

A new approach to Grid execution environments

CERN openlab II quarterly review
June 19, 2007

Xavier Grehant

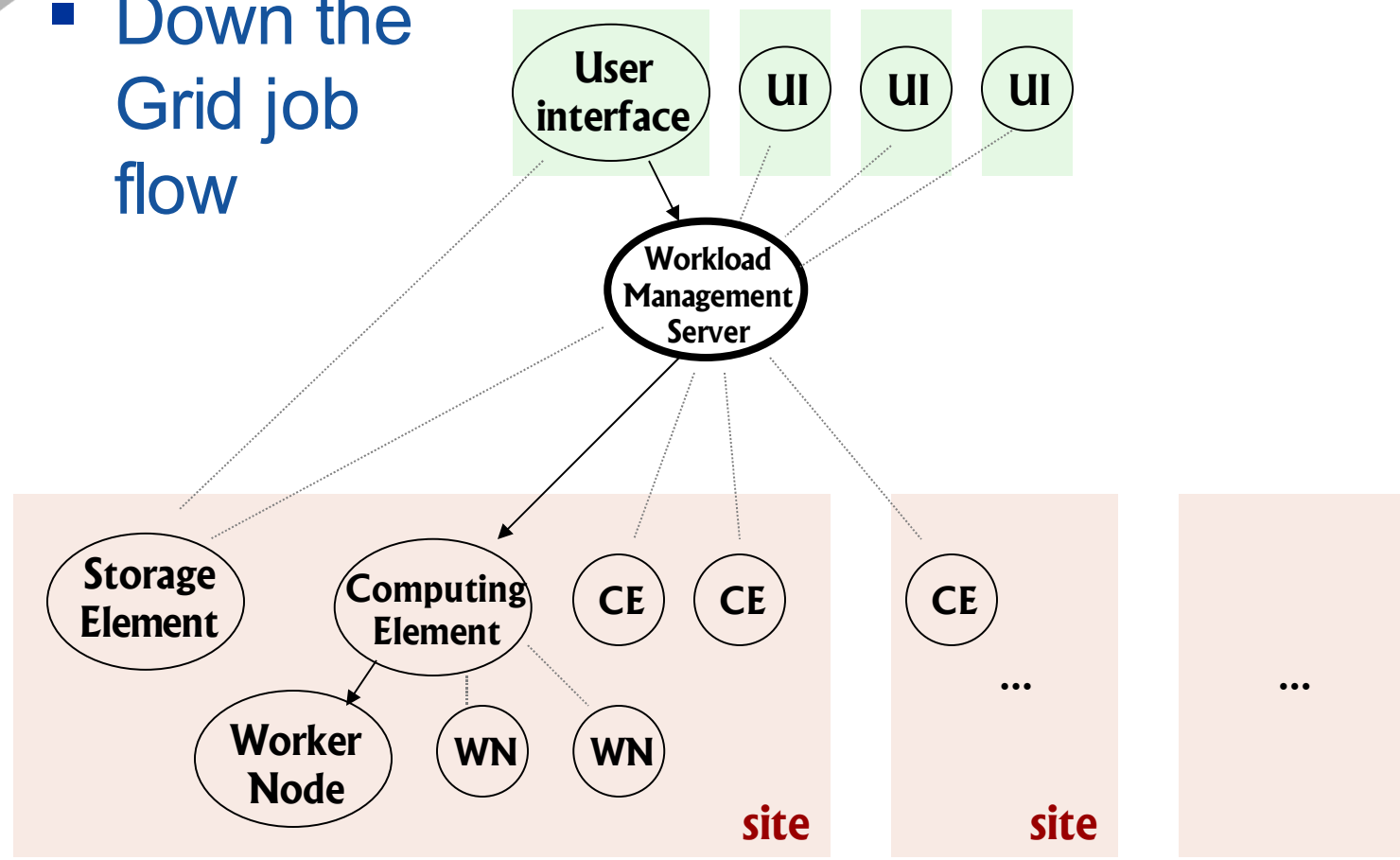


- Facts on Grid execution environments
- Choice on who is in control
- A better compromise is possible

- New trends

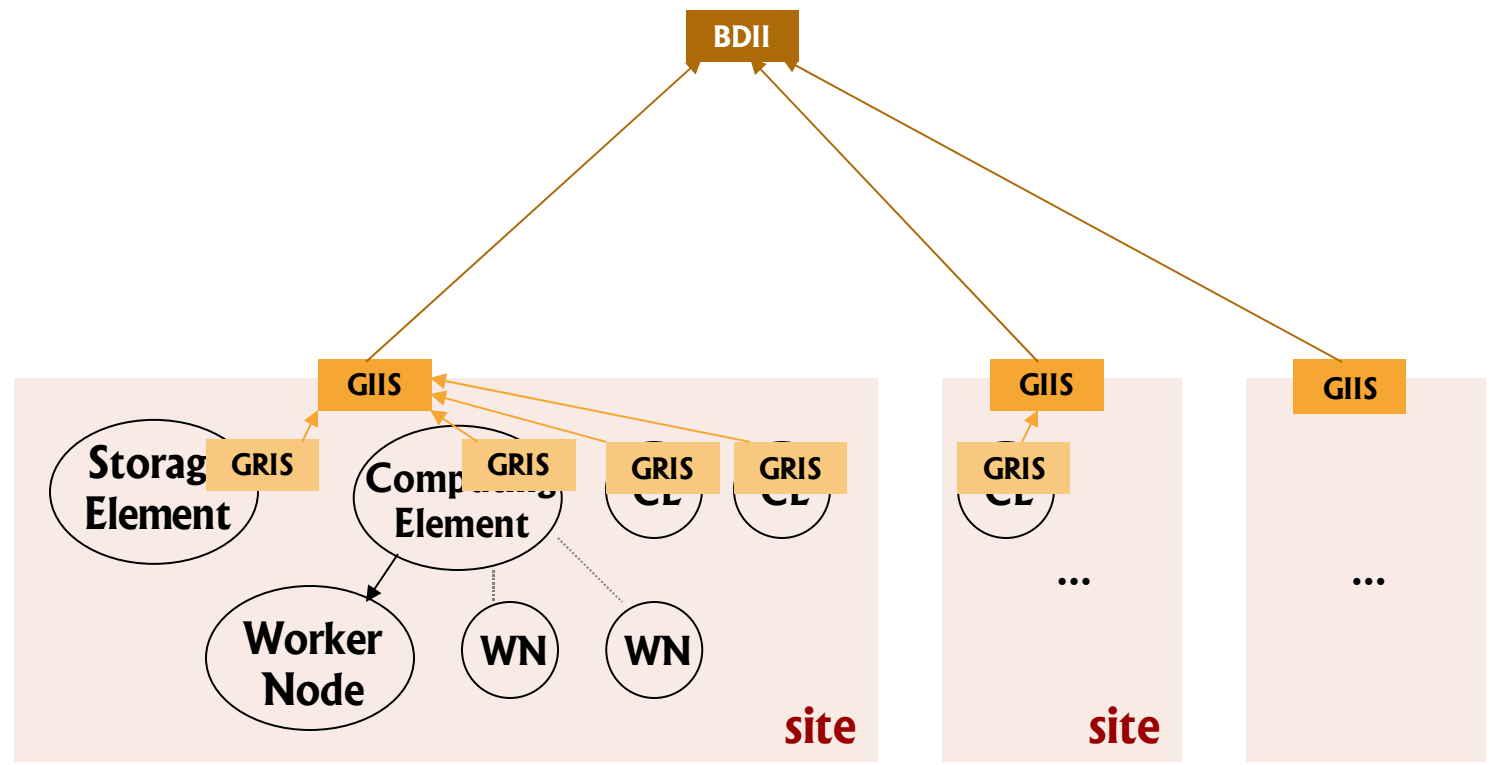
- Grid execution environments:
 - Machines that run the jobs
 - Run a Worker Node (WN)
 - Under a Computing Element (CE)
 - Run monitoring agents
 - Run software from a Virtual Organization (VO)
 - Run the job

- Down the Grid job flow

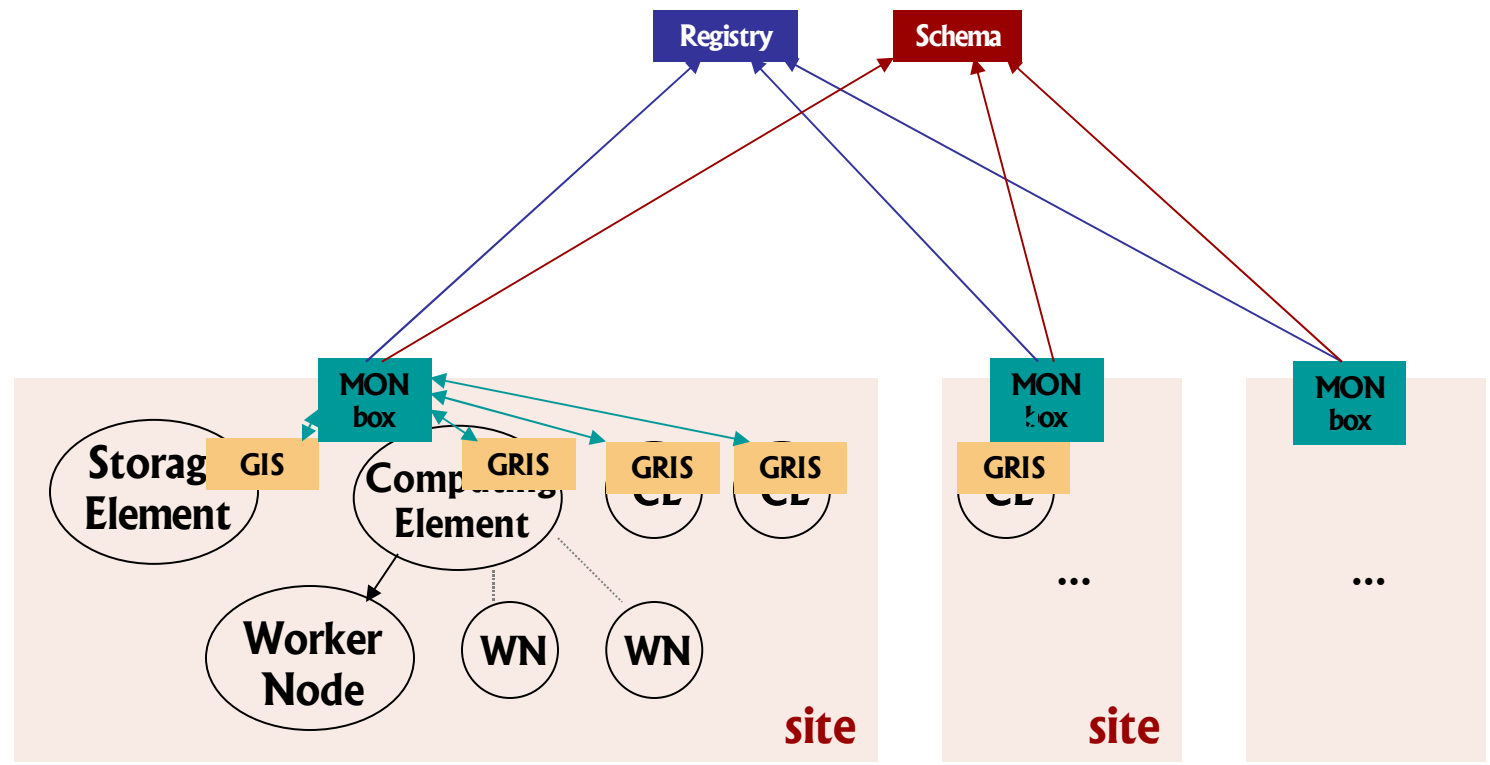
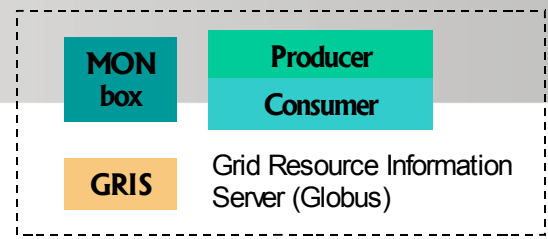


BDII	Berkeley Database Information Index
GIIS	Grid Information Index Server
GRIS	Grid Resource Information Server (Globus)

- Feeds the Grid information flow



- Feeds the Grid information flow





CERN openlab

Facts

INFN - GridICE - Grid Monitoring Service - Mozilla Firefox

File Edit View History Bookmarks Tools Help

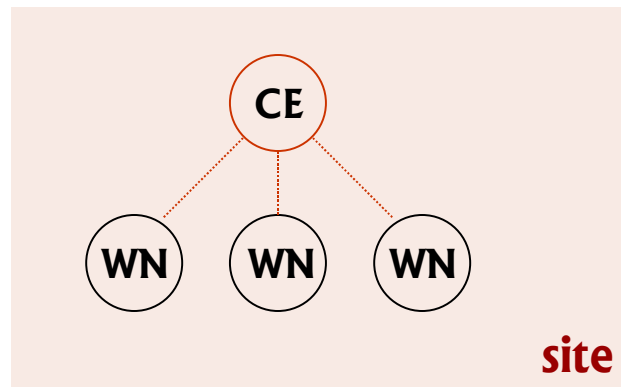
http://gridice2.cnaf.infn.it:50080/gridice/site/site.php

Site ▼	Region	GK#	Q#	RunJob	WaitJob	Computing Resources					Storage Resources			
						JobLoad	Power	WN#	CPU#	CPUload	Available	Total	%	MH#
AMD64.PSNC.PL	CentralEu	1	8	0	4	-	-	-	-	-	4.1 TB	4.3 TB	6%	-
AUVERGRID	France	3	26	75	17	-	-	-	-	-	3.6 TB	3.9 TB	9%	-
Australia-UNIMELB-LCG2	World	1	40	10	0	-	-	-	-	-	2.3 TB	10.2 TB	78%	-
BEIJING-CNIC-LCG2-IA64	CERN	1	7	0	0	0%	19K	8	32	0%	7.4 GB	62.8 GB	68%	11
BEIJING-LCG2	CERN	1	8	0	1	0%	41K	14	28	0%	979.6 GB	2.1 TB	55%	8
BEgrid-KULeuven	NorthEu	1	5	20	22	-	-	-	-	-	-	-	-	-
BEgrid-UGent	NorthEu	1	9	160	0	72%	0	75	208	72%	249.2 GB	367.2 GB	32%	78
BEgrid-ULB-VUB	NorthEu	1	10	65	135	0%	0	1	2	0%	2.9 TB	33.2 TB	91%	3
BG-INSRE	SEE	-	-	-	-	4%	0	14	27	0%	-	-	-	3
BG01-IPP	SEE	2	11	2	294	67%	0	6	9	5%	905.8 GB	1 TB	12%	7
BG02-IM	SEE	1	6	5	6	-	-	-	-	-	11.9 GB	32.9 GB	64%	-
BG04-ACAD	SEE	2	24	30	1544	59%	0	40	34	47%	553.9 GB	637.6 GB	13%	43
BG05-SUGrid	SEE	1	9	6	0	33%	32K	8	24	8%	26.8 GB	83.5 GB	68%	2
BIFI	SWE	1	6	0	0	8%	0	13	26	0%	63.6 GB	103.5 GB	49%	17
BMEGrid	CentralEu	1	9	7	0	-	-	-	-	-	340.9 GB	371.5 GB	8%	-
BUDAPEST	CentralEu	1	9	97	0	93%	0	77	102	63%	6 TB	6.7 TB	10%	7
BelGrid-UCL	NorthEu	1	3	120	0	6%	0	19	72	0%	-	-	-	1
CERN-PROD	CERN	2	32	0	14222208	-	-	-	-	-	647.7 GB	1.7 TB	64%	-
CESGA-EGEE	SWE	1	9	49	0	51%	0	20	120	43%	27.6 GB	68.1 GB	59%	16
CGG-LCG2	France	1	11	56	3	-	-	-	-	-	688.7 GB	952.3 GB	28%	-
CIEMAT-LCG2	SWE	2	12	319	75	77%	455K	105	334	85%	7.3 TB	34.2 TB	79%	65
CNB-LCG2	SWE	1	7	0	71	0%	0	7	14	0%	12.4 GB	67.7 GB	82%	10
CNR-ILC-PISA	Italy	1	6	2	0	25%	3K	2	4	25%	754 GB	763.2 GB	1%	4
CREAM-PADOVA	Italy	1	1	0	0	-	-	-	-	-	-	-	-	-
CSC	World	1	4	0	496	-	0	0	0	-	65.9 GB	67.8 GB	3%	2
CSCS-LCG2	Ger/Swi	1	7	114	106	99%	0	39	122	2%	34.3 TB	54.5 TB	47%	5
CY-01-KIMON	SEE	1	8	9	0	14%	0	31	74	15%	2.2 TB	3 TB	28%	35
CY-03-INTERCOLLEGE	SEE	1	3	8	0	-	-	-	-	-	-	-	-	-
CYFRONET-IA64	CentralEu	1	13	36	13	-	-	-	-	-	267.3 GB	2 TB	67%	-
CYFRONET-LCG2	CentralEu	1	17	60	1	65%	251K	133	266	73%	16.5 TB	29.9 TB	45%	136

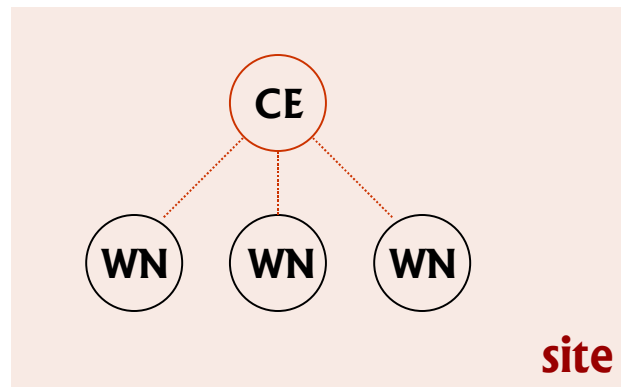
start | C:\Document... | Microsoft Po... | INFN - GridI... | Inkscape -- O... | EN | 3:38 PM

- Worker nodes behind the CE
- Two-steps job scheduling, push model
- Resource control by node manager

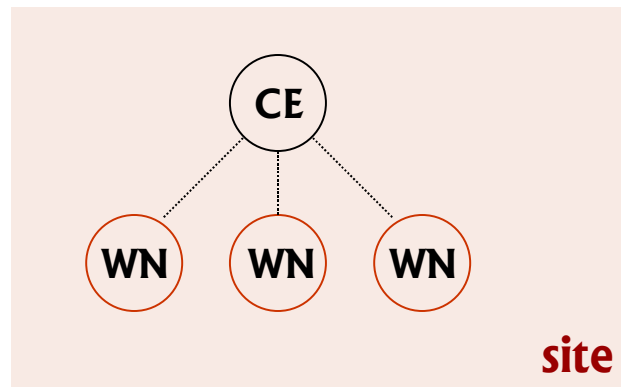
- Batch system choice by node manager
 - Various batch system flavours and fonctionnalités (LSF, PBS, SGE, Condor)
 - Grid-global scheduling (Condor-G) intersects the fonctionnalités



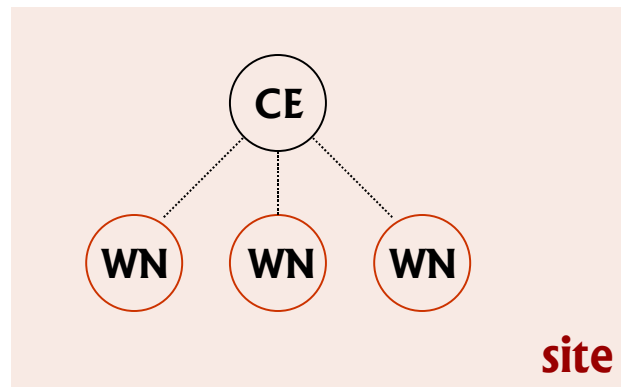
- Allocation policies by node manager
 - Various local resource allocation policies
 - No consistent Grid-global end-resource allocation strategy



- Configuration by node manager
 - Node manager does user software maintenance work
 - User has no guarantee that configuration is appropriate (Black holes)

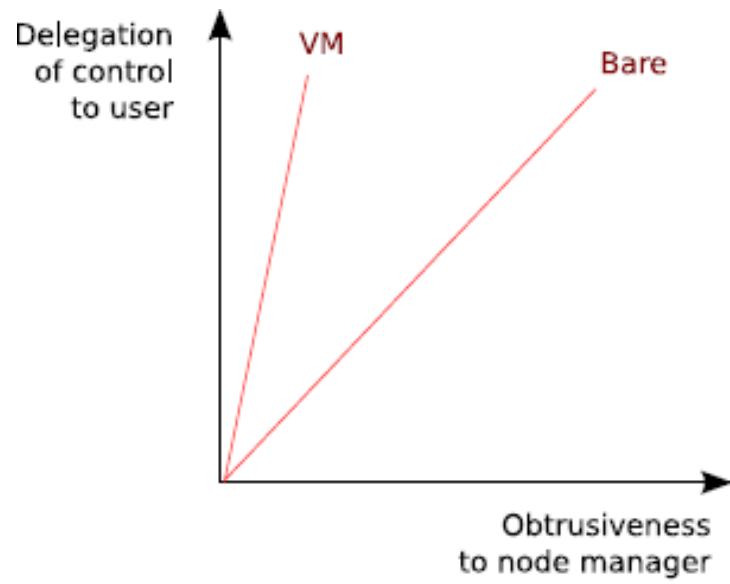


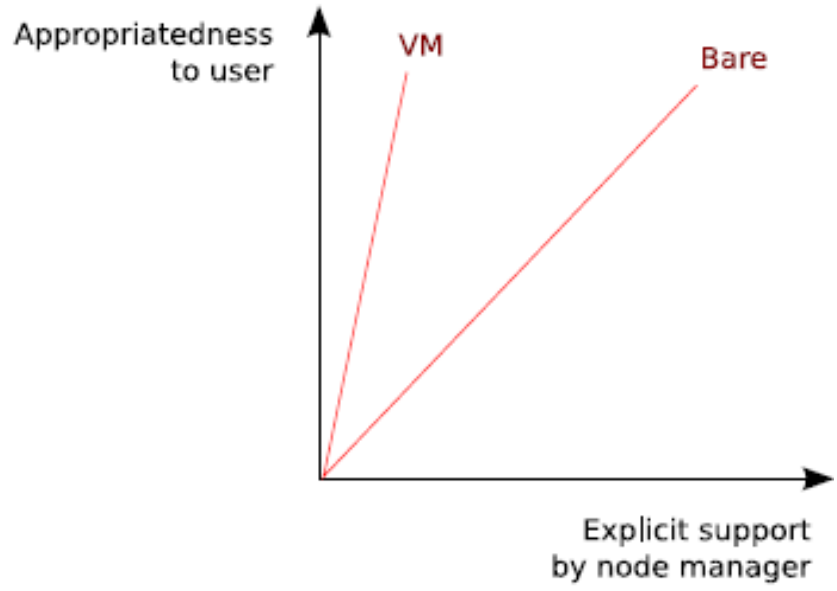
- Interventions by node manager
 - Service disruption for site maintenance (security, upgrades)
 - Or poor site maintenance efficiency
 - Node manager intervenes on user problems



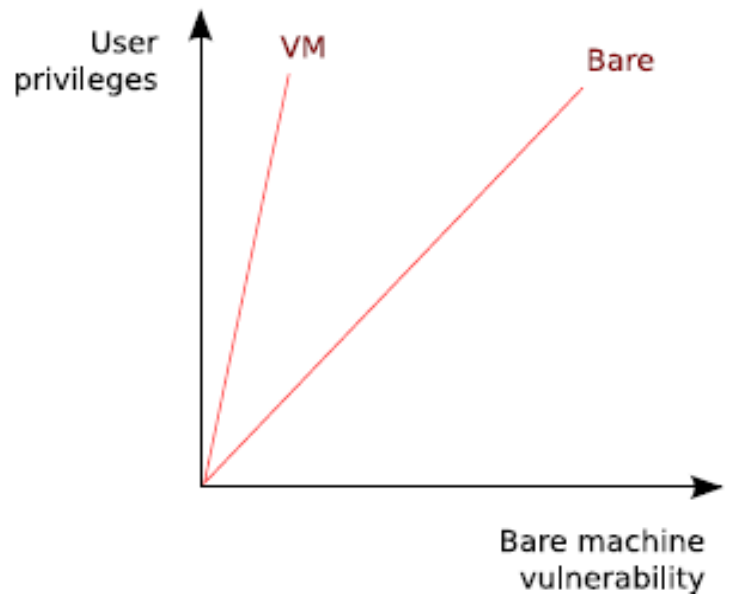
- No resource delegation to Grid user
 - “This is production Grid
 - Service quality is not a problem
 - Sysadmins are dedicated to us
 - Security is a problem”
- Provide resource without resource control?
 - Unbalanced solution to a compromise
- Note: a look elsewhere
 - On PlanetLab, same compromise, different solution: the user gets full control

Improve compromise





Improve compromise



- Cronus: users care about control
 - Atlas resource allocation system on long-reserved nodes
 - Constantly around 2500 nodes on EGEE, OSG, NorduGrid

- Thank you for your attention
- Questions?