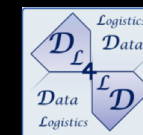# Digital Marketplaces Using Novel Infrastructure Models.!
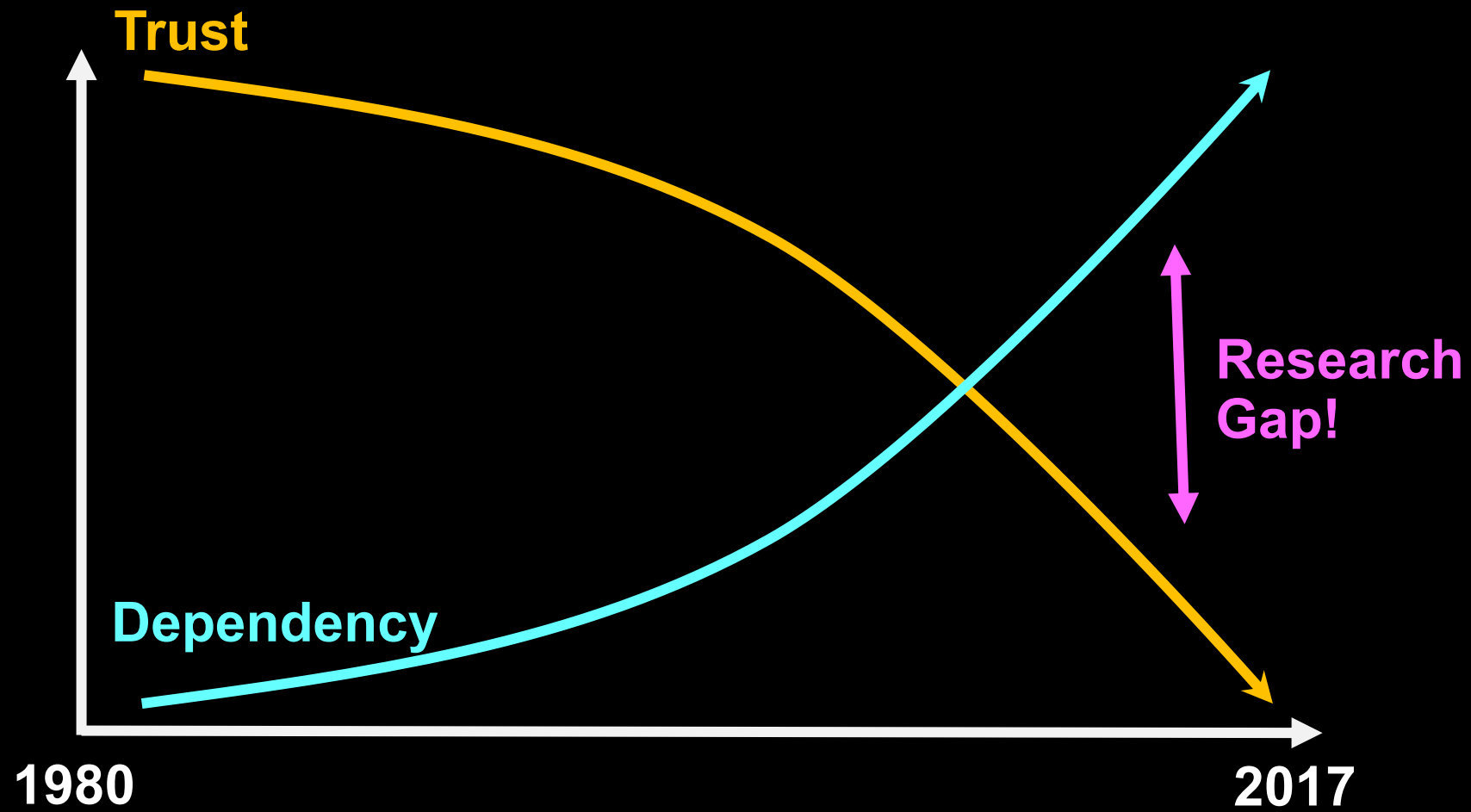
The Global Big Data Hub infrastructure inspired by PRP

Cees de Laat

Systems & Network Laboratory
University of Amsterdam

Main problem statement

- Organizations that normally compete have to bring data together to achieve a common goal!

- The shared data may be used for that goal but not for any other!

- Data may have to be processed in untrusted data centers.
  - How to enforce that using modern Cyber Infrastructure?
  - How to organize such alliances?
  - How to translate from strategic via tactical to operational level?
  - What are the different fundamental data infrastructure models to consider?

Big Data Sharing use cases placed in airline context

Global Scale

National Scale

City / regional Scale

Campus / Enterprise Scale

Aircraft Component Health Monitoring (Big) Data
NWO **CIMPLO project**
4.5 FTE

Cargo Logistics Data
(C1) DaL4LoD
(C2) Secure scalable
policy-enforced
distributed data
Processing
(using blockchain)

NLIP iShare project

iSHARE
powered by NLIP

Cybersecurity Big Data
NWO COMMIT/
SARNET project
3.5 FTE

System and Network Engineering

AIR FRANCE KLM

Approach

- Strategic:
  - Translate legislation into machine readable policy
  - Define data use policy
  - Trust evaluation models & metrics
- Tactical:
  - Map app given rules & policy & data and resources
  - Bring computing and data to (un)trusted third party
  - Resilience
- Operational:
  - TPM & Encryption schemes to protect & sign
  - Policy evaluation & docker implementations
  - Use VM and SDI/SDN technology to enforce
  - Block chain to record what happened (after the fact!)

Strategic Level

Tactical Level

Operational Level

Secure Digital Market Place Research
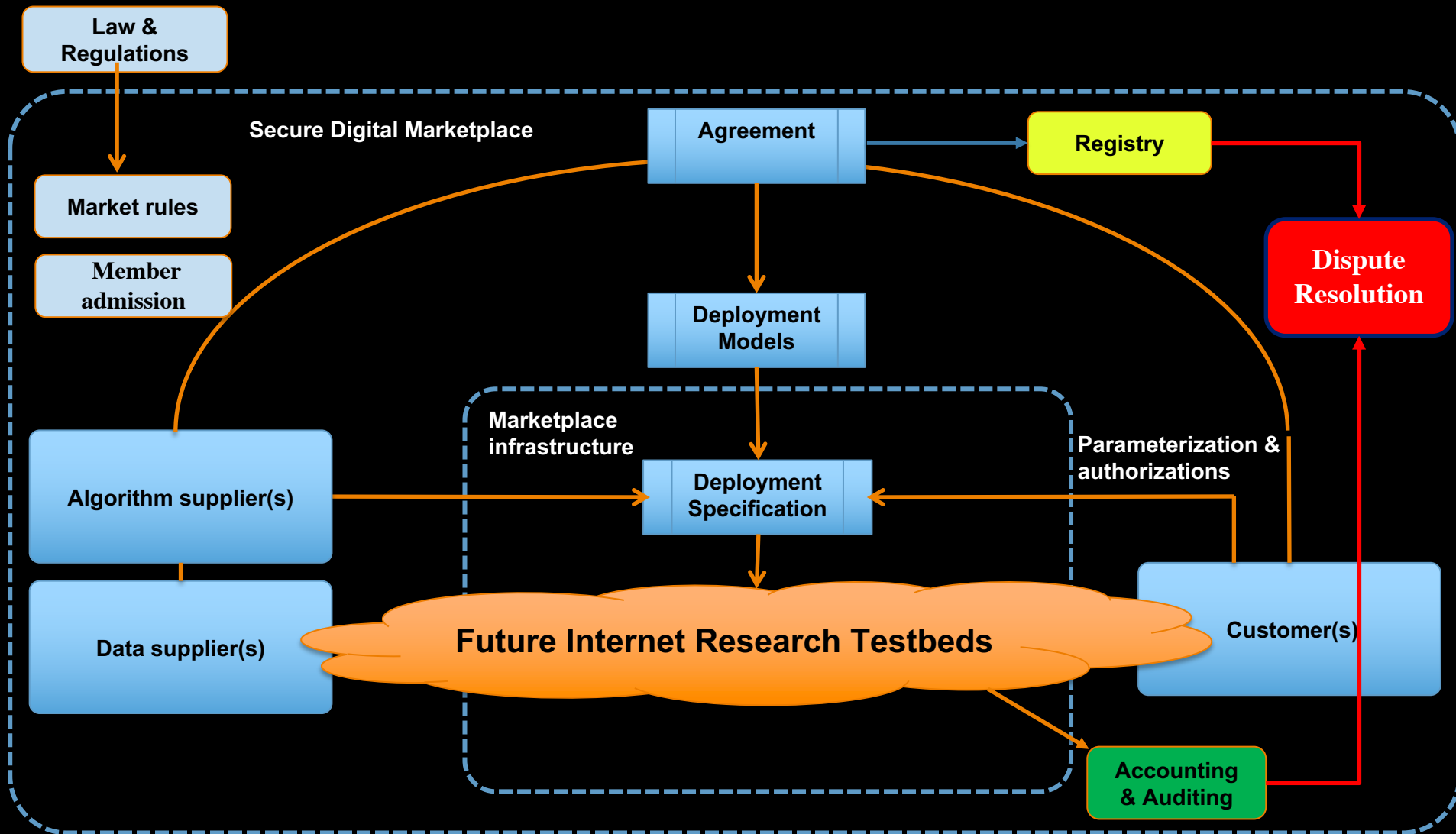
SC16 Demo

DockerMon Sending docker containers with search algorithms to databases all over the world.

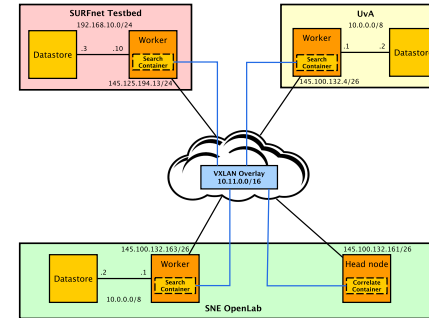http://sc.delaat.net/sc16/index.html#5

Container-based remote data processing

UNIVERSITEIT VAN AMSTERDAM

Łukasz Makowski, Daniel Romão, Cees de Laat, Paola Grosso
System and Networking Research Group, University of Amsterdam

System and Network Engineering

Problem Description

- Scientific datasets are usually made publicly available
  ….but data cannot always leave the organization premises

- On-site data processing can be challenging because of incompatibility of systems or lack of manpower

- Can a container-based system perform remote on-site data processing efficiently?

- What are the networking issues to solve?

Problem

Approach

Data leaving the organization

Container

Send container to the organization

Underlay and Overlay

SURFnet Testbed
192.168.10.0/24
Datastore
Worker
Search Container
145.125.194.13/24

UvA
10.0.0.0/8
Worker
Search Container
Datastore
145.100.132.4/26

VXLAN Overlay
10.11.0.0/16

145.100.132.163/26
145.100.132.161/26
Datastore
Worker
Search Container
Head node
Correlate Container
10.0.0.0/8
SNE OpenLab

Main features:

- Networked containers

- VXLAN overlay

- Containers that perform data retrieval and computation

- Containers built on-demand

- On-site data processing

- Distributed data source

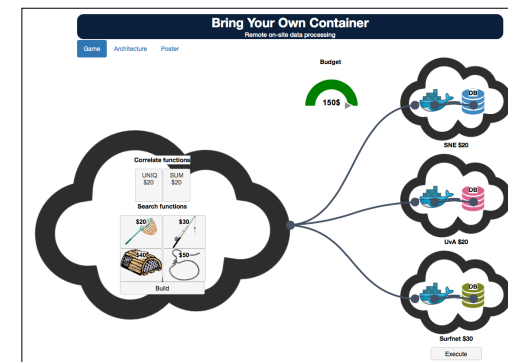- Multiple sites with datasets

The Game

Our SC16 demo is a gamification of the remote dataset processing architecture.

How many different animal species can you find?
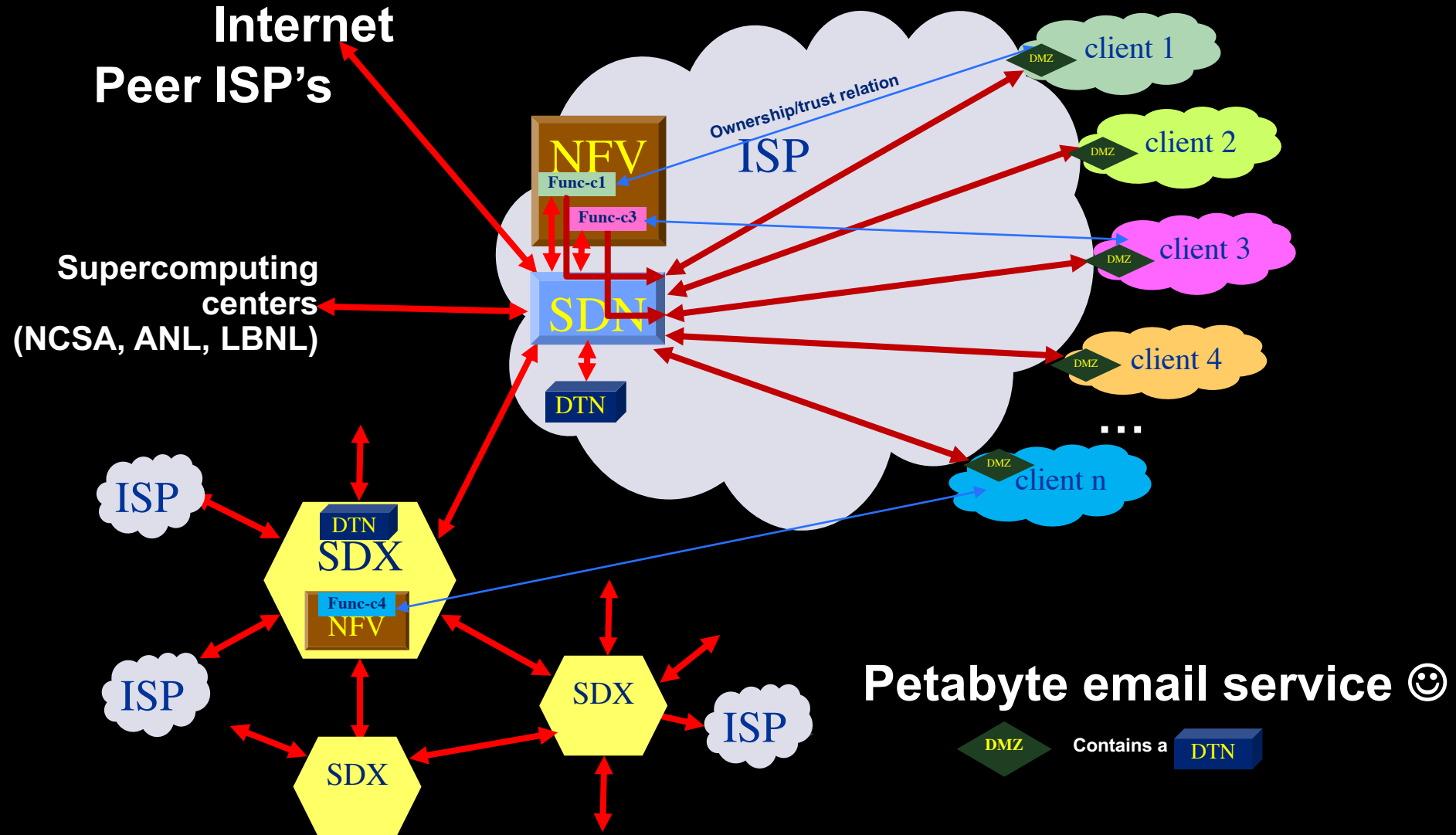You have a fixed budget and each function and processing will cost you money!

In our game you will:

- Select a correlate function to combine the results of the different sites.
- Pick different search functions, represented as tools, to find animals in the remote datasets.
- Build containers with the search and correlate functions.
- Execute the containers on the sites of your choice.

Will you have the best score?

Bring Your Own Container
Remote on-site data processing
Game   Architecture   Poster
Budget
150$
Correlate functions
UNIQ $20   SUM $20
Search functions
$20 $30
$40 $50
Build
SNE $20
UvA $20
Surfnet $30
Execute

SURF NET

Networks of ScienceDMZ's & SDX's

Program

- 10h45 Cees de Laat, UvA
  - Intro
- 10h55 Leon Gommans, Air France KLM & UvA
  - Exploring Digital Marketplaces.
- 11h15 Panel of stakeholders Flash talks (~3 min each):
- David Groep (NIKHEF):
  - Trust & Science, the need for Data control.
- Craig Waldrop (EQUINIX):
  - Enabling the Data Economy & Avoiding the Pitfalls.
- Rodney Wilson (CIENA):
  - Data Markets in the Fog, IOT & 5G
- Leon Gommans (KLM).
- 11h25 Panel discussion moderated by Cees de Laat
- 11h45 End of session.
- More information:
  - http://delaat.net/dl4ld