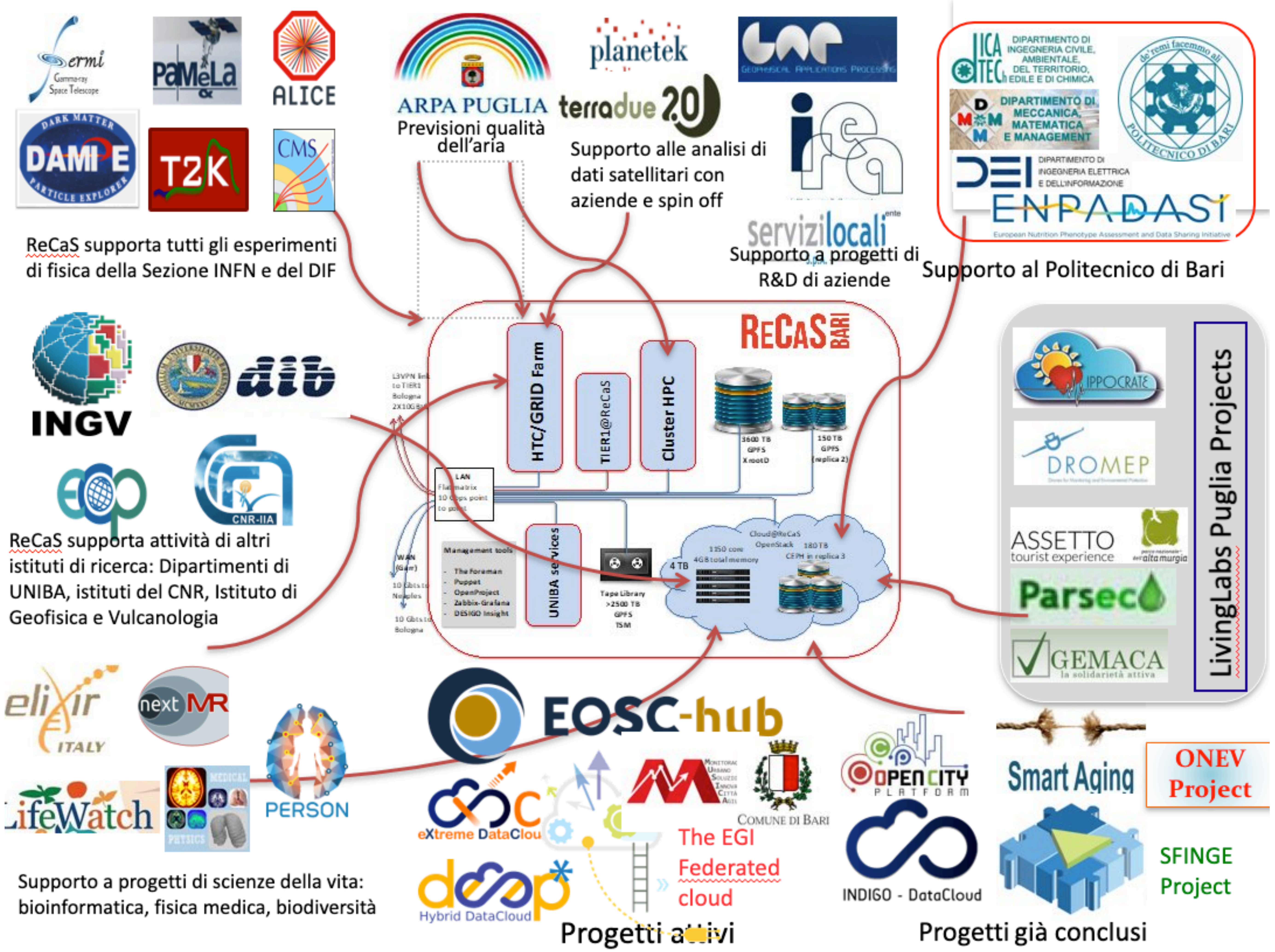
Bari Cloud site evolution: Introducing the architecture and implementation of the upgraded IaaS infrastructure

Alessandro Italiano on behalf of:
Antonacci Marica, Donvito Giacinto, Nicotri Stefano, Perniola Michele, Sguera Ruggiero, Renna Luigi, Valentini Roberto

context:ReCas-[IBisCO]-Bari Datacenter



ReCaS supporta tutti gli esperimenti di fisica della Sezione INFN e del DIF

ARPA PUGLIA Previsioni qualità dell'aria

Supporto alle analisi di dati satellitari con aziende e spin off

Supporto a progetti di R&D di aziende

Supporto al Politecnico di Bari

ReCaS supporta attività di altri istituti di ricerca: Dipartimenti di UNIBA, istituti del CNR, Istituto di Geofisica e Vulcanologia

LivingLabs Puglia Projects

Supporto a progetti di scienze della vita: bioinformatica, fisica medica, biodiversità

The EGI Federated cloud

Progetti attivi

Progetti già conclusi

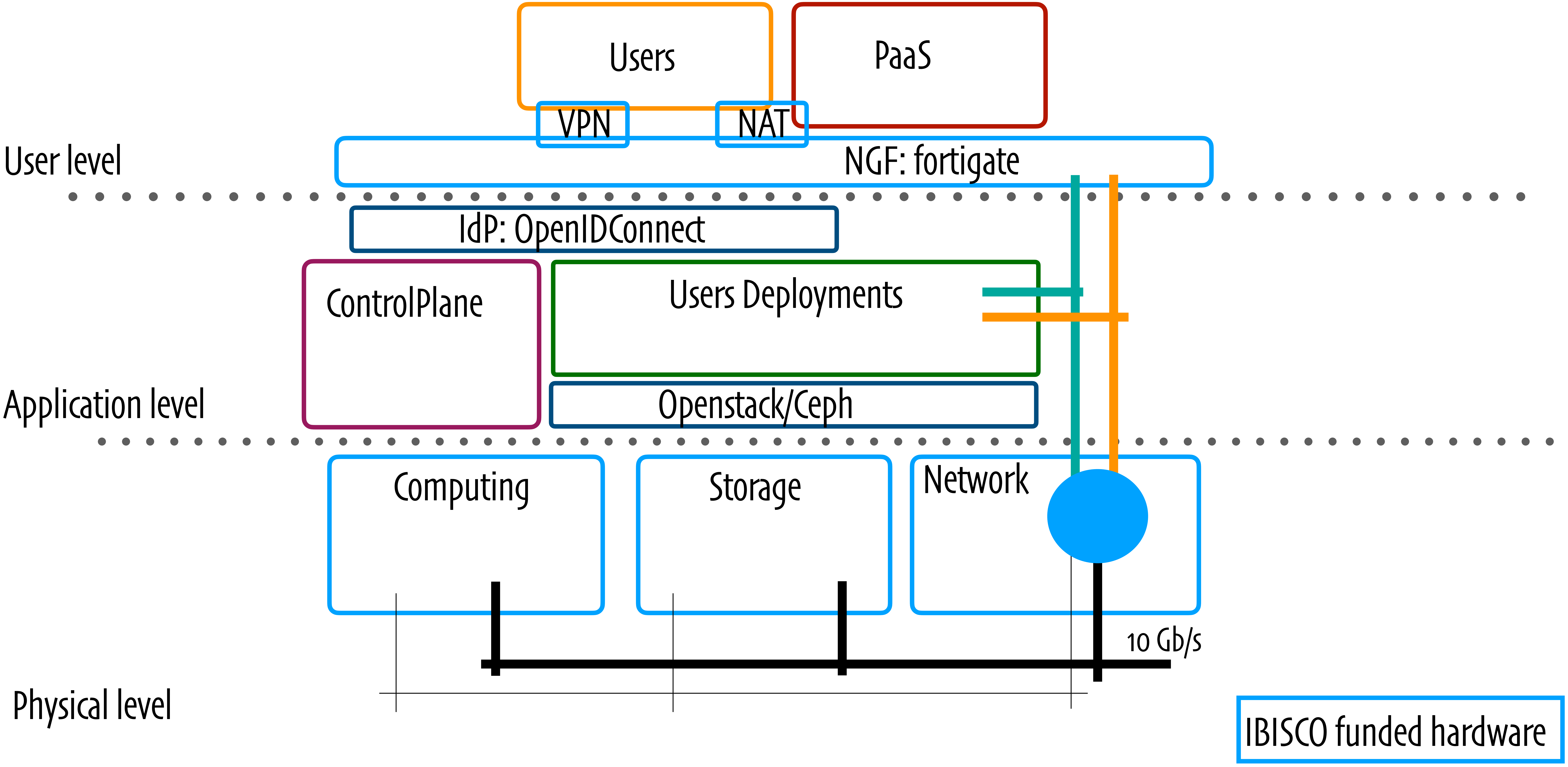
the upgraded IaaS infrastructure

- One of the DataCenter assets
- built on top
  - resources founded by IBISCO project
  - Public Openstack, Open Source Cloud Software
  - experience gained from previous cloud deployments [INFN-CLOUD/DATACLOUD, Cloud@ReCaS]
  - Automatic deployment and management provided by puppet

the upgraded IaaS infrastructure

- Exploited by different scientific community users to meet, on demand resources tailored to their computing requirements
- federated with DATACLOUD, EGI and other distributed computing infrastructures
- provides more resources than the previous cloud instance in terms of
  - cores available +40%
  - TeraBytes available +300%
  - new StorageType available based on SSD disks
  - saving space in terms of Cores/TeraBytes per 1U

Architecture



Computing Details

- 29 servers, 4224 cores available
- CPU ratio: 2 Memory ratio: 1.2
- Nova compute [libvirt] use ceph rbd as local storage backend
- live migration enabled so a ComputeNode can be drain for maintanance
- No GPUs available now

Networking Details

- all the networks service running on IBISCO funded hardware
- Flat public and private network deployed on top of local networks. No Overlay network
- One private network/subnet per tenant created at the tenant bootstrap
- Linuxbridge agent used to attach VM to the local network
- NGF[Fortigate] provides:
    - VPN service in order to let users access their tenant private network
    - NAT service lets tenant private subnet to reach Internet
    - IPS for all the InBound traffic
        - All the inbound[outbound] public traffic is inspected by the NGF dropping malicious traffic

Storage Details

- 43 ceph host-osd storage servers (and 3 ceph-mon)
- 826 OSDs up and in
  - OSDs running on the same host share one SSD disk hosting db e wal FileSystem
- 3.3 PiB available automatically splitted in 2 disk classes, HDD: 3.2 PiB and SSD: 175 TiB
- 2 root bucket-types:
  - the default one called DEFAULT
  - one dedicated for RadosGateWay [Object storage data for OpenStack Swift APIs] called RGW

Storage Details

- ceph pools are created on top of the two available buckets and disks classes
- well defined rules are created in order to map a pool to the desired bucket and disck class
- several standard and erasure-coded pools have been created since the bootstrap
    - standard pool has a replica size of 3
    - erasure-coded use the following profile
        - k=5 [data chucks], m=2 [coding chucks]
        - overhead factor[k+m/k] = 1.4 [1 GiB data use 1.4 GiB if disk space]
        - crush-failure-domain = host

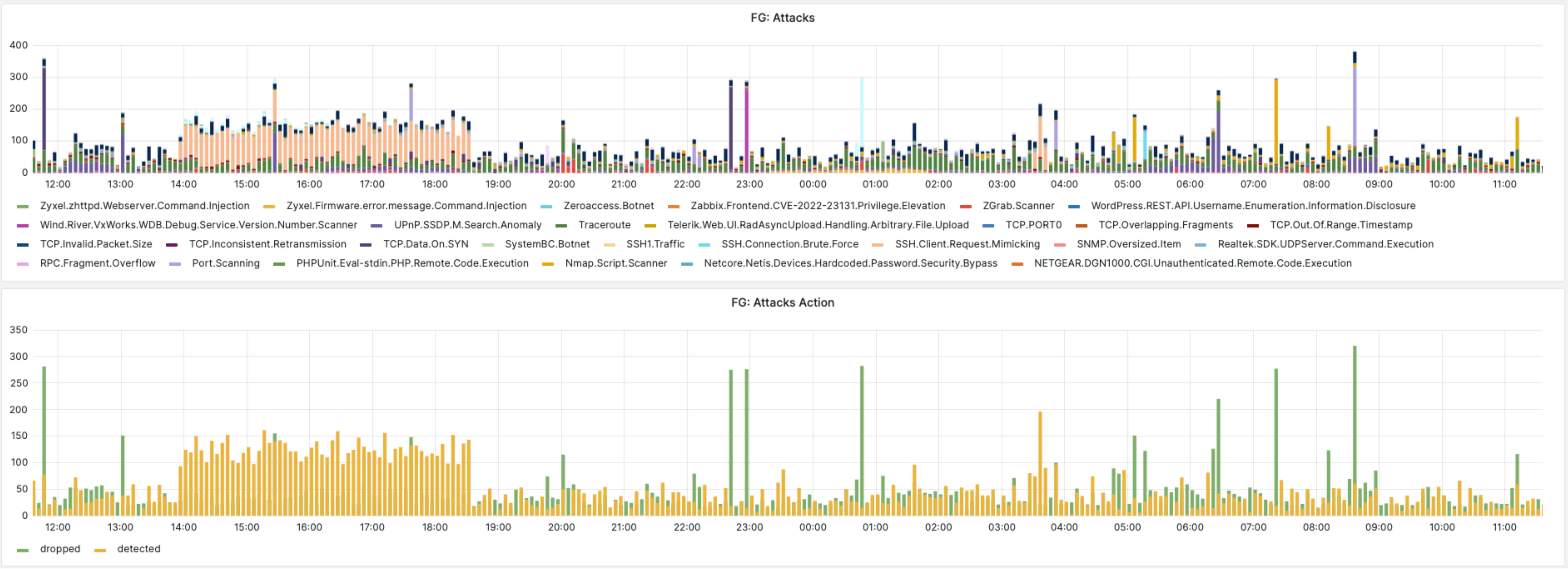
Authentication

- Cloud User Authentication process is manly based on OpenIDConnect protocol
- User Digital Identity is not locally managed, actually it is delegated to an IdentityProvider[IdP]
- Multiple IdPs can be enabled/configured to allow users to access the cloud
- Users without an already Digital Identity available can exploit a local IdP after a standard user registration process
- Due to a configuration weakness of the Apache OIDCAuth module, multiple IdPs can be used only through an “Esaco” instance which acts as gateway for token validation and introspection.

backup slides

Networking Details

Fortigate detects several malicious traffic based on signatures



Fortigate drops the traffic on per policies base

OpenStack Storage Backend

- Cinder volumes: 2 pools in replica 3 for writing data to either HDD or SSD if performance is needed
- Cinder backups: 1 erasure-coded pool using HDDs under "default" bucket-type
- Glance: 1 pool in replica 3 using HDDs under "default" bucket-type
- Swift: Ceph rados-gateway is used for Swift APIs with 2 pools:
  - Erasure-coded pool under the “rgw" bucket
  - 1 pool for metadata under the "default" bucket using SSD
- Nova: 1 pool in replica 3 using HDDs under "default" bucket-type