LHCONE POINT- TO-POINT SERVICE
WORKSHOP SUMMARY

Artur Barczyk/Caltech

Atlas Technical Interchange Meeting

Tokyo, May 2013

Background

- LHCONE data transport services
  - Routed IP service aka “VRF” service
    - virtualized routing environment, implemented
  - (Dynamic) Point to point service
    - being worked out, to be ready for LHC restart in 2015

- 1st LHCONE P2P workshop was held in December 2012
- 2nd workshop held May 2013 in Geneva

- This presentation aims at giving a representative summary of the May meeting
  - no claim to completeness
- URL: https://indico.cern.ch/conferenceDisplay.py?confId=241490

The Workshop

- Initial starting point: Whitepaper draft

- ~40 participants, representing
  - R&E Networking community
    - ESnet, NORDUnet, GEANT, Internet2, US LHCNet, others
  - LHC computing

Agenda

CALTECH HEP NETWORKING

Thursday, 2 May 2013

10:00 - 10:01 Meeting start

10:01 - 10:10 Welcome, Agenda bash *9'*
Speakers: Artur Jerzy Barczyk (California Institute of Technology (US)), Edoardo Martelli (CERN)

10:10 - 10:50 Introduction, Updates on activities since last Workshop *40'*
Speaker: Lars Fischer (NORDUnet)

10:50 - 11:20 LHCONE/OPN traffic patterns *30'*
Speaker: Ronald van der Pol

11:20 - 12:30 Understanding and definition of requirements *1h10'*
Speaker: William (Bill) Johnston (ESnet)

12:30 - 14:00 Lunch break

14:00 - 14:50 Application Perspe
Speaker: Dr. Tony Wil

14:50 - 15:30 LHCONE objective
Speaker: Dale Finkels

15:30 - 16:00 Coffe break

16:00 - 16:50 SA2 "Testbed as a
Speaker: Jerry Sobies

16:50 - 17:30 Whitepaper dicussi
status, open discussion
Speaker: Lars Fischer

17:30 - 17:31 Meeting end for the

Friday, 3 May 2013

09:00 - 09:01 Meeting re-start

09:01 - 10:30 Definition of possible scenarios, opportunities for a trial *1h29'*
Speaker: Inder Monga (ESnet)

10:30 - 11:00 Coffee break

11:00 - 12:30 Next steps *1h30'*
Speakers: Lars Fischer (NORDUnet), Artur Jerzy Barczyk (California Institute of Technology (US))

12:30 - 12:31 Meeting end

14:00 - 16:30 SDN/Openflow Discussion
Convener: Artur Jerzy Barczyk (California Institute of Technology (US))

14:00 **Introduction** *10'*
Speaker: Artur Jerzy Barczyk (California Institute of Technology (US))

14:10 **Openflow - possible directions** *10'*
Speaker: Inder Monga (ESnet)

Application Perspective (Tony Wildish, CMS)

- PhEDEx and “Bandwidth on Demand”
  - https://twiki.cern.ch/twiki/bin/view/Main/PhEDExAndBoD
- Data-placement in CMS
  - T0 -> T1: custodial data
    - Primary use-case for investigation/prototyping
  - T2 -> T1: harvest MC production
  - T1->T2, T2->T2: placement for analysis
  - #nodes, time-profile, concurrency vary considerably
- Long-duration transfers (TB, hours)
- What is the right service interface?

Application Perspective (Tony Wildish, CMS)

- Basic motivation: improved performance in long transfers
- Basic questions to address:
  - Cope with call blocking
  - must be able to prioritise requests
  - request parameters?
    - Min. & Max. Bandwidth
    - Min/Max. data volume

- **Request needs to return options**
  - decision taken by requester
  - binary replies not useful

CMS use cases (Tony Wildish), cont.

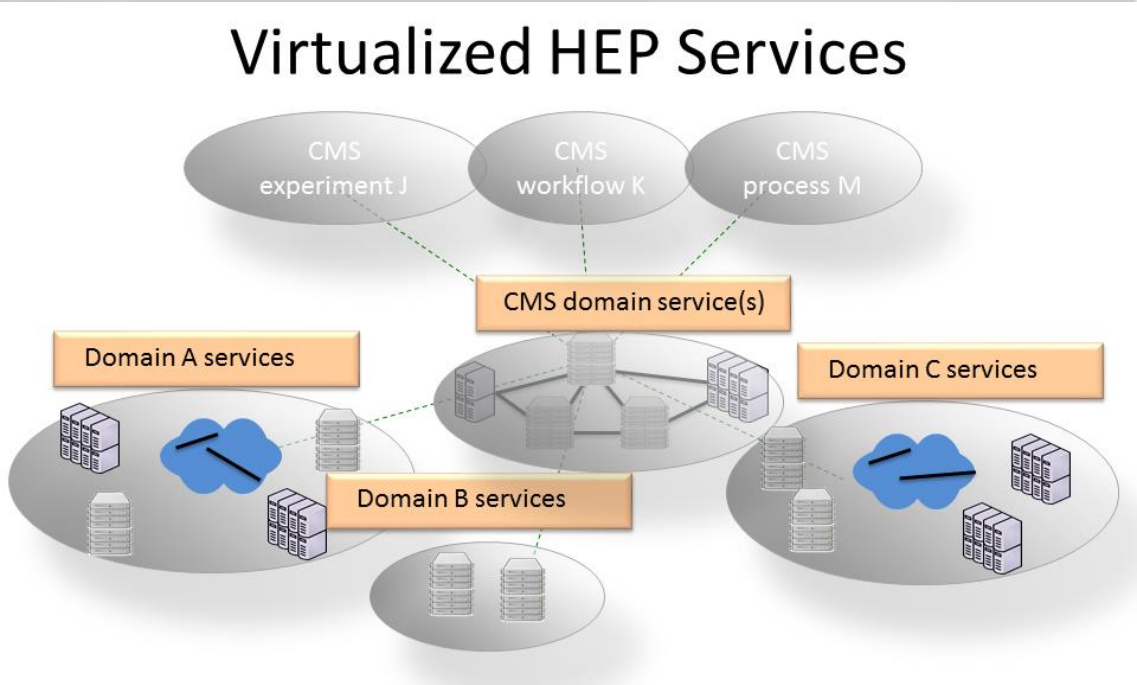
- PhEDEx has 3 use cases with different features
  - #circuits, topologies, time-evolution
  - Scales: hours, TB, nothing smaller
  - Start with T0 -> T1s
  - Ultimate goal is to support analysis flows too
- RESTful service
  - Augment existing capabilities with circuits
  - Expect occasional failure or refusal from service
  - Need priority (& ownership?)
  - Budget/share-management? Who provides that?

Testbed as a Service (Jerry Sobieski, NORDUnet)

- A new “Testbed as a Service” activity in GN3+
- Objective: Provide rapid prototype network testbeds to the network research community
  - Draws from numerous predecessor projects: FEDERICA, Network Factory, OFELIA, NOVI, MANTICHORE, GEYSERS, DRAGON, GENI, PLANET LAB, ...

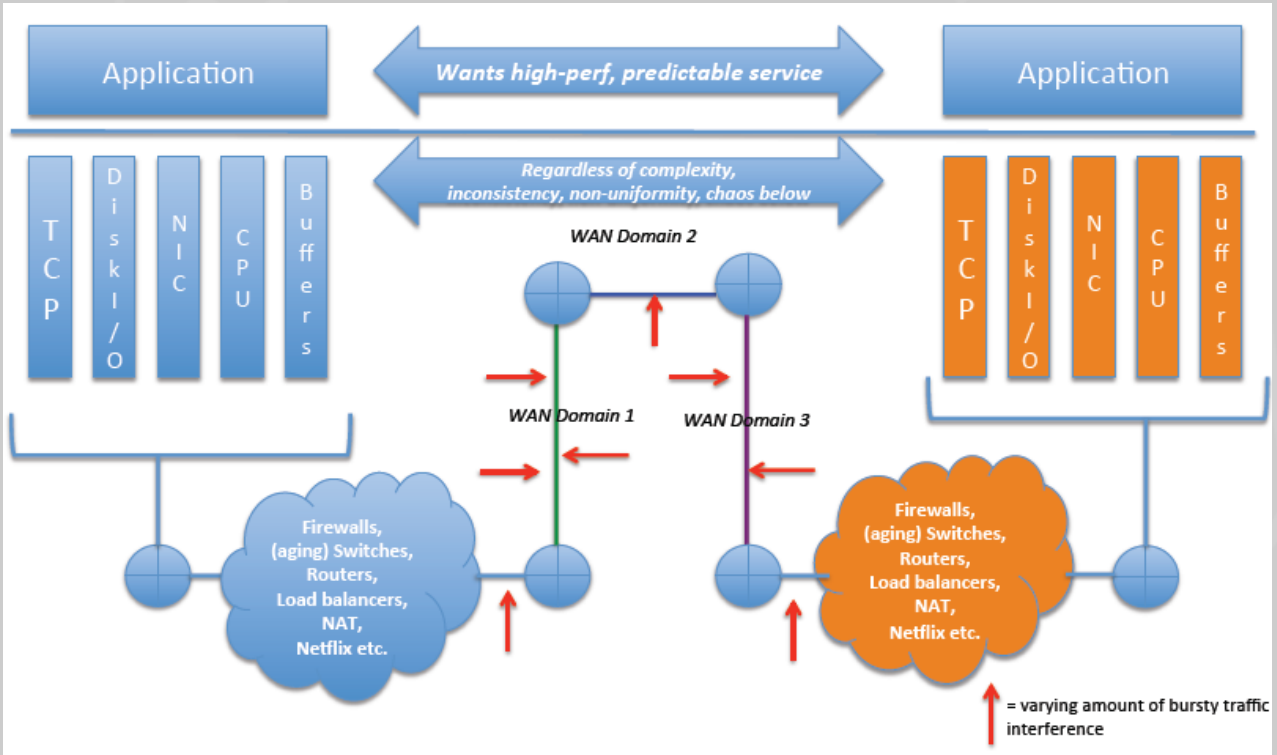
- Potential platform for R&D in the LHCONE context?

- Also note GEANT Open Calls



Networks have hot been an issue for LHC so far because people desired “better than best [effort service]”

- Network complexity (exists!) is hidden from users
- If we want “better than best”, we need a well defined interface
- NSI provides consistent guaranteed bandwidth across multiple WAN
- domains
- “Dynamic” is not a requirement



Point-to-point Demo/Testbed

CALTECH HEP NETWORKING

- Proposed by Inder Monga (ESnet)
  1) Choose a few interested sites
  2) Build static mesh of P2P circuits with small but permanent bandwidth
  3) Use NSI 2.0 mechanisms to
     - Dynamically increase and reduce bandwidth
     - Based on Job placement or transfer queue
     - Based or dynamic allocation of resources
- Define adequate metrics!
  - for meaningful comparison with GPN or/and
- Include both CMS and ATLAS
- Time scale: TDB (“this year”)
- Participation: TDB (“any site/domain interested”)
- More discussion at the next LHCONE/LHCOPN meeting in Paris (June 2013)

Point-to-point Demo/Testbed, cont.

- Site selection:
  - Set of Tier1, and Tier 2 sites, maybe Tier0
  - DYNES/ANSE projects in the US
  - want sites in Europe, Asia
- Use bandwidth from the VRF?
  - might be necessary where circuit infrastructure not existing. But: will shared BW not taint the results?
  - transatlantic: 2 dedicated links for circuits existing
  - (include 100G transatlantic wave?)
- API to the experiments' SW stacks: here's where ANSE is particularly important

Other items discussed

- Much focus on network monitoring
  - as basis for decision taking in the experiments' software
  - choices need to be provided
- SURFnet starts a project aiming at analysing LHCOPN and LHCONE flows
  - interesting basis for discussion
  - caution
- Remote I/O
  - Different traffic pattern to what networks were assuming until now (many small vs "few" large flows)
  - Will this become significant? (not likely)
    - but will it be critical (i.e. require better than best effort?)

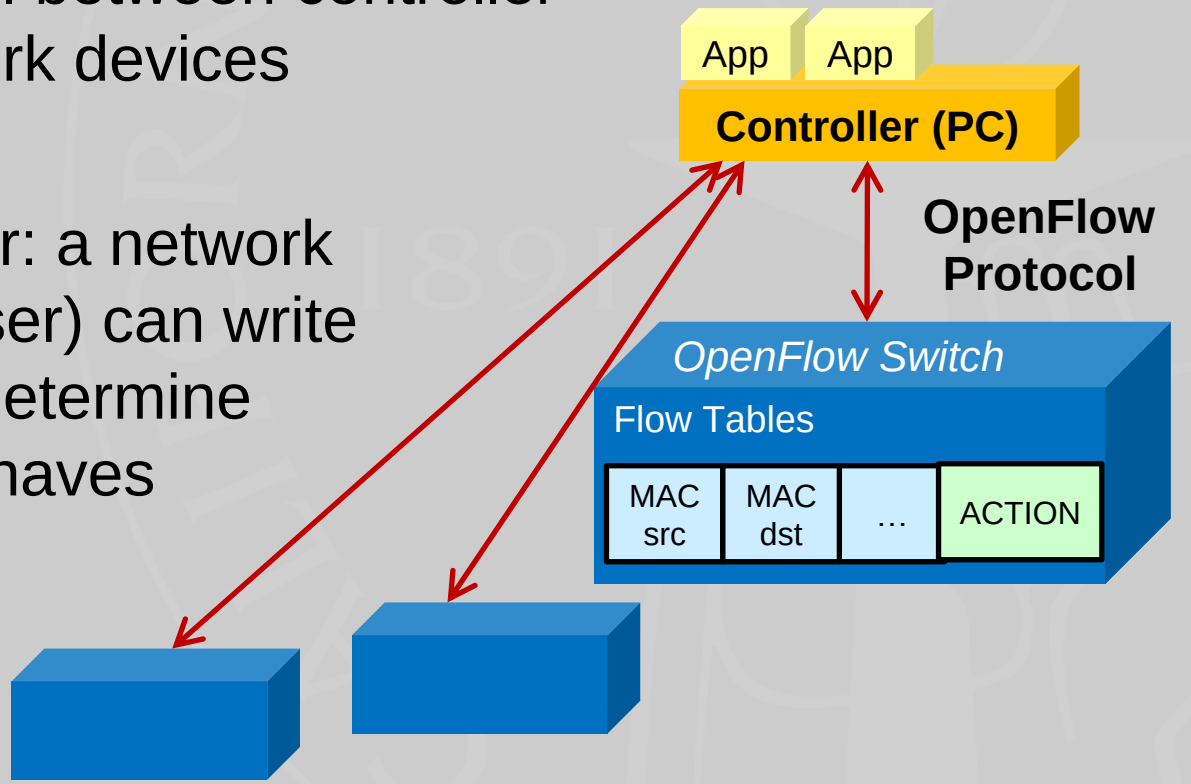
SDN/OPENFLOW MEETING

One slide of introduction

- Software Defined Networking (SDN):
  Simply put, **physical separation of control and data planes**

- Openflow: a protocol between controller entity and the network devices

- The potential is clear: a network operator (or even user) can write applications which determine how the network behaves



SDN - Brief summary

CALTECH HEP NETWORKING

SDN/Openflow could enable solutions to problems where no commercial solution exists

Identify possible issues/problems Openflow could solve, for which no other solution currently exists… e.g.:

- Multitude of transatlantic circuits makes flow management difficult
  - Impacts the LHCONE VRF, but also the GPN
  - No satisfactory commercial solution has been found at layers 1-3
  - Problem can be easily addressed at Layer2 using Openflow
  - Caltech has a DOE funded project running, developing multipath switching capability (OLiMPS)
  - We'll examine this for use in LHCONE
- Second use case: ATLAS is experimenting with OpenStack at several sites.
  - Openflow is the natural virtualisation technology in the network. Could be used to bridge the data centers.
  - Needs some more thought to go into this, interest in ATLAS

Summary

- Good discussion between R&E networking and LHC experiments
- Approaching understanding of
  - what networks can or could provide
  - what experiments (may) need
- Point-to-point service: we'll construct a trial infrastructure
  - with participation of several Tier1/Tier2 sites
  - ideally on at least 2 continents
  - Comparison with GPN and/or VRF service
  - More details will be discussed at the next LHCONE/LHCOPN meeting in Paris June 17/18
    - http://indico.cern.ch/conferenceDisplay.py?confId=236955
- SDN: identified potential use cases
  - flow load balancing within the networks
  - elastic virtual data center
- Need to continue the dialog, and keep the momentum.