Building a LAN to Support Multiple Lightpath Projects

Ronald van der Pol

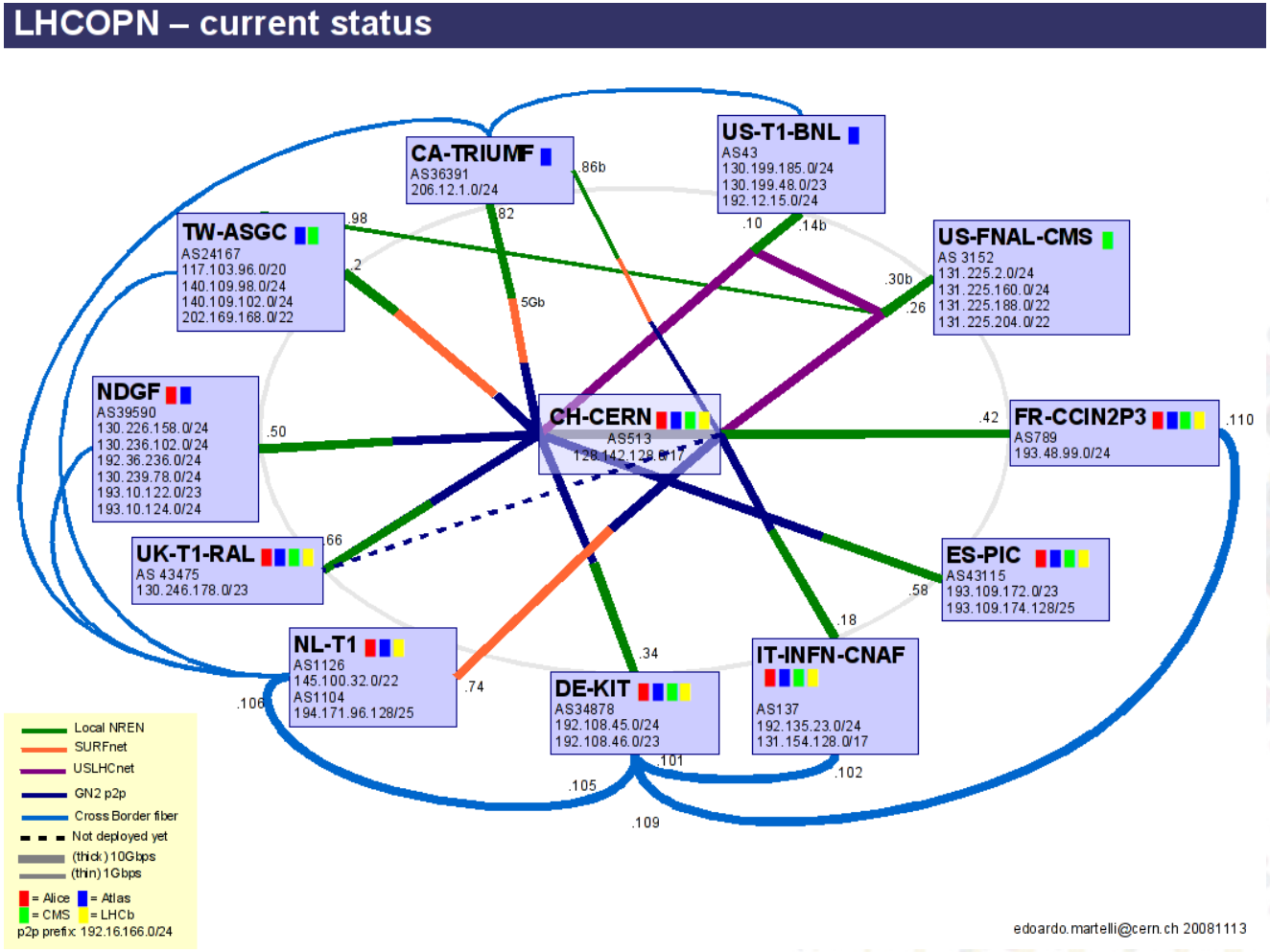
<rvdp@sara.nl>

About SARA

- Computing and Networking services
- Houses and operates national supercomputer Huygens
- Houses and operates national cluster Lisa
- LightHouse (joint lab of SARA, UvA and SURFnet for optical networking experiments and demos)
- SURFnet's subcontractor for SURFnet6 NOC
- SURFnet's subcontractor for NetherLight NOC
- One of the co-location sites of the AMS-IX
- CERN LHC Tier-1 site
- LOFAR Tier-1 site

rvdp@sara.nl

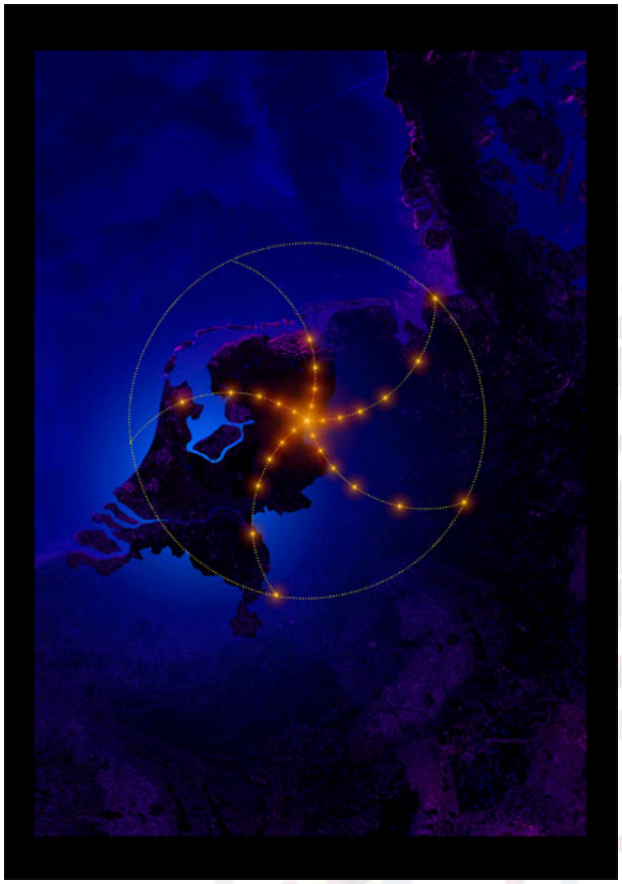
LHC OPN Tier-1 Site



rvdp@sara.nl

LOFAR Tier-1 Site

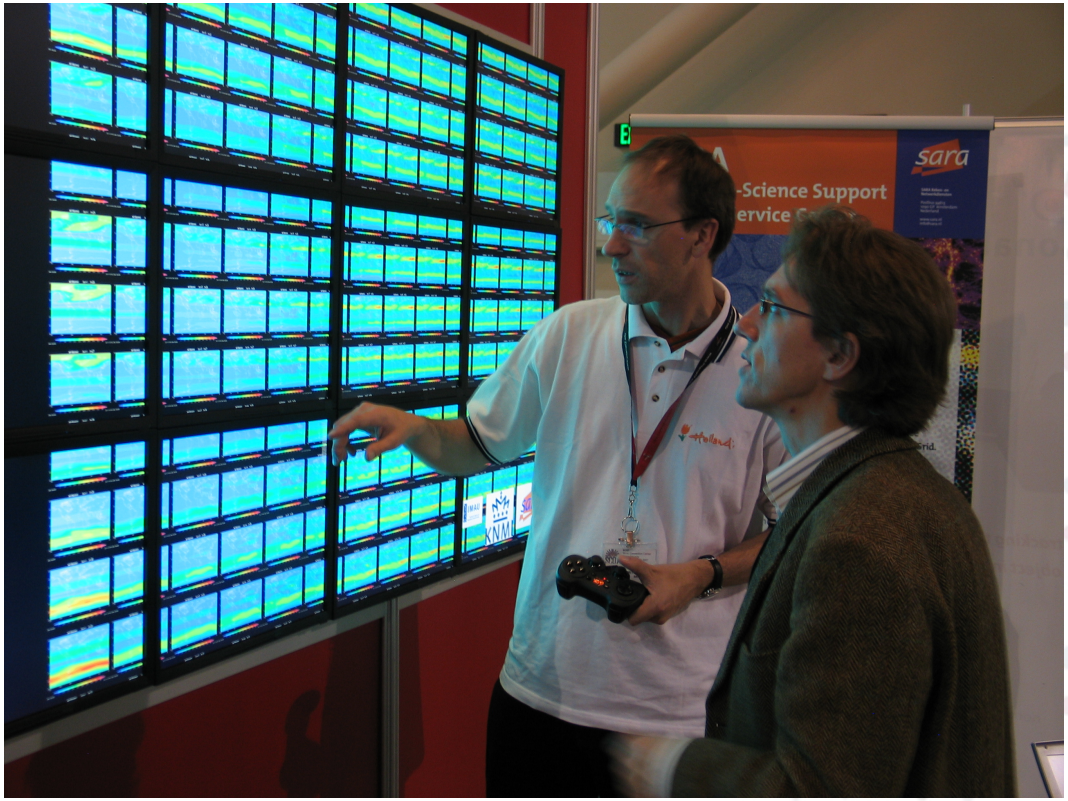
- LOw Frequency ARray
- Radiotelescope
- Consists of Sensor Fields
- Data Storage @ SARA



rvdp@sara.nl

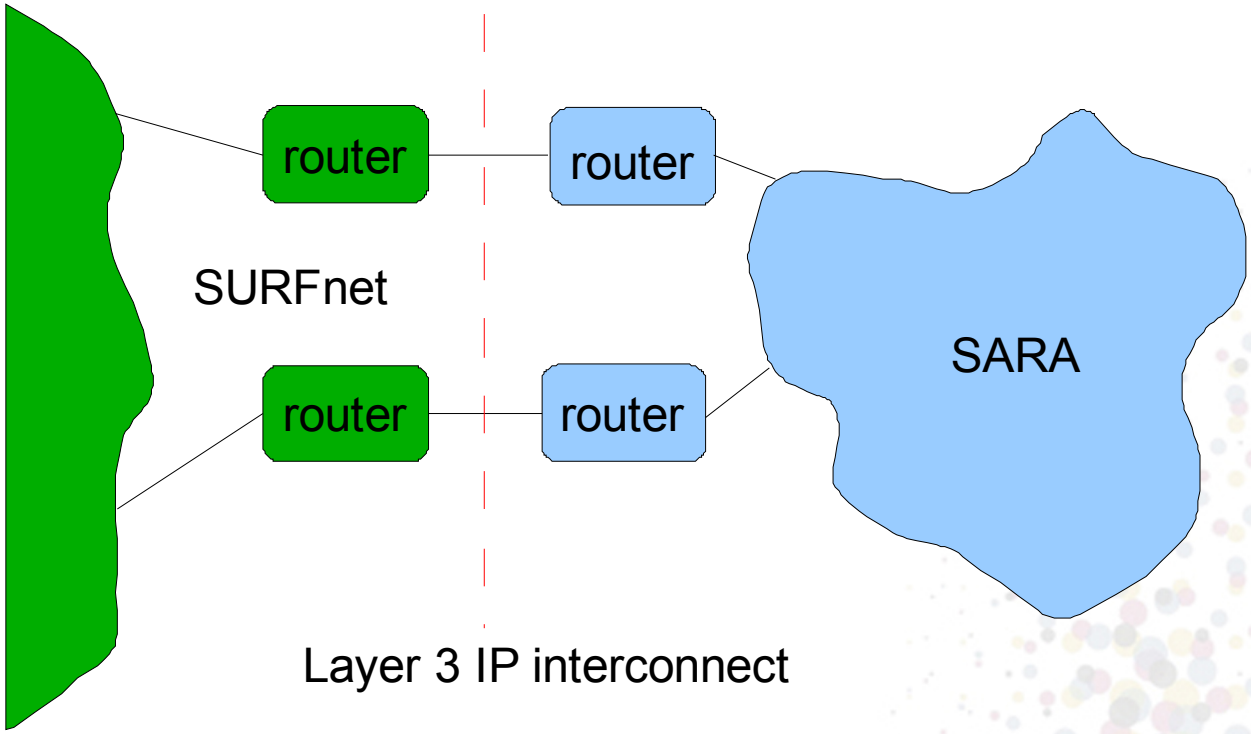
IMAU Climate Model

- Rendering at SARA
- Visualization at IMAU
- Connected with a SURFnet6 1G lightpath



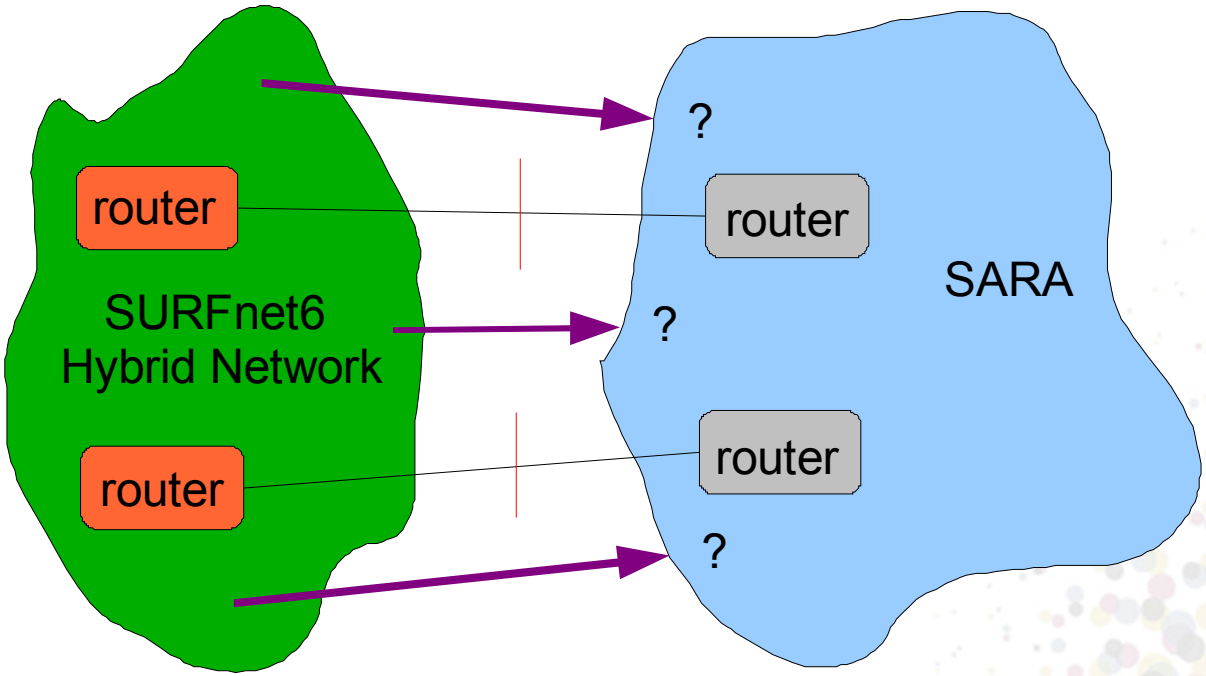
rvdp@sara.nl

Traditional ISP Connection



rvdp@sara.nl

Introduction of Lightpaths



rvdp@sara.nl

Lightpath Challenges

- Interconnect sites at L2 or at L3?
- How to handle security?
- How to handle addressing?
- How to protect against configuration errors and accidents at other site?

rvdp@sara.nl

L2 versus L3

- L2 pros
  - Cheap Ethernet switches
- L2 cons
  - No IP ACLs
  - Mixing of administrative domains
    - One broadcast domain, one IP subnet
- L3 pros
  - Well-known (we know how to do this between sites)
  - Supports ACLs and firewall
  - Easier fault resolution
    - Ping, traceroute, router reachability
- L3 cons
  - Routers (and L3 switches) usually more expensive

rvdp@sara.nl

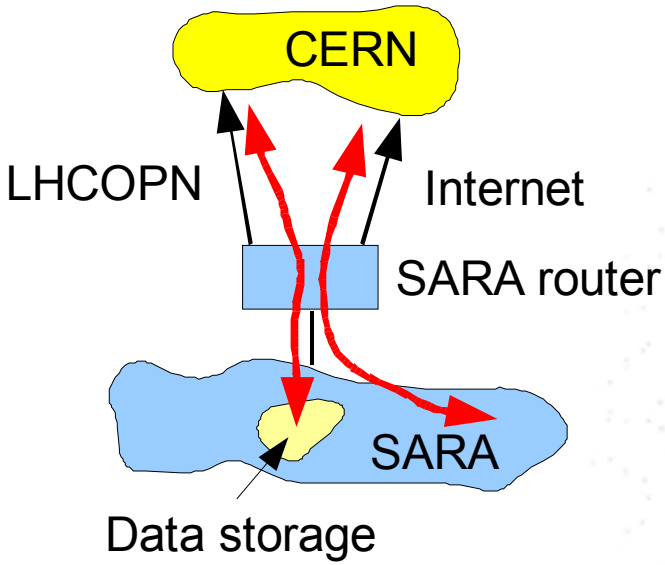
SARA's Requirements

- Keep services separated
  - Access to one service does not mean access to another service, unless explicitly allowed
- No (accidental) connectivity between lightpaths via SARA
- No (accidental) Internet connectivity via SARA
- Solution must scale to multiple services and multiple lightpath peer sites
- Solution must support multiple 10G connections
- No big routing tables on the servers
  - Only a default gateway
- Segmenting the routing tables
  - e.g. No LHCOPN prefixes in global routing table

rvdp@sara.nl

Problems Encountered in LHCOPN

- Only storage servers traffic allowed on the LHCOPN
- Other hosts and servers must reach CERN via Internet
- Traditional destination based routing does not work
- We needed to find a good, scalable solution



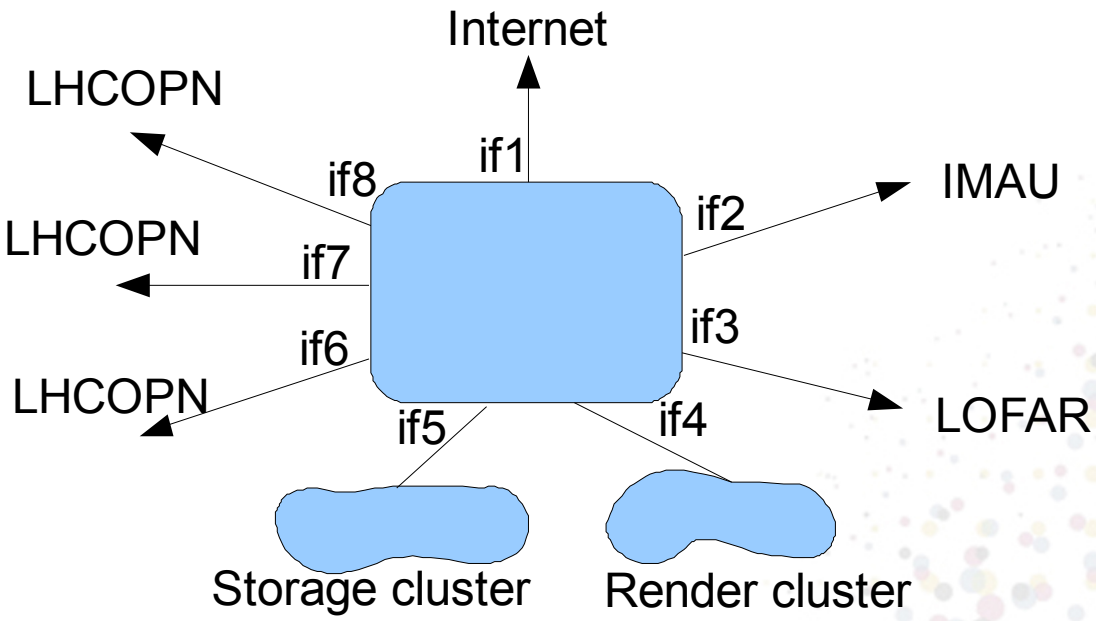
rvdp@sara.nl

SARA's Choices

- Interconnect at L3
  - L2 only for few very simple cases
- BGP routing
  - BGP detects when peer is unreachable
  - BGP needed when there are multiple paths
- Routing segmentation
  - Put each lightpath project in its own virtual router
  - Good way to keep projects and services separated

rvdp@sara.nl

Virtual Routing



Global Table: if1, if4, if5
VR1 (LHCOPN): if6, if7, if8
VR2 (IMAU): if2
VR3 (LOFAR): if3

rvdp@sara.nl

Virtual Router Solution

- *Virtual routing is a scalable way to keep services and lightpath peers separated*
- Problem with traditional destination based routing + ACLs:
  - ACLs are difficult to maintain
  - Not a scalable solution
  - Configuration errors mean unwanted access
- Problem with policy based routing:
  - Only 1 next hop, does not work with multiple links
  - Next hop is specified as specific interface
  - Does not use BGP, no route information exchange
  - No link failure detection when switches in path

rvdp@sara.nl

Problems Encountered

- Often little BGP knowledge at peer sites
- Many peer sites do not have a global AS
- Most routers have insufficient Virtual Routing capabilities
- We had to gain knowledge of virtual routing
- Detecting of link failures often difficult
  - Link failures do not propagate through Ethernet switches
  - (BGP session, 802.1ag, BFD, ...)

rvdp@sara.nl

Conclusions

- Supporting multiple lightpaths and multiple services is *not a trivial task*
- Virtual routing is a relatively simple way to handle the routing and separation requirements
- Routing requirements often result in the choice for BGP

rvdp@sara.nl

Thank You

Ronald van der Pol

rvdp@sara.nl