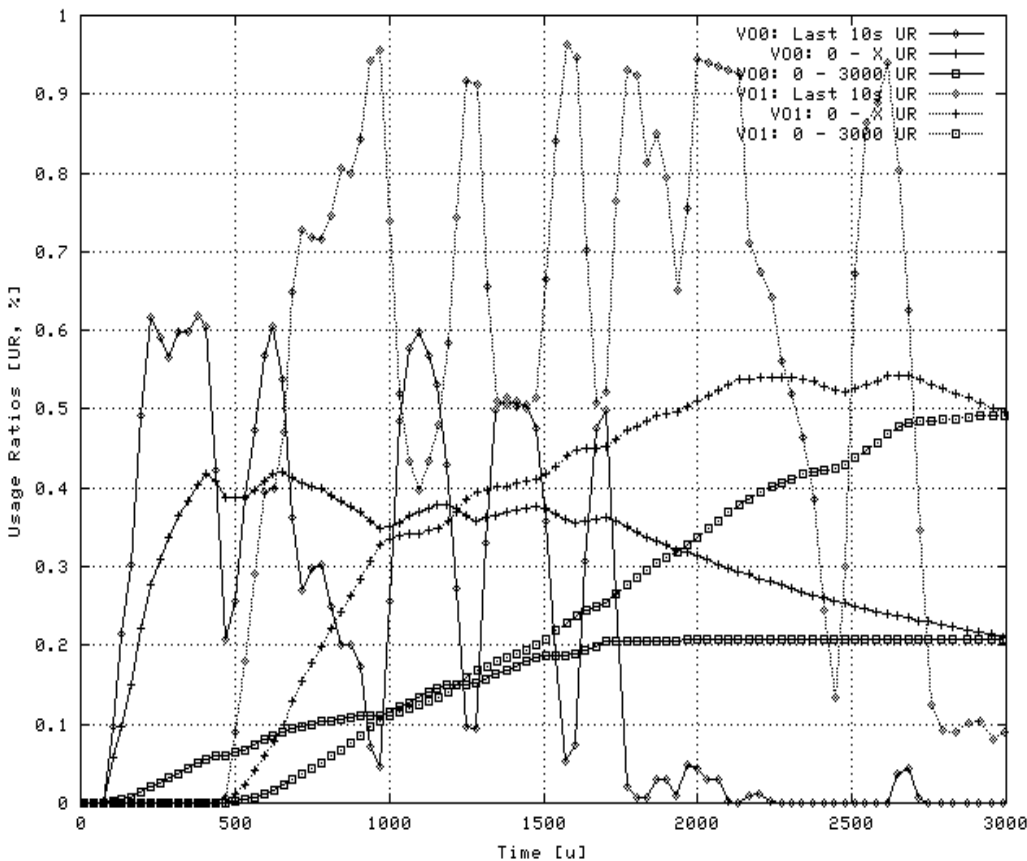
► Gateway Approach



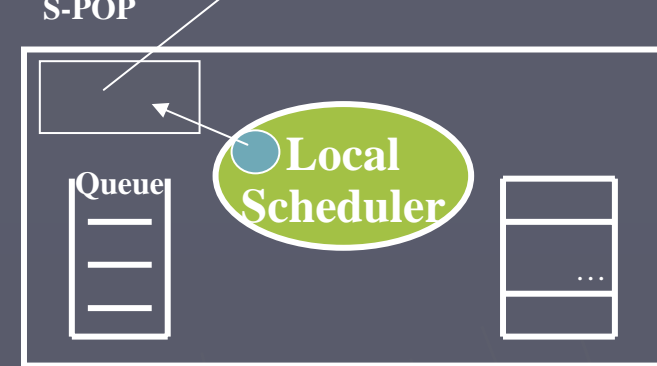
Continuous Monitoring



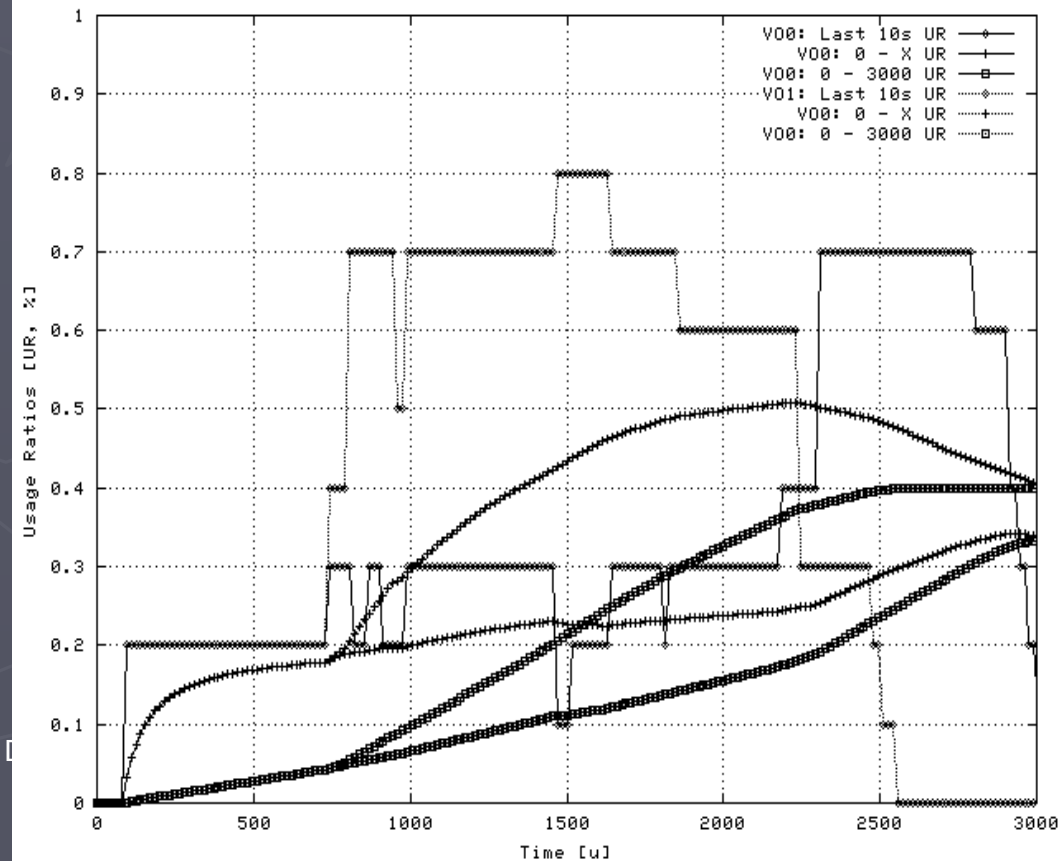
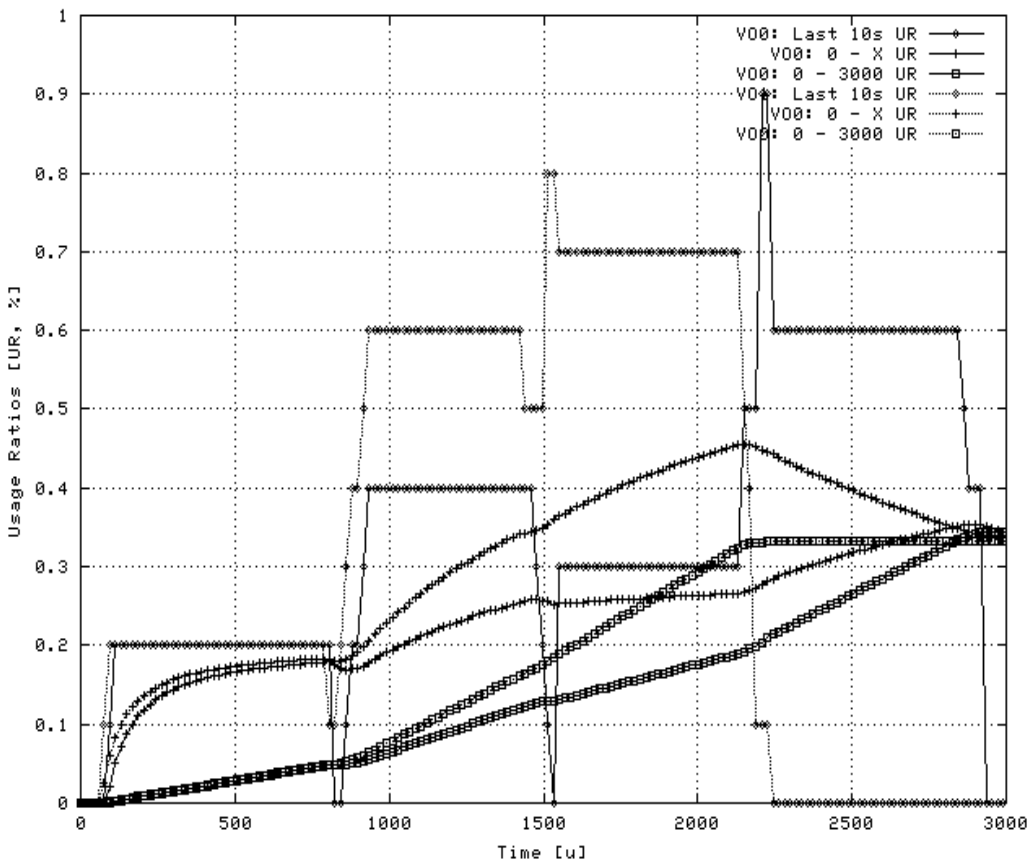
- ▶ RM independent behavior
- ▶ Complex and pluggable Usage SLAs



Gateway Approach



- ▶ Better performances
- ▶ Already familiar to site administrators



Evaluation Methodology

➤ Analytical:

- x Aggregated Response Time (**ART**)
 - x time interval from submission to start
- x Aggregated Site Utilization (**ASU**)
 - x Total of CPU hours burned over the Grid
- x Aggregated Job Completion (**AJC**)
 - x Total number of jobs completed over the Grid

➤ Experimental:

- Real Workloads completion time: **Bio, Atlas, CMS, bTeV.**

Gains

- Additional information that give grid schedulers hints about where to submit jobs – for example, when a site is busy executing a VO's work which had grabbed all resources when they were free
- Time-based entitlement to resources – VOs are guaranteed under different FS policies that they can use resources when they need them instead of maintaining constant workloads

Policy\Metrics	ART	ARU	AJC
RA /NP	97.01	0.027	548
RR /NP	114.99	0.027	549
RA /SC	126.43	0.035	695
RR /SC	130.70	0.036	655

Monitoring Approach - Details

➤ MP: Manager Policies

- the description provided by the person in charge about how resources must be used [*Site Level*]

➤ SC: Administrator Policies

- the technical description written by site's administrator, in short, the RM's configuration files [*Site Level*]

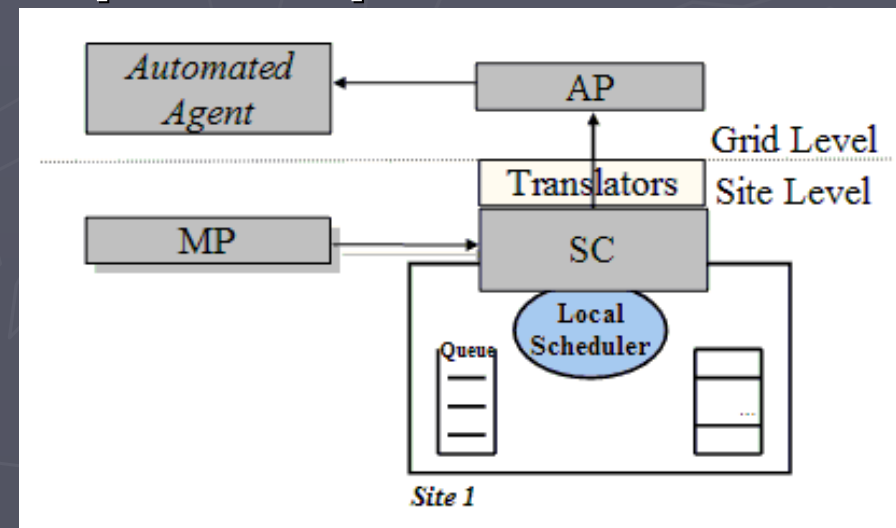
➤ AP: Abstract Policies

- Grid level *understanding* of site policies [*Grid Level*]

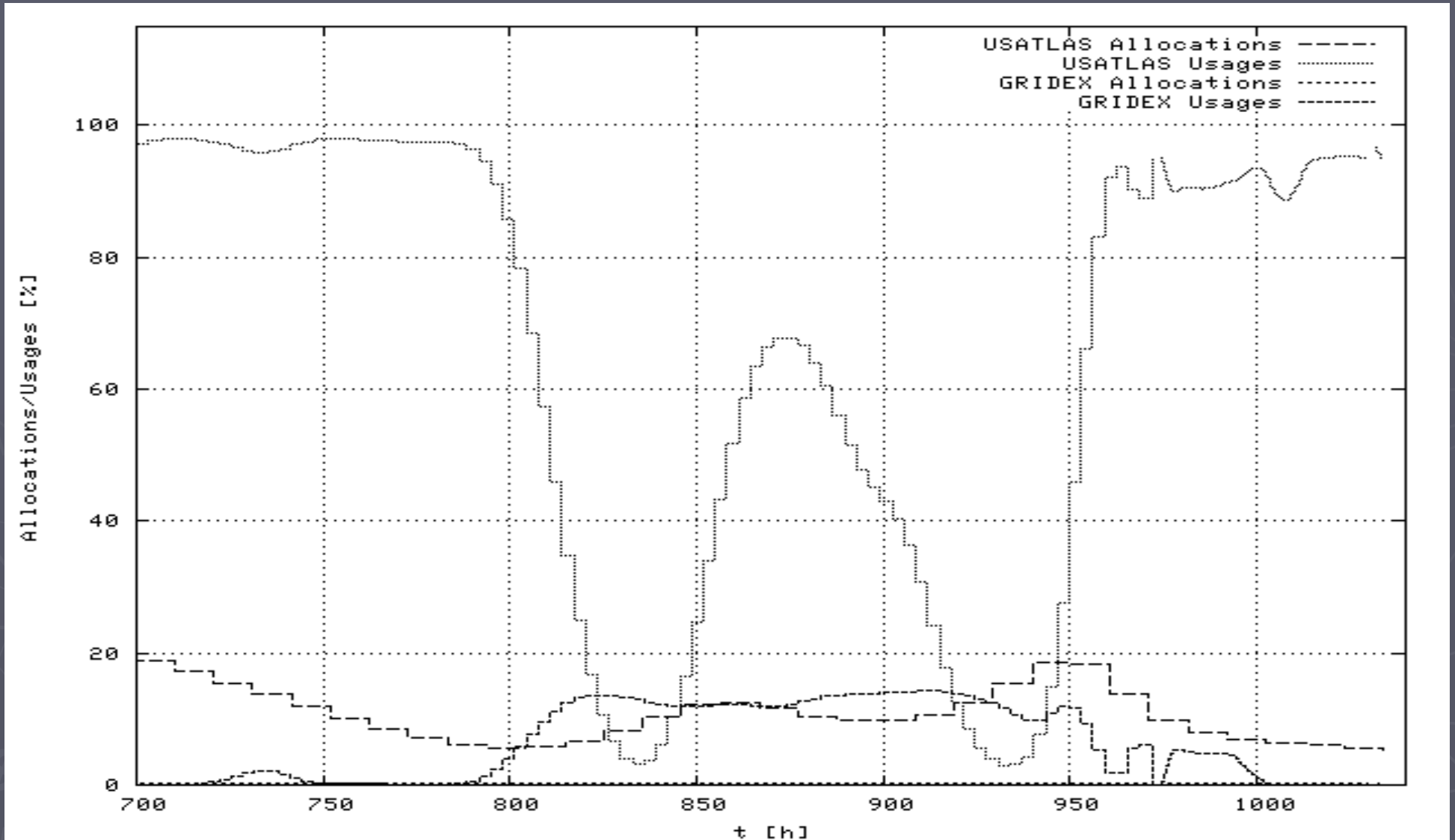
➤ Translators:

- convert SC to AP
- try to abstract to a generic RM model

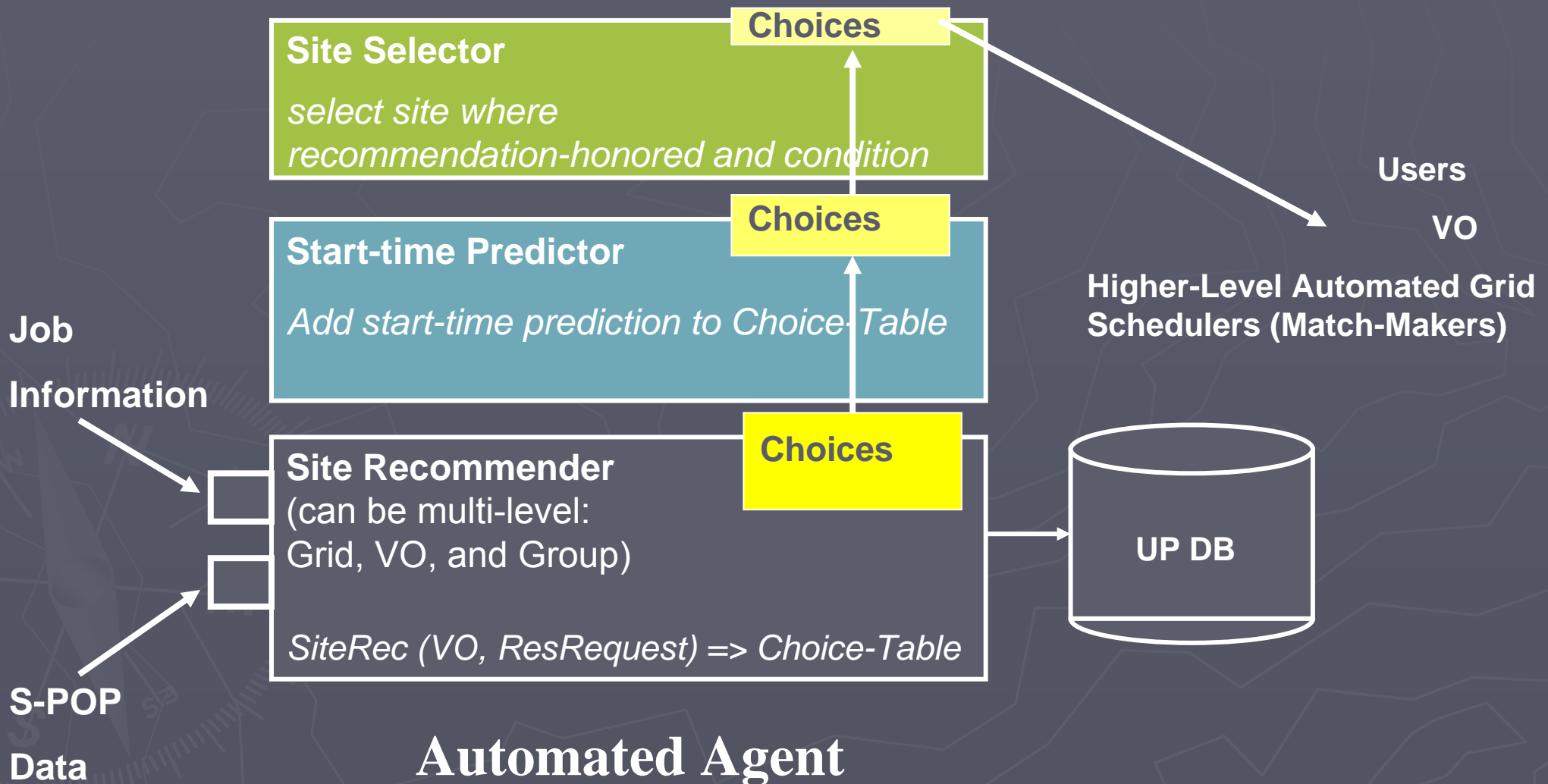
➤ AP + translators + others = Grid Scheduler



Real Site Example



S-PEP Usages and Applications



S-POP -> SiteSel Example

VO	Target	Current	Demand	Level	Accuracy
<i>USCMS</i>	60	50	50	OK	65%
<i>USATLAS</i>	20	15	30	Under	
<i>bTEV</i>	10	10	100	OK	
<i>LIGO</i>	5	3	3	OK	
<i>SDSS</i>	5	22	50	Over	

Resource allocation cases:

- █ UP quota
- █ Under-allocation due to external causes
- █ Over-allocation without contention
- █ Over-allocation with contention
- █ Under-allocation due to internal causes



Open Questions

- Is Usage SLA resource allocation necessary?
- Are there other alternatives?
- Is the proposed model good enough to achieve better resource scheduling?

Thanks!

Questions?