egee

Geographical failover for the EGEE-WLCG Grid collaboration tools

CHEP 2007
Victoria, Canada, 2-7 September

Alessandro Cavalli, Alfredo Pagano

(INFN/CNAF, Bologna, Italy)

Cyril L'Orphelin, Gilles Mathieu, Osman Aidel

(IN2P3/CNRS Computing Centre, Lyon, France)

Rafal Lichwala (PSNC, Poznan, Poland)
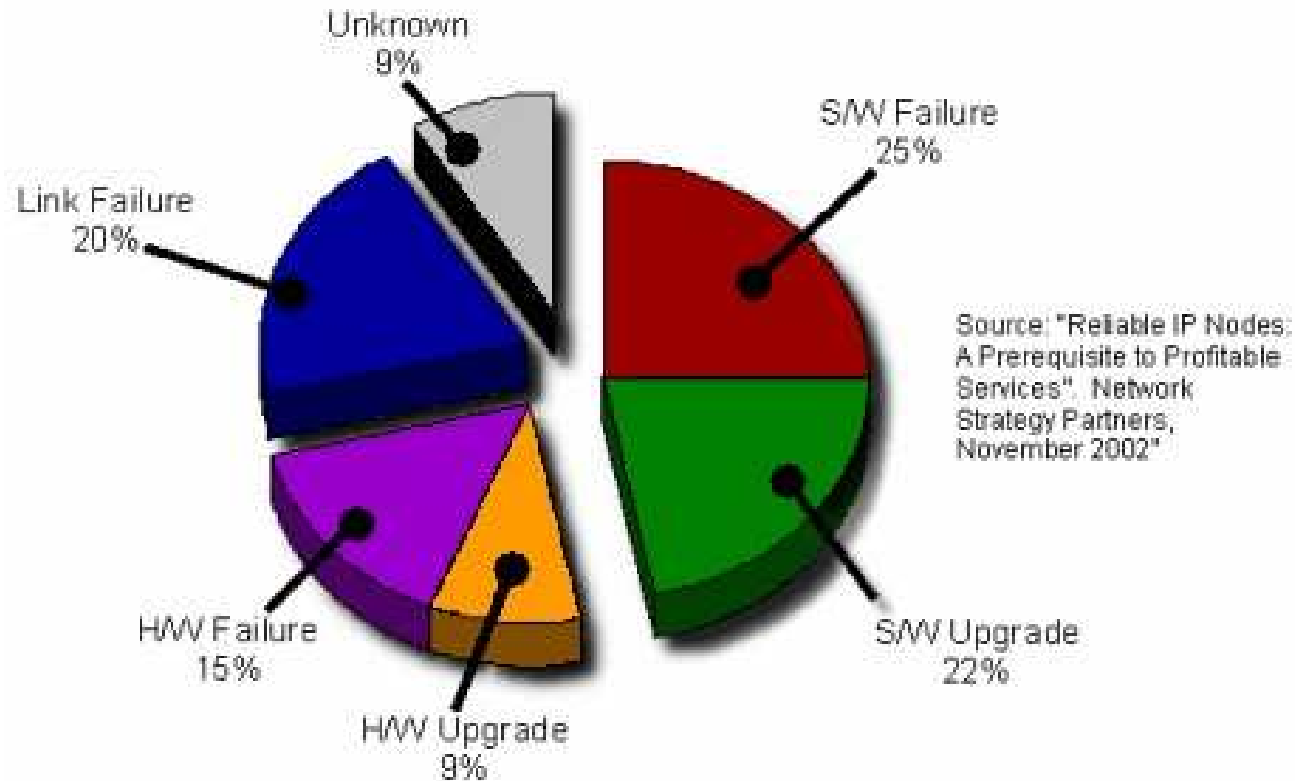
Information Society

- **The Failover System**

- **Technical solution**

    – the DNS and the new domain

    – www.gridops.org

    – Geographical failover examples

- **Use cases**

    – CIC Portal

    – GOCDB, SAM

    – GSTAT, GRIDICE

    – SAMAP (SAM Admin's Page)

- **Future plans and improvements**

    – Oracle replication

    – Distributed agents and Monitoring system

Enabling Grids for E-sciencE

- **Failover is an important fault tolerance function of mission-critical systems that rely on constant accessibility.**

- **Failover automatically and transparently to the user redirects requests from the failed or down system to the backup system that mimics the operations of the primary system.**

- **How much availability must we guarantee ?**

	Availability	Downtime/Year	Examples
1	90.0%	36 days, 12 hours	Personal Computers
2	99.0%	87 hours, 36 min	Entry Level Business
3	99.9%	8 hours, 45.6 min	ISPs, Mainstream Business
4	99.99%	52 min, 33.6 sec	Data Centers
5	99.999%	5 min, 15.4 sec	Banking, Medical
6	99.9999%	31.5 seconds	Military Defense

Enabling Grids for E-sciencE

Unknown 9%

Link Failure 20%

HW Failure 15%

HW Upgrade 9%

S/W Failure 25%

S/W Upgrade 22%

Source: "Reliable IP Nodes: A Prerequisite to Profitable Services", Network Strategy Partners, November 2002"

- **Magic words are:**
  - Redundancy
  - Remove Single Points of Failure (SPOF)

- **Propose, implement and document failover procedures for the collaboration, management and monitoring tools used in EGEE/WLCG Grid.**

  – The mentioned tools (listed later in this talk) are daily and heavily used by COD teams, regional and sites operators and other user categories, for grid management and control purposes.

  – **These are the reasons for an availability requirement that is high and which tend to become higher in future.**

- **Born as EGEE SA1 Operations COD task**

- **Reminder: who are the "CODs" ?**
  - Teams provided by EGEE federations, working in pairs (one lead and one backup) on a weekly rotation
  - Role:
    - Watch the problems detected by the grid monitoring tools
    - Problem diagnosis
    - Report these problems (GGUS tickets)
    - Follow and escalate them if needed (well defined procedure)
    - Provide help, propose solutions
    - Build and maintain a central knowledge database (WIKI)

Operational tools dependencies map

- http://goc.grid.sinica.edu.tw/gocwiki/Failover_mechanisms/Optools Map
- **A revealing survey done to have the clear view of how each service, when working or in failure, can influence the whole system**
- **An important input to the failover work**

	CIC PORTAL	eNOC	FCR	GGUS	GOCDB	GSTAT	SAM	TOP BDIIs	WMS FOR SAMAP
BDII CONF GEN					CRIT DB				
CIC PORTAL				CRIT WS	CRIT DB	FAIR.CRIT HTTP	CRIT WS	CRIT LDAP	
FCR					FAIR.CRIT DB				
GOCDB	NON-CRIT DB-neceive (TODO)	NON-CRIT DB-receive							
GRIDICE					NON-CRIT DB				
GSTAT					FAIR.CRIT DB		FAIR.CRIT WS		
SAM	NON-CRIT HTTP(S)?				FAIR.CRIT DB				
SAM ADMIN					CRIT DB		CRIT WS		CRIT wms/gsiftp
TOP BDIIs			FAIR.CRIT HTTP						
	OTHER TOOLS								
CENTRAL EUROPE NAGIOS					FAIR.CRIT DB		FAIR.CRIT HTTPS		

egee

Enabling Grids for E-sciencE

- **DNS choice:**
  - Well supported by local staff at our institutes
  - Easy to understand how to exploit its features
  - Very stable and consolidated (born in 1983)
  - Widely used as element for failover solution by ISPs and IT companies

- **the DNS approach consist in:**
  - mapping the service name to one or more destinations
  - update this mapping whenever some failure is detected

- **this must be coupled with procedures that:**
  - keep data in sync where it is needed
  - kill unnecessary processes on the system in failure
  - enable needed processes on the replacing system.

- **ISP caching policies**
  - Some provider could have caching policy longer than our TTL
  - The institutes participating to our Grid in general shouldn't
  - So user-service and service-service connections should be generally able to benefit from short TTLs, except rarely for some roaming user

- **Caching at OS level**
  - Local resolvers on MS Windows and Mac OS X provide caching, but they take into account possible shorter TTLs on DNS records

- **Caching by the web browsers**
  - MS Internet Explorer: 30 minutes
  - Mozilla Firefox: 60 seconds
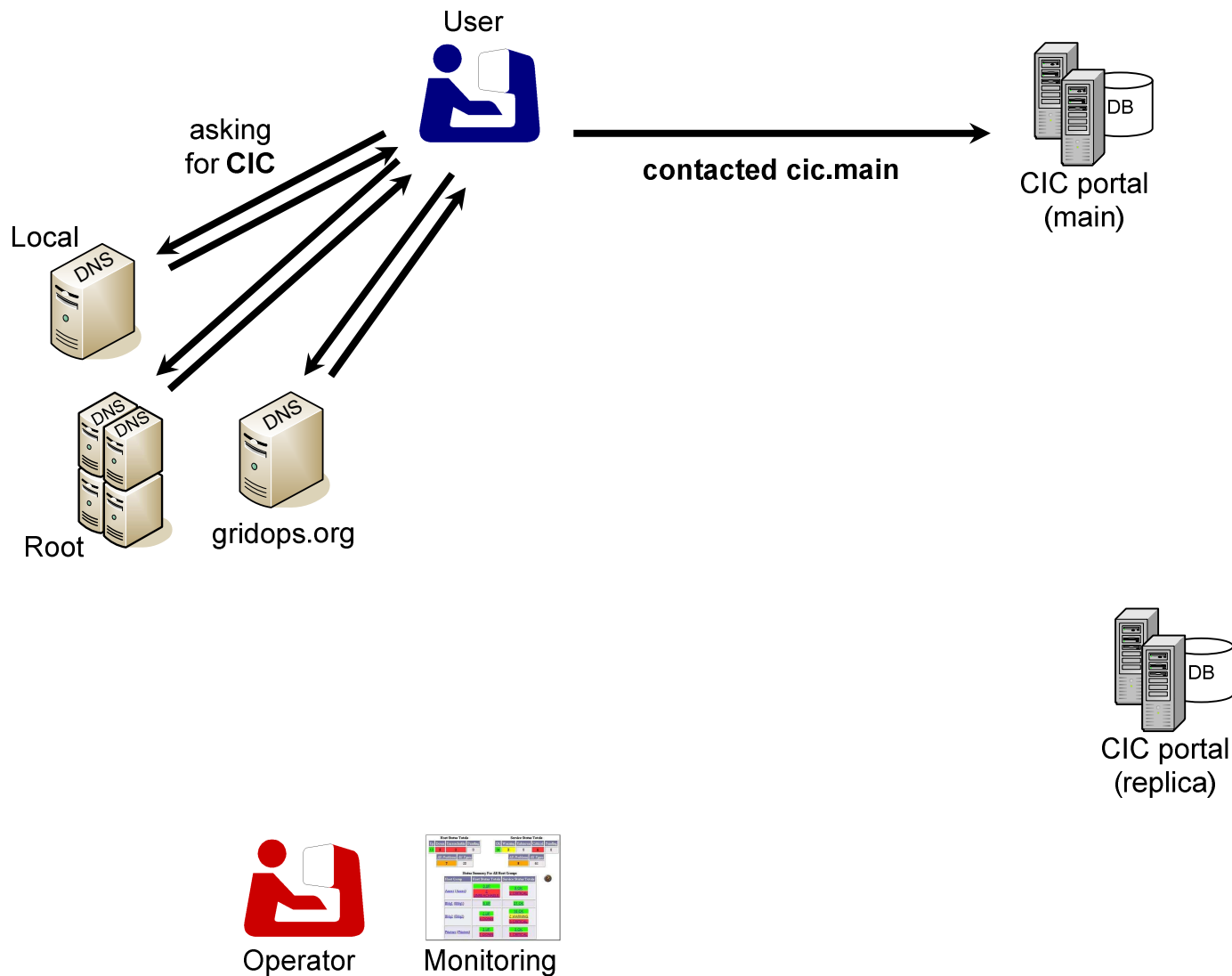 - A shorter time would be preferable for MSIE, but tolerable

- **A new _gridops.org_ domain has been registered by CNAF**
- **Redundant master & slave DNS provided by CNAF & GRNET**
- **All the replicated services' names inserted as CNAMEs:**
  - cic.gridops.org, cic2.gridops.org;
  - goc.gridops.org, goc2.gridops.org;
  - etc…
- **Default "`$TTL 60`" imposed to the gridops _zone_**
- **Names in gridops securely updated via `nsupdate` using `dnssec-keygen` generated keys**
- **X509 certificates with main or alternative names for the gridops names have been requested**
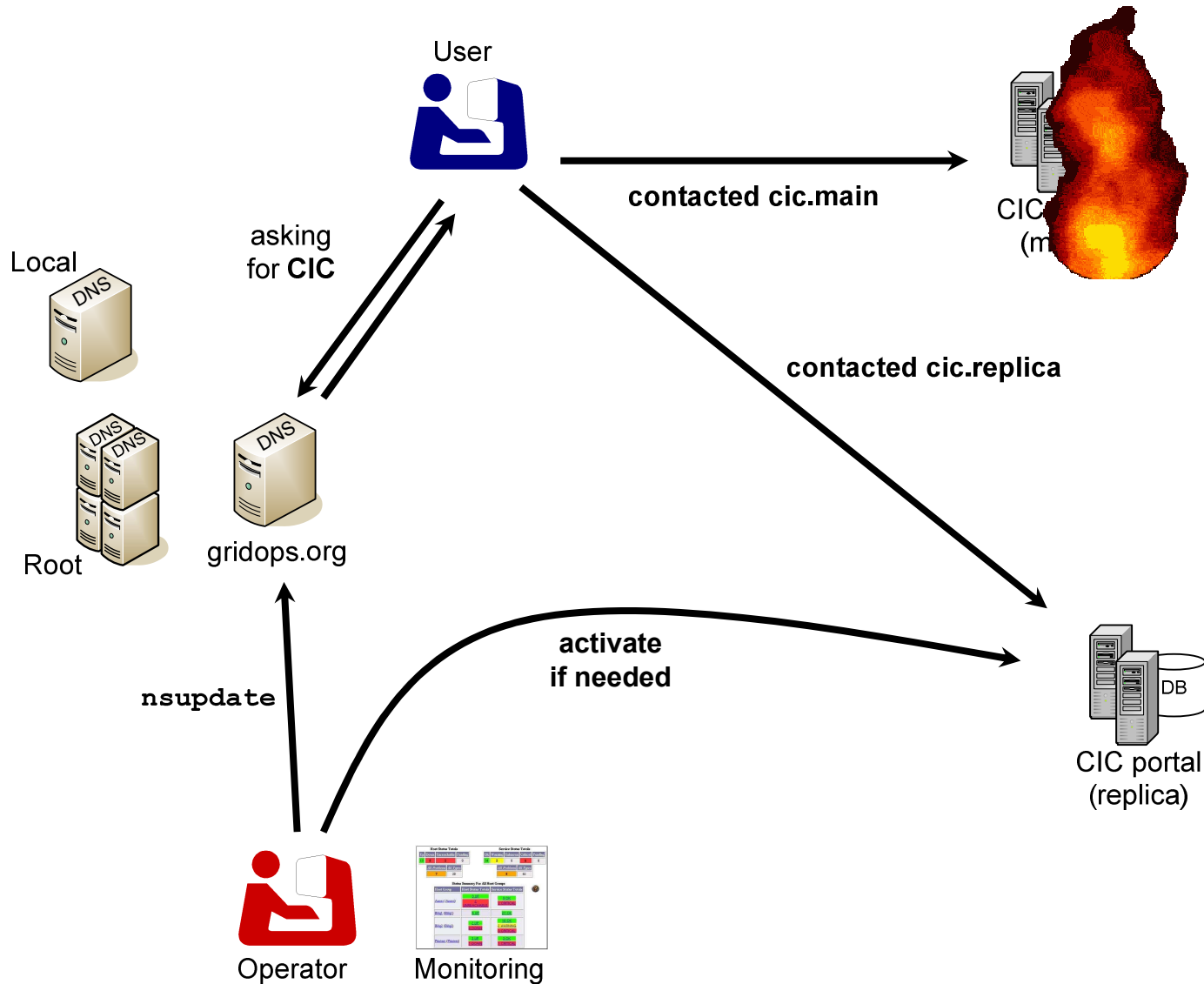
- **CIC Portal**
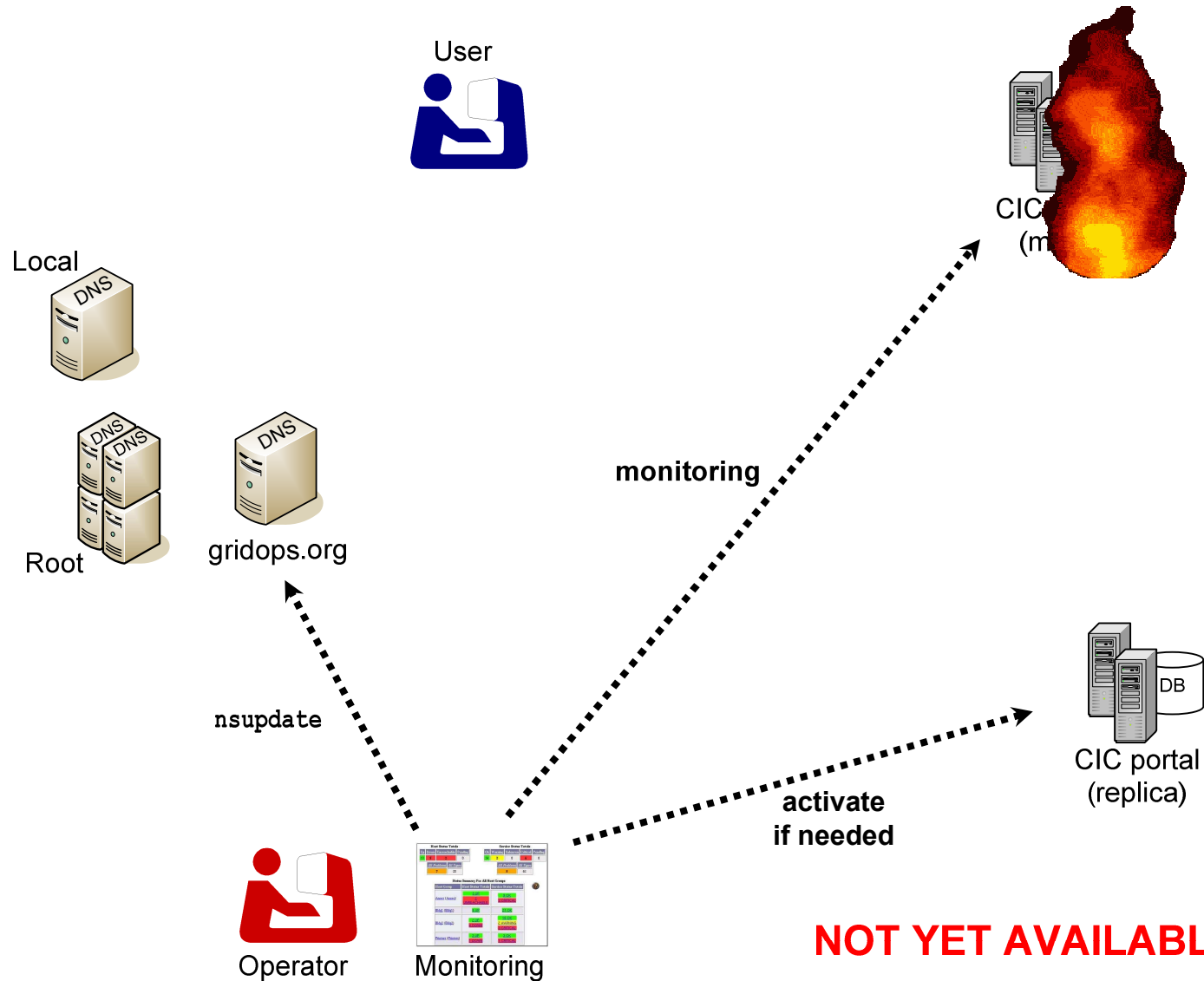- **GOCDB**
- **GRIDICE**
- **GSTAT**
- **SAM**
- **SAMAP**

- **Available operations tools are listed on www.gridops.org failover page. The portal is at present the collector of the proper links to the main and replica operations tools.**
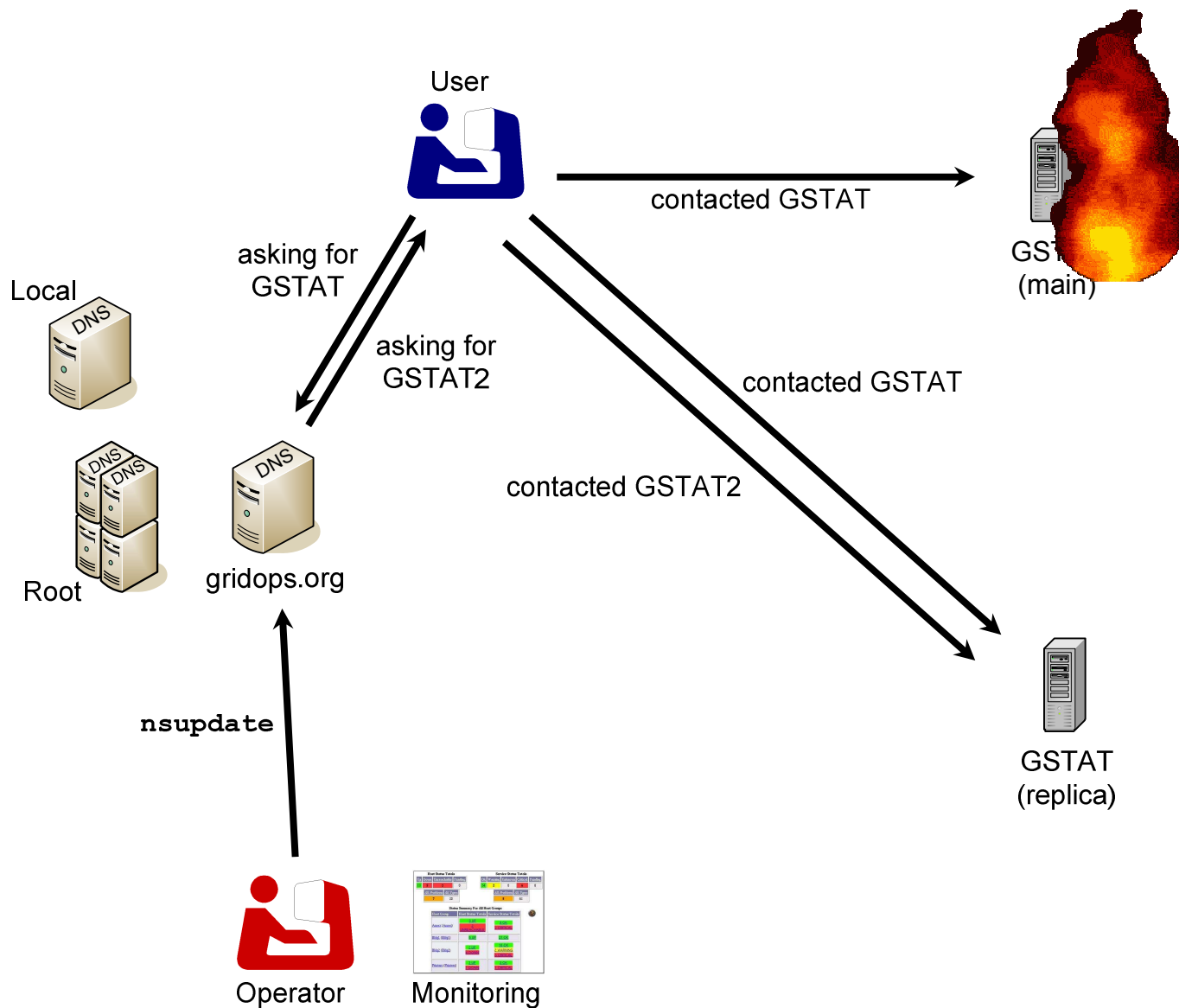

EGEE SA1 Failover

LCG-EGEE operations tools

- **CIC Portal** (IN2P3CC-FR)
  - CIC Portal (CNAF-IT)
- **GOCDB3 Portal** (RAL-UK)
  - GOCDB3 Portal (ITWM-DE)*
- **GRIDICE** (CNAF-IT)
  - GRIDICE (GRNET-GR)
- **GSTAT** (ASGC-TW)
  - GSTAT (CNAF-IT)
- **SAM** (CERN-CH)
- **SAM-PPS** (CERN-CH)
- **SAM Admin's Page** (IN2P3CC-FR)
  - SAM Admin's Page (CNAF-IT)

* Frontend replicated. Work in progress about Oracle backend.

Wiki page: http://goc.grid.sinica.edu.tw/gocwiki/Failover_mechanisms

Contact: sitemanager@cnaf.infn.it

project-egee-sa1-failover@cern.ch (only CERN/SIMBA users allowed to post)

User

contacted cic.main

CIC
(m

Local
DNS

asking
for **CIC**

contacted cic.replica

DNS
DNS

DNS

Root   gridops.org

activate
if needed

DB

nsupdate

CIC portal
(replica)

Operator   Monitoring

User

Local
DNS

DNS
DNS

DNS

Root

gridops.org

CIC
(m

monitoring

`nsupdate`

DB

CIC portal
(replica)

**activate
if needed**

Operator   Monitoring

NOT YET AVAILABLE

User

contacted GSTAT

GST...
(main)

Local

DNS

asking for
GSTAT

asking for
GSTAT2

contacted GSTAT

Root

DNS
DNS
DNS

DNS

contacted GSTAT2

gridops.org

GSTAT
(replica)

nsupdate

Operator   Monitoring

Enabling Grids for E-sciencE

- **Replication added early on the list**
  - Highly critical tool
  - Planned or unexpected service outages could break the continuity of daily activity

- **First switch successfully done in December 2006**

  - Replica instance used in production during one whole week

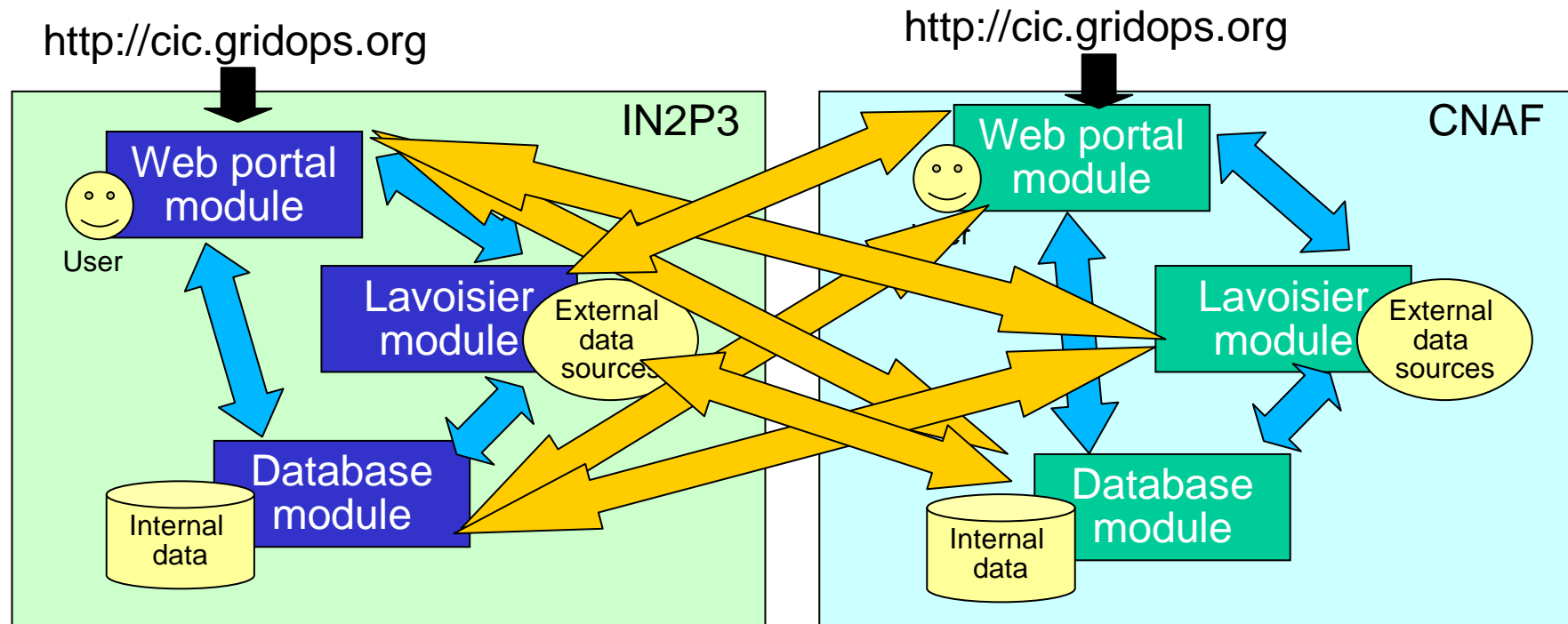 - Normal use of the portal during this time

  - No problem reported

- **The CIC portal is based on three distinct components :**
  - A web portal module (php and html, css files)
  - A database module (Oracle)
  - A data processing system (named Lavoisier)
- **Each component can work with any of the other, master or replica: 8 possible scenarios !**

- **Web module replication**
  - Portable code
  - Environment & configuration on replica (Apache, PHP, libs)
  - Host certificate for the replica

- **Data processing system (Lavoisier) replication**
  - Environment and configuration on replica (Apache ANT, Globus toolkit, java web services)
  - Deployment of a second instance of Lavoisier
  - Settings on replica (e-mail address for local error reporting…)

- **Database replication**
  - Dump of master DB exported to replica
  - Well established procedure, involving 2 persons and an interruption of service
  - We are working on better solutions

- **General considerations:**
  - GRIDICE and GSTAT are very different, but similar in one aspect: they collect a lot of information mainly from the Grid Information System. The same one source in both cases.
  - Therefore for their replication:
    - we basically installed another instance
    - we waited for the historical data to be long enough to be usable
    - we accept some small inconsistencies between the instances, because this has the positive aspect to give different views of the monitored resources
- **GRIDICE**
  - Main instance installed and administered by CNAF and GridICE teams
  - Secondary instance recently installed at FORTH-ICS Greek site and administered with the support of GridICE experts
- **GSTAT**
  - Main instance in ASGC-Taiwan
  - Secondary instance running at CNAF for more than one year

- **SAMAP(SAM A**dmin's **P**age) **- web-based monitoring tool for submitting SAM Framework test jobs on demand**

- **SAMAP architecture divided into two independent parts: Portal part and UI (grid User Interface) part**
- **Portal part integrated with the CIC Portal**
- **UI part installed on dedicated UI machine**
- **SAMAP installed in two independent instances and linked to proper DNS domain entries**
- **synchronization of instances via CVS repository**
- **two geographically remote WMS servers available for both instances**
- **easy switch from main to backup instance by DNS entry modification**
- **full transparency for end users**

- **GOCDB**
  - From the GOCDB3 release, introduced Oracle backend: more complex to replicate
  - GOCDB3 web front-end promptly replicated in ITWM German site
    - Release synchronized via RPM and apt-get
  - Oracle DB failover plans:
    - Short term: weekly DB dump to CNAF (in progress)
    - Medium term: 2 separate sites with Oracle Streams in UK (in progress)
    - Medium-long term: another replica in TW, when UK instructions are ready
- **SAM**
  - Complex framework:
    - Web, WS, Oracle, UI(2), RB(2), WMS(2), BDII(2)
    - Need to focus what to replicate and how to take maximum benefit from the replication effort
  - DB size and growth (order):
    - 100.000.000 rows, 100 GB TOTAL
    - 100.000 – 1.000.000 rows, 100 MB – 1 GB DAY
  - Call for volunteers:
    - Implies Oracle license(s)
    - Good results from the tests done by CYFRONET (Poland), where SAM is already known, installed and used. It is likely that replica will be there

Enabling Grids for E-sciencE

- **We are working on… *automatic* failover!**

  – Automatize the ORACLE backend synchronization

  - Stream, DataGuard and materialized views

  - This will improve <u>CIC Portal</u>, <u>GOCDB</u> and <u>SAM</u> failover

  – Automatize the failover process through a crafty monitoring system

  - A distributed monitoring system that checks for the availability of the different tools and performs the DNS switch
  - Based on Nagios
  - Decision made upon the results of the distributed agents

- **LCG-3D**

  - Some differences

    - Our failover activity deals with operational tools

    - LCG-3D deals with database replication and data transfers.

  - Some similarities and shared goals

    - work on databases replication and switches

    - Same concern in disaster recovery solutions

    - Often involves the same (or at least related) teams.

  - discussions and working sessions engaged

- **Other failover-related topics**

  - GSLB (Global Server Load Balancing): a DNS based failover offered by some IT companies

  - Multihoming (BGP, IPv6): announcing the same IP(s) from multiple places

Enabling Grids for E-sciencE

- **EGEE Failover web:**
  - www.gridops.org

- **…and Wiki:**
  - http://goc.grid.sinica.edu.tw/gocwiki/Failover_mechanisms

- **LCG 3D:**
  - https://twiki.cern.ch/twiki/bin/view/PSSGroup/LCG3DWiki

- **GSLB:**
  - http://www.networkcomputing.com/showitem.jhtml?docid=1605ws1
  - http://www.tenereillo.com/GSLBPageOfShame.htm

- **Multihoming/BGP:**
  - http://en.wikipedia.org/wiki/Multihoming
  - http://tools.ietf.org/wg/shim6/