eGee

Enabling Grids for E-sciencE

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

egee

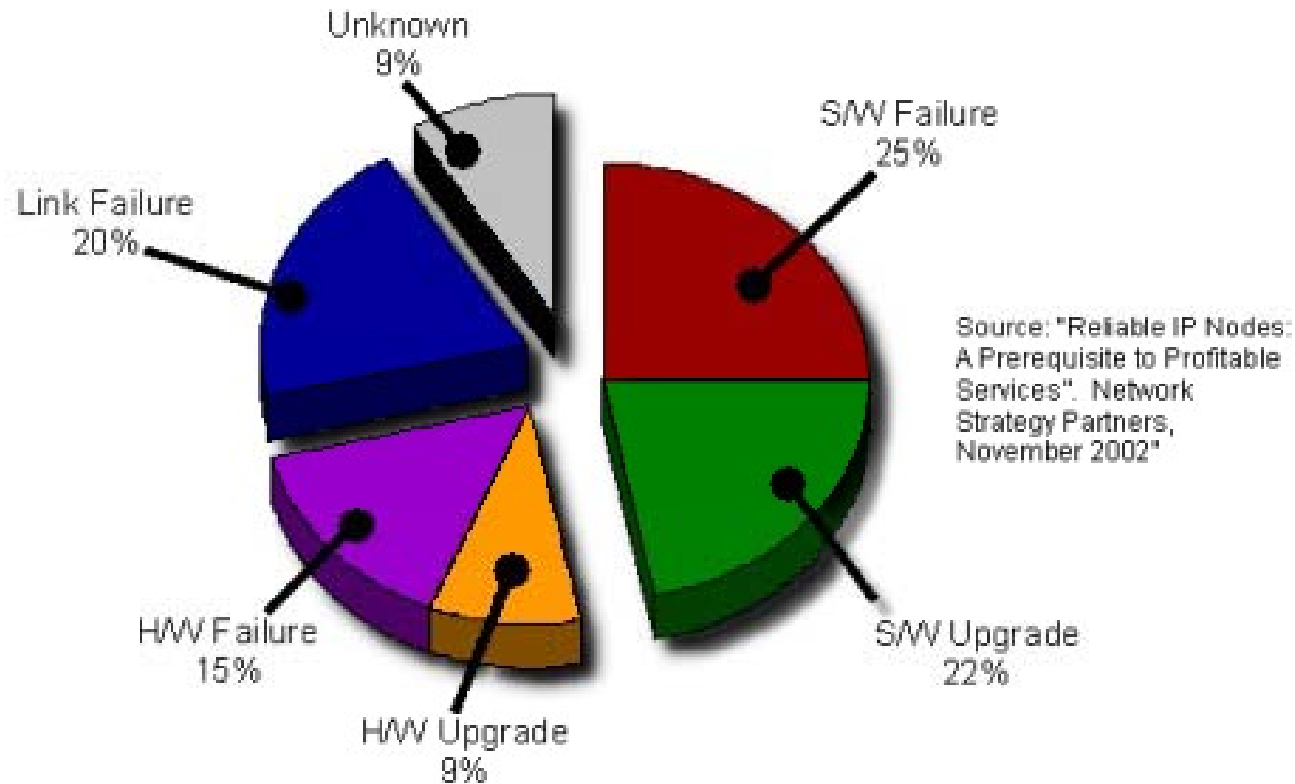Enabling Grids for E-sciencE

- **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

Failover: definition and principle

- **A backup operation that automatically switches to a standby database, server or network if the primary system fails or is temporarily shut down for servicing.**

- **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

Unknown 9% · S/W Failure 25% · S/W Upgrade 22% · H/W Upgrade 9% · H/W Failure 15% · Link Failure 20%

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)

- **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 examples**

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 the '70s)
  - 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
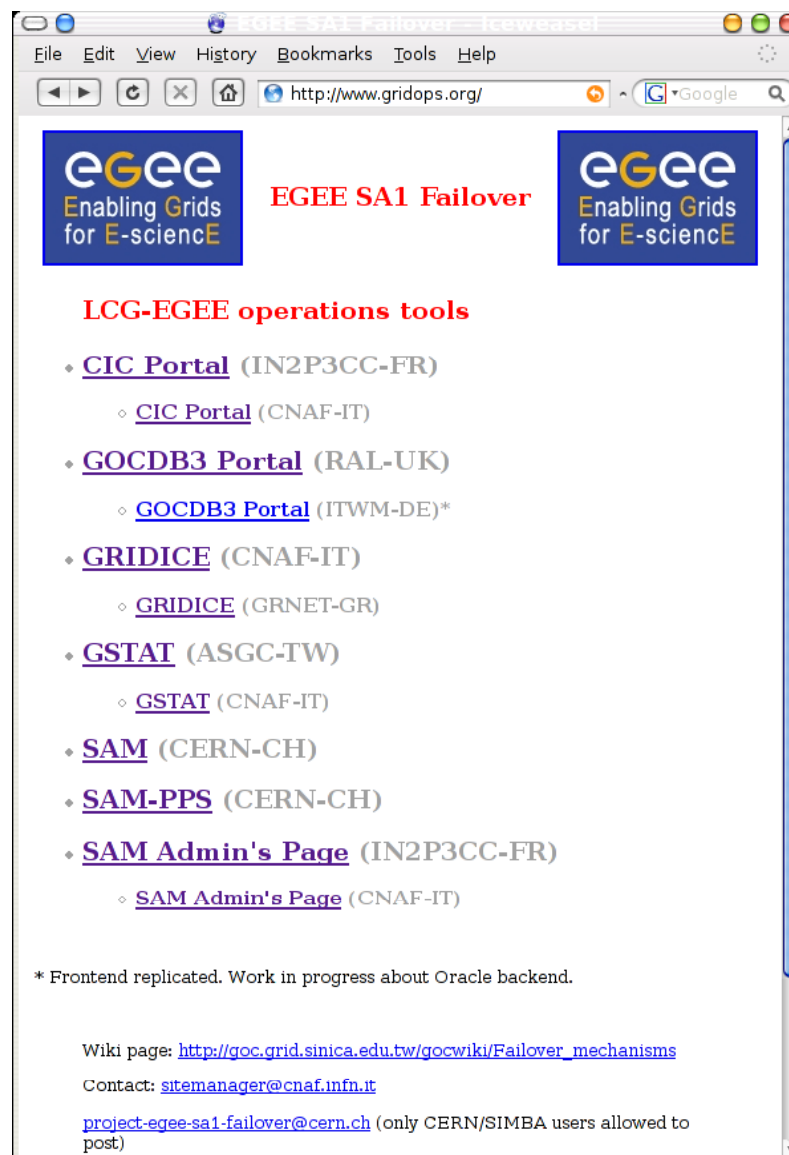 - enable needed processes on the replacing system.

- **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
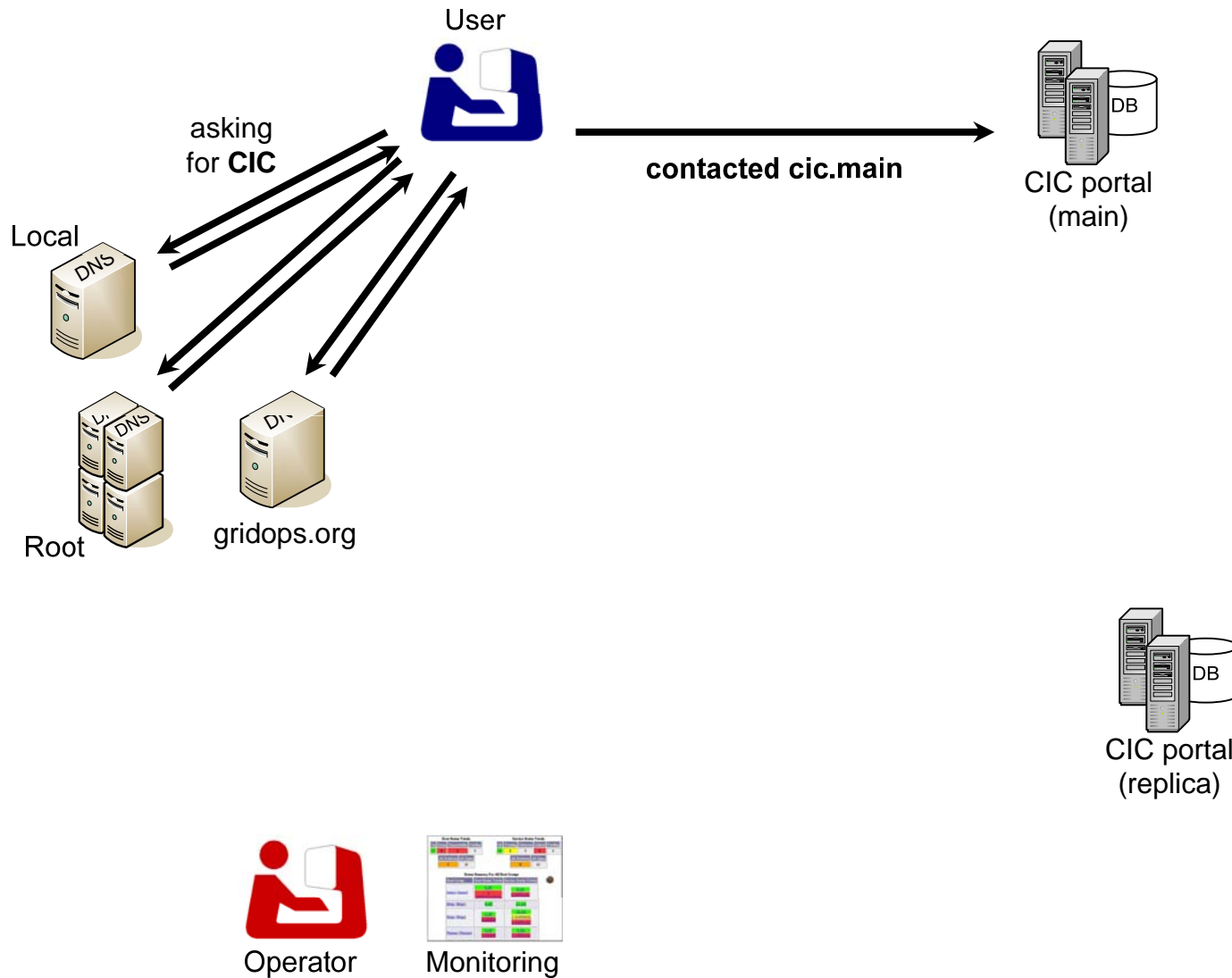 - A shorter time would be preferable for MSIE, but tolerable

- **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**
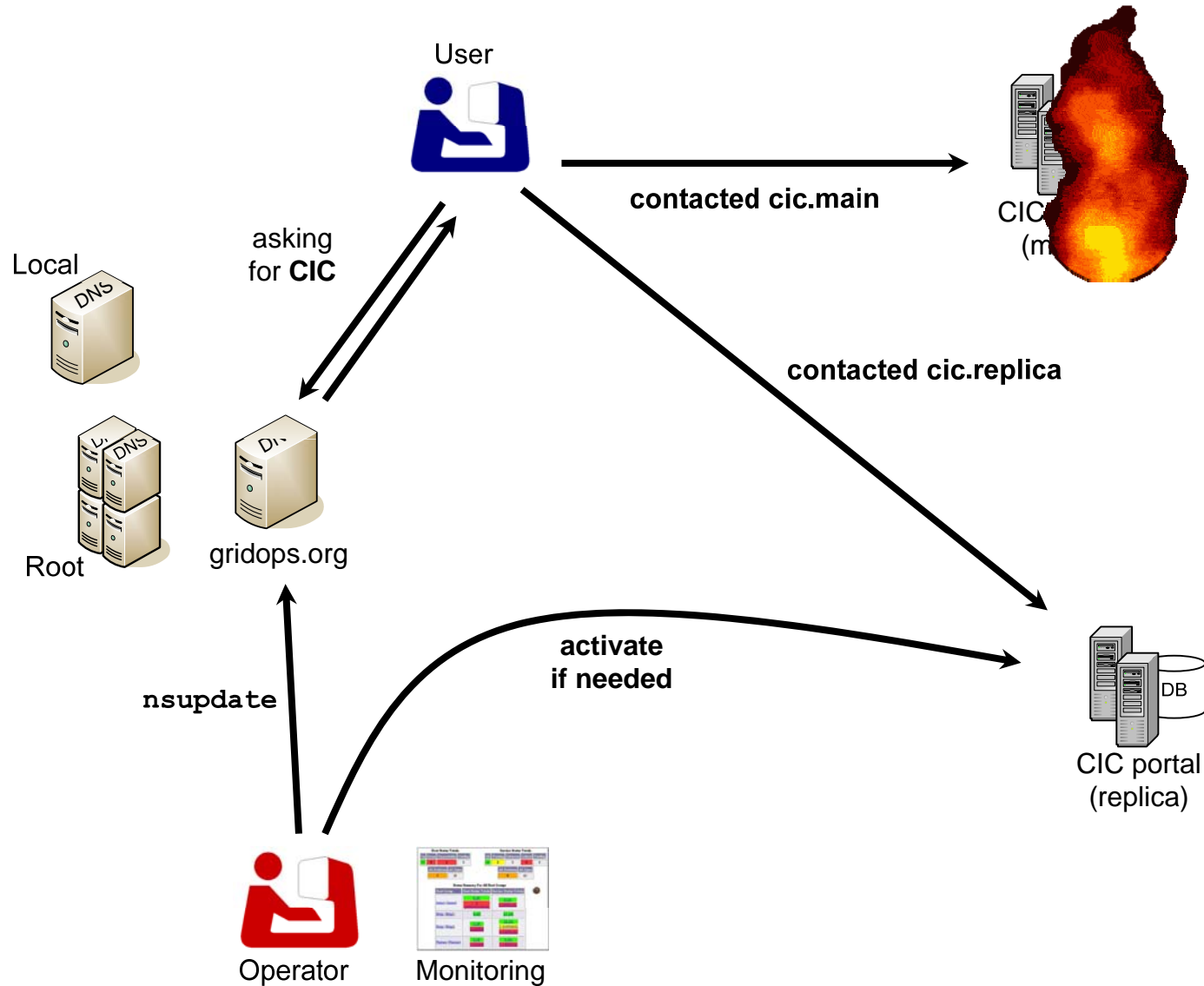
eGee
Enabling Grids for E-sciencE

- **CIC Portal**
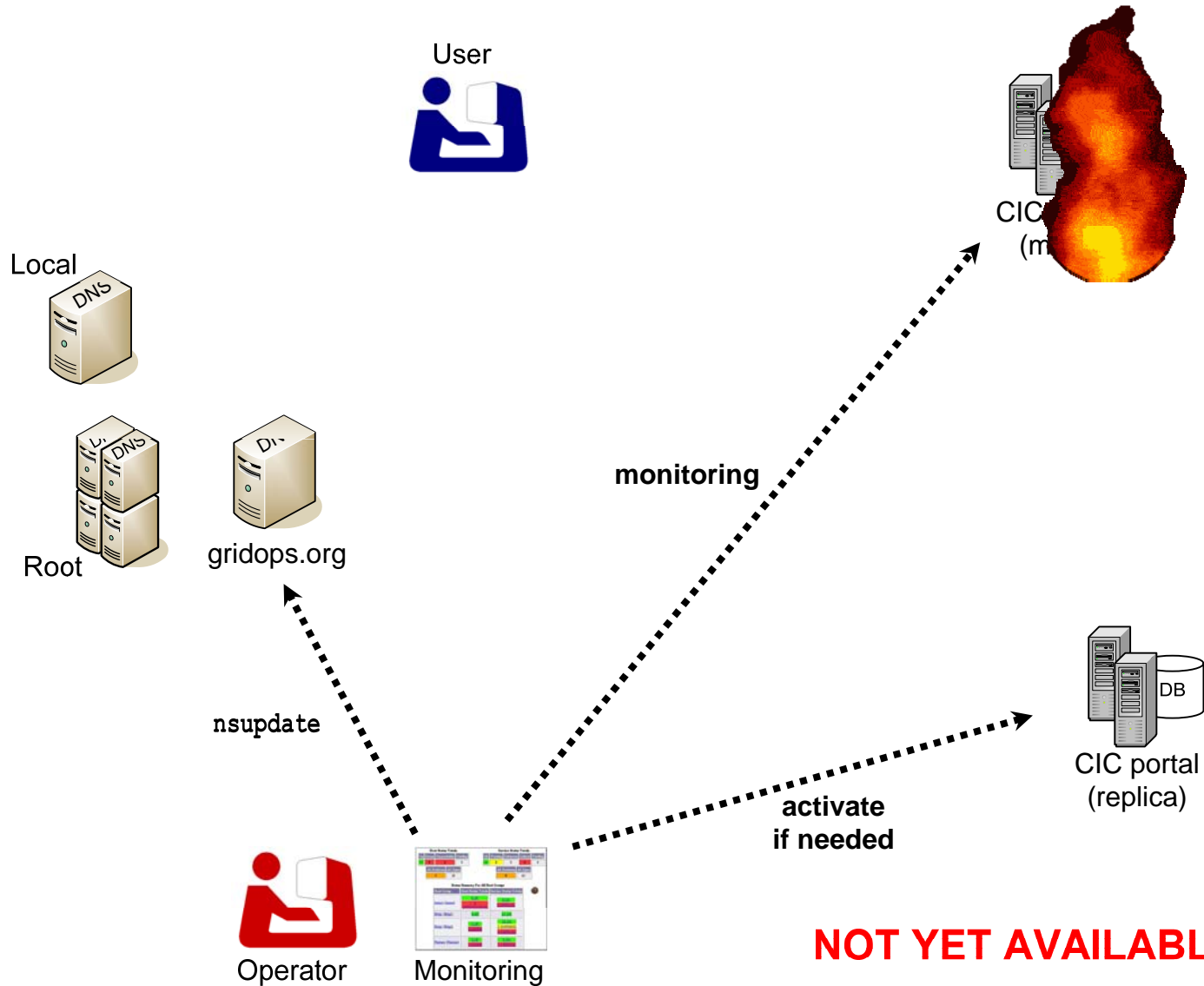- **GOCDB**
- **GRIDICE**
- **GSTAT**
- **SAM**
- **SAMAP**

- **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)

Enabling Grids for E-sciencE

User

asking for **CIC**

contacted cic.main

CIC portal (main)

Local DNS

Root DNS

gridops.org DNS

CIC portal (replica)

Operator   Monitoring

User

contacted cic.main

CIC
(m

asking
for **CIC**

Local

contacted cic.replica

Root

gridops.org

activate
if needed

`nsupdate`

CIC portal
(replica)

DB

Operator   Monitoring

User

Local

Root

gridops.org

CIC
(m

monitoring

nsupdate

DB

CIC portal
(replica)

**activate
if needed**

Operator

Monitoring

NOT YET AVAILABLE

User

contacted GSTAT

GST
(main)

asking for
GSTAT

Local

asking for
GSTAT2

contacted GSTAT
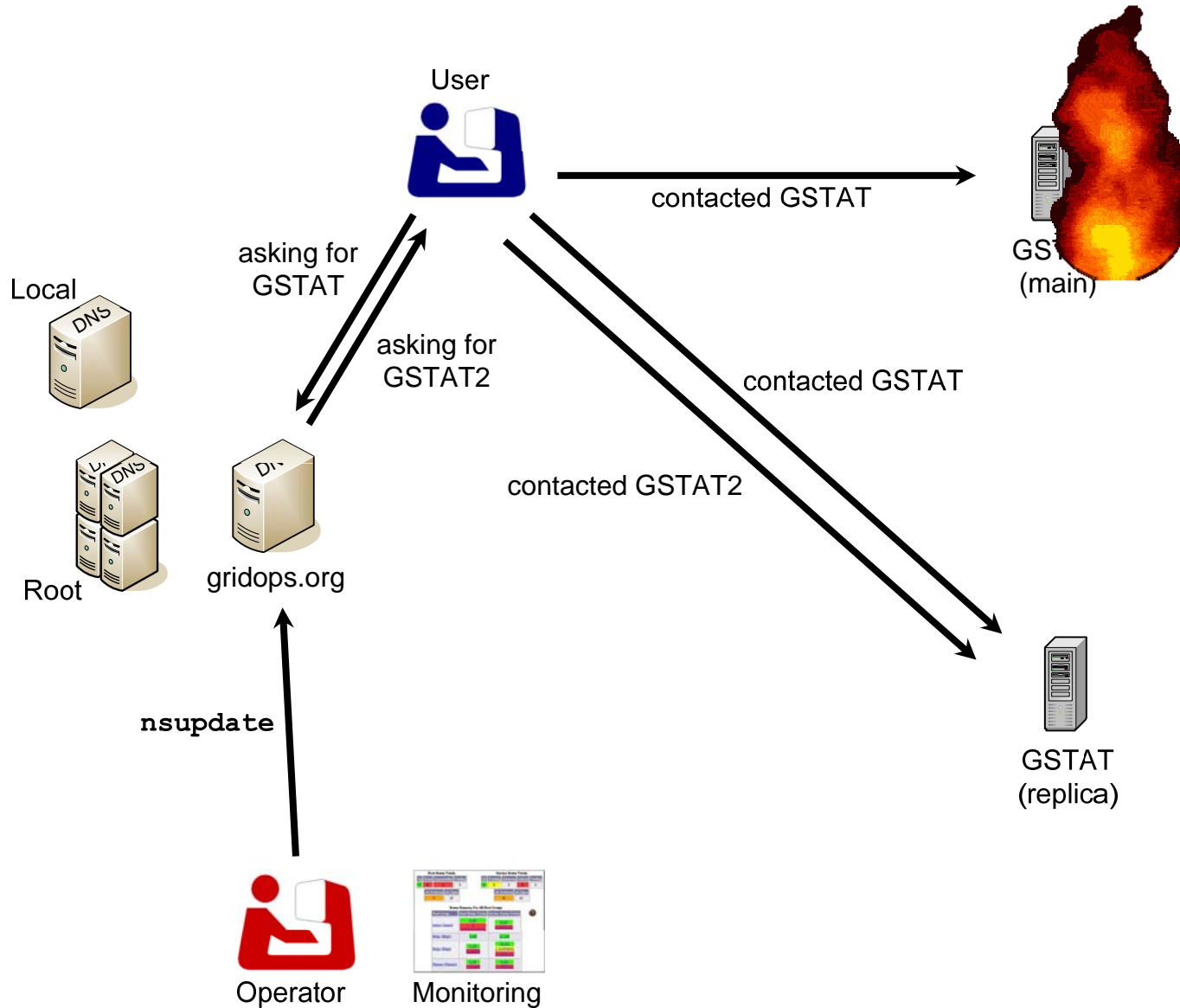
Root

gridops.org

contacted GSTAT2

GSTAT
(replica)

nsupdate

Operator   Monitoring

- **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

CNAF - BOLOGNA

CC-IN2P3 - LYON

CCIN2P3   CIC PORTAL Failover   INFN CNAF

- **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 !**

http://cic.gridops.org   http://cic.gridops.org

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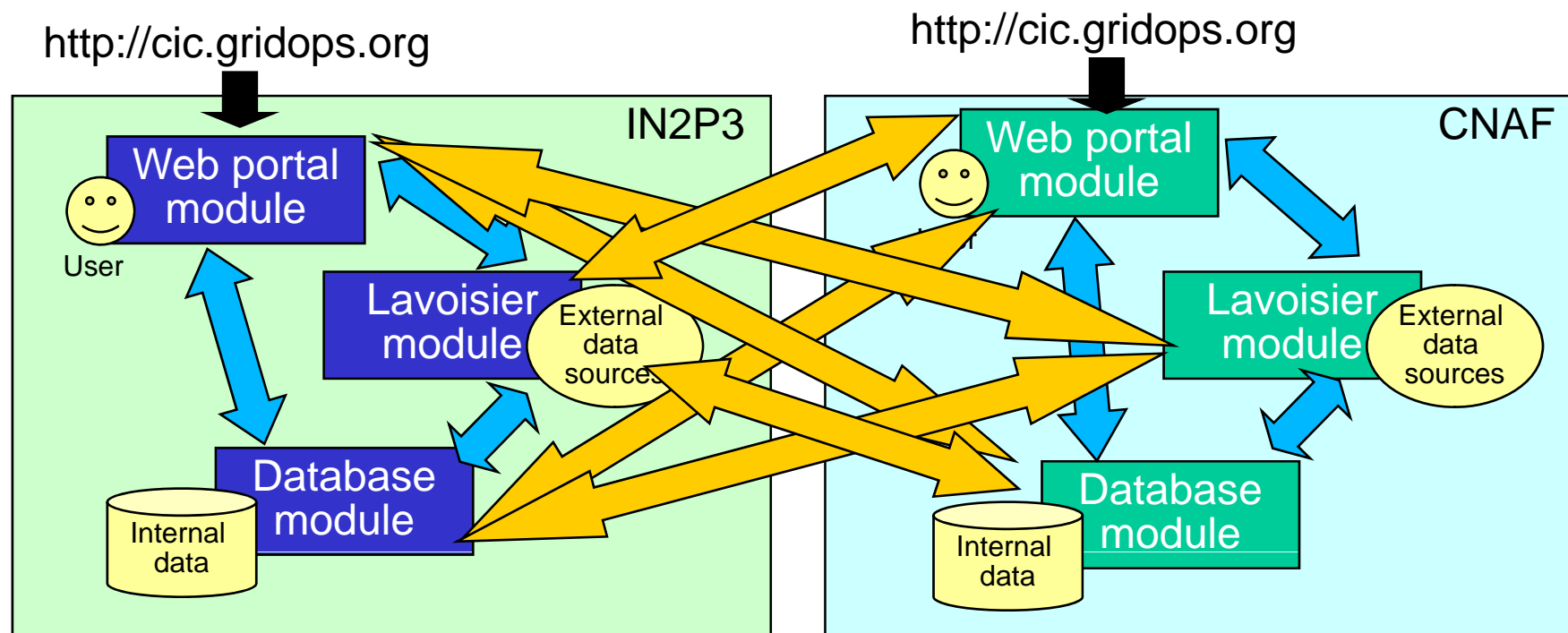
- **Web module replication**
  - Portable code
  - Environment & configuration on replica (Apache, PHP, libs)
  - Host certificate for the replica

- **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

- **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

- **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 Admin's Page) - web-based monitoring tool for submitting SAM Framework test jobs on demand**

- **based on SAM-client (Site Availability Monitoring Framework in EGEE)**

- **additional functionality implemented in response to the site administrator's needs**

- **provided functionality:**

  – SAM job submission on demand

  – check status of the running test jobs

  – cancel submitted SAM jobs

  – publish SAM job results

  – show logging info of the running test jobs

  – schedule regular SAM job submission (cron task management)

Enabling Grids for E-sciencE

- **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**

egee

Enabling Grids for E-sciencE

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

  – Automatize the ORACLE backend synchronization

    - Stream, DataGuard and materialized views

  – 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

Enabling Grids for E-sciencE

- **EGEE Failover web: www.gridops.org (wiki link inside)**