LHCOPN operations Presentation and training CERN's session II

1- Goals and general overview of operational model

Guillaume Cessieux (FR-CCIN2P3, EGEE SA2)

CERN, 2009-06-16

Agenda

- Goal
- Overview
- Actors
- Information repositories
- Events management
  - Incident
  - Maintenance
  - Change
- Grid interactions
- Processes tools

Remark

- **Everything documented and maintained on**
  - https://twiki.cern.ch/twiki/bin/view/LHCOPN/OperationalModel

Goal of the ops model

- **Smartly manage LHCOPN at L3 delivering best network service as possible to WLCG**

- **LHCOPN objectives**
  - T0 – T1 traffic
    - T1 – T1 traffic as best effort
      - *T1-T1 links primary goal: T0-T1 backups links*
    - \+ Backup through generic IP

- **LHCOPN is a key block of infrastructure around WLCG**

MoU (1/3)

- **No particular MoU on LHCOPN operations, part of WLCG MoU signed by T1s**
  - http://lcg.web.cern.ch/LCG/MoU/Goettingen/MoU-Goettingen-18MAR09.pdf
  - Page A.3.2 (T0), A.3.4 (T1s)
  - For T0:

[5] (time running)/ (scheduled up-time)

Service	*Maximum delay in responding to operational problems*			*Average availability*[5] *measured on an annual basis*	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Raw data recording	4 hours	6 hours	6 hours	99%	n/ a
Event reconstruction or distribution of data to Tier-1 Centres during accelerator operation	6 hours	6 hours	12 hours	99%	n/ a
Networking service to Tier-1 Centres during accelerator operation	6 hours	6 hours	12 hours	99%	n/ a
All other Tier-0 services	12 hours	24 hours	48 hours	98%	98%

MoU (2/3)

For T1s:

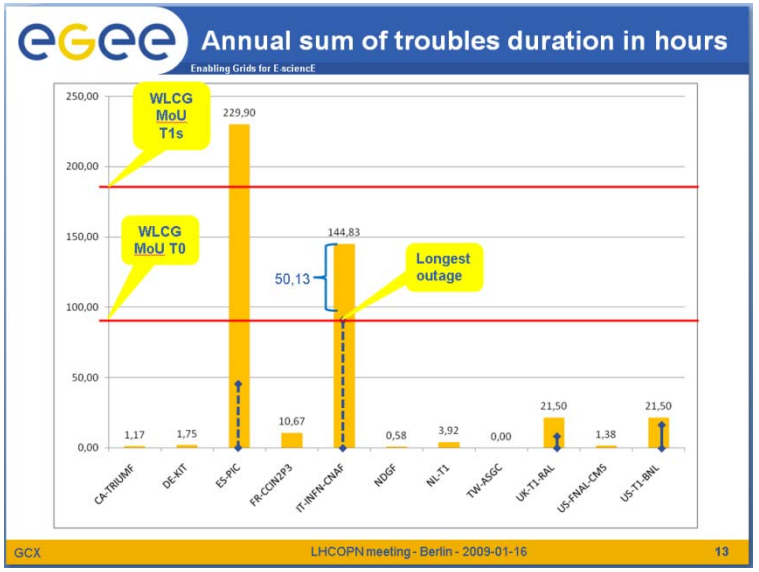
Service	*Maximum delay in responding to operational problems*			*Average availability*[5] *measured on an annual basis*	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Acceptance of data from the Tier-0 Centre during accelerator operation	12 hours	12 hours	24 hours	99%	n/ a
Networking service to the Tier-0 Centre during accelerator operation	12 hours	24 hours	48 hours	98%	n/ a
Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres outwith accelerator operation	24 hours	48 hours	48 hours	n/ a	98%
All other services[6] – prime service hours[9]	2 hour	2 hour	4 hours	98%	98%
All other services[6] – outwith prime service hours[9]	24 hours	48 hours	48 hours	97%	97%

[5] (time running)/ (scheduled up-time)

MoU (3/3)

- **Raw conclusion**
  - T0:
    - Response delay: 6 hours
    - Unexpected downtimes: 3.65 days/year ~ 87 hours
  - T1s
    - Response delay: 12 hours
    - Unexpected downtimes: 7.3 days/year ~ 175 hours

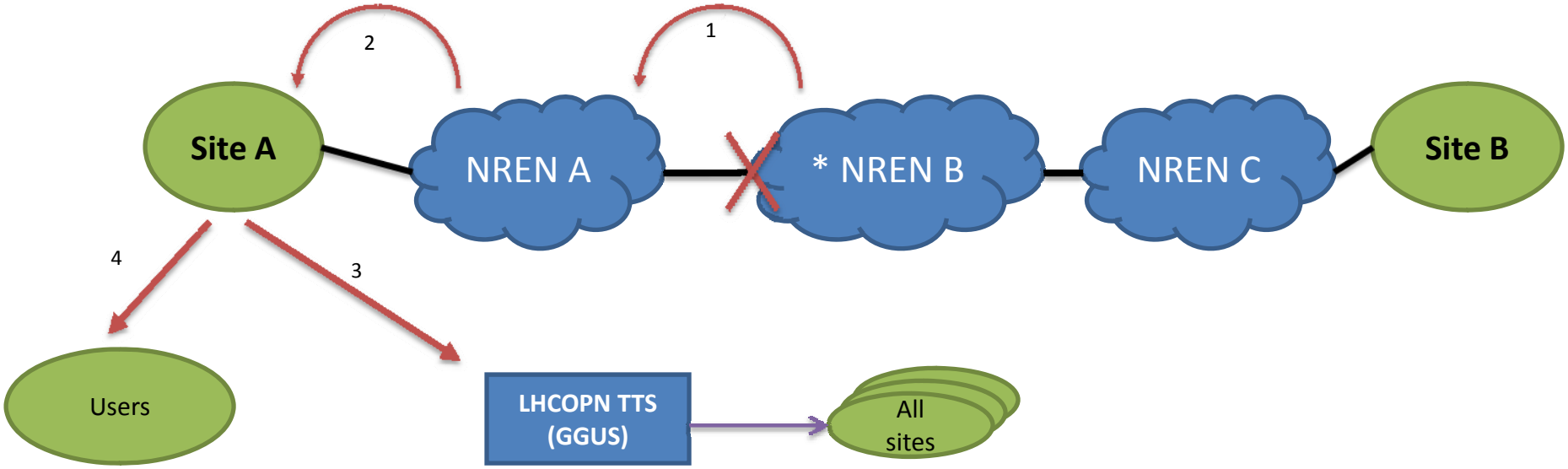
- **This seems really achievable**
  - Cf. https://edms.cern.ch/document/982588/
    - But true scheduled downtimes previously not correctly handled
  - Delays in announcements to be respected...



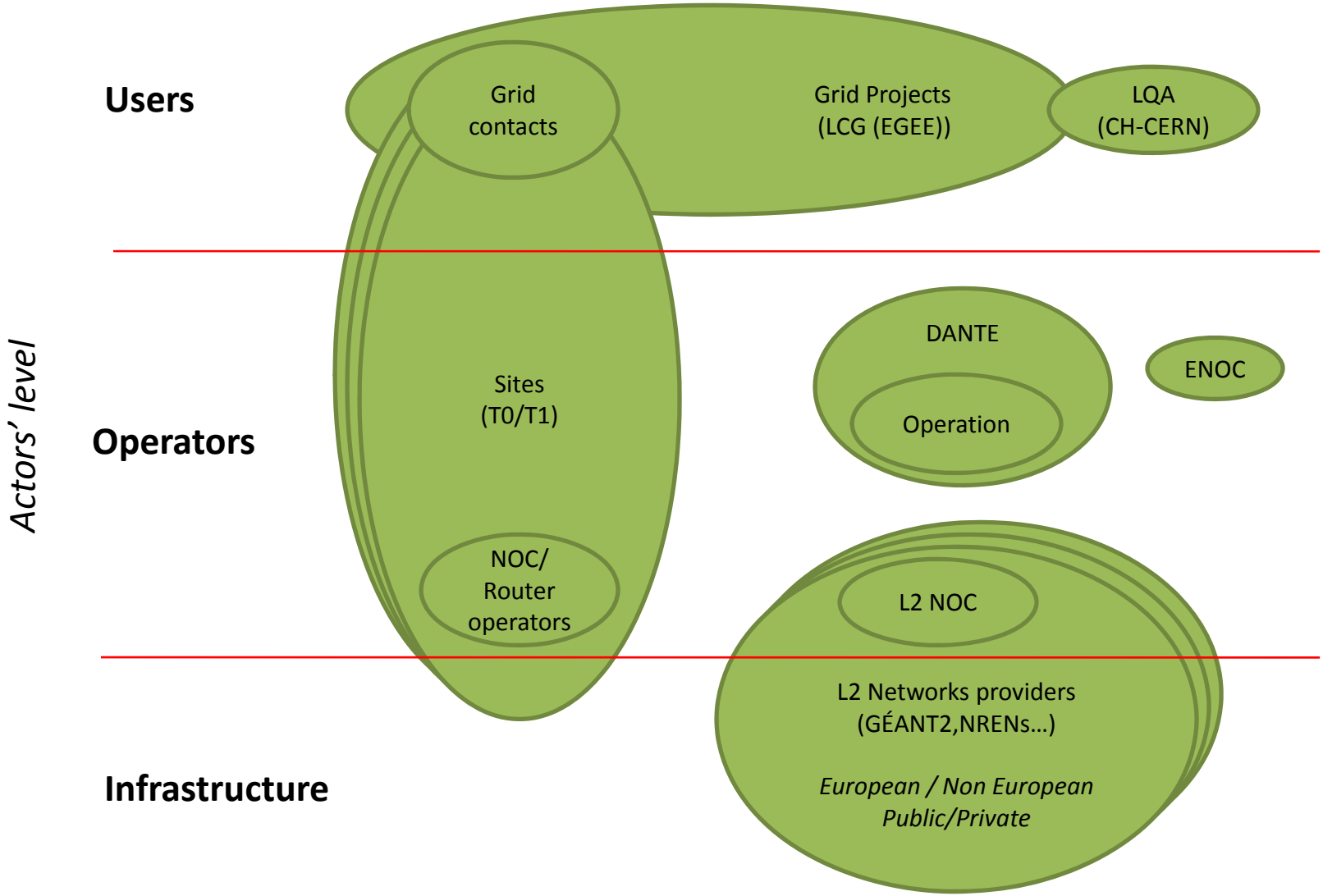
Overview (1/2)

- **Federated operational model with key responsibilities on sites**
  - Interaction with network providers
  - Management of network devices on sites
  - Interaction with the Grid
- **Some information centralised**
  - Serialisation of fault resolution and avoid duplicated information
  - TTS, web repository…

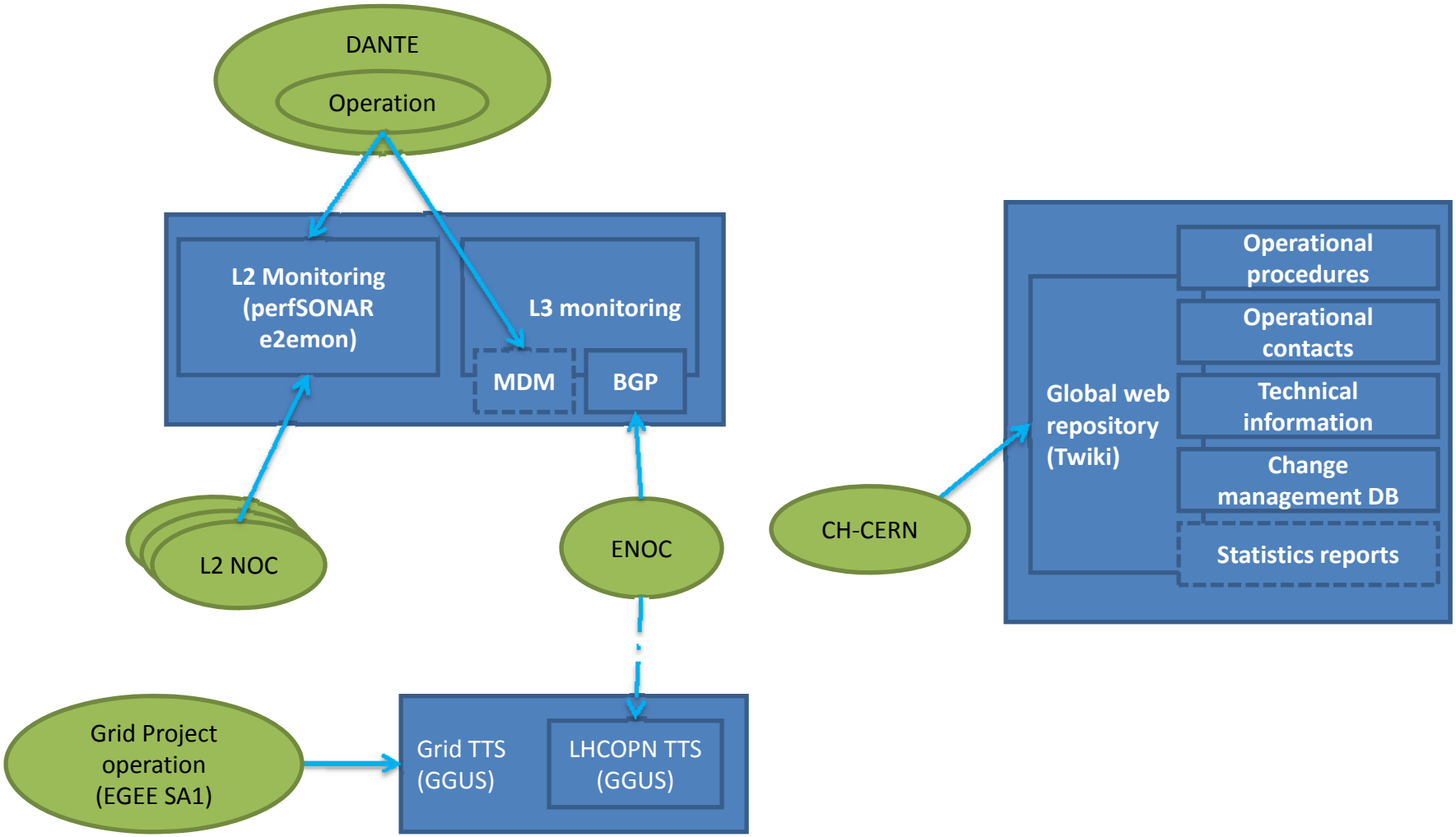
Overview (2/2)


2
1
Site A
NREN A
* NREN B
NREN C
Site B
4
3
Users
LHCOPN TTS
(GGUS)
All
sites

Actors


Actors' level
Users
Operators
Infrastructure
Grid contacts
Grid Projects (LCG (EGEE))
LQA (CH-CERN)
Sites (T0/T1)
NOC/ Router operators
DANTE
Operation
ENOC
L2 NOC
L2 Networks providers (GÉANT2,NRENs…)
European / Non European
Public/Private

Information repositories



Threshold for processes

- **Any events**
  - Lasting more than 1 hour
  - or occuring more than 5 times an hour
  - Should have a ticket in the TTS
- **Otherwise could be silently handled**
  - But good to report them (statistics, cross checking…)

3 kinds of events

- **Incident**
  - Unscheduled event
  - Generic process when cause and location unknown
- **Maintenance**
  - Scheduled event
- **Change**
  - Scheduled change on the infrastructure
  - Implemented by a maintenance if it impacts!

L2 vs L3

- **LHCOPN built as L2 paths ending on sites**
  - True, some exceptions…
- **Shortcuts**
  - L2: OFF-SITE: optical level, fibre cuts in NREN, etc.
  - L3: ON-SITE: Router down, power cut, BGP flaps, filtering, IOS upgrade etc.

Processes

- **6 key processes to handle 3 kinds of event**

Complexity

- **Incident management**
  - Global Problem management processes
  1) L3 incident management
  2) L2 incident management
  - *Escalated incident management*
    - *(~ trouble > 1 week)*
- **Maintenance management**
  3) L3 maintenance management
  4) L2 maintenance management
- **Change management process**
  5) L3 change management
  6) L2 change management

Unscheduled
(Minimum for on duty people…)

Scheduled

Change VS Maintenance

- **Change to broadcast and document the change**
- **Any change with a impact should be implemented with an associated maintenance**

Sample events

- **Incident**
  - L2: Dark fibre outage
  - L3: Router down, BGP filtering, bad routing
- **Maintenance**
  - L2: Fibre rerouted, fibre to be cleaned
  - L3: Scheduled power cut on site, IOS upgrade
- **Major change**
  - L2: New LHCOPN link
  - L3: New IP adresses, prefixes, filtering

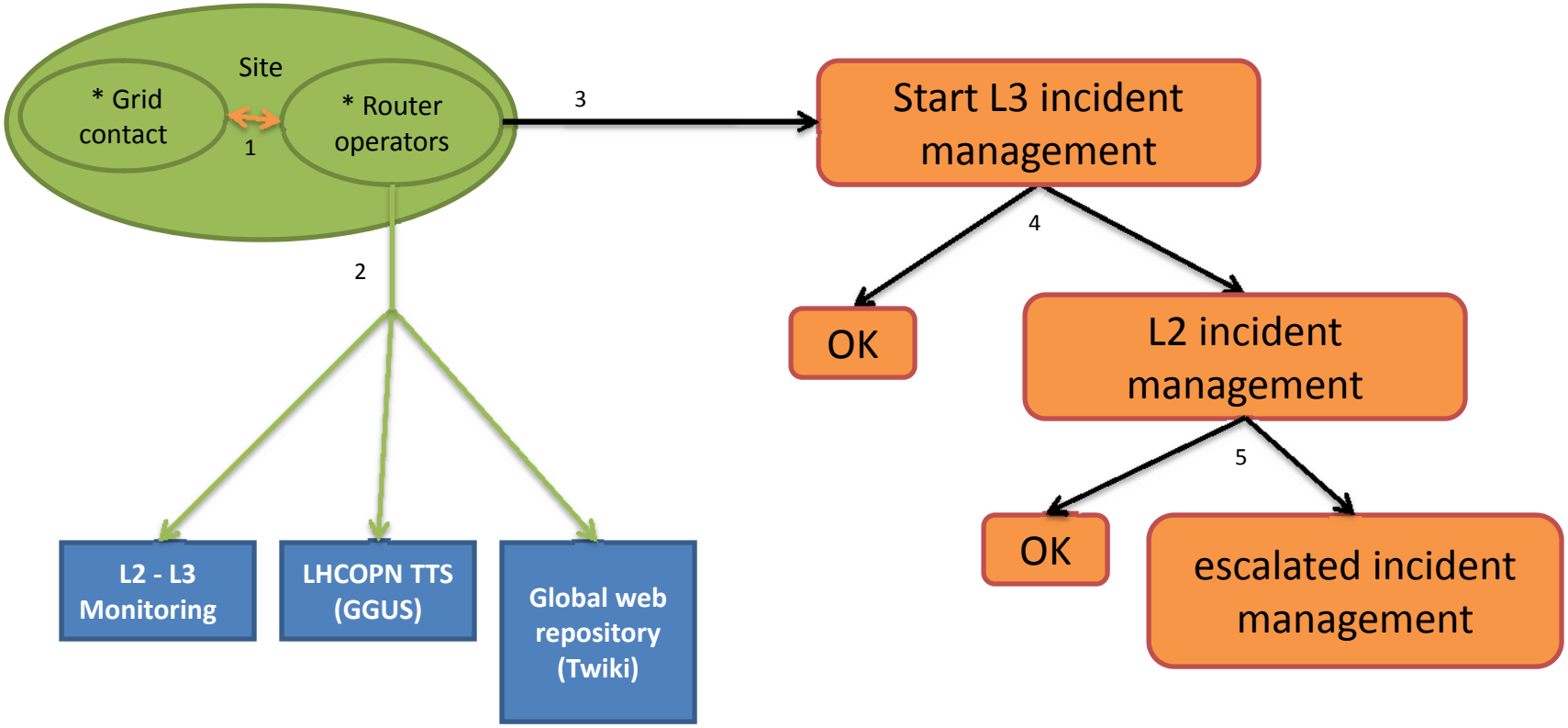
Who is who

- **Router operator**
  - People acting on sites' network devices = You
- **Network provider**
  - NRENs, GÉANT2 etc.
- **Grid contact**
  - Role supported by each sites
  - Typicaly FTS and Dcache managers etc.

Responsibilities

- **Outages on links between T0 and T1 are of responsibility of T1s (who ordered the link)**
- ***Responsibility for outages on T1-T1 links are being investigated***
- **Responsibility for GGUS' ticket is on the site which the ticket is assigned to**
  - Only one entity responsible at any time
    - Avoid the no one move effect

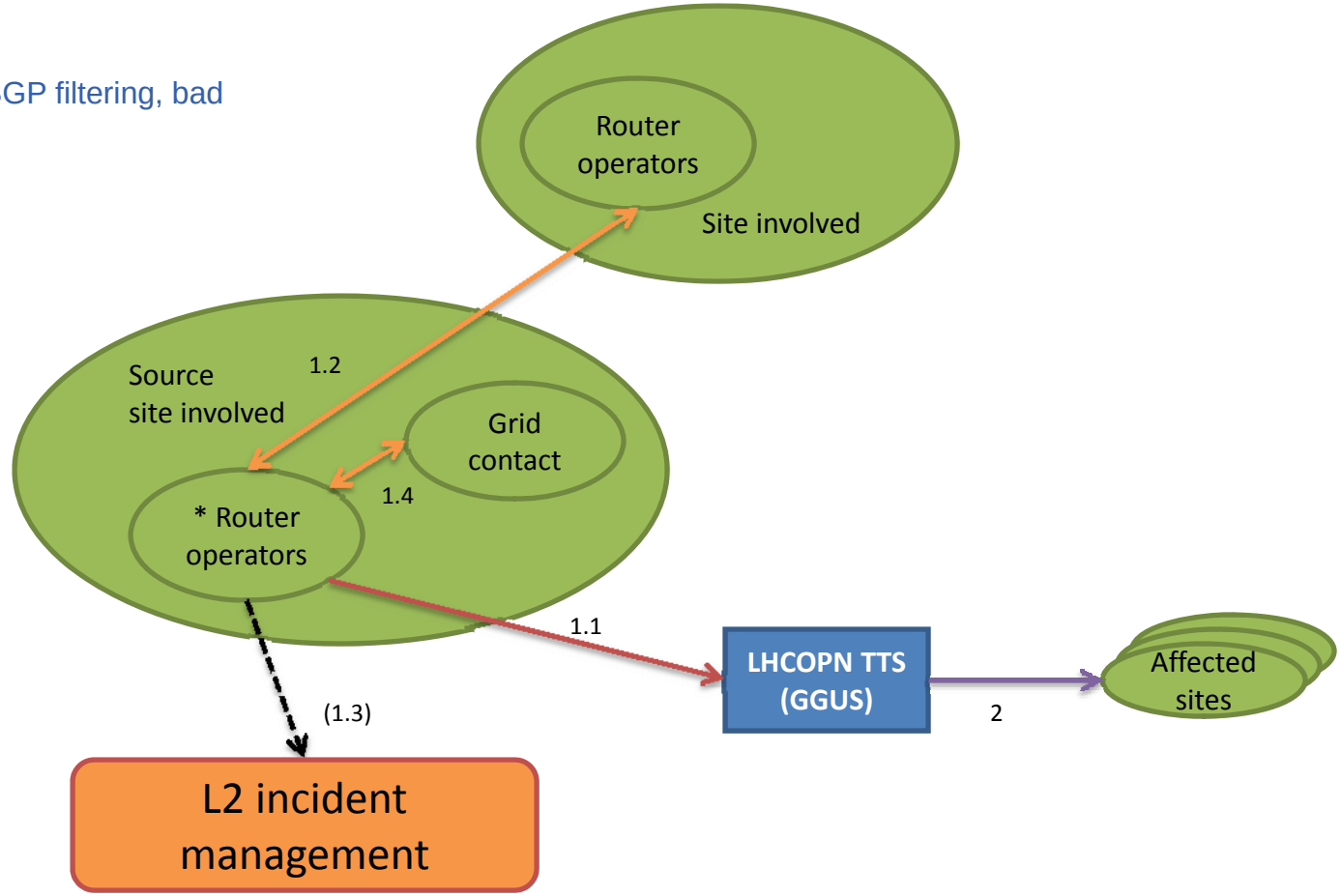
Generic process for incident



1.1 Incident management

L3 incident management process

Scope: Router down, BGP filtering, bad routing...

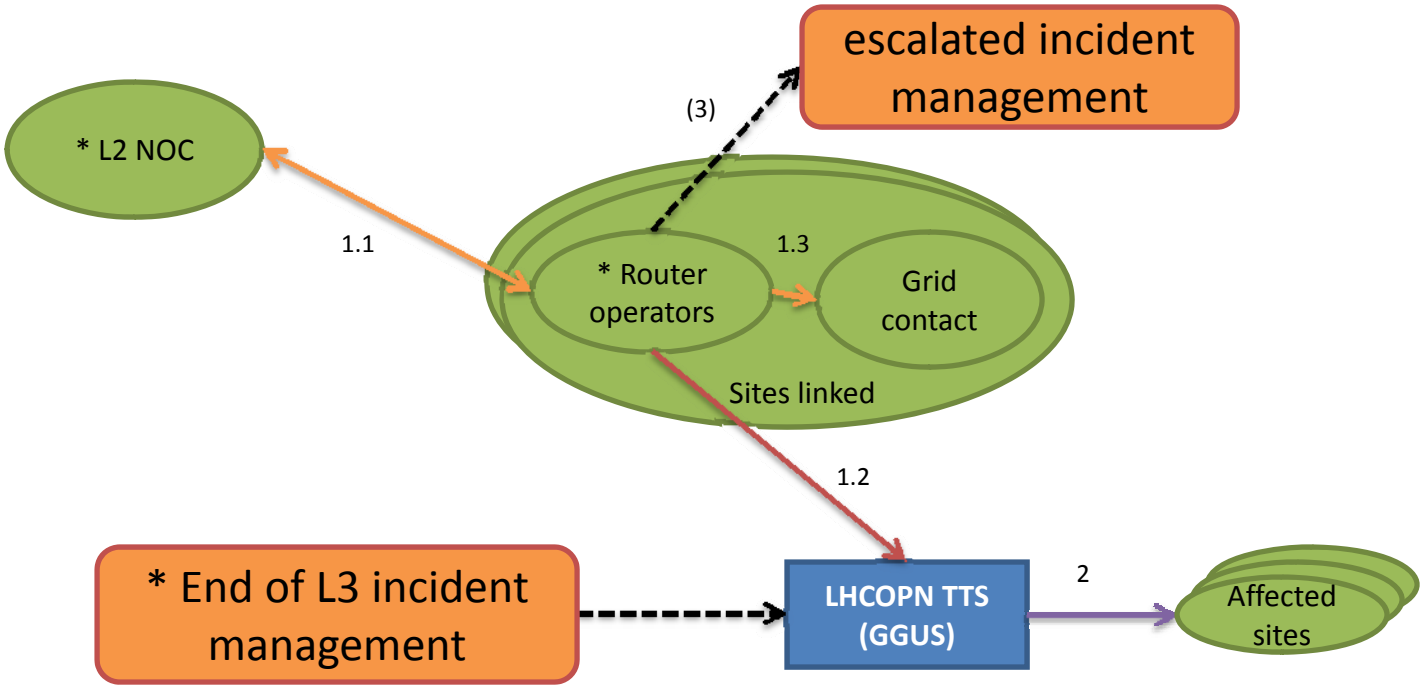


L3 incident management process
Sample use case: Power cut at FR-CCIN2P3

1. Incident registration: Put a GGUS ticket into the TTS
2. Warn Grid contact and give them reference of network ticket
3. Update it
4. Close it

L2 incident management process

Scope: Dark fibres outages...



L2IM: use case: RENATER fibre cut CERN-IN2P3-LHCOPN-001 down

1. **Start Generic process**
2. **Start L3 incident management**
   - Nothing at CH-CERN, should be L2 related
3. **Then go to L2 incident management**
   1. See with RENATER NOC what happens
      - Maybe open a ticket to their NOC
   2. Put a ticket in the LHCOPN TTS
   3. Warn Grid contact (and give them ticket #)
   4. Follow

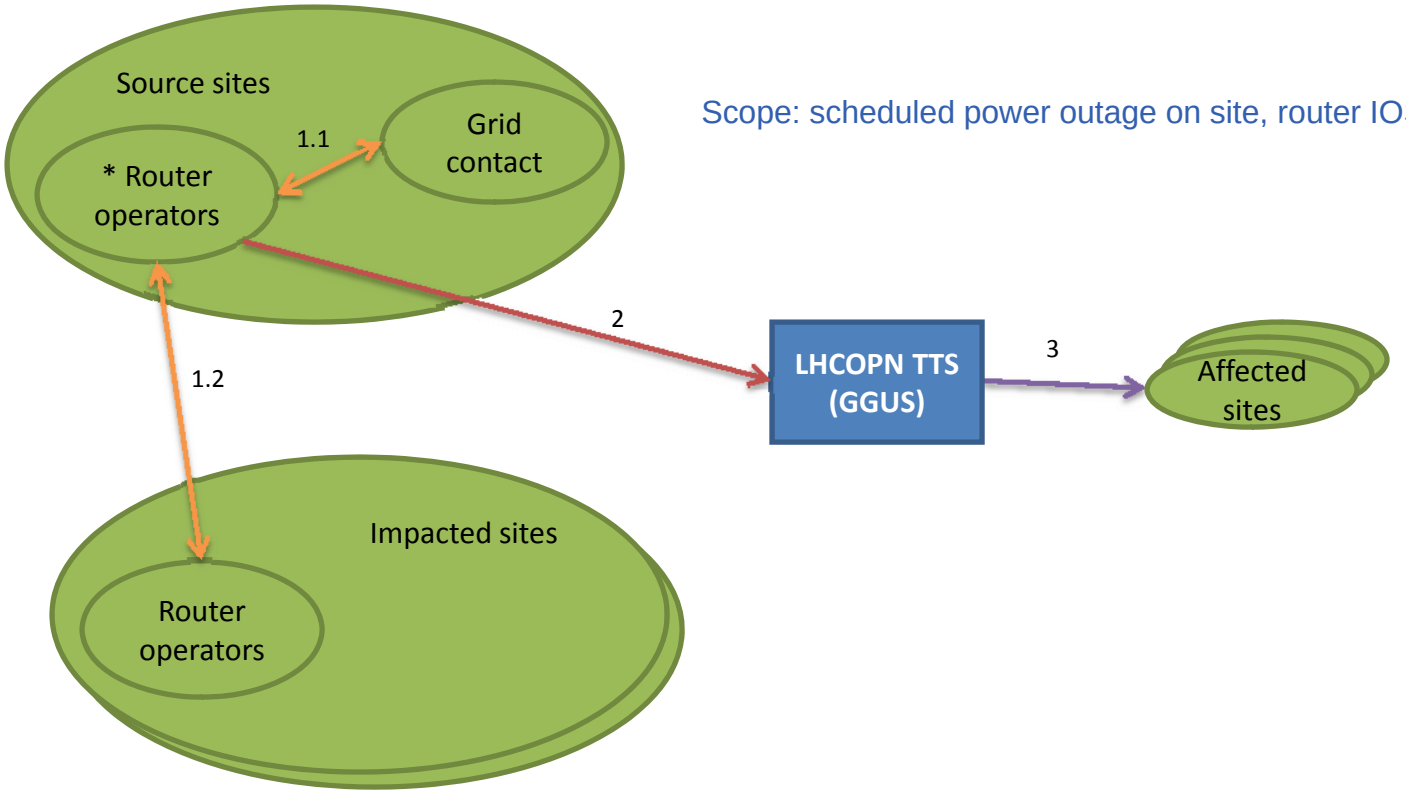
1.2 Maintenance management

Maintenance notice delay

Impact duration	Notice window
More than 1 hour	1 week
Less than 1 hour	2 days
No impact	1 day

Otherwise events might be computed in statistics as Incident...

L3 Maintenance management



Scope: scheduled power outage on site, router IOS upgrade, ...

A → B A interacts with B

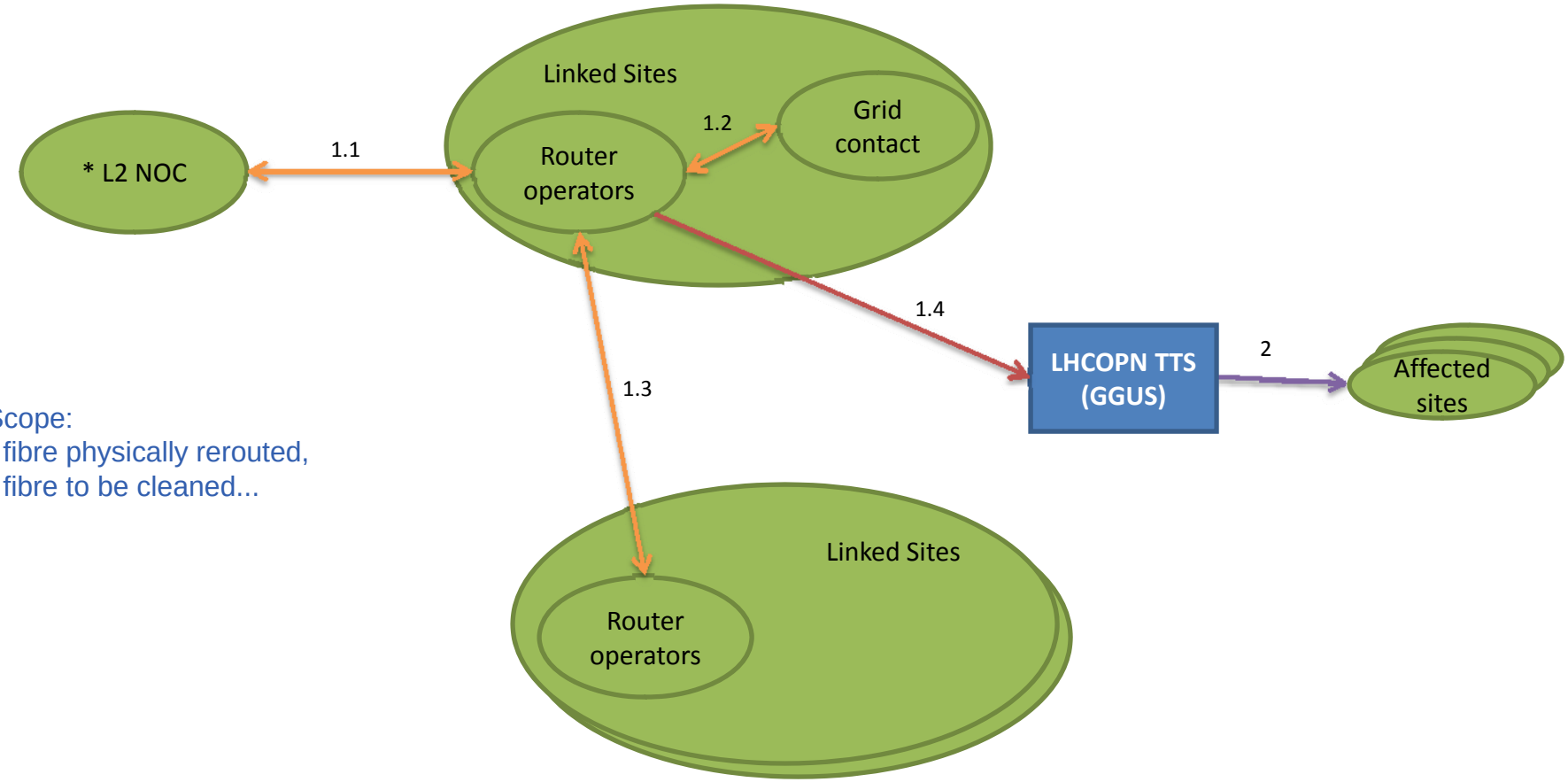
A → B A notifies B

A → B A reads and writes B

L3M: use case: Scheduled power cut in network area at FR-CCIN2P3

- Service Impact > 1h, maintenance window = 1w
- Warn Grid contact and see if ok
- (Ask DE-KIT and CH-CERN if no overlaping event foreseen)
- Put a ticket about in the TTS
  - Yes one week in advance
  - Give ticket # to Grid contact
- Update, follow, close ticket the D-day

L2 maintenance management



Scope:
- fibre physically rerouted,
- fibre to be cleaned...

A → B A interacts with B

A → B A notifies B

A → B A reads and writes B

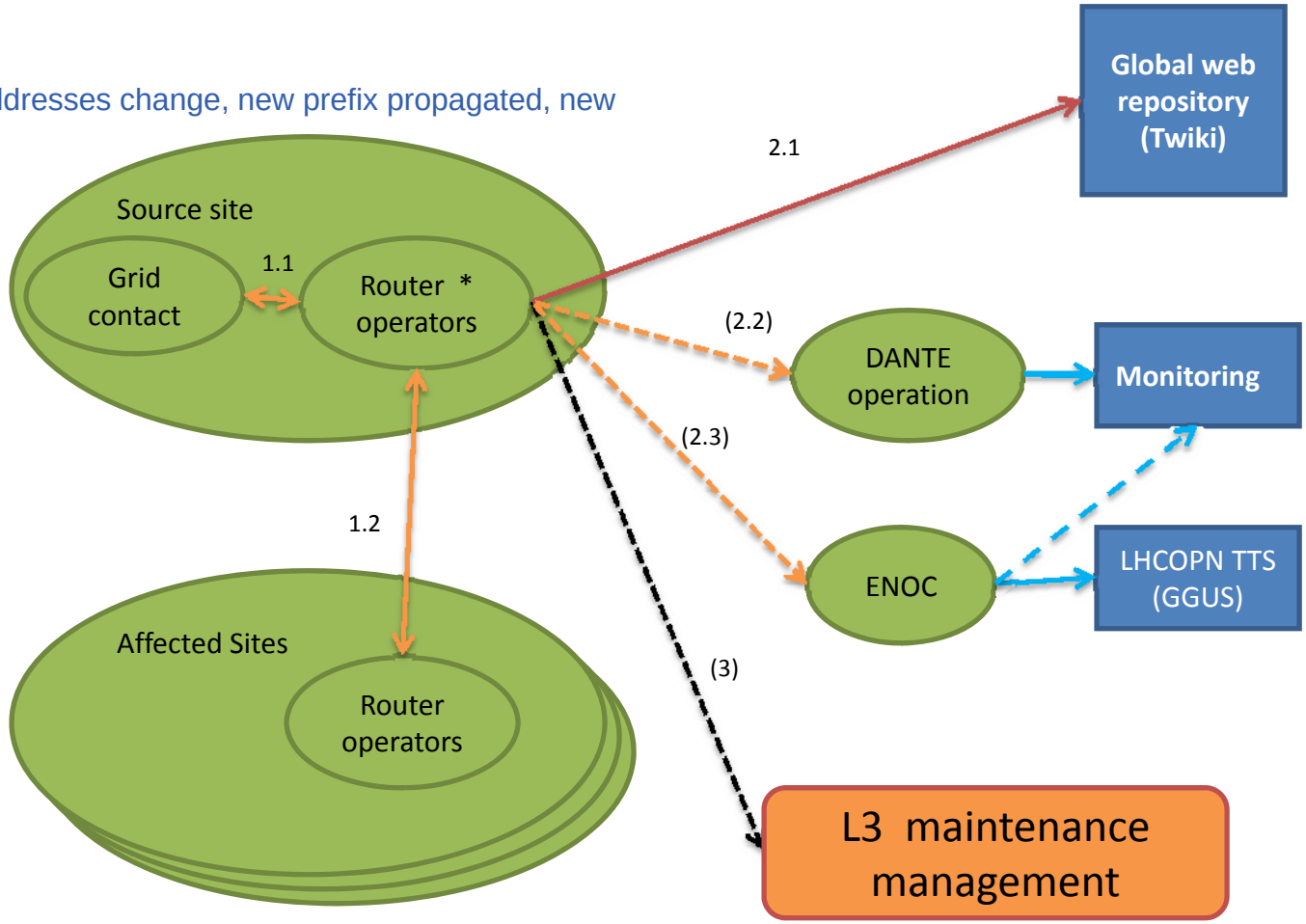
L2M: use case : RENATER scheduled work on CERN-IN2P3-LHCOPN-001

- **Received ticket from RENATER**
- **Link will be down 6 hours**
  - No impact on service to be confirmed with DE-KIT and CH-CERN
  - See also with Grid contact as this may impact performance
- **Put a ticket at least 1d before the event**
  - Give reference to Grid contact
- **Update and follow ticket**

1.3 Change management

L3 Change Management

Scope: IP addresses change, new prefix propagated, new filtering



A → B A interacts with B

A → B A reads and writes B

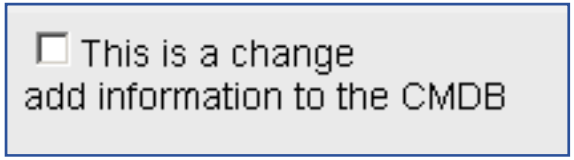
A → B A notifies B

A → B A is responsible for B

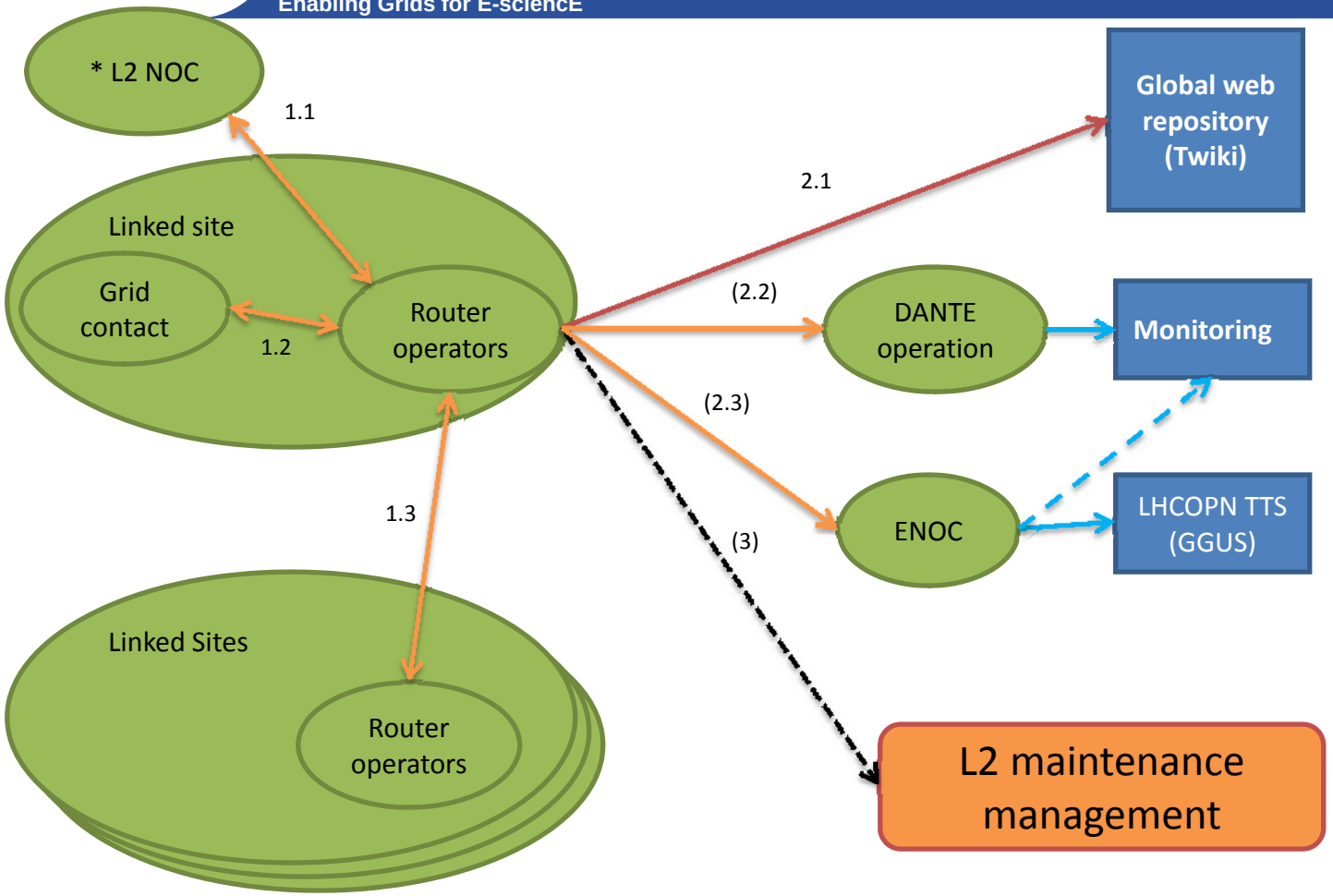
L3C: use case: Change of p2p IP adresses

Simplified

- **No change on service delivered**
  - Not of interest for Grid contact
- **Discuss with CH-CERN and DE-KIT about the change**
- **Document the scheduled change on twiki and update technical informations**
  - https://twiki.cern.ch/twiki/bin/view/LHCOPN/ChangeManagementDatabase
  - https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome part "Technical Information"
- **Deal with DANTE Ops (e2emon) & ENOC to have monitoring adapted**
  - In the background as they should not act on the TTS
    - operations AT dante.org.uk;enoc.support AT cc.in2p3.fr
- **Implement (=commit) the change with a maintenance**
  - Tick the box



L2 Change Management



A → B A interacts with B

A → B A reads and writes B

A → B A notifies B

A → B A is responsible for B

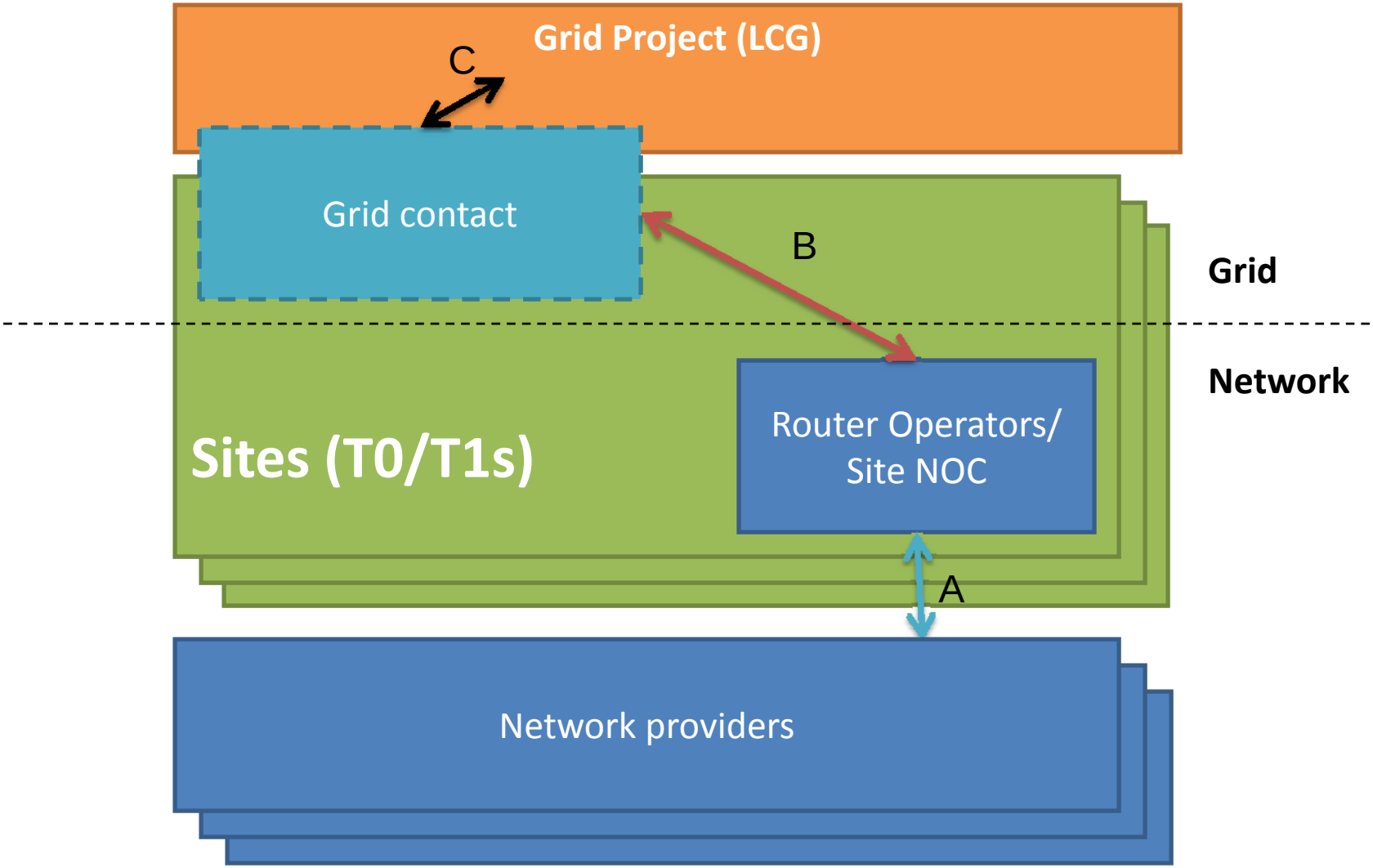
Scope: New LHCOPN L2 link, L2 link with new physical path, change of L2 network provider for a segment...

L2C: use case: New link CERN-IN2P3-LHCOPN-002 provided by SWITCH

Simplified

- See with SWITCH NOC details
- See with CH-CERN, DE-KIT new p2p IPs and routing policy
- Warn Grid contact: New bandwidth and backup possiblities for the project
- Document the scheduled change and update technical informations
  - https://twiki.cern.ch/twiki/bin/view/LHCOPN/ChangeManagementDatabase
  - https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome part "Technical Information"
- Deal with DANTE Ops (e2emon) & ENOC to have monitoring adapted
- No change on infrastructure without tickets
- Put a L2 maintenance ticket to commit changes
  - This is the root event, even if L3 changes are also performed
    - IP adresses, routing and testing period before production use

Grid interactions (1/4)



Grid interactions (2/4)

On going work

A. Daily operational workflow

- Scheduled and unscheduled outages – what to practically do?
  - Try to also avoid overlap of network events & Grid events
- Each site is responsible – No central entity

→ Use simple existing things in place

- Existing tools, processes and communication channel to be used
  - EGEE broadcasting tools etc.
  - Standard GGUS system to reference LHCOPN TT
- Grid contacts could also report in the daily WLCG phoneconf when needed
  - https://twiki.cern.ch/twiki/bin/view/LCG/WLCGOperationsMeetings
  - Can be done offline: e-mails, reading minutes etc.
    - *But phone turned on when needed*
  - Key point: VOs and experiments are reached here

Grid interactions (3/4)

B. Upper level and long term interactions

- Regular problems, improvement and change requests, global assessment of the service delivered etc.

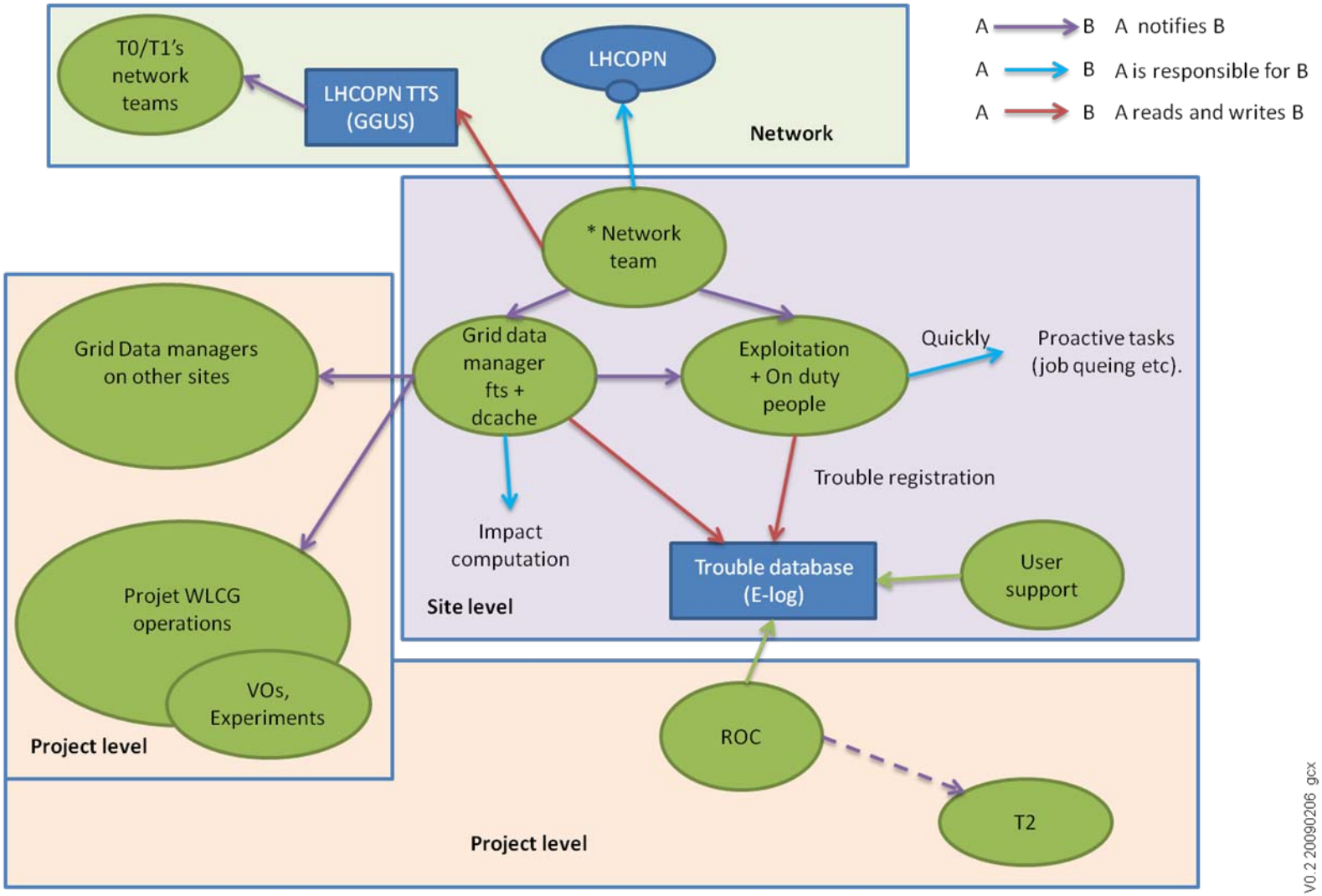
On going work

→ **A LHCOPN representative will be the exchange point between LHCOPN and Grid**

- Report to Grid from quarterly LHCOPN network ops phoneconf
  - Global view of infrastructure and ops
  - Quality assessment, key incident report etc.
- Import items from Grid on the agenda
- Write conclusions into some quarterly reports

Grid interactions (4/4)

Sample FR-CCIN2P3 implementation:



Operation review

- **LHCOPN ops phoneconf**
  - Globally review infrastructure and ops behaviour
    - Goal: To improve ops and infrastructure
  - Focus on issues
  - Each 3 months
  - 15:30 UTC is most convenient
    - But no timeslot easy for everyone
  - Supported by CH-CERN
  - Next is **2nd of July**

Backup tests

- **“Any resilience possibility should be regularly verified”**
  - ***Each 6 months***
- **Incident demonstrating backup efficiency counts as a backup test - if accurately reported**
  - **Reported on the twiki** https://twiki.cern.ch/twiki/bin/view/LHCOPN/LhcopnBackupTests
- **Unexpected backup test are encouraged if not impacting or disturbing**
  - Also testing Ops

Conclusion

- **Only 2 incident management processes to be fully known**
- **This is light?**
  - Change management processes really simplified!
- **Model should be flexible enough for site dependant implementation**
  - From huge layered NOCs to single guy
- **Open to improvements!**