

Eugen Mudnić
Technical university Split -FESB

Grid simulation
(AliEn)

Outline

- GRID simulation
- Simulation tool – Ptolemy (Berkeley)
- Simulation architecture
- Simulation software environment
- Some results

GRID simulation

- Why GRID simulation:
 1. Validating grid architecture and scalability
 2. System resources planning (storage size, network throughput, processing power)
 3. Simulating different job placement and file transfer scheduling scenarios
 4. GRID services tuning
 5. Designing and testing of user strategies (how to efficiently use the system)

Modeling and simulation

- **Top-down** or Bottom-up ?
- Simulated system is very large and complex (and still evolving)
- Appropriately high degree of abstraction
- Focusing on the Grid key components
 - Storage elements
 - Network
 - Computing elements
- Method : Discrete event simulation (DES)

DES tool - Ptolemy

- Ptolemy Classic (Berkley)

- Pros:

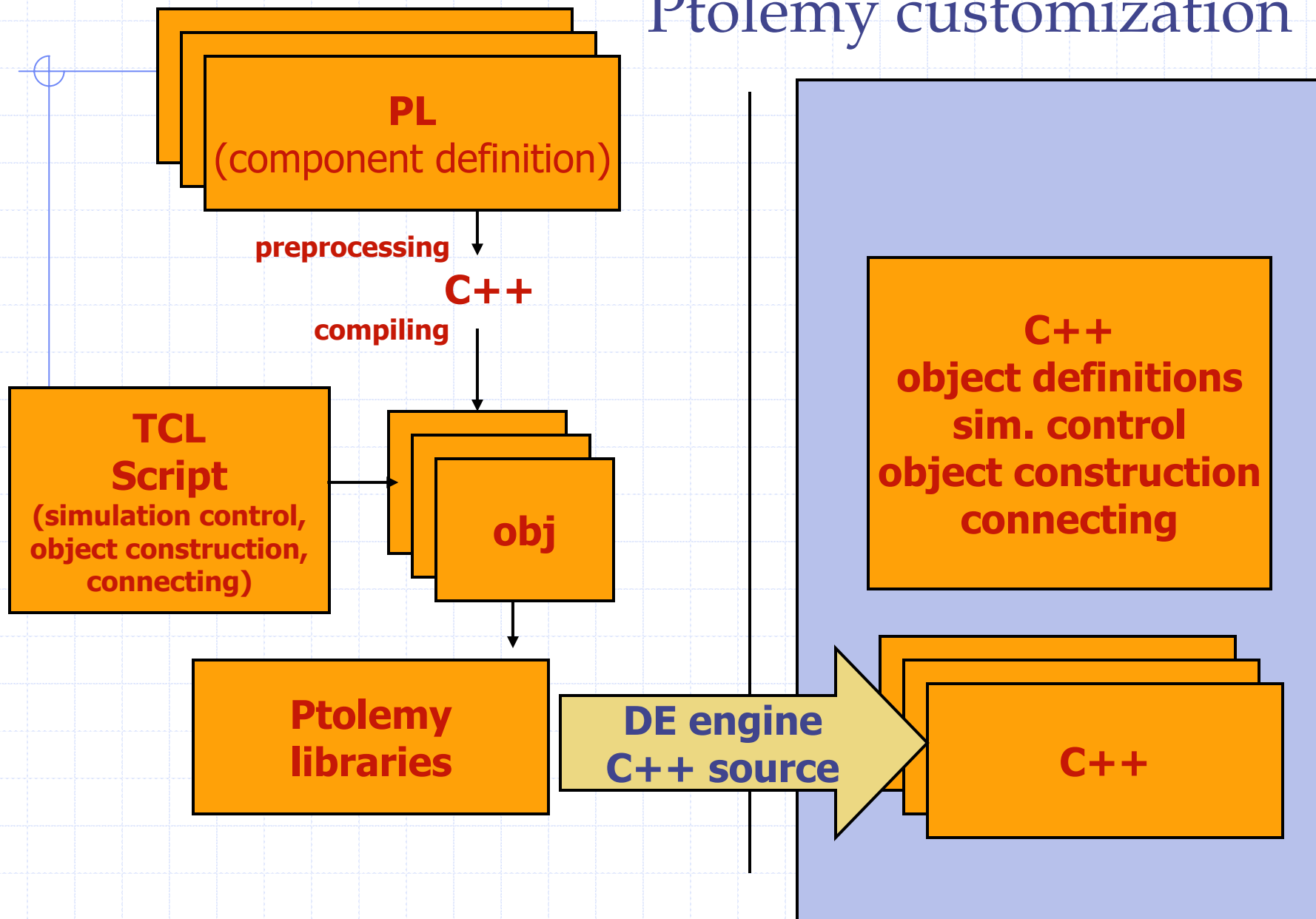
- ◆ Fast (C++), stable
- ◆ Good DE library (component and messaging system)
- ◆ Free , source code available

- Cons:

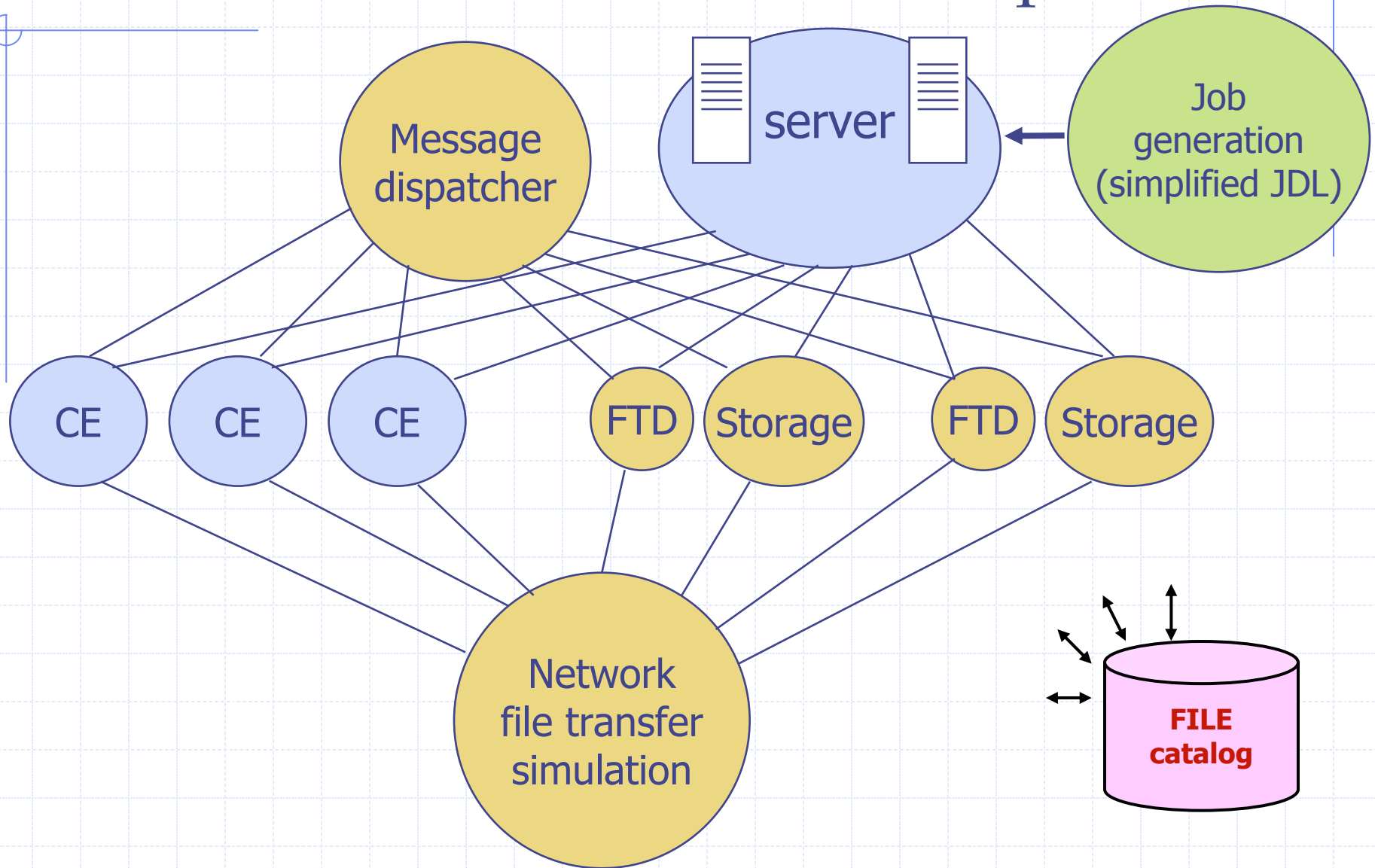
- ◆ Complicated (CPP,TCL,PL(Ptolemy language),...)
- ◆ Resulting simulation needs Ptolemy installation
- ◆ Ask other Ptolemy users



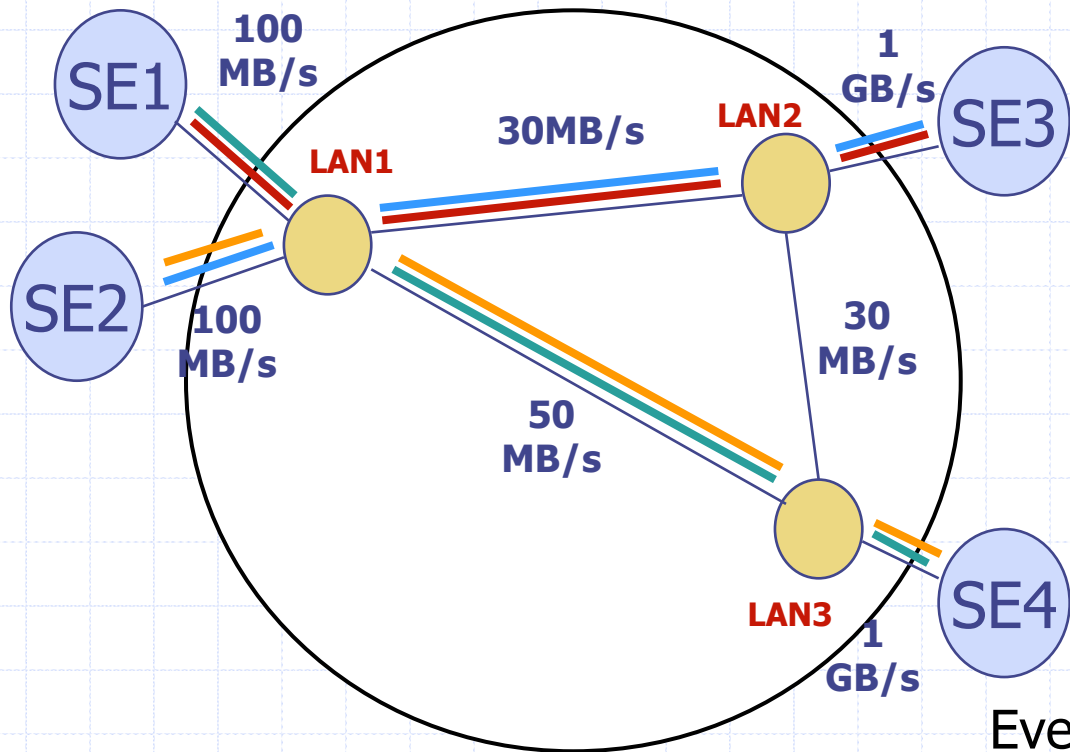
Ptolemy customization



AliEn DES model & DE components



Network file transfer calculation component



links with capacity (C_1, \dots, C_L)

N streams , every stream has a predefined transfer route $r_i \subset \{1, \dots, L\}$

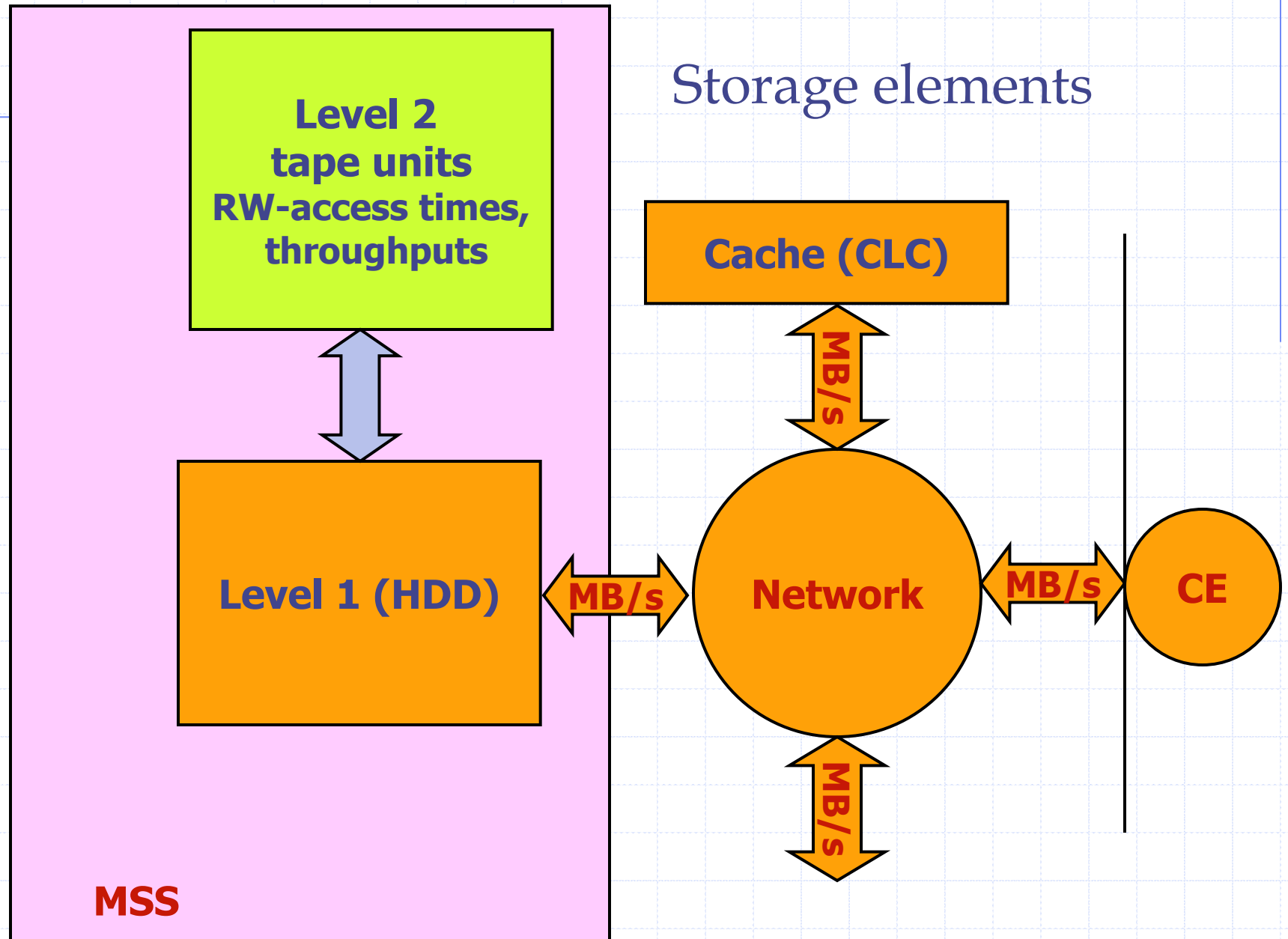
Every stream has equal priority

Network as a set of links shared by changeable number of data streams

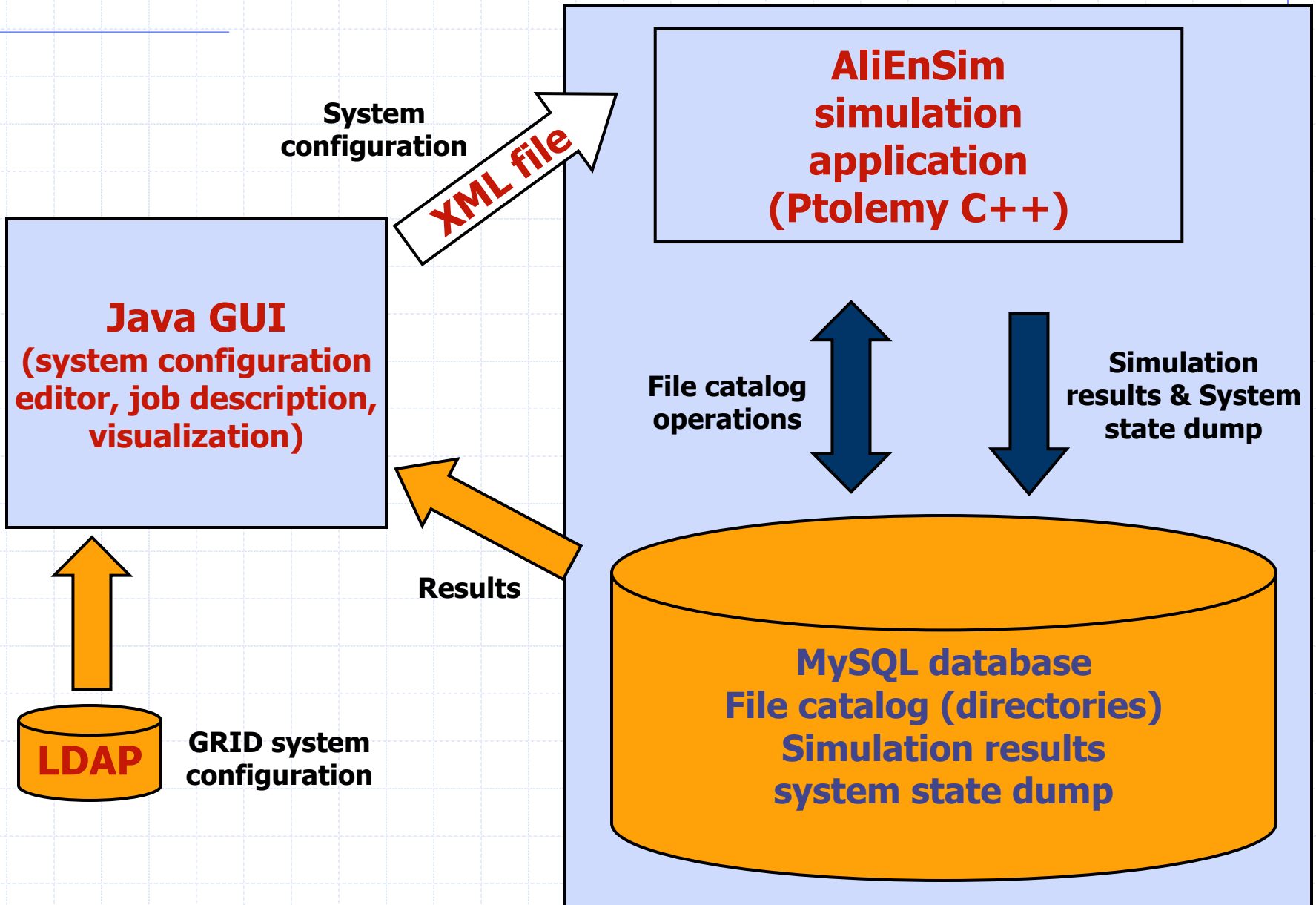
Stream bandwidth allocation must conform to:

$$\sum_{r_i \ni l} \phi_i(x) \leq C_l$$

Storage elements



Software environment



Java GUI – grid configuration

The screenshot shows the AliEnSim Java GUI. The window title is "AliEnSim". The menu bar includes "File", "Look & Feel", "Themes", and "Tool Tips". The toolbar contains buttons for "SimCon", "Par", "CE", "SE", "FTD", "CE->SE", "FTD->SE", "FTG", "JG", "LAN", "Link", "TASK", "TRAN", and "Visual". Below the toolbar is a set of navigation buttons: left arrow, right arrow, plus, minus, checkmark, close, and refresh. The main area displays a table with the following data:

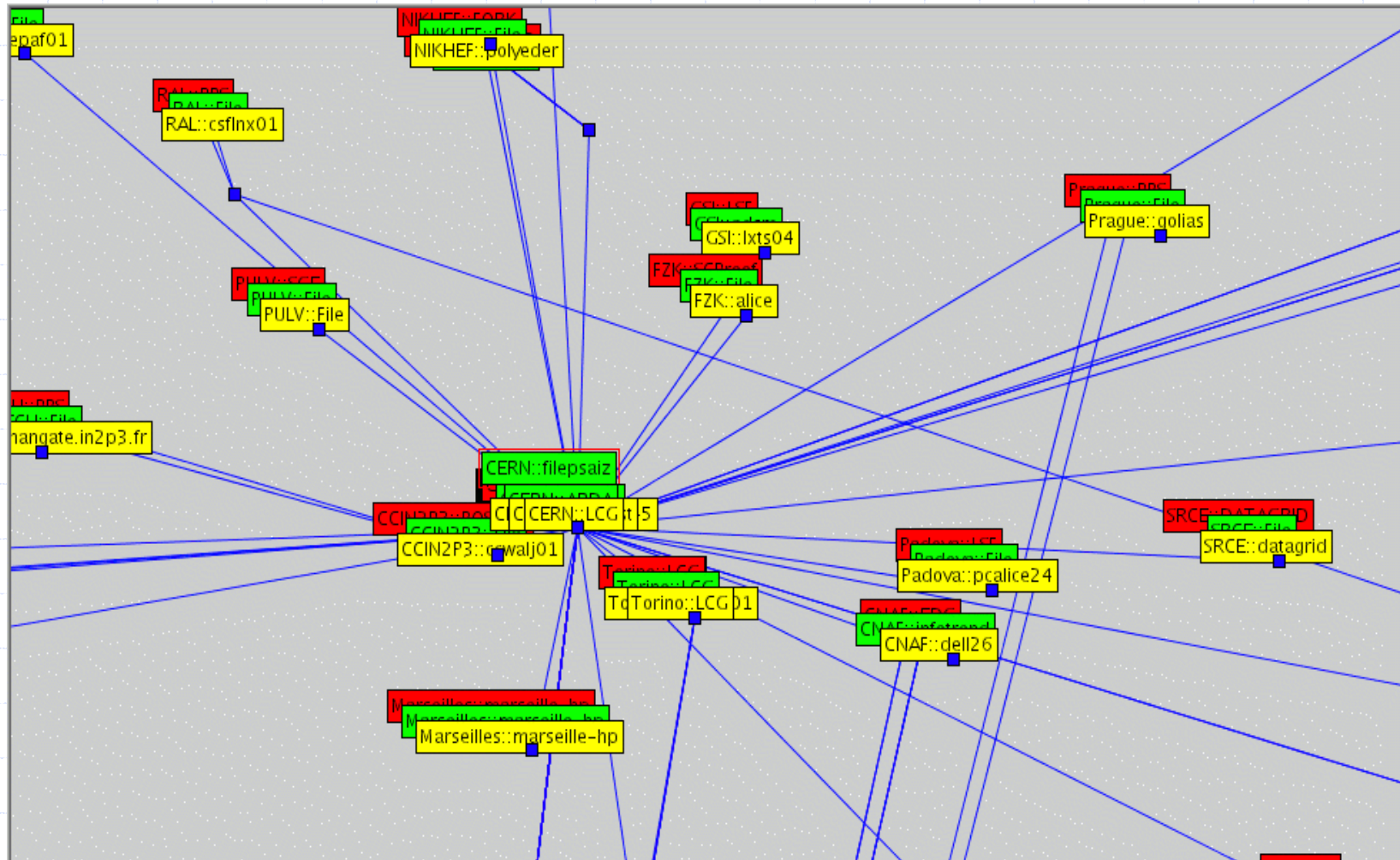
	CEId	name	maxJobs	active	host	LAN	bandwidth	execute
1	403	Alice::FZK::PBS	250	1	alice.fzk.de	fzk.de	1,000	EX1
2	404	Alice::Catania::PBS	60	1	grid012.ct.infn.it	ct.infn.it	1,000	EX1
3	407	Alice::LBL::LSF	100	1	pdsflx002.nersc.gov	nersc.gov	1,000	EX1
4	405	Alice::CCIN2P3::BQS	200	1	ccwali48.in2p3.fr	in2p3.fr	1,000	EX1
5	413	Alice::CNAF::PBS	200	1	wn-04-10-16-a	cr.cnaf.infn.it	1,000	EX1
6	416	Alice::SRCE::DATAGRID	3	1	dgrid-1.srce.hr	srce.hr	1,000	EX1
7	451	Alice::Bari::PBS	8	1	alicegrid2.ba.infn.it	ba.infn.it	1,000	EX1
8	448	Alice::Torino::LCG	300	1	ui-alien.to.infn.it	to.infn.it	1,000	EX1
9	447	Alice::CERN::LCG	800	1	ln1175.cern.ch	cern.ch	1,000	EX1
10	438	Alice::Cyfronet::PBS	20	1	zeus05.cyf-kr.edu.pl	cyf-kr.edu.pl	1,000	EX1
11	422	Alice::Prague::PBS	36	1	goliath.farm.particle.cz	farm.particle.cz	1,000	EX1
12	420	Alice::OSC::PBS	40	1	alice.osc.edu	osc.edu	1,000	EX1
13	401	Alice::JINR::PBS	30	1	lxpub02.jinr.ru	jinr.ru	1,000	EX1
14	400	Alice::Torino::PBS	60	1	alifarm01.to.infn.it	to.infn.it	1,000	EX1
15	399	Alice::Padova::LSF	20	0	pcalice24.pd.infn.it	pd.infn.it	1,000	EX1
16	402	Alice::GSI::LSF	16	0	lxts04.gsi.de	gsi.de	1,000	EX1
17	411	Alice::CCIN2P3::BQSTest	10	0	ccwali48.in2p3.fr	in2p3.fr	1,000	EX1
18	408	Alice::CERN::LXSHARE	200	0	lxplus005.cern.ch	cern.ch	1,000	EX1
19	406	Alice::CapeTown::PBS	20	0	qgp3.phy.uct.ac.za	phy.uct.ac.za	1,000	EX1

At the bottom of the window, there are two buttons: "LDAP import" and "SetLan".

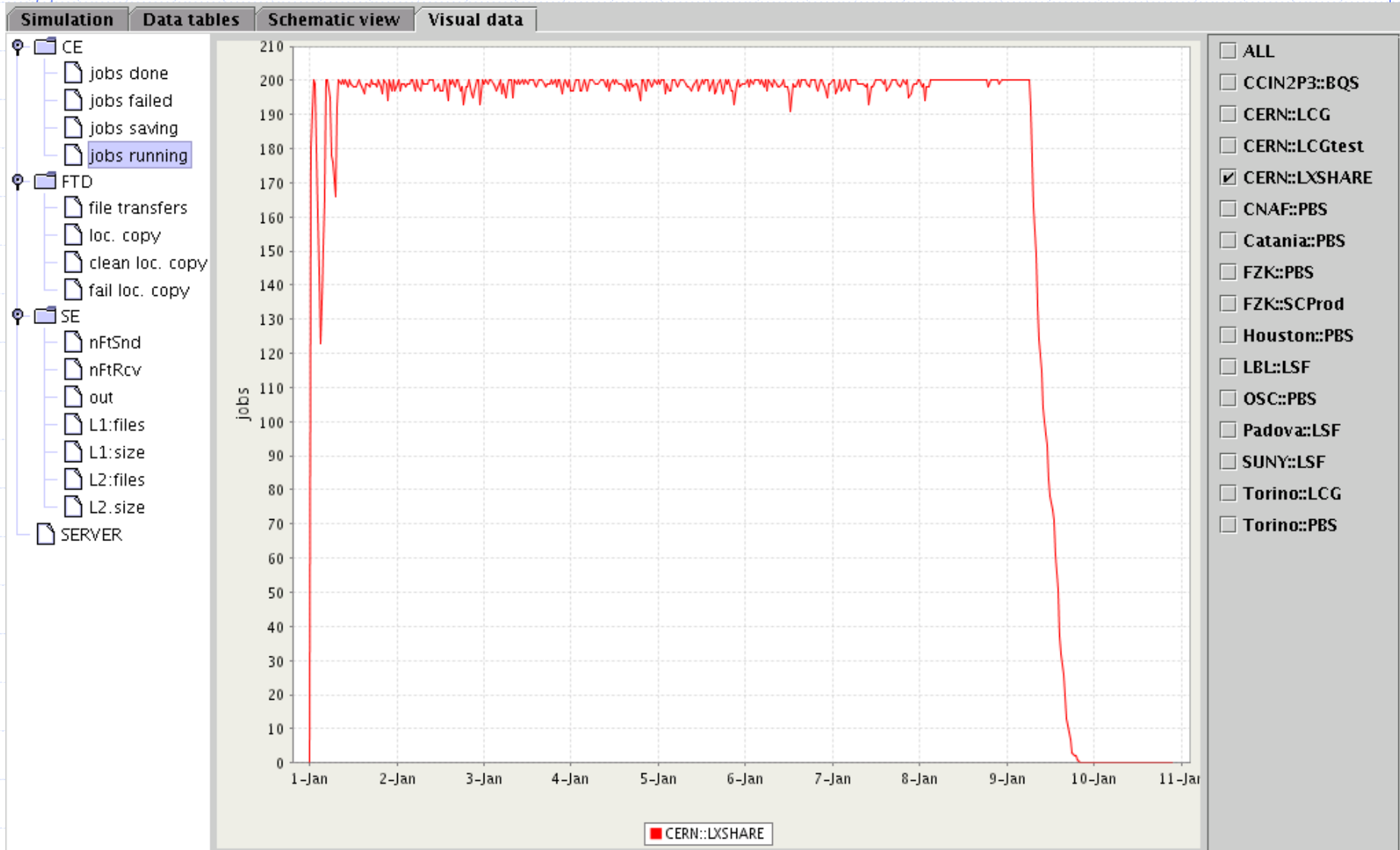
Java GUI – job description

Simulation		Data tables				Schematic view			Visual data		
PARAM	CE	SE	FTD	CLC	LAN	LANLINK	JG	FG	FileType		
id	name	split				outSEID1	outFilelds1	outSEID2	outFilelds2	events	
1	phase1-cent	PRODUCTION:RUN:EVENT				CERN::Castor	1,2,3,4,5,6,7,...	CERN::scratch	25,26,27,28,...	1-500	
3	phase1-per1	PRODUCTION:RUN:EVENT				CERN::Castor	1,2,3,4,5,6,7,...	CERN::scratch	25,26,27,28,...	1-500	
4	phase1-per2	PRODUCTION:RUN:EVENT				CERN::Castor	1,2,3,4,5,6,7,...	CERN::scratch	25,26,27,28,...	1-500	
5	phase1-per3	PRODUCTION:RUN:EVENT				CERN::Castor	1,2,3,4,5,6,7,...	CERN::scratch	25,26,27,28,...	1-500	
6	phase1-per4	PRODUCTION:RUN:EVENT				CERN::Castor	1,2,3,4,5,6,7,...	CERN::scratch	25,26,27,28,...	1-500	
7	phase1-per5	PRODUCTION:RUN:EVENT				CERN::Castor	1,2,3,4,5,6,7,...	CERN::scratch	25,26,27,28,...	1-500	
2	phase2	FILE					1,7,9,20			1	

Java GUI – grid visualization(control)



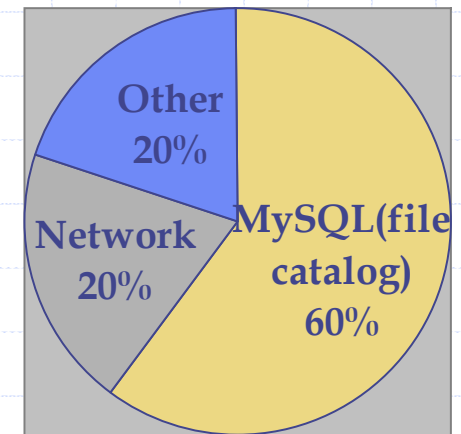
Results example (jobs running)



Some performance testing results

- **Simulation of Alice PDCIII-Phase I:**
 - 104000 jobs execution , output is 3.8mil files (36 per job)
 - 15 sites (total 2100 CPU)
 - central storage at CERN::Castor(MSS) & CERN::scratch
- **Simulated real time: 10days**
- **Simulation time (P4-2GHz):**
 - ~1h

Simulation time



Future work ?

- Test and enhance component models
 - Identify and use different system measuring
 - ◆ **Network – throughput**
 - ◆ SE (MSS) : throughput, R/W access times
- Optimize,optimize,...
 - Faster real system (more computing elements) <-> slower simulation
- Simulate and compare with Alice data processing PDCIII phase I -phase III