Deployment & Experiment Integration

Flavia Donno & Markus Schulz
LCG

LCG Review
17 November 2003

Overview

- Release History
- Deployment status
- Release and Deployment Procedures
- Information System
- Experiment Integration
  - Support
  - Software distribution
- Problems
- Resources Deployment/Experiment Support
- Summary

History

- First set of reasonable middleware on C&T Testbed end of July (PLAN April)
  - limited functionality and stability

middleware was late

- Deployment started to 10 initial sites
  - Focus on establishing procedures (not functionality)
  - Training sites (we sent to two sites a support person for 2 days)
- End of August only 5 sites in
  - Underestimation of the effort and dedication needed
    - Complexity of the middleware, installation and configuration
    - Lack of experience with install/config tool
- First certified version LCG1-1_0_0 release September 1st (PLAN in June)
  - Limited functionality, improved reliability
  - Training paid off -> 5 sites upgraded (reinstalled) in 1 day
  - Last after 1 week….
- Security patch LCG1-1_0_1 first not scheduled upgrade took only 24h.
- Sites need between 3 days and several weeks to come online
  - All sites using the fabric management tool for service nodes

Estimation of duration of the deployment process was correct

Release History

Overview:

Tag	Date	Comment
lcg1_20030717_1455	17 Jul	Pre-release:CNAF, CERN
LCG1-1_0_0	01 Sep.	EDG pre 2.0, several fixes by LCG
LCG1-1_0_1	19 Sep.	Security Patch (10sites)
LCG1-1_1_0	24 Oct.	Fixes, new WLM (17 sites)
LCG1-1_1_1	05 Nov.	Experiment SW
LCG1-1_1_2	Xx Nov.	Experiment SW (20 sites)

LCG-1 Deployment Status

- Up to date status can be seen (here) expect >20 by end of 2003

- –PIC-Barcelona (RB)
  - •IFIC Valencia (RB)
  - •Ciemat Madrid
  - •UAM Madrid
  - •USC Santiago de Compostela
  - •UB Barcelona
  - •IFCA Santander (RB)
- –BNL
- –Budapest (RB)
- –CERN (RB)
- –CNAF (RB)

- –FNAL
- –FZK
  - •Krakow
- –Moscow (RB)
- –Prague
- –RAL (RB)
  - •Imperial c.
- –Taipei (RB)
- –Tokyo

Total number of CPUs ~150 WNs

(current focus on # of sites)

CPUs added on experiments request

Sites to enter soon

Lyon, CSCS Switzerland, Cavendish,

Several tier2 centres in Italy

Sites preparing to join

Pakistan, Sofia,

Users:

EDG: Experiment independent testers

Experiments: Alice, Atlas, CMS LHCb

LCG-1 Deployment Status Overview

Site's Status Page	CVS	Site Tag in CVS	LCG tag	Installation Status		Operation	Comment
PIC-BARCELONA(pic.ifae.es)	Yes	LCG1-1_0_1_PIC_20030923_1836	LCG1-1_1_1	Installed	S	Running	LCG-ES Coordinator South MDS
IFIC-VALENCIA(ific.uv.es)	Yes	LCG1-1_1_1_IFIC_20031105_1800	LCG1-1_1_1	Installed	S	Running	LCG-ES
CIEMAT-MADRID(ciemat.es)	Yes	LCG1-1_1_1_CIEMAT_20031105_1800	LCG1-1_1_1	Installed	S	Running	LCG-ES
UAM-MADRID(ft.uam.es)	Yes		LCG1-1_1_1	Installed	S	Running	LCG-ES
USC-SANTIAGO-DE-COMPOSTELA(usc.cesga.es)	Yes	LCG1-1_1_1_USC_20031105	LCG1-1_1_1	Installed	S	Running	LCG-ES
UB-BARCELONA(ecm.ub.es)	Yes	LCG1-1_1_1_UB_20031106_0001	LCG1-1_1_1	Installed	S	Running	LCG-ES
IFCA-SANTANDER(ifca.unican.es)	No		LCG1-1_1_1	Installed	S	Running	LCG-ES
BNL(bnl.gov)	Yes	LCG1-1_1_1_BNL_20031107_1525	LCG1-1_1_1	Installed	W	Running	-
BUDAPEST(kfki.hu)	Yes	LCG1-1_1_1_BUDAPEST_20031110_1535	LCG1-1_1_1	Installed	E	Running	
CERN(cern.ch)	Yes	LCG1-1_1_1_CERN_20031107_0857	LCG1-1_1_1	Installed	E	Running	East MDS
CSCS-SWITZERLAND(cscs.ch)	initial		LCG1-1_1_0	Installing	E	Offline	
CNAF(cr.cnaf.infn.it)	Yes	LCG1-1_1_1_CNAF_20031105_1630	LCG1-1_1_1	Installed	S	Running	South MDS
FNAL(fnal.gov)	Yes	LCG1-1_1_1_FNAL_20031106_1130	LCG1-1_1_1	Installed	W	Running	
FZK(fzk.de)	Yes	LCG1-1_1_1_FZK_20031105_1907	LCG1-1_1_1	Installed	E	Running	Primary site
KRAKOW(cyf-kr.edu.pl)	Yes		LCG1-1_0_1	Installed	E	Running	Secondary site
LYON	Yes		LCG1-1_1_1	Installed	E	Offline	Missing host certificates, web page
MOSCOW(sinp.msu.ru)	Yes	LCG1-1_0_0_MOSCOW_20030903_1946	LCG1-1_1_1	Installed	E	Running	
PRAGUE(farm.particle.cz)	Yes	LCG1-1_1_1_PRAGUE_20031107_1317	LCG1-1_1_1	Installed	E	Running	
RAL(gridpp.rl.ac.uk)	Yes	LCG1_1_1_1_RAL_20031106_1437	LCG1-1_1_1	Installed	W	Running	West MDS P-site
IMPERIAL (hep.ph.ic.ac.uk)	Yes	LCG1_1_1_1_IMPERIAL_20031110_1715	LCG1-1_1_1	Installed	W	Running	S-site, dead link
CAVENDISH (hep.phy.cam.ac.uk)	Yes	LCG1-1_1_1_CAVENDISH_20031111_1220	LCG1-1_1_1	Installing	W	Offline	S-site, wrong format web page
TAIPEI(grid.sinica.edu.tw)	Yes		LCG1-1_1_1	Installed	E	Running	East MDS
TOKYO(icepp.s.u-tokyo.ac.jp)	Yes	LCG1-1_1_0_TOKYO_20031027_2012	LCG1-1_1_1	Installed	E	Running	

LCG-1 Site Layout

SiteName	PIC-LCG1

Contact	lcg.support@pic.ifae.es

Domain	ifae.es

PIC
port d'informació
científica

Date	Message of the Day
27/11/2003	Update to LCG1-1_1_0
23/09/2003	change published siteName to PIC-LCG1
22/09/2003	add RB (+BDII) machines at PIC
22/09/2003	tag LCG1-1_0_1 (security updates) installed
02/09/2003	tag LCG1-1_0_0 installed
28/08/2003	start installing LCG-1

Type	Host	System	Middleware	Install	Status Install	Status Operation	Comment	Local Comment
LCFG	grid-lcfgng-73	RH7.3edg	-	manual	configured	running		
CE	grid-w1	RH7.3edg	LCG1-1_1_0	LCFGng	configured	online	gateway to PIC's LCG farm	
SE	grid-s1	RH7.3edg	LCG1-1_1_0	LCFGng	configured	online		
UI	grid-ui	RH7.3edg	LCG1-1_1_0	LCFGng	configured	online		
WN	grid-w3	RH7.3edg	LCG1-1_1_0	LCFGng	configured	online		
WN	grid-w5	RH7.3edg	LCG1-1_1_0	LCFGng	configured	online		
RB	grid-rb	RH7.3edg	LCG1-1_1_0	LCFGng	configured	online		
BDII	grid-bdii	RH7.3edg	LCG1-1_1_0	LCFGng	configured	online		

LCG-1 Site Config in CVS

File	Rev.	Age	Author	Last log entry
Attic/ [show contents]				
budapestWorkerNode.h	1.2	2 months	gdebrecz	First version using LCG1-1_0_0. Simple LDAP quieries, gridftp and simple job su...
cfgdir-cfg.h	1.2	2 months	gdebrecz	First working configuration !
do_mkxprof.sh	1.1	2 months	emanuele	First version of the config files for Budapest site
grid100	1.5	2 hours	gdebrecz	Slight modifications in UI config.
grid109	1.2	2 months	gdebrecz	First version using LCG1-1_0_0. Simple LDAP quieries, gridftp and simple job su...
grid110	1.4	2 hours	gdebrecz	Slight modifications in UI config.
grid111	1.2	2 months	gdebrecz	First version using LCG1-1_0_0. Simple LDAP quieries, gridftp and simple job su...
grid112	1.2	2 months	gdebrecz	First version using LCG1-1_0_0. Simple LDAP quieries, gridftp and simple job su...
grid113	1.2	2 months	gdebrecz	First version using LCG1-1_0_0. Simple LDAP quieries, gridftp and simple job su...
grid114	1.2	2 months	gdebrecz	First version using LCG1-1_0_0. Simple LDAP quieries, gridftp and simple job su...
grid115	1.2	2 months	gdebrecz	First version using LCG1-1_0_0. Simple LDAP quieries, gridftp and simple job su...
grid116	1.2	2 months	gdebrecz	First version using LCG1-1_0_0. Simple LDAP quieries, gridftp and simple job su...
grid117	1.1	2 hours	gdebrecz	New working node.
local-cfg.h	1.2	2 months	gdebrecz	First working configuration !
nfsmount-cfg.h	1.3	2 months	gdebrecz	First working configuration !
site-cfg.h	1.5	2 hours	gdebrecz	Slight modifications in UI config.

Show files using tag: – Non-branch tags –

Show

For primary sites first version is provided by CERN deployment team. For secondary sites the supporting sites help with the initial config.

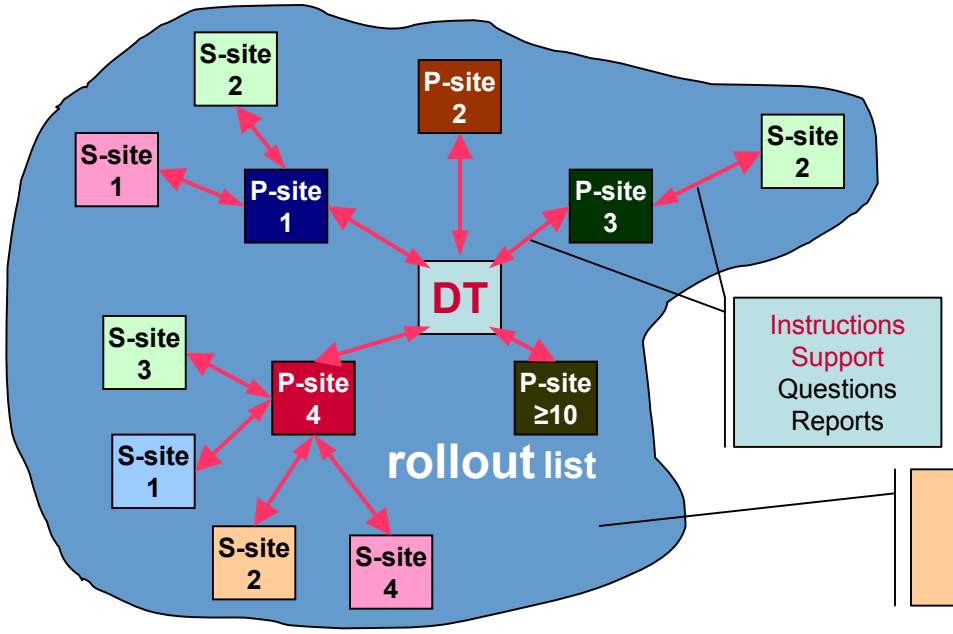
Introducing a new release

- Well established procedure (C&T presentation)
  - Software first assembled on the Certification & Test Testbeds (functional test)
  - Software handed to the Deployment Team
    - Adjustments in the configuration
    - Update of documentation (in CVS)
      - Questionnaire,Release Notes, Installation Instructions, Introduction material
    - More installation tests
- How do we deploy?
  - Service Nodes (RB, CE, SE …)
    - LCFGng (fabric management tool from EDG),
    - We provide for new sites config files based on a questionnaire
  - Worker nodes and UI

    Install by tool and manual
- **Communication:** list **LCG-rollout@rl.ac.uk** (~10 mails/day)

Work intensive, limited to <10 sites

Deployment Hierarchy



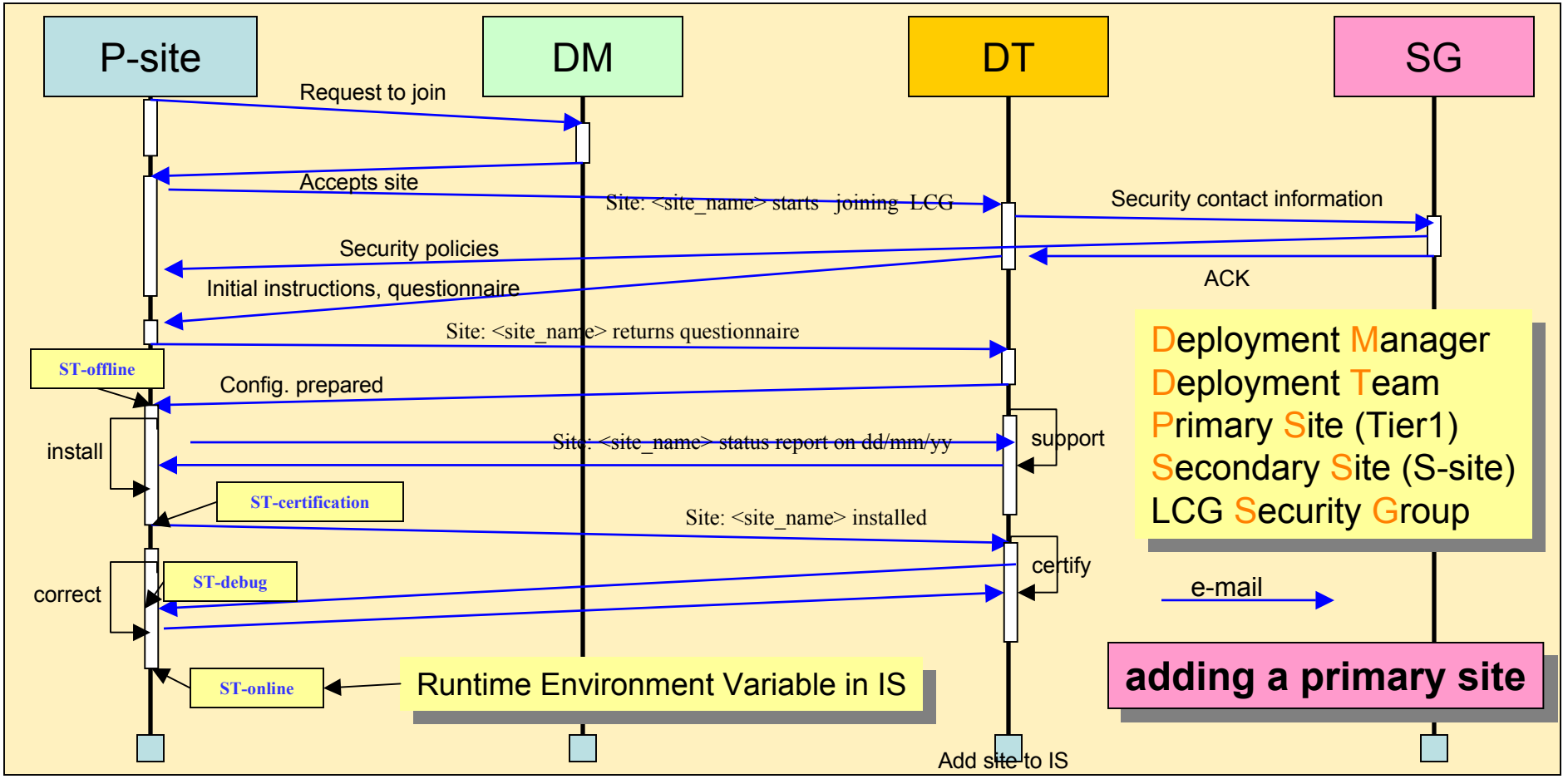
Deployment Team	Prepares releases, deploys first
Primary site	Experienced site with resources to support some sites
Secondary site	Inexperienced site (can have more computing than P-site)

- Too many sites to support centrally
- DT supports P-sites (creates initial config. Files, provides assistance)
- P-sites support S-sites and escalate problems to DT
- DT interacts with S-sites through their P-sites
  - Spread knowledge
  - Limits load on DT
- "Broadcast" style communication via the rollout list
- For emergency actions DT contacts sites directly

Procedures

Draft of the procedures for adding primary/secondary site, software upgrades:

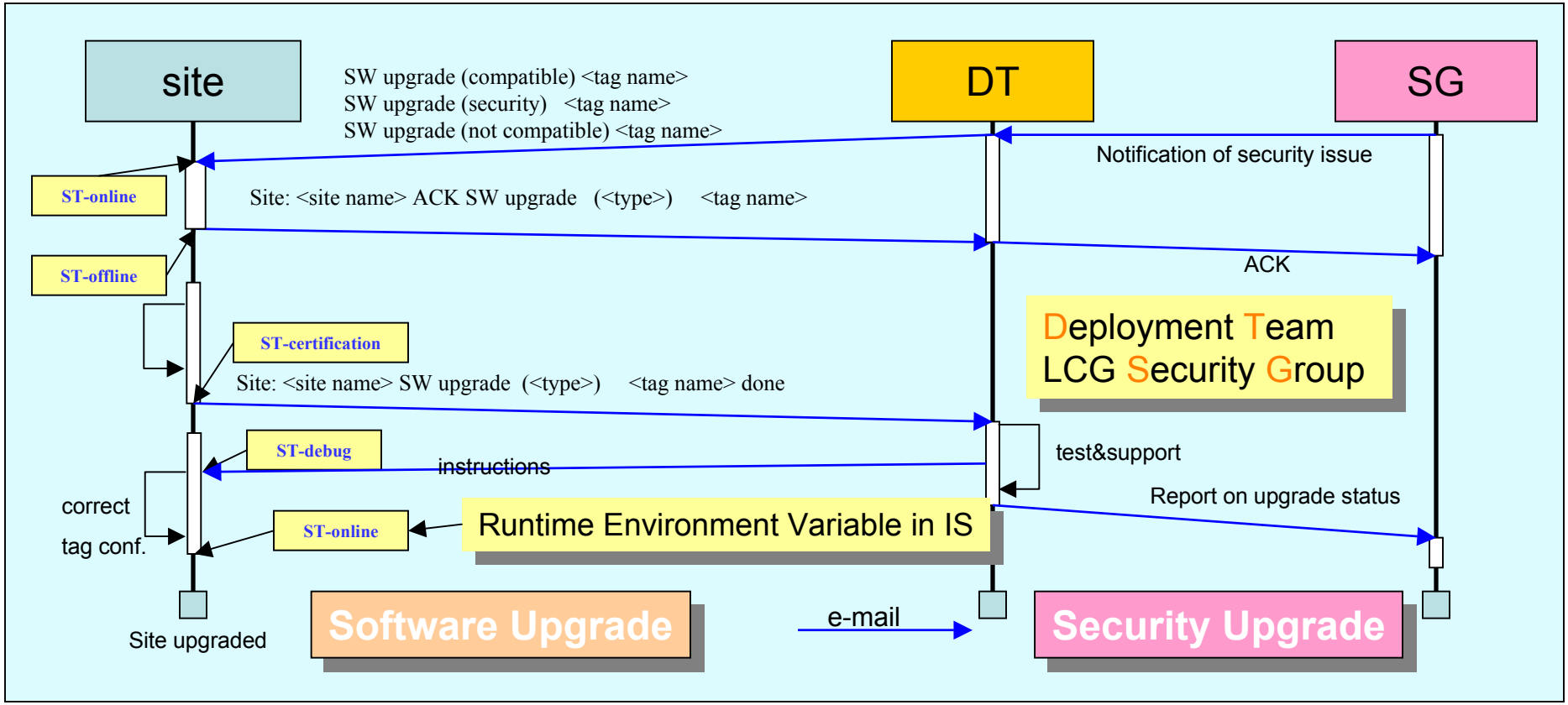
- http://cern.ch/markusw/JoiningLCG.doc
- http://cern.ch/markusw/JoiningLCG.pdf
- http://cern.ch/markusw/JoinLCG.html



Procedures

Adding Secondary sites:

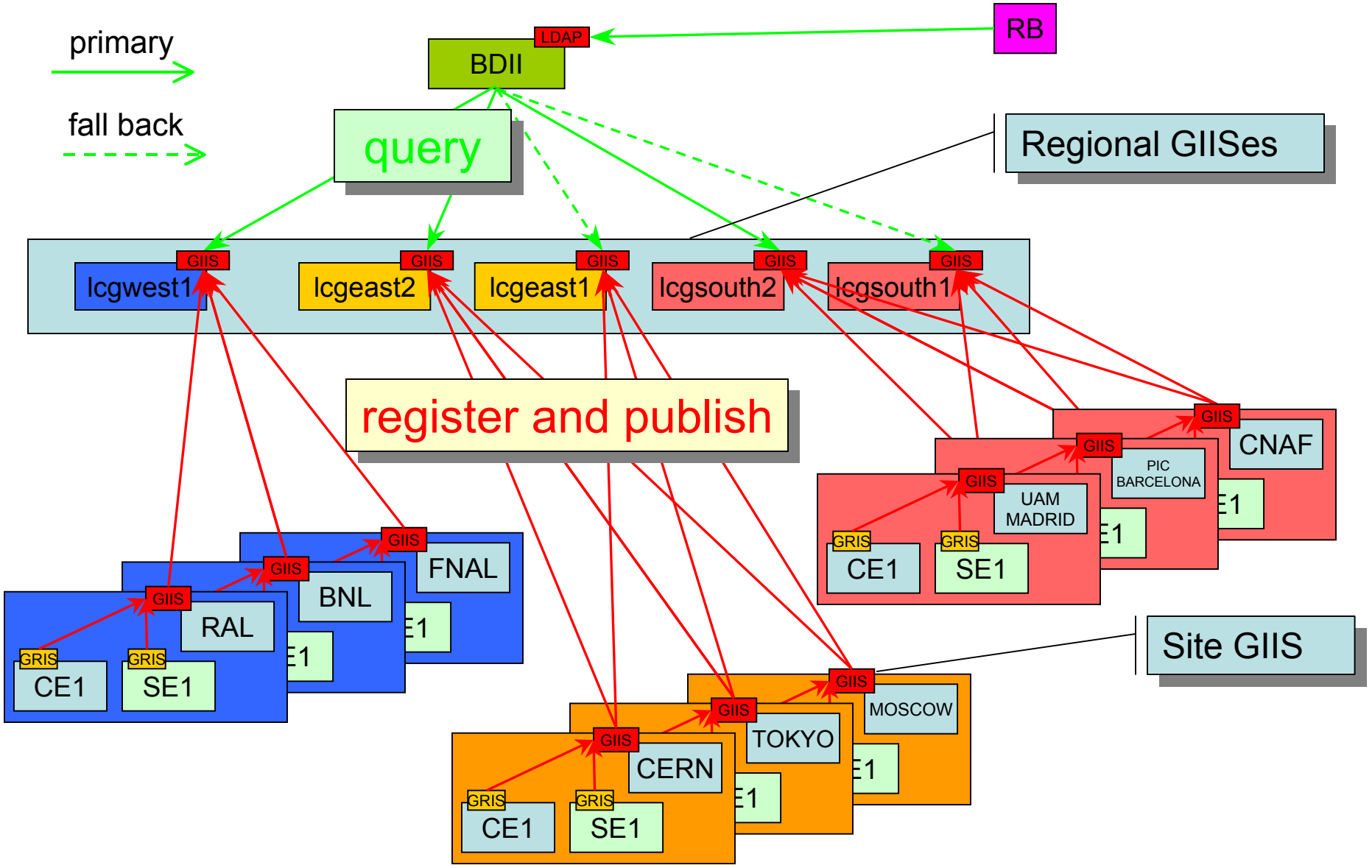
- Primary sites assume deployment team's role towards secondary sites
- Problems that P-site can't solve are escalated to the deployment team and rollout list
- Communication through P-site



LCG1 Information System

- The Information System (IS) is the nervous system of LCG
  - Used by almost all services (RB, Replica Manager, RLS, …) to
    - Discover resources and their properties ***(static and dynamic)***
  - Based on Globus MDS (based on LDAP, publish/subscribe)
  - Know scalability problems with MDS
    - Number of sites
    - Amount of data
    - Fatal handling of failures that propagate through he hierarchy
- Hardening of MDS
  - EDG-BDII replacing top level MDS
    - BDII == database + LDAP server + perl script to query MDS
    - LCG improved version: no stale information, redundant sources
    - Flat hierarchy (Site, Region, BDII)
    - Partitioning in regions: less load/region, confinement of problems

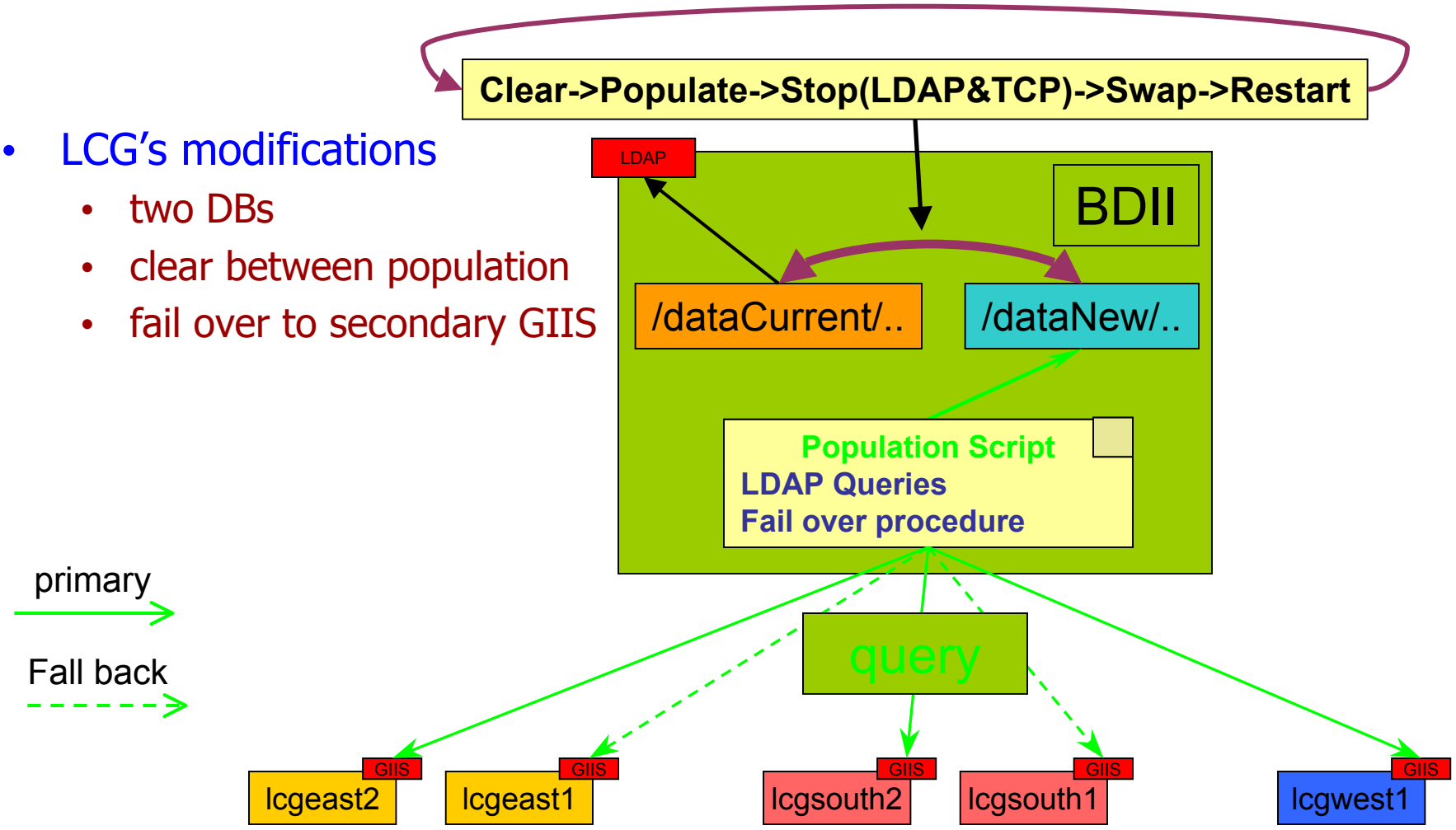
LCG1 IS Hierarchy



LCG1 BDII

Berkeley DB based Information Index (NIKHEF for EDG)

- LCG's modifications
  - two DBs
  - clear between population
  - fail over to secondary GIIS



Security (related to deployment)

- LCG Security Group (Dave Kelsey (RAL))
  - Define policies
- CERN LCG-GIS
  - LCG registration http://lcg-registrar.cern.ch/
  - LCG Certification Authority
    - http://lcg-registrar.cern.ch/pki_certificates.html
  - Tools for VO management
  - Host one VO
  - Distribution of security policies to sites
  - Maintains security contacts

Experiment Integration

Goal: Help experiments integrating their production and analysis environment with LCG Grid.

- One person assigned to each experiment. But global scope.
- Deep understanding of the middleware described in Guides and Manuals, Tutorials.
  - Interface definition for EDG Workload, EDG Data Management and POOL software for LCG-1
    (https://edms.cern.ch/file/384019/0.4/WP1-WP2.doc)
  - The LCG-1 Information System
    (https://edms.cern.ch/file/384587/0.2/LCG-1_Information_System.pdf)
  - LCG-1 User Guide (https://edms.cern.ch/file/412777/1/LCG-1-UserGuide.pdf)
  - Experiment Software Installation on LCG-1
    (https://edms.cern.ch/file/412781/1/SoftwareInstallation.pdf)
  - LCG-1 Tutor Manual Installation Guide
    (https://edms.cern.ch/file/412774/1/LCG-Manual-Installation.pdf)

Experiment Integration

- Providing solutions/testbed to exercise/integrate new middleware features after understanding experiment requirements:
  - ALICE: AliEn tests on LCG-1 (https://wwwlistbox.cern.ch/earchive/alice-support-lcg-eis)
  - Significant effort to create the CMS LCG-0 testbed: real production done and produced 2 million events (http://cmsdoc.cern.ch/cms/LCG/LCG-0/)
  - ATLAS exercises with software installation via PACMAN in the new proposed Experiment Software Installation LCG Tools; Integration with Grid3 (https://wwwlistbox.cern.ch/earchive/support-eis)
  - CMS integration with POOL. Exercise with usage of catalogues. (http://server11.infn.it/archive-cms-lcg-edt/)
- Identifying missing functionality and proposing solutions: exercises with SRB and D-cache. RLS Usage.

Experiment Integration: Areas of work

- Access to Mass Storage
- Usage of POOL
- I/O to remote Grid files
- Coherent set of API to interface with GRID services
- Adequate user monitoring and tracking tools
- Interoperability with US tools (integration until a better solution is found): SRB experiments, Grid3, ...
- Installation and Configuration of experiment software and dependencies

Experiment Software Installation

- Satisfy experiment requirements to be able to control software installation, validation, and publication at a site
- Accomodate the different configuration: shared vs non-shared experiment software areas in a coherent way
- Tools provided and under test.
- Need to solve the problem of dependencies and triggering of global installation on all Worker Nodes (next release of the tools)

SEE **Experiment Software Installation on LCG-1**
https://edms.cern.ch/file/412781/1/SoftwareInstallation.pdf

Experiment Status on LCG

- Alice, Atlas, CMS, and LHCb on LCG-1
- Basic functionality of the facility tested.
- Good list of problems reported which helps out in a better definition of a validation procedure for sites.
- Preparing for Data Challange
- EIS is providing first user support. Grid problems forwarded to LCG Deployment or LCG Certification

Support-eis@cern.ch

Problems (deployment)

- Sites without LCFGng have severe problems getting it right
  - middleware's dependencies too complex
  - Support only for WNs and Uis manual procedures
  - (Instant Grid would be good)
- Debugging site configurations
  - Discovery of the remote site's setup is hard
    - Especially hard if VO depended
  - History of the components, many config files
  - Changes take a long time (fly by e-mail)
  - Misleading error messages during installation
  - Site certification procedures not adequate
- Some sites are in contact with grids for the 1st time
  - "Beginners Guide to Grids" needed
- Time zones slow down the propagation of changes

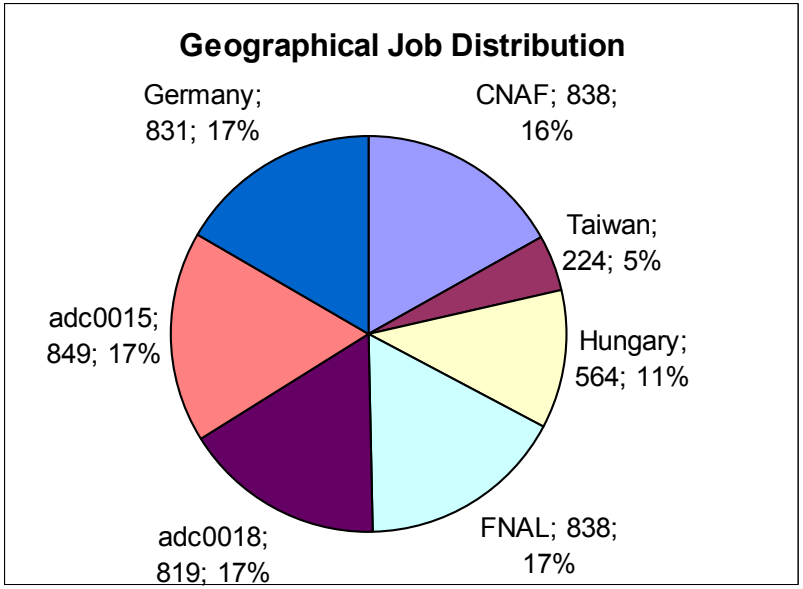
LCG

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

GRID

Stability&Operation

- Running jobs has greatly improved
- Only few MDS related problems (improvements under test)
- Focus now on: Hardening services for production
  - Jobs with realistic workload (long running, complex data access)
  - Chaotic usage test (multiple users, burst)
  - Integration with local production fabric (to add CPUs)
  - Operate services for extended periods
    - Do they "age" or "pollute" the platforms they are running on?
  - Capture "state" of services to restart them without loosing active jobs
  - Learn how to upgrade services (RMC, LRC…) without stopping
    - LCG1 can't be drained for upgrading
  - Integration of new components (GFAL, MSS, managed storage)



Test

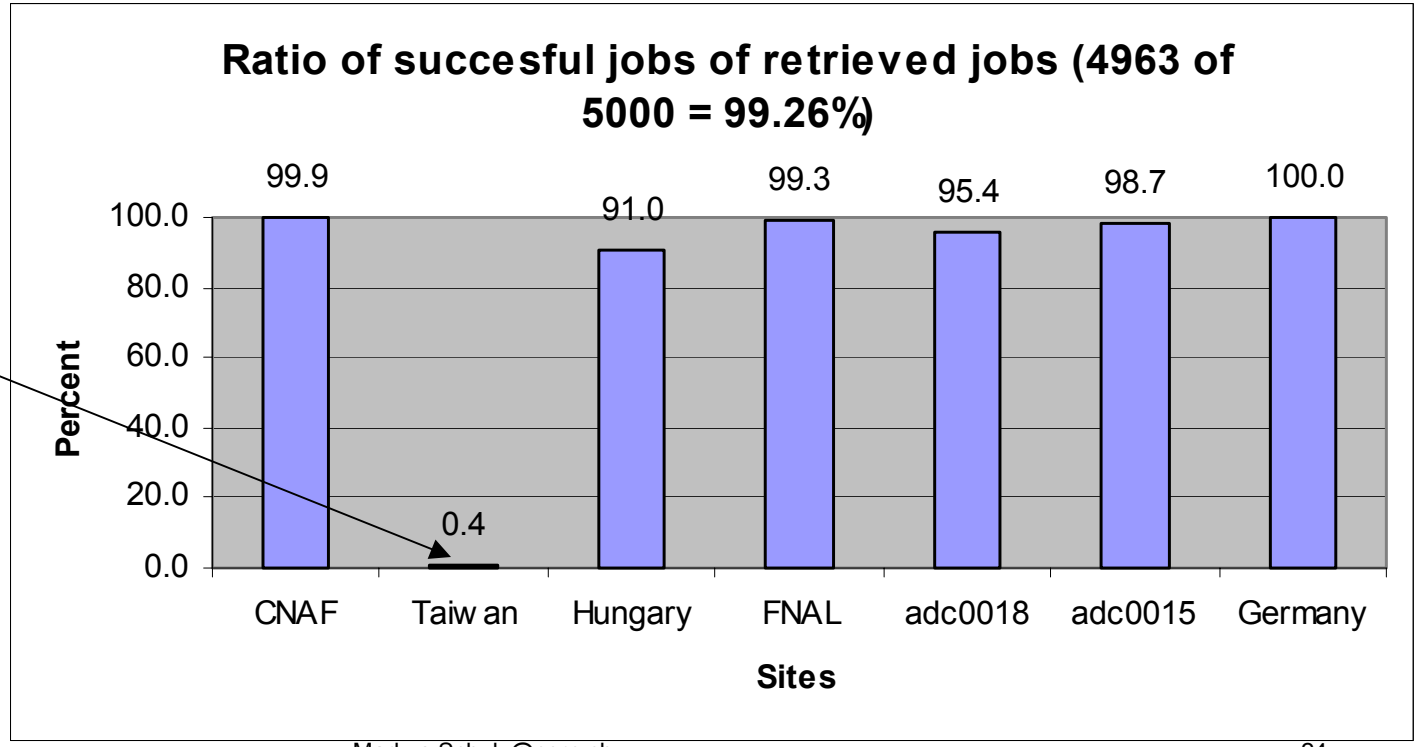
LCG 1.0 Test (19./20. Sept. 2003):

- 5 streams
- 5000 jobs in total
  - Input and OutputSandbox
  - Brokerinfo query
  - 30 sec sleep

Ingo Augustin

EDG WP8

Taipei changed configuration during the test.



Resources

- Hardware:
  - In addition to nodes used for service: ~20 nodes for verification
- Personnel
  - Deployment (grew within the last 3 month from 3 to 8 persons)
    - Security related activities 2
    - Installation procedures 1.5
    - Release preparation 1
    - Fault detection, site certification 1.5
    - Hardening of services 1
    - Running of local services 0.5
    - Support for first users 0.5
  - Experiment Integration Support (grew from 2 to 4)
    - One support person per experiment
    - No strict separation between tasks

Summary

- Middleware was 3 months late
  - Less: functionality, tests, experience with operation
- Number of sites now at scale foreseen (20 sites)
  - Deployment process seems to work
  - Need better site certification
- Experiments are testing the system
  - good end user documentation
  - discover problems (config. errors)
  - SW- distribution process implemented, needs testing/acceptance
- Very little time to turn this into a real production system
  - Critical components are just arriving (SE)
  - Has to be done incrementally on the running service
- Deploying the software at new sites is not always easy
  - various reasons (attitude, complexity, priorities, acceptance of tools)