Collecting infrastructure data in virtualized environments

predict
prioritise
prevent
TREsPASS

Axel Tanner
IBM Research – Zürich
axs@zurich.ibm.com

ADaCoR Workshop
University Luxembourg
2016-04-19

SEVENTH FRAMEWORK PROGRAMME

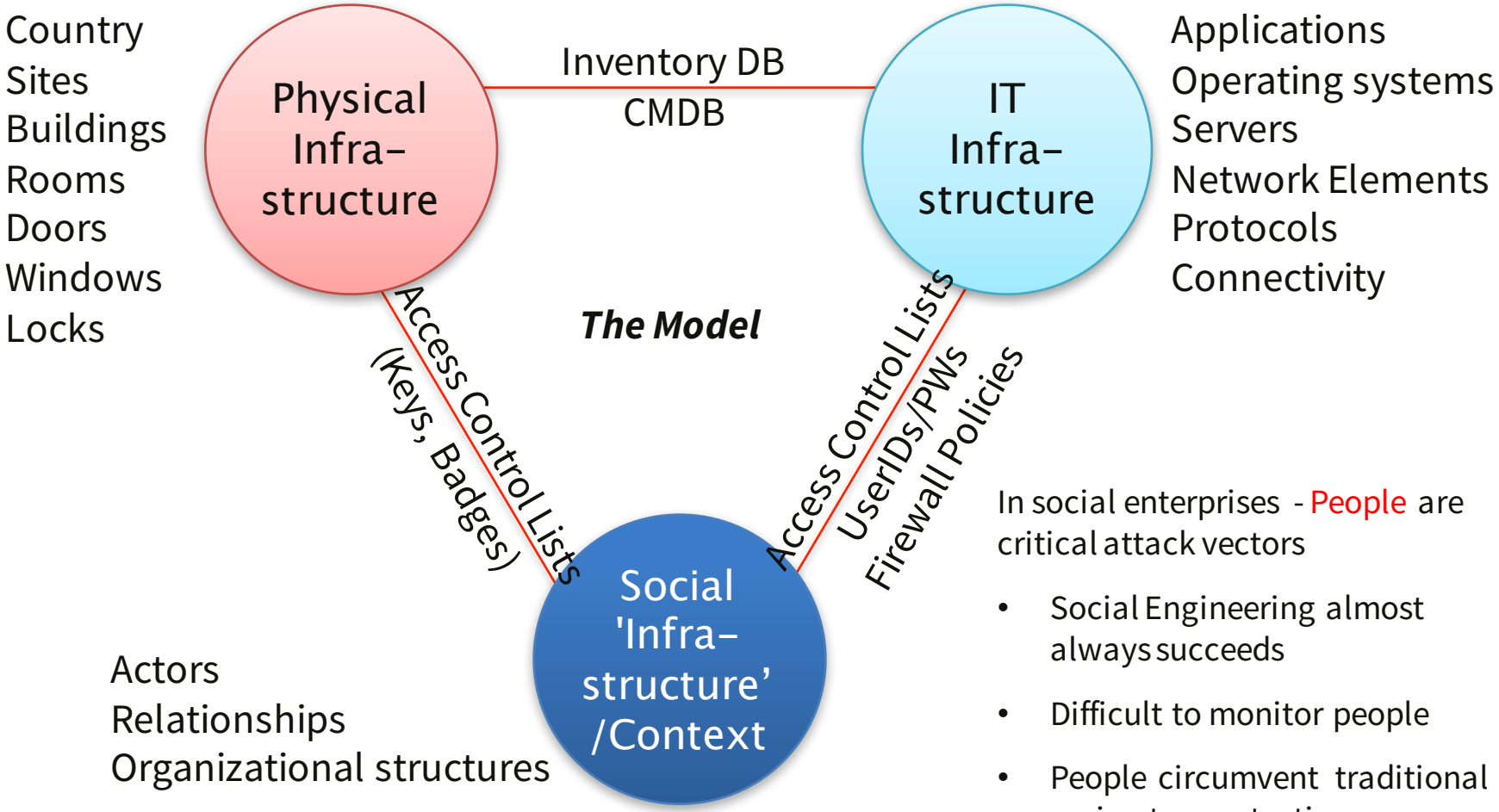
European Commission

IBM

Overview

- Context and Cloud specialties
- Data extraction from virtualized infrastructures
- Application examples

Environment Dimensions for Risk Analysis

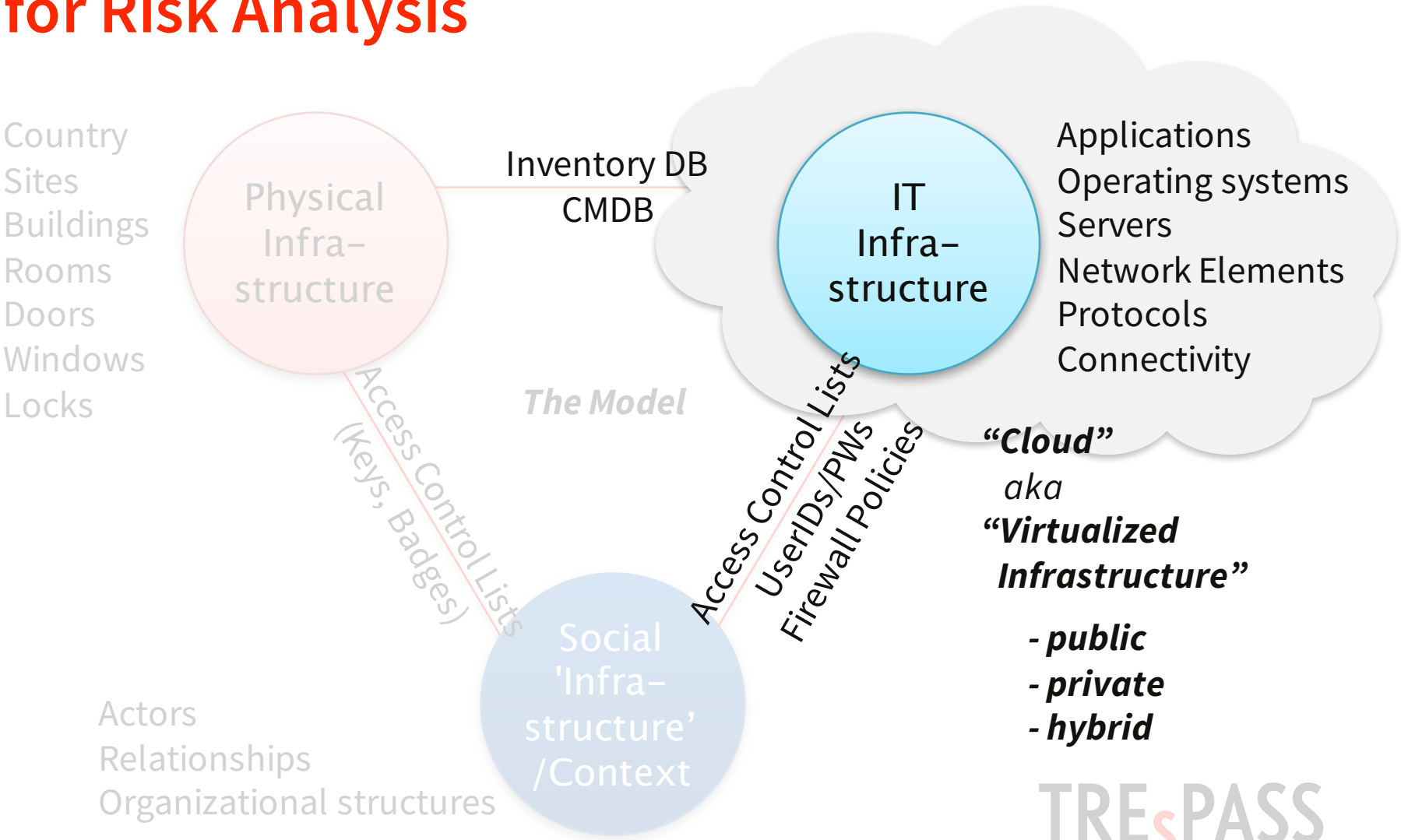


In social enterprises - People are critical attack vectors

- Social Engineering almost always succeeds
- Difficult to monitor people
- People circumvent traditional perimeter protection

Environment Dimensions for Risk Analysis

TREsPASS Use Case: Cloud



Disruptive Innovation

Virtualization

The logical abstraction of physical computing resources (OS, application, switches, storage, networks).

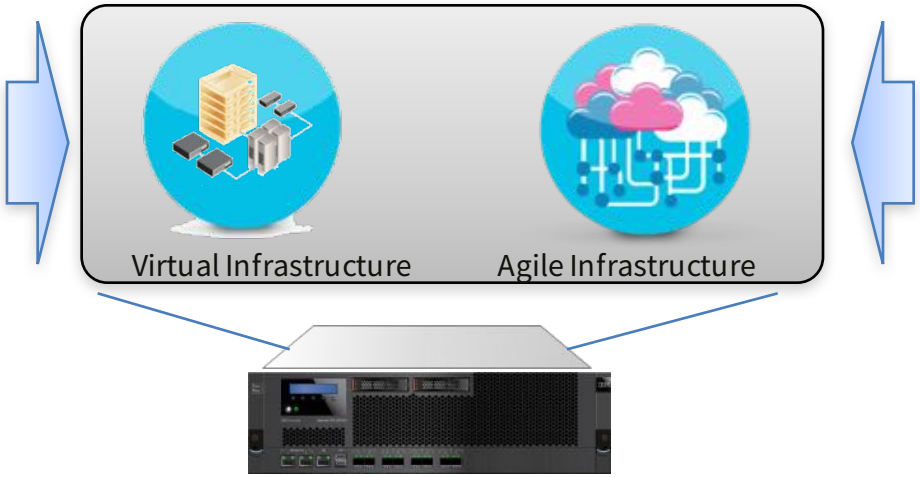
Computing environments that are not restricted by physical configuration or implementation.

Software Defined Environments (SDE)

Environments that optimize compute, storage and networking infrastructure based on workload.

Shared software management tools dynamically assign and manage workloads.

Physical Infrastructure



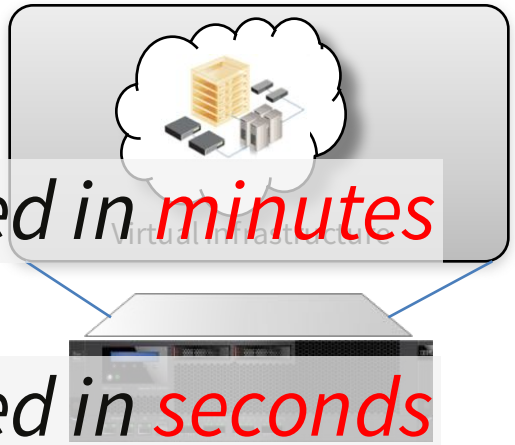
SDE

Disruptive Innovation

- Before
  - Order and install servers
  - Order network Connectivity
  - Install operating system
  - Configure Network/Storage
  - Install middleware/Applications
    - Web Servers/Databases/Authentication/Messaging Bus/Monitoring/…
  - Configure Applications

Measured in days/hours

- Today
  - Select template solution (Virtual Application Pattern)
  - Click 'create' button

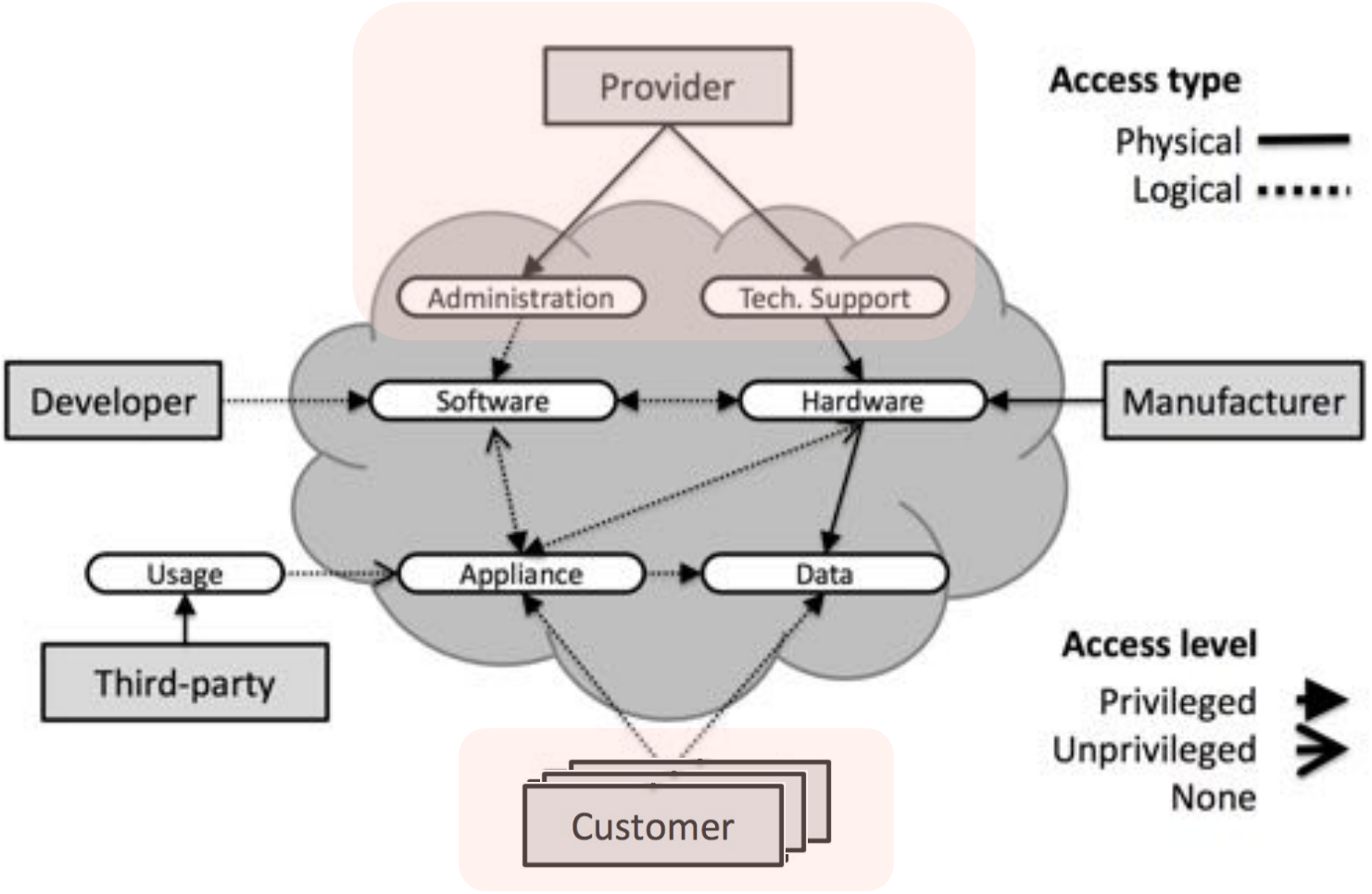


Measured in minutes

- Failover or scaleout to other infrastructure

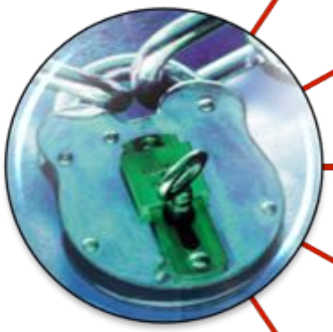
Measured in seconds

Cloud: involved entities



"Defining the Cloud Battlefield" Sören Bleikertz, Toni Mastelić, Sebastian Pape, Wolter Pieters, Trajce Dimkov; IEEE International Conference on Cloud Engineering (IC2E 2013)

Cloud: Additional Risk Areas



Virtual Images

- Storing snapshots
- Moving snapshots
- Crypto Replay Attacks

Sharing Resources

- Multi-tenancy
- Side Channel Attacks
- Storage Co-location

Roles

- Cloud provider in addition
- Fewer, more powerful roles
- Spear Fishing/Watering Hole Attacks

Software

- Community/Open Source
- Interpreted vs. Complied
- DevOps

Networks

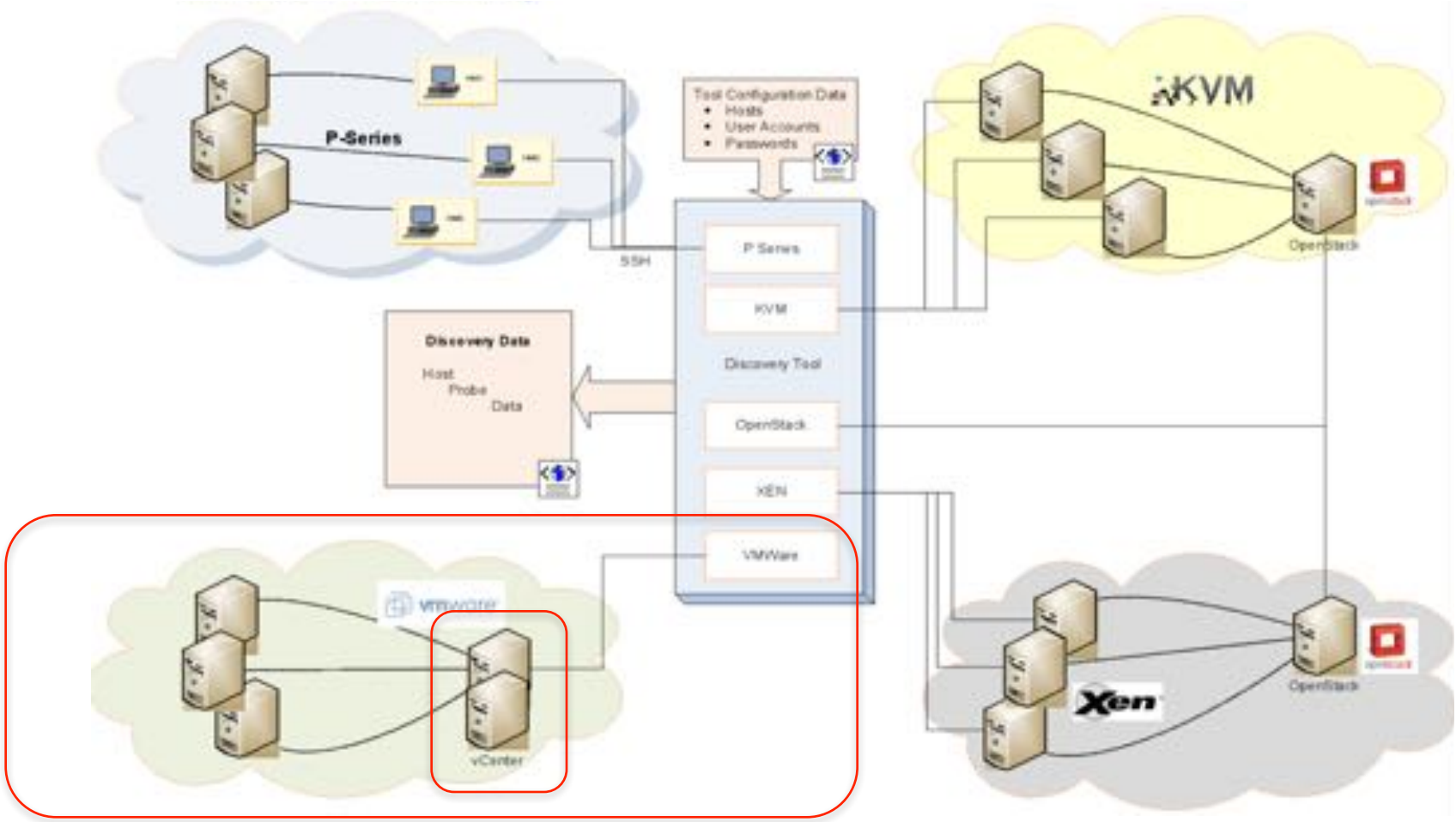
- Flatter Networks
- Better Visibility

… but there are also advantages

- As ’software defined’
  → easier to access and understand current setup

- Most often centralized administration and management

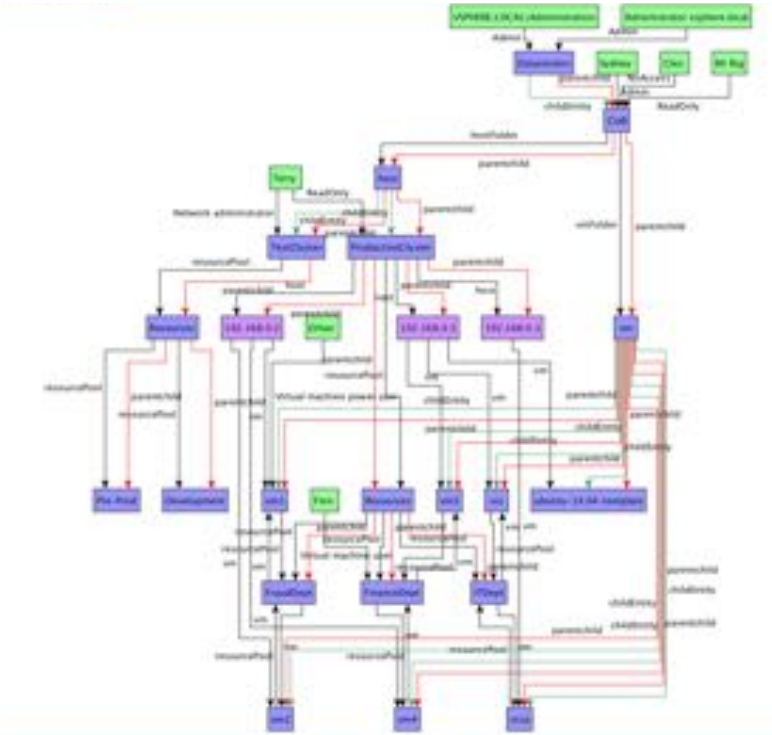
Different cloud technologies



Looking at VMware in particular

- Central management component VMware vSphere
  - official VMware Web Services API available
  - access with (read-only) credentials
  - [also Python bindings pyVmomi]
- Data to extract
  - enumeration and properties of the infrastructure components
    - host systems, virtual machines, network definitions (like physical network interfaces of the hosts, as well as virtual network interfaces, connectivity and VLAN details), storage components
  - Data related to the structuring/grouping of infrastructure components
    - container information like DataCenters, folders and resource pools
    - relationships between components, e.g., virtual machines contained in a host, storage components and networks used by hosts and virtual machines
  - data related to access control
    - definition of existing roles as combinations of base privileges in the system
    - lists of existing user ids, lists of existing groups of user ids
    - lists of permissions of user ids or groups on specific parts of the infrastructure

Using RDF as representation



```
@prefix ldap: <http://kb.trespass.demo/ciab/resource/ldap/> .
@prefix meta: <http://kb.trespass.demo/ciab/resource/meta/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix vim: <http://kb.trespass.demo/ciab/resource/vmware/> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

vim:entity_vim.VirtualMachine_vm-51 a vim:Entity ;
    rdfs:label "vm1" ;
    vim:datastore vim:entity_vim.Datastore_datastore-41 ;
    vim:entityStr "vim.VirtualMachine:vm-51" ;
    vim:ipAddress "192.168.0.11" ;
    vim:network vim:entity_vim.Network_network-43 ;
    vim:resourcePool vim:entity_vim.ResourcePool_resgroup-60 .

vim:entity_vim.HostSystem_host-49 a vim:Entity ;
    rdfs:label "nuc2" ;
    vim:datastore vim:entity_vim.Datastore_datastore-41 ;
    vim:entityStr "vim.HostSystem:host-49" ;
    vim:ipAddress "192.168.0.2" ;
    vim:network vim:entity_vim.Network_network-42,
        vim:entity_vim.Network_network-43,
        vim:entity_vim.Network_network-45 ;
    vim:vm vim:entity_vim.VirtualMachine_vm-101,
        vim:entity_vim.VirtualMachine_vm-102,
        vim:entity_vim.VirtualMachine_vm-82 .

vim:principal_VSPHERE.LOCAL_terry a vim:Principal ;
    rdfs:label "Terry" ;
    vim:group false .

[] a vim:Permission ;
    vim:entity vim:entity_vim.ClusterComputeResource_domain-c28 ;
    vim:group false ;
    vim:principal vim:principal_VSPHERE.LOCAL_terry ;
    vim:propagate true ;
    vim:role vim:role_9 ;
    vim:roleId 9 .

vim:role_9 a vim:Role ;
    rdfs:label "Network administrator" ;
    vim:privilege "Network.Assign",
        "System.Anonymous",
        "System.Read",
        "System.View" ;
    vim:roleId 9 ;
    vim:system false ;
    rdfs:comment "Network administrator" .

...
```

Namespaces

Information in form of "triples":

subject – predicate – object

corresponding to edge in directed graph

Resource Description Framework
Semantic Web Standards
http://www.w3.org/RDF

IBM

Using RDF as representation

Flexible to

- combine data from multiple cloud instance
- combine with additional data, e.g. Data from vulnerability scanners

```
@prefix ldap: <http://kb.trespass.demo/ciab/resource/ldap/> .
@prefix meta: <http://kb.trespass.demo/ciab/resource/meta/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix vim: <http://kb.trespass.demo/ciab/resource/vmware/> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

vim:entity_vim.VirtualMachine_vm-51 a vim:Entity ;
    rdfs:label "vm1" ;
    vim:datastore vim:entity_vim.Datastore_datastore-41 ;
    vim:entityStr "vim.VirtualMachine:vm-51" ;
    vim:ipAddress "192.168.0.11" ;
    vim:network vim:entity_vim.Network_network-43 ;
    vim:resourcePool vim:entity_vim.ResourcePool_resgroup-60 .

vim:entity_vim.HostSystem_host-49 a vim:Entity ;
    rdfs:label "nuc2" ;
    vim:datastore vim:entity_vim.Datastore_datastore-41 ;
    vim:entityStr "vim.HostSystem:host-49" ;
    vim:ipAddress "192.168.0.2" ;
    vim:network vim:entity_vim.Network_network-42,
        vim:entity_vim.Network_network-43,
        vim:entity_vim.Network_network-45 ;
    vim:vm vim:entity_vim.VirtualMachine_vm-101,
        vim:entity_vim.VirtualMachine_vm-102,
        vim:entity_vim.VirtualMachine_vm-82 .

vim:principal_VSPHERE.LOCAL_terry a vim:Principal ;
    rdfs:label "Terry" ;
    vim:group false .

[] a vim:Permission ;
    vim:entity vim:entity_vim.ClusterComputeResource_domain-c28 ;
    vim:group false ;
    vim:principal vim:principal_VSPHERE.LOCAL_terry ;
    vim:propagate true ;
    vim:role vim:role_9 ;
    vim:roleId 9 .

vim:role_9 a vim:Role ;
    rdfs:label "Network administrator" ;
    vim:privilege "Network.Assign",
        "System.Anonymous",
        "System.Read",
        "System.View" ;
    vim:roleId 9 ;
    vim:system false ;
    rdfs:comment "Network administrator" .

...
```

Namespaces

Information in form of "triples":

subject – predicate – object

corresponding to edge in directed graph

Resource Description Framework
Semantic Web Standards
http://www.w3.org/RDF

IBM

Caveats

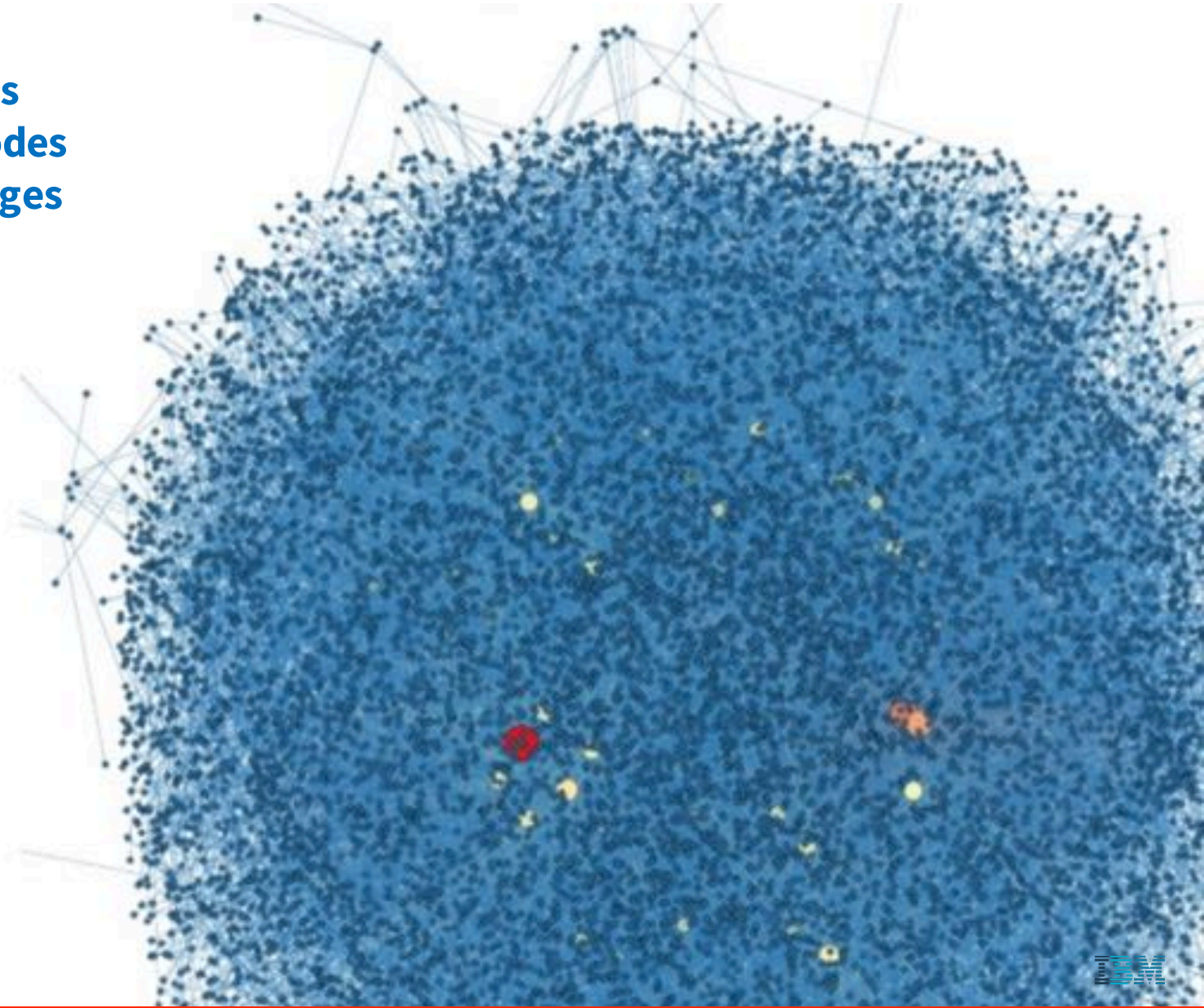
- **Scope of access control** access control in VMware vSphere can be based on vSphere-wide user ids, but there also can be local user ids on individual host VMware ESXi systems. This means that although one can list and track all user ids in the vSphere environment itself, individuals still could have or get access through host-local user ids (unseen from the central console)
  Solutions
  - Use lockdown modes (*normal* and *strict*) which disallow access to the individual hosts of the vSphere environment
  - Query all VMware ESXi hosts in addition and combine information

- **Network IP addresses** vSphere knows the IP addresses of host systems, but it can only know IP addresses of virtual machines while they are running answered only in case they have the VMware tools installed inside the guest operating system.
  Solution
  - Use additional information from network level

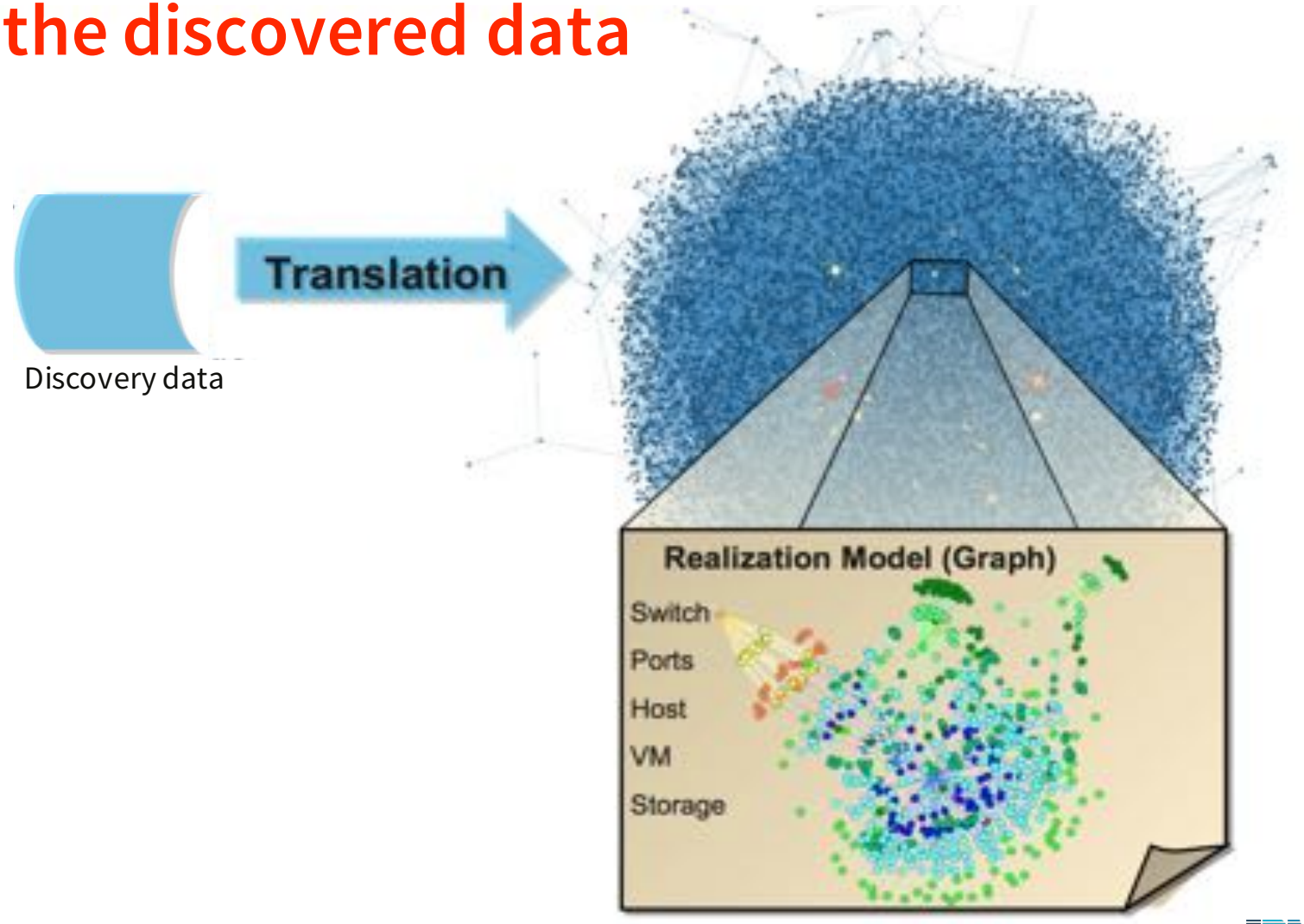
How can that be useful?

- Analyze and compare with policies
- Visualize to get better understanding
- Put data into context

1’300 VMs
25’000 Nodes
30’000 Edges



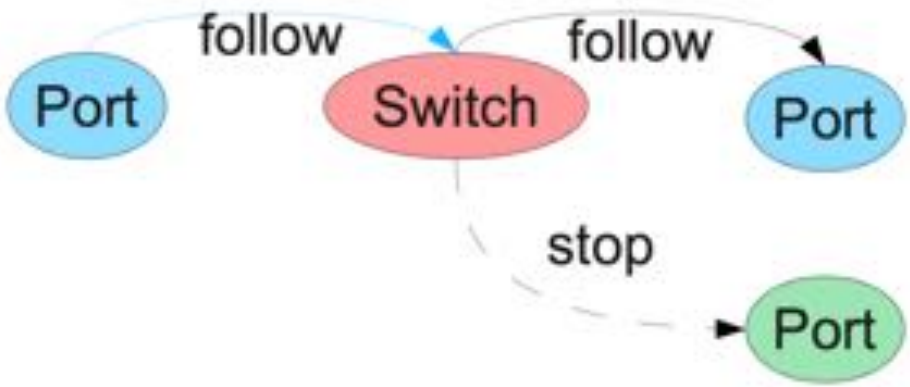
Create a Realization Model out of the discovered data



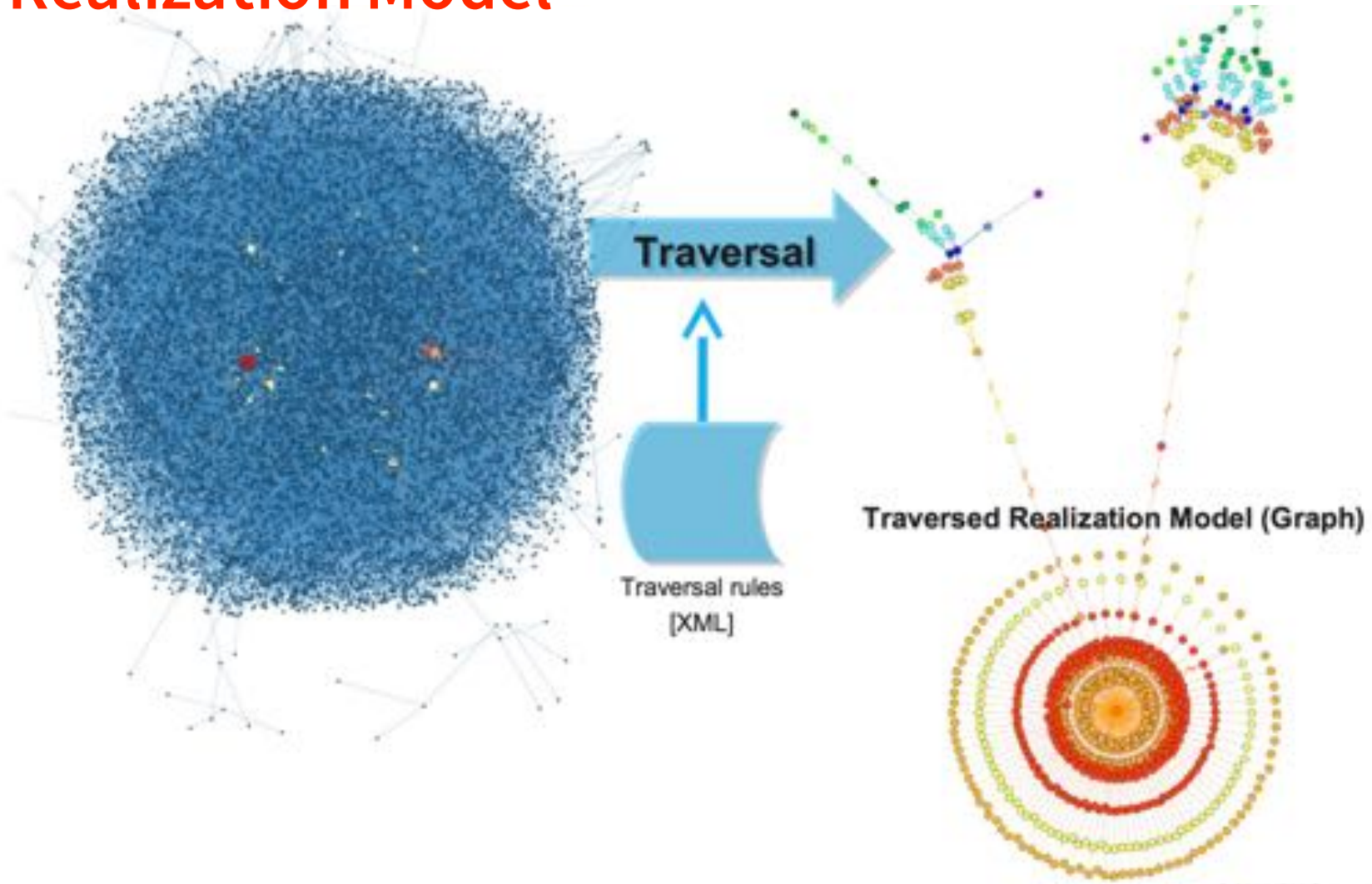
Traversal Rules: Trust and Isolation Assumptions

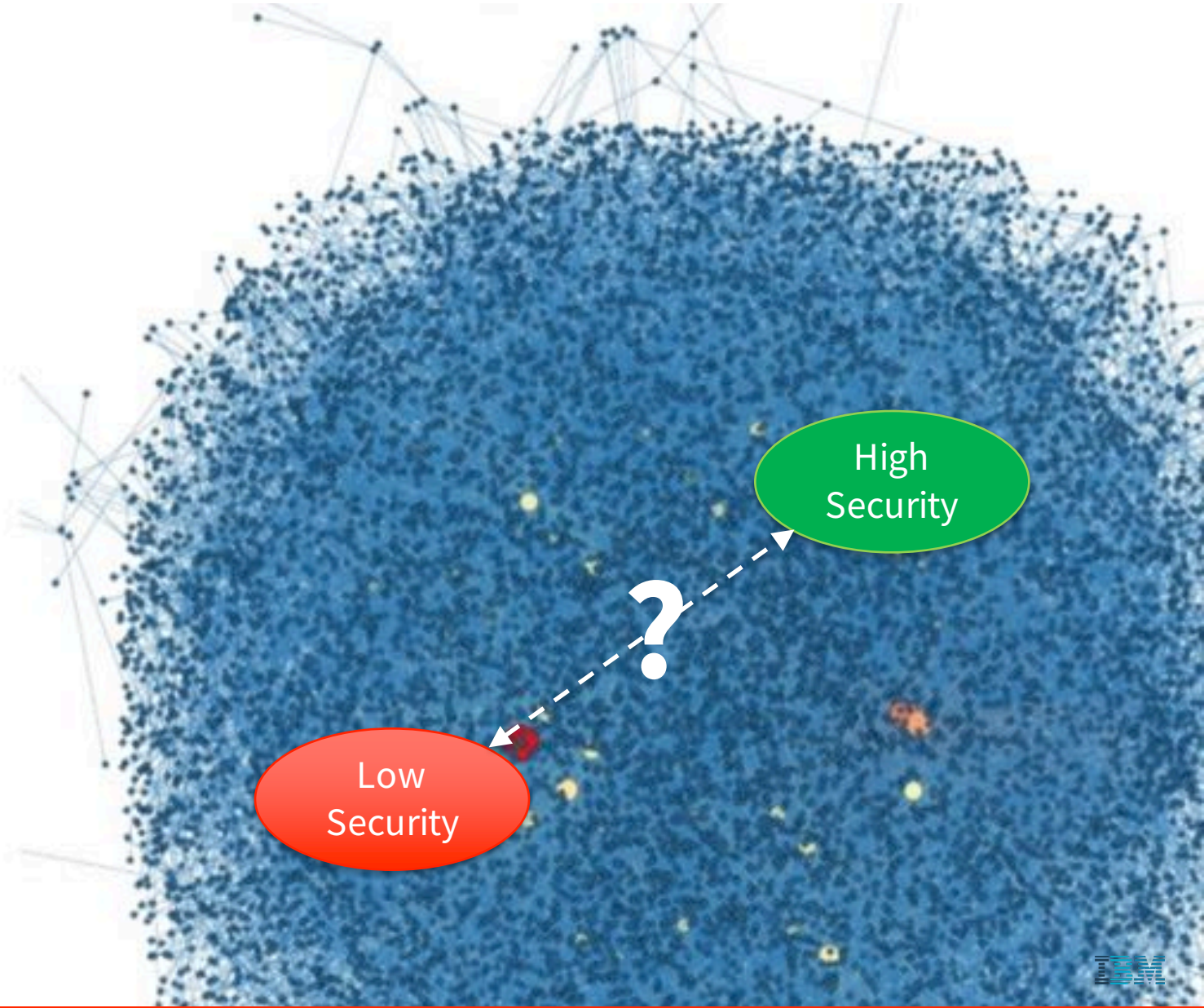
- Edges are potential information flow
  → now decide on "actual" flow

- Trusted and isolating components: stop rules
  - Secure hypervisor: no covert channels
  - Secure management OS
  - Firewalls
  - VLANs

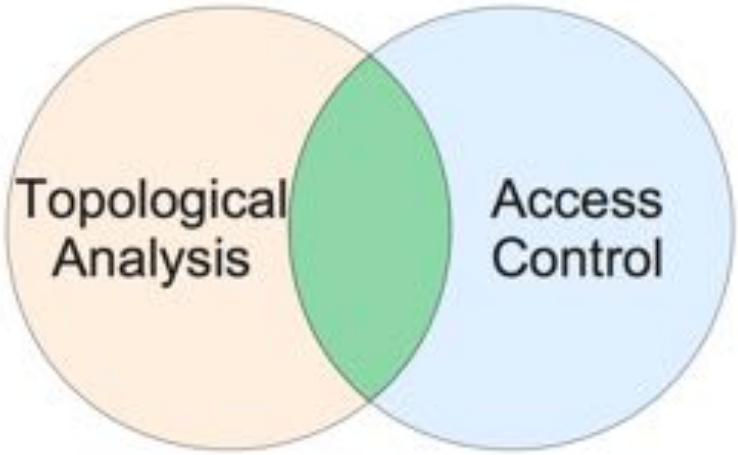


Analyze Traversal Flow in Realization Model



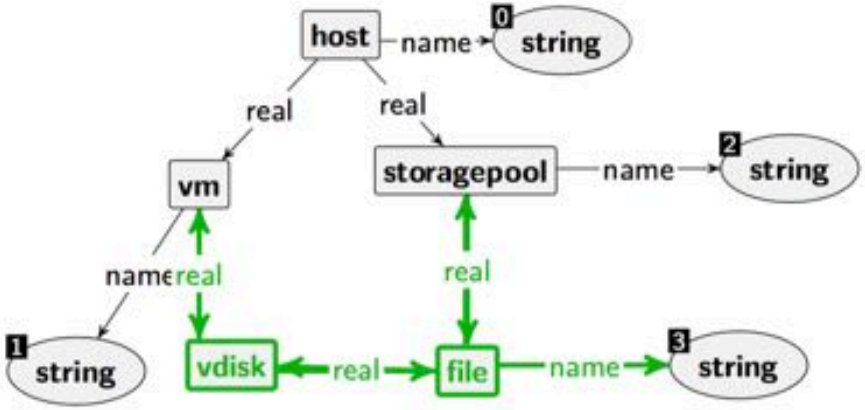

High
Security
?
Low
Security

Including Access Control data …

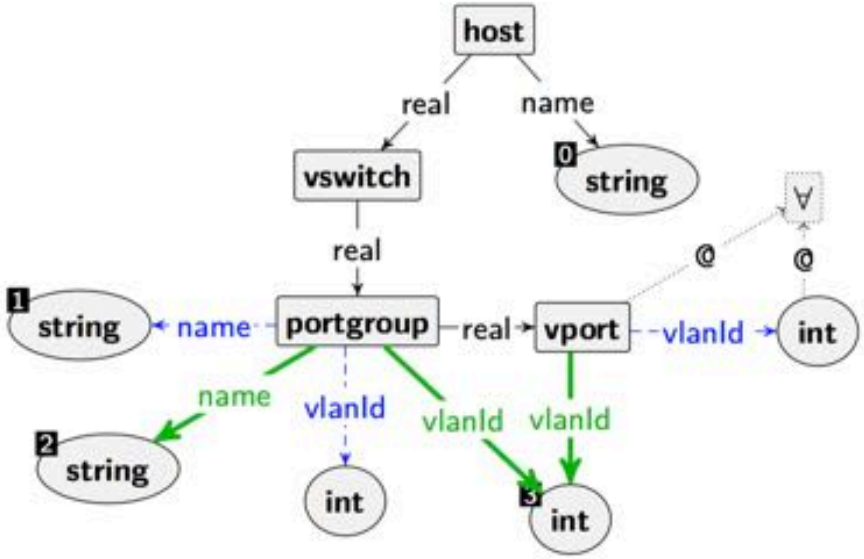


- Given a set of privileges, can an admin transform the infrastructure into an insecure state?
- Use-cases: detecting potential insider attacks, verifying separation of duty and privilege minimizations

Potential Actions correspond to graph transformations



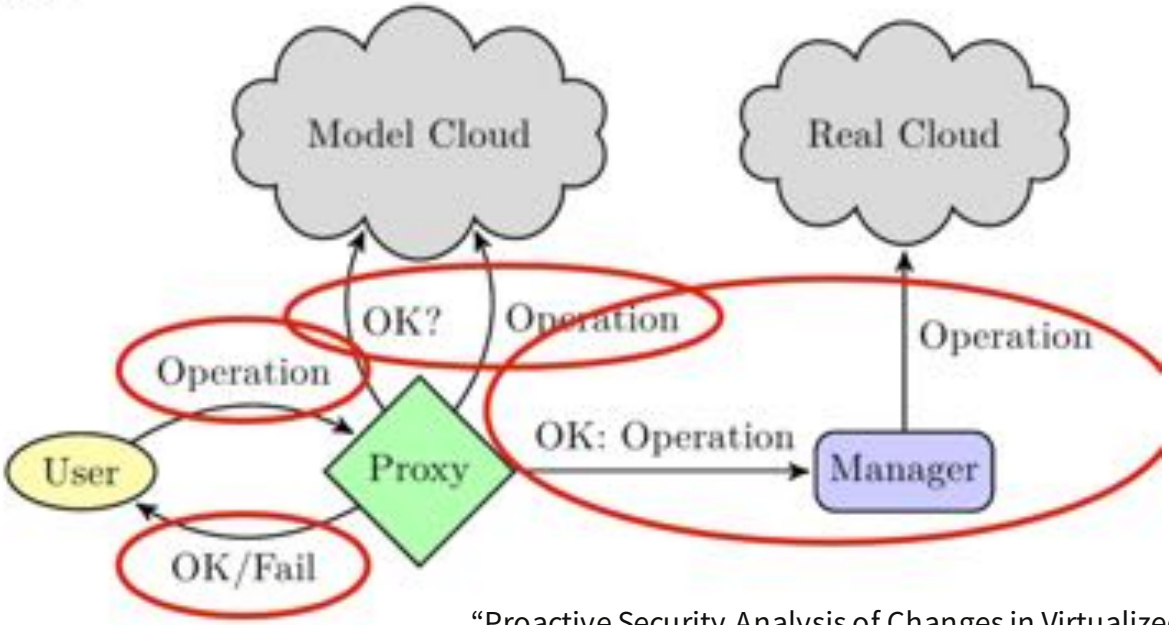
(a) AddVirtualDisk



(b) UpdatePortGroup

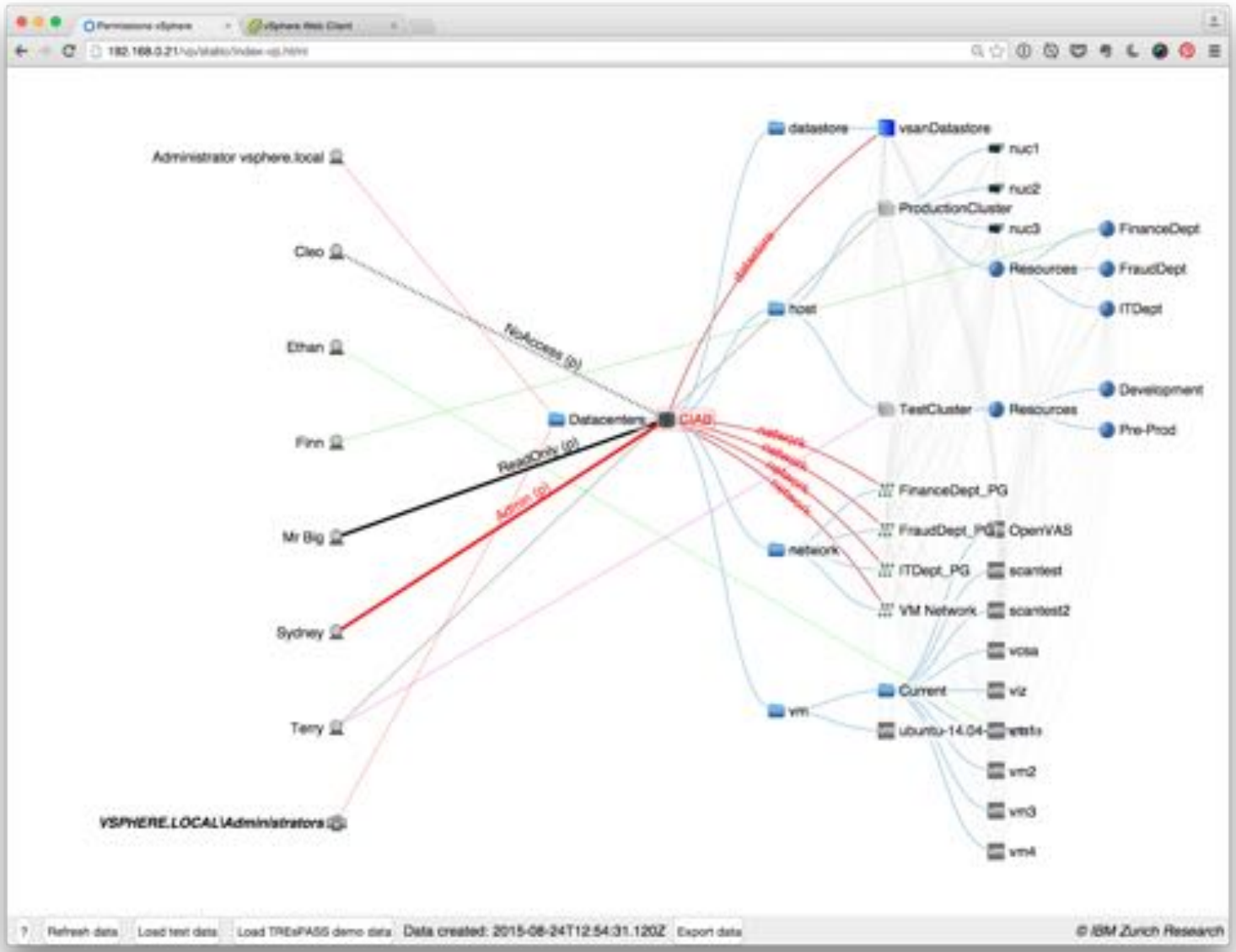
One more step: being proactive

- *Analyze* cloud operations before actually deployed
- Prevent / highlight misconfigurations by administrators
- Suggesting corrections?

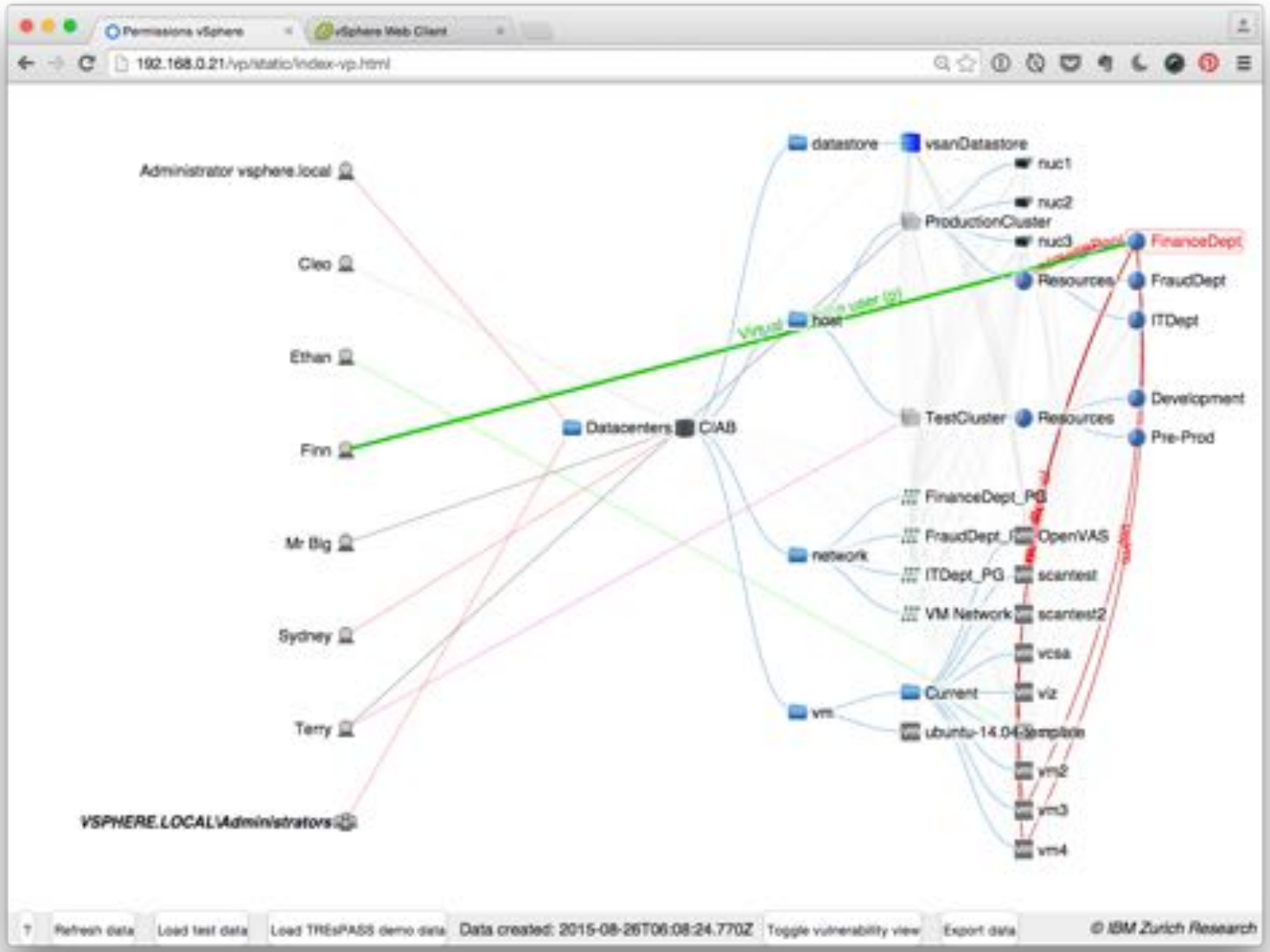


"Proactive Security Analysis of Changes in Virtualized Infrastructures" Sören Bleikertz, Thomas Groß, Sebastian Mödersheim, and Carsten Vogel; Annual Computer Security Applications Conference (ACSAC 2015)

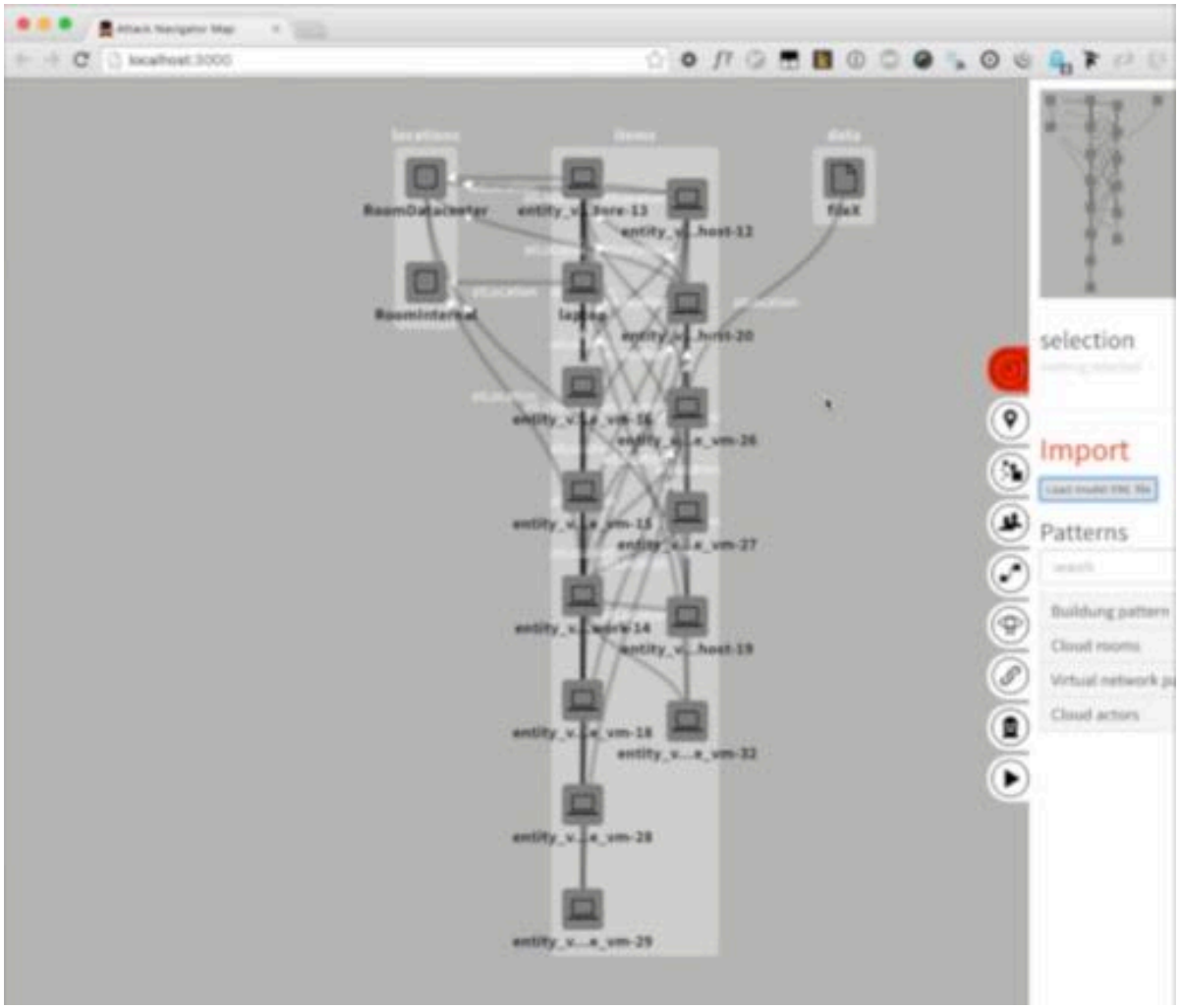
Gaining visual insight …



Gaining visual insight …



… back to TREsPASS: putting the cloud in perspective



Concluding

- "Cloud" technologies, i.e. Software Defined Environments, are disruptive changes
  → *that nearly everyone is exploiting*
- Security-wise
  - Much more agile – security solutions have to follow
  - Changing and additional risks
    - Cross-tenant / sharing of infrastructure
    - Cloud provider as additional party
    - New technologies for virtualization
  - Easier to discover and monitor due to (mostly) central management and software definitions
    - Possible to monitor changes in real-time

Selected References and Links

- VMware vSphere Web Services SDK http://www.vmware.com/support/developer/vc-sdk
- VMware vSphere API Python Bindings pyVmomi http://github.com/vmware/pyvmomi
- Resource Description Framework – Semantic Web Standards http://www.w3.org/RDF
- "Defining the Cloud Battlefield" Sören Bleikertz, Toni Mastelić, Sebastian Pape, Wolter Pieters, Trajce Dimkov; IEEE International Conference on Cloud Engineering (IC2E 2013)
- "Proactive Security Analysis of Changes in Virtualized Infrastructures" Sören Bleikertz, Thomas Groß, Sebastian Mödersheim, and Carsten Vogel; Annual Computer Security Applications Conference (ACSAC 2015)
- "Tool-based Risk Assessment of Cloud Infrastructures as Socio-Technical Systems "; Nidd M., Ivanova M.G, Probst C.W, Tanner A. 2015. in "The Cloud Security Ecosystem" :495–517