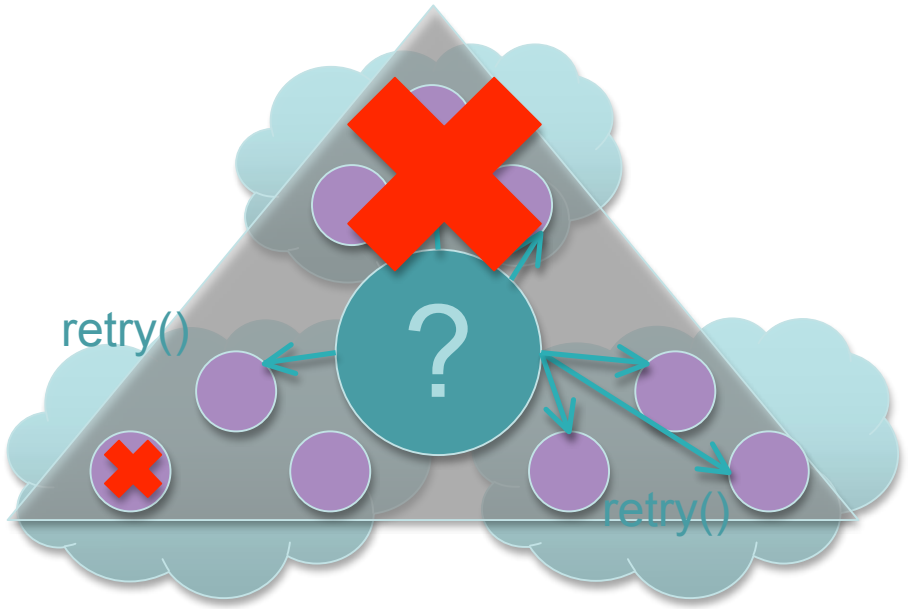
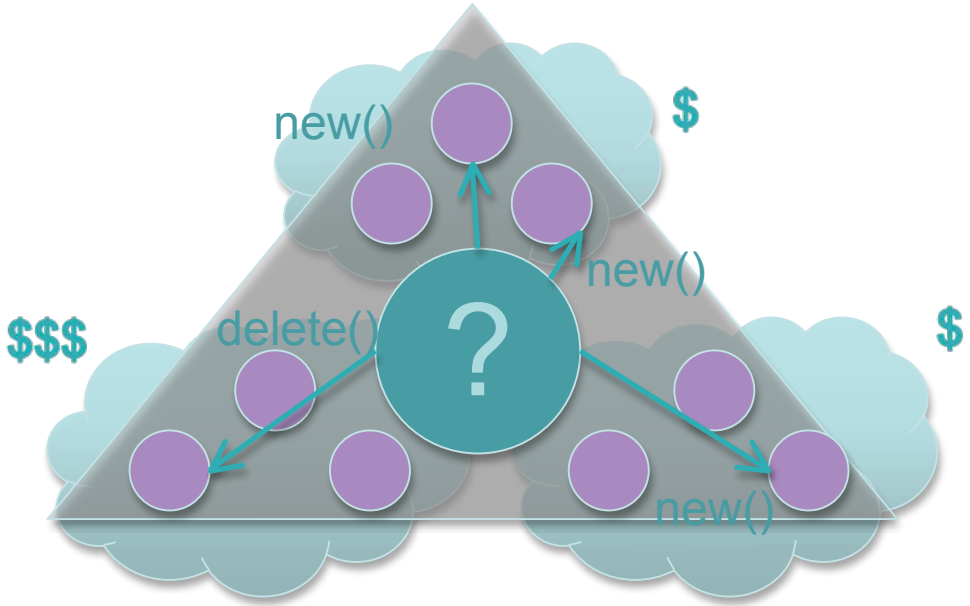
Fault-tolerance

- Handle faults
  - node failures
  - process failures
  - network failures
  - …



Adaptability to resource dynamicity

- Resource status changes over time
  - Resource availability
  - Resource price (e.g. EC2 spot instances)



Sky Computing Toolkit

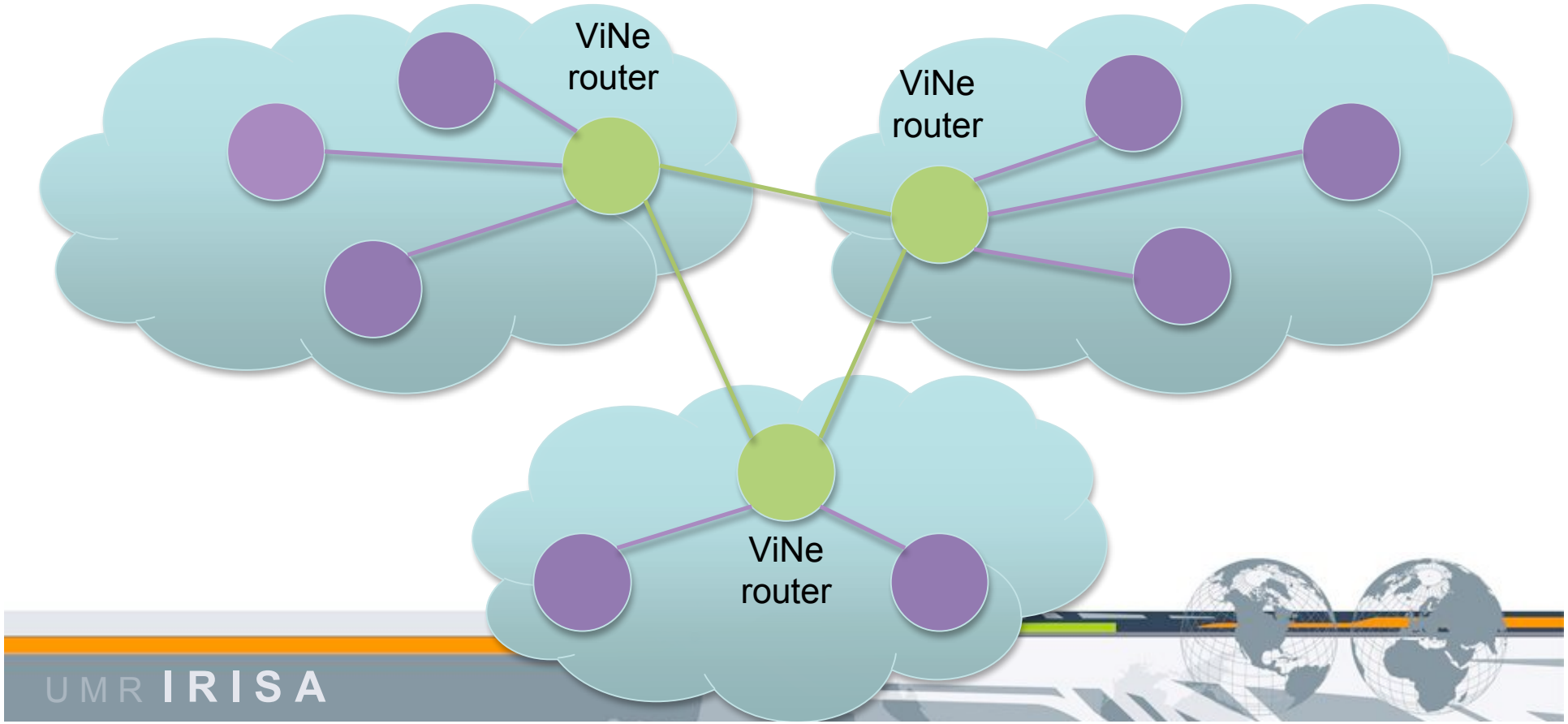
- Nimbus
  - Resource management
  - Contextualization
- ViNe
  - All-to-all connectivity
- Hadoop
  - Task distribution
  - Fault tolerance
  - Resource dynamicity

Nimbus

- Cloud computing Toolkit
  - Open source Infrastructure-as-a-Service implementation (Amazon EC2-compatible)
  - Higher-level tools, e.g. Cluster contextualization
    - Allocation of a complete cluster with different roles
    - Example: Hadoop master + Hadoop slaves
    - Targets Clouds for Science
  - Cumulus (Amazon S3-compatible storage cloud)
  - Fast propagation mechanisms (HDFS, etc.)

ViNe

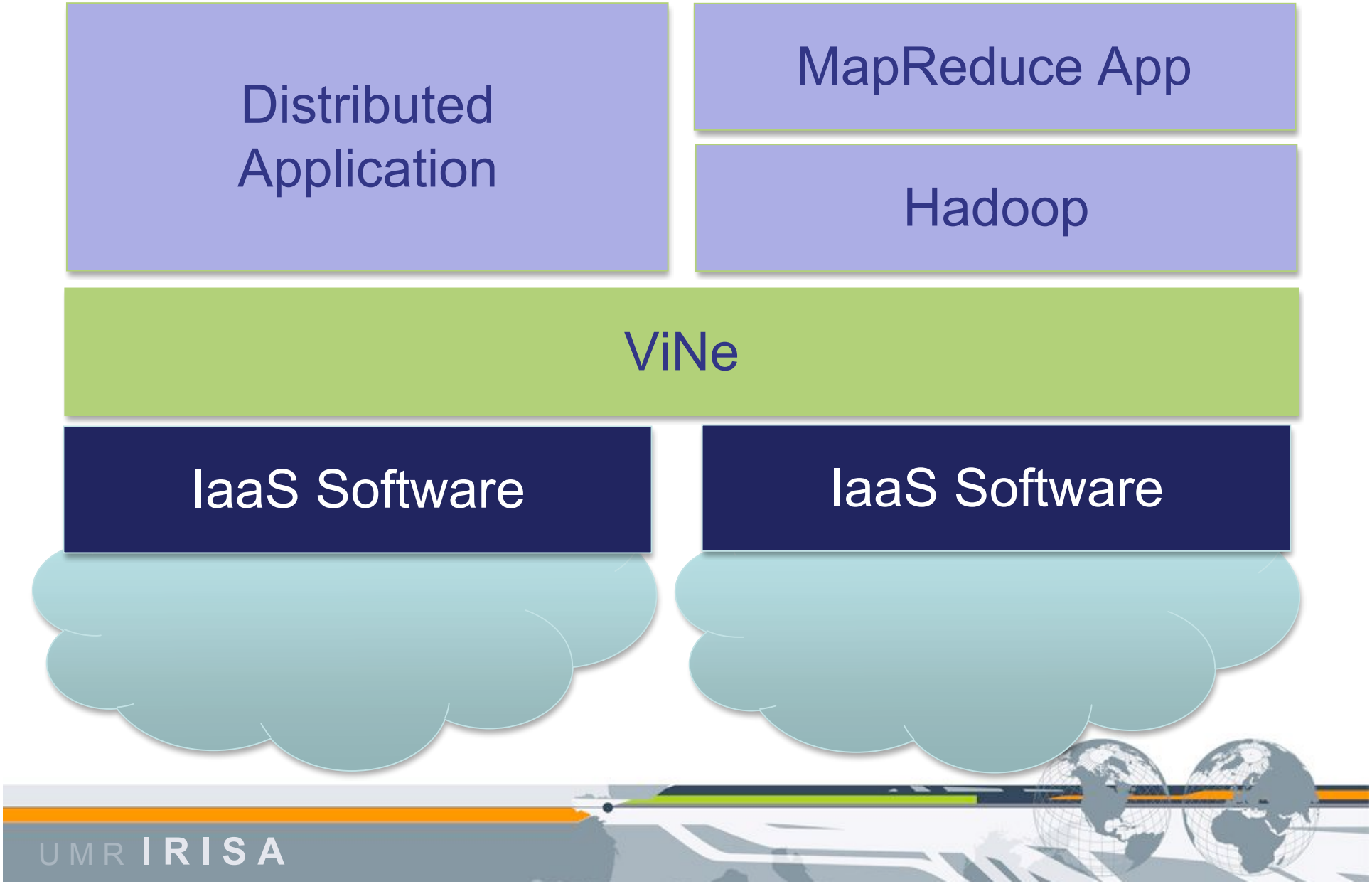
- High performance virtual network
- All-to-all connectivity



Hadoop

- Open-source MapReduce implementation
- Used by many companies (Yahoo…)
- Efficient framework for distribution of tasks
- Fault-tolerance
- Distributed file system (HDFS)

Sky Computing Architecture



LARGE-SCALE EXPERIMENTS

Large-scale Sky Computing

- Large-scale experiments using Nimbus
  - Grid'5000 2010 Spring School Challenge (April 2010)
  - OGF-29 demo (June 2010)
- Using experimental testbeds
  - Grid'5000
  - FutureGrid

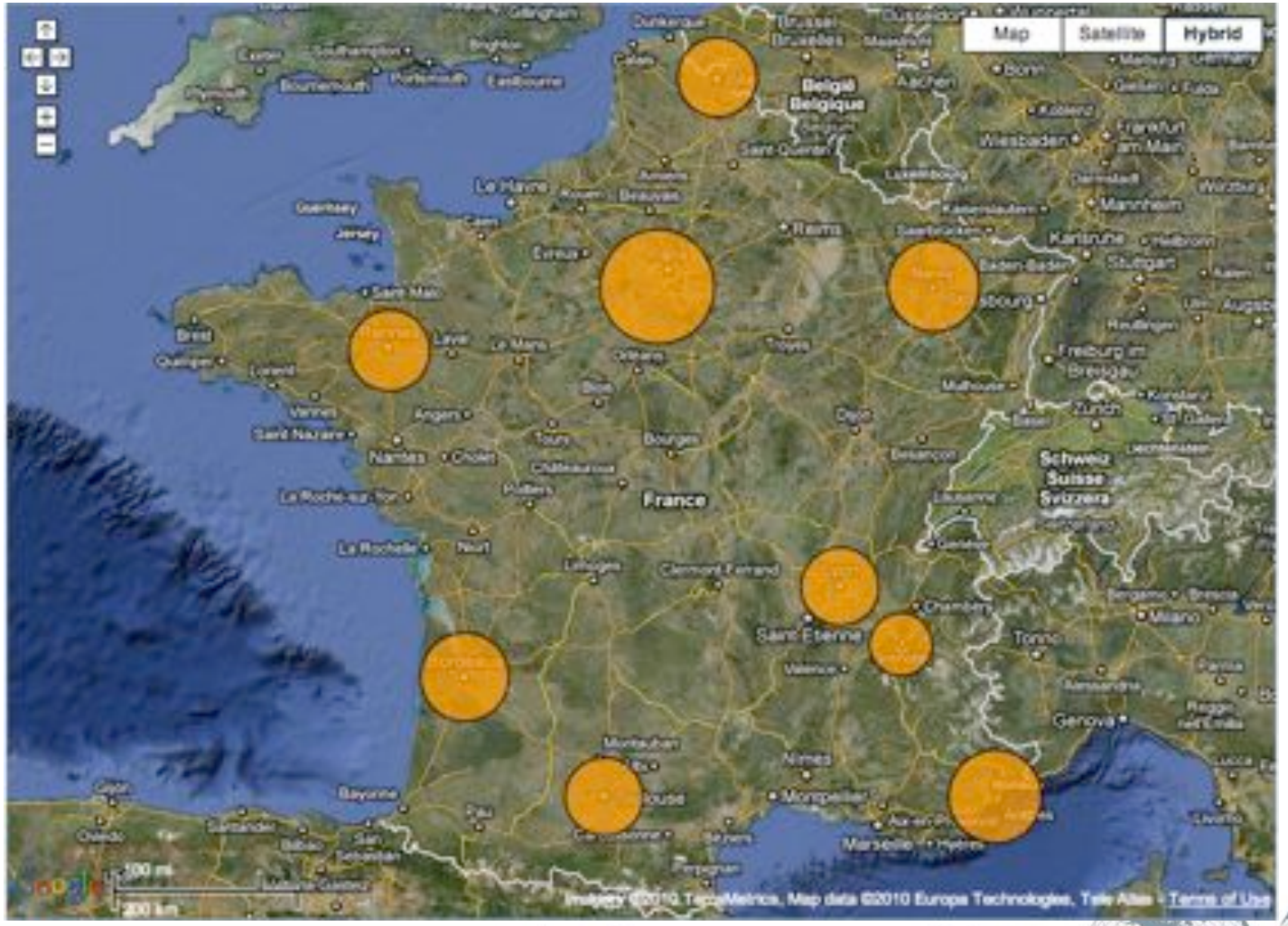
Experimental Testbeds

- Large-scale (many nodes/many cores)
- Distributed (several/many sites)
- Good for research in distributed & parallel
  - Systems
  - Algorithms
  - Middleware
  - …

Grid'5000 Overview

- Distributed over 9 sites in France
- ~1500 nodes, ~5500 CPUs
- Study of large scale parallel/distributed systems
- Features
  - Highly reconfigurable
    - Environment deployment over bare hardware
    - Can deploy many different Linux distributions
    - Even other OS such as FreeBSD
  - Controlable
  - Monitorable (metrics access)
- Experiments on all layers
  - network, OS, middleware, applications

Grid'5000 Node Distribution

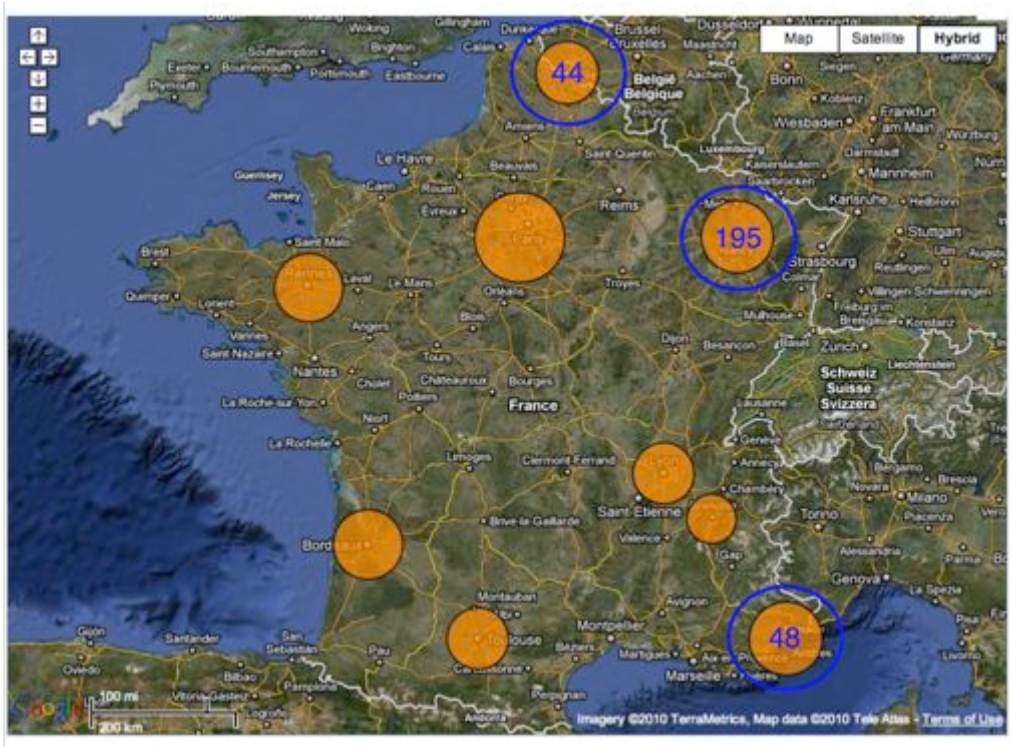


Grid’5000 2010 Spring School Challenge

- Concentrated on quickly deploying Nimbus clouds on Grid’5000 resources
- Used 3 sites to create 3 separated Nimbus clouds
- Deployment + Configuration: 30 minutes
  - Reserve nodes
  - Deploy images
  - Configure roles (using Chef)

Grid'5000 2010 Spring School Challenge

- 278 VMMs
  - 1628 virtual cpus
  - 2 124 GB of memory



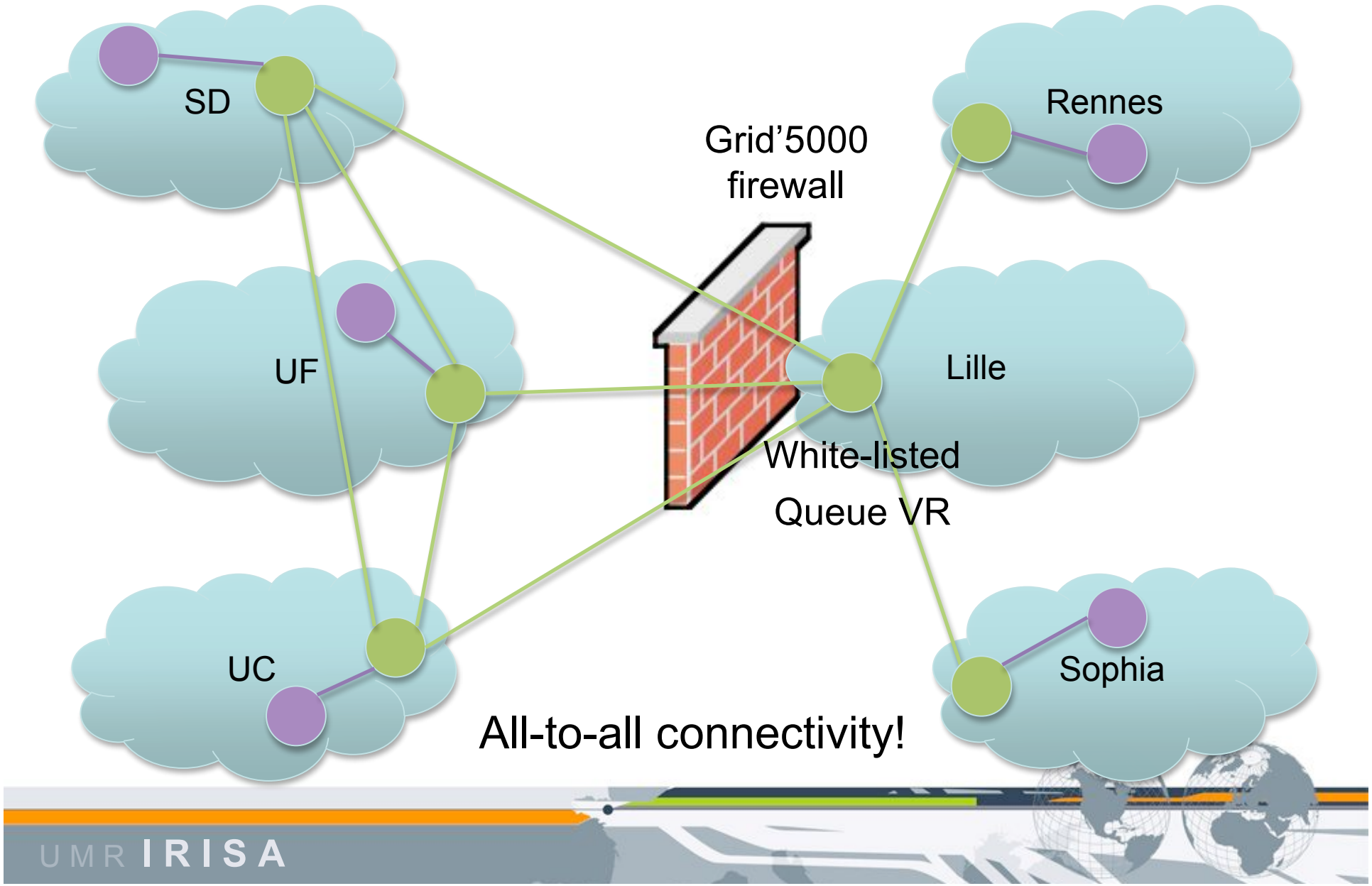
FutureGrid: a Grid Testbed

- NSF-funded experimental testbed
- ~5000 cores
- 6 sites connected by a private network

Demo: Resources

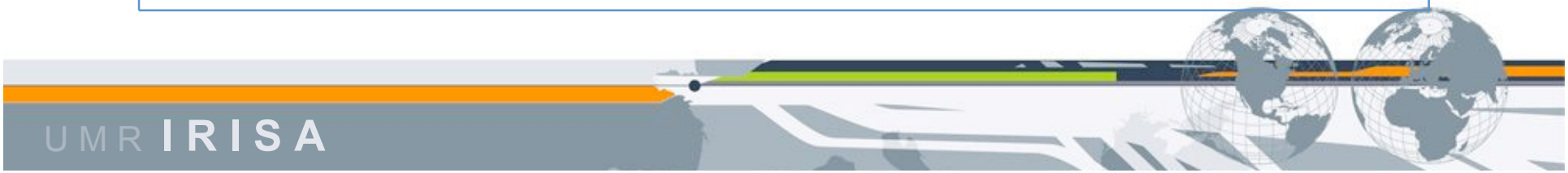
- 3 FutureGrid sites (US)
  - UCSD (San Diego)
  - UF (Florida)
  - UC (Chicago)
- 3 Grid'5000 sites (France)
  - Lille (contains a white-listed gateway to FutureGrid)
  - Rennes
  - Sophia
- Grid'5000 is fully isolated from the Internet
  - Got one machine white-listed to access FutureGrid
  - ViNe queue VR (Virtual Router) for other sites

ViNe Deployment Topology



Experimental setup

	SD San Diego (US)	UF Florida (US)	UC Chicago (US)	Grid'5000 Rennes (FR)	Grid'5000 Sophia (FR)
Xen VMM	3.1.2	3.1.2	3.1.2	3.2.1	3.2.1
CPU architecture	Intel Xeon Harpertown	Intel Xeon Gainestown	Intel Xeon Gainestown	Several clusters	Several clusters
CPU clock	2.5 GHz	2.27GHz	2.93GHz	Several clusters	Several clusters
CPU cache	6 MiB	8 MiB	8 MiB	Several clusters	Several clusters
Number of VMs	71	90	100	97	100
VCPUs per VM	2	2	4	2	2
Memory per VM	3.5GiB	3.5GiB	3.5GiB	1GiB	1GiB
Networking	Public	Private	Public	Private	Private
Name	guest*	nx*	hotel*	rennes-vm*	sophia-vm*
IP	198.202.120.*	172.31.10.*	149.165.148.*	10.156.*	10.144.*



Demo Outline

- Hadoop sky computing virtual cluster already running in FutureGrid (SD, UF, UC)
- Launch BLAST job
- Start virtual cluster on Grid'5000 resources
- Automatically extend the Hadoop cluster
  - Number of nodes increases
    - TaskTracker nodes (Map/Reduce tasks execution)
    - DataNode nodes (HDFS storage)
  - Hadoop starts distributing tasks in Grid'5000
  - Job completes faster!