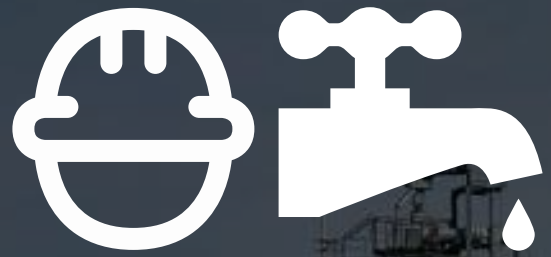
OPEN NETWORKING DANS UN MESOCENTRE



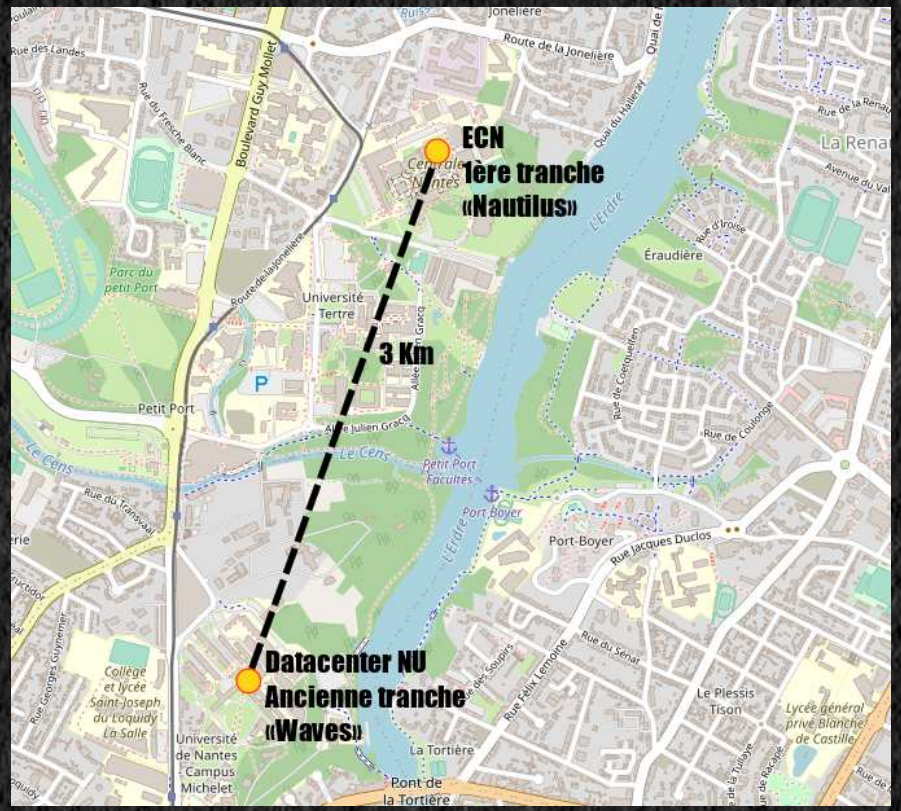
GLiCID

université angers · CENTRALE NANTES · Nantes Université · Le Mans Université

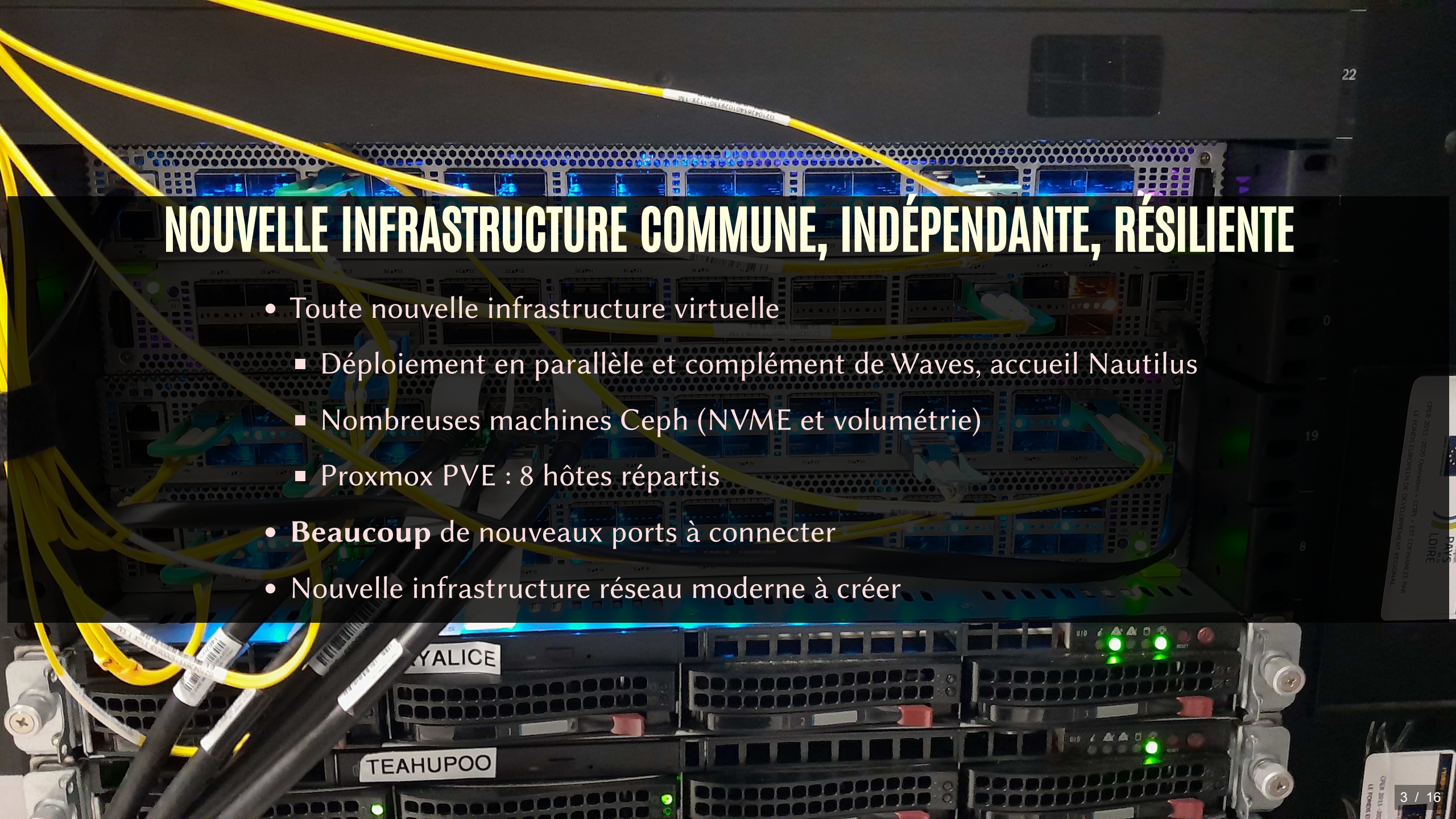
Yann Dupont <Yann.Dupont@univ-nantes.fr>

DACAS / GLICID : CALCULATEUR RÉGIONAL.

- 3 volets CPER DACAS :
  Datacentre (2025-26),
  Réseau régional (06/2023)
  HPC (GLiCID), plusieurs clusters :
  - Waves (existant, migrera fin 2023)
  - Nautilus (tranche 1, 06/2023)
  - Tranches 2 et 3 à venir
    (installation nouveau datacentre)



- À cheval sur 2 salles (jusqu'à 2025/2026)
- Liaison unique 100 Gb/s depuis 11/2021

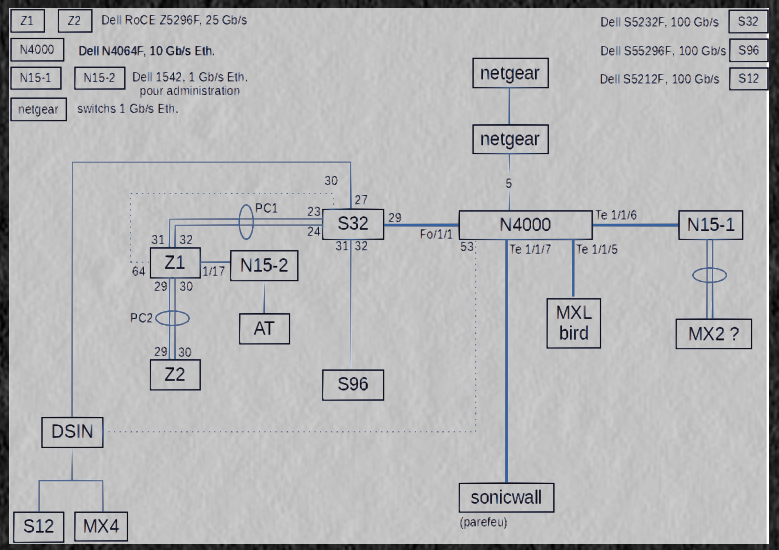


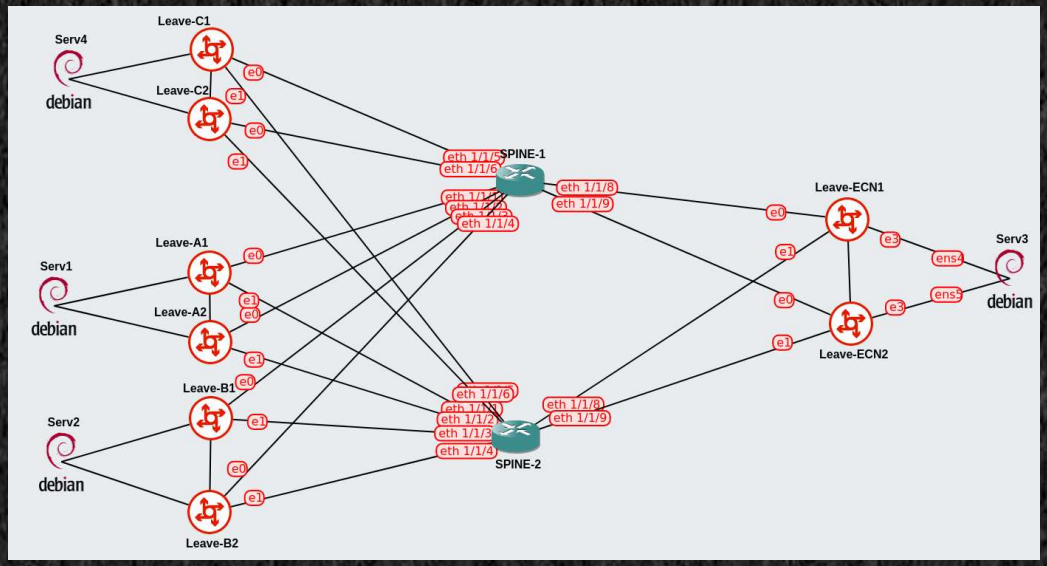
NOUVELLE INFRASTRUCTURE COMMUNE, INDÉPENDANTE, RÉSILIENTE

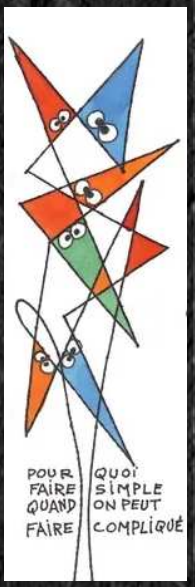
- Toute nouvelle infrastructure virtuelle
  - Déploiement en parallèle et complément de Waves, accueil Nautilus
  - Nombreuses machines Ceph (NVME et volumétrie)
  - Proxmox PVE : 8 hôtes répartis
- **Beaucoup** de nouveaux ports à connecter
- Nouvelle infrastructure réseau moderne à créer

MIGRATION PROGRESSIVE VERS *FABRIC ETHERNET*

- Réseau original hétérogène peu résilient, utilisation HPC (RoCE, faible latence)
  Réorganisation impossible avec l'existant, plus assez de ports libres
- *Fabric Ethernet* : pas de points de faiblesse, évolutif, routage décentralisé







- Nativement, Ethernet ne permet pas bouclages et chemins multiples
  - «underlay» physique L3 routé (BGP IPv6) (schéma) + «overlay» virtuel (VxLAN / EVPN).
- Processus long et complexe, migration partie existant, ajout de liens DC ⟶ ECN (⟶ 01/2024)

OPEN NETWORKING ?

- Beaucoup de matériel à acheter, intégration existant ?
- *Fabric* multi-constructeurs : interopérabilité difficile (protocoles incomplets/propriétaires…)
  - Dépendance *Network OS* ⟶ constructeur unique ⟶ captivité
- Compatibilité **ONIE** : installeur standard de *Network OS* pour commutateurs *Open Networking*
  - 7 de nos switches sont déjà compatibles **ONIE**
- *Whitebox* : commutateur nu, générique, avec **ONIE**
  - Constructeurs et matériels (Asics) divers, OS identique ⟶ interopérabilité, indépendance

⚠ Changer de *Network OS* : séduisant … Mais peu de retour de la communauté, pas notre spécialité

Entreprise risquée : (maturité, fiabilité, pas de support, support RoCE, consommateur de temps)

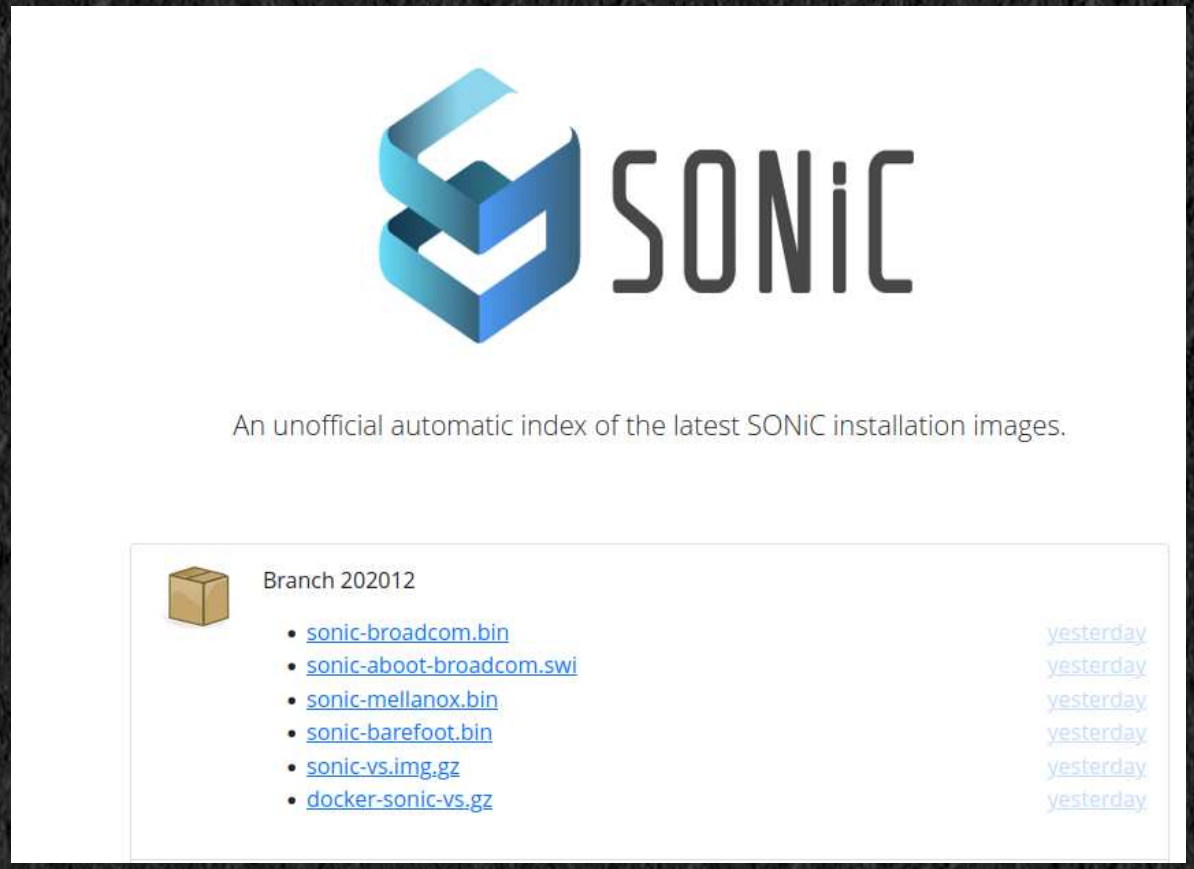
Switches existant donnant satisfaction, support, garantie de bon fonctionnement

idée initiale : rester chez le même constructeur, veille technologique sur les *Network OS*

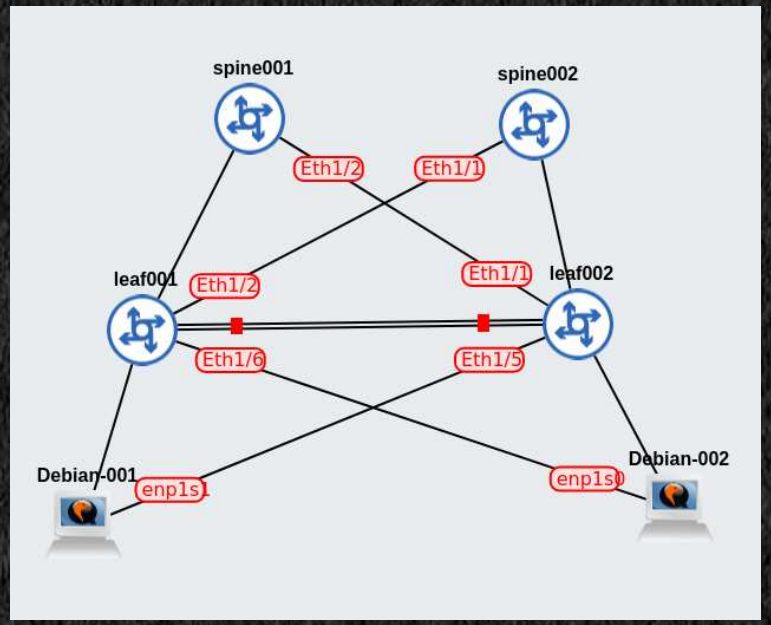
NETWORK OS COMPATIBLES ONIE

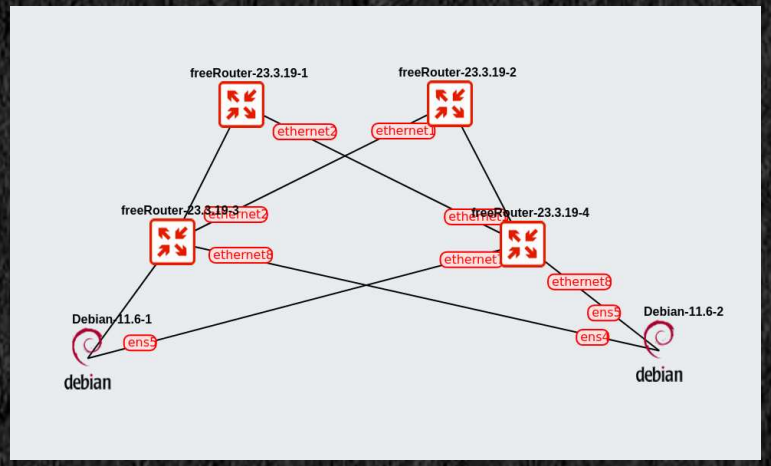
- Commerciaux : OS10, SONiC Entreprise, Cumulus
- Opensource, multi-constructeurs/chipsets
  - SONiC, RARE
- Simulation : GNS3
- Tests : SONiC et SONiC entreprise sur S5212F
  - Installation ONIE automatique via clé USB

```
dupont-y@UN-5CG6460BMY:/media/dupont-y/9555-E351$ ls -1hs
total 11G
1008M Enterprise_SONiC_OS_3.5.0_Enterprise_Standard.bin
 1,6G onie-installer-old-x86_64.bin
 1,4G onie-installer-x86_64.bin
 777M PKGS_OS10-Enterprise-10.5.3.2.89buster-installer-x86_64
 488M rare-old.bin
 1,4G sonic-barefoot-202205.bin
 1,6G sonic-barefoot-master.bin
 1,1G sonic-broadcom-202205.bin
 1,2G sonic-broadcom-master-182904.bin
```



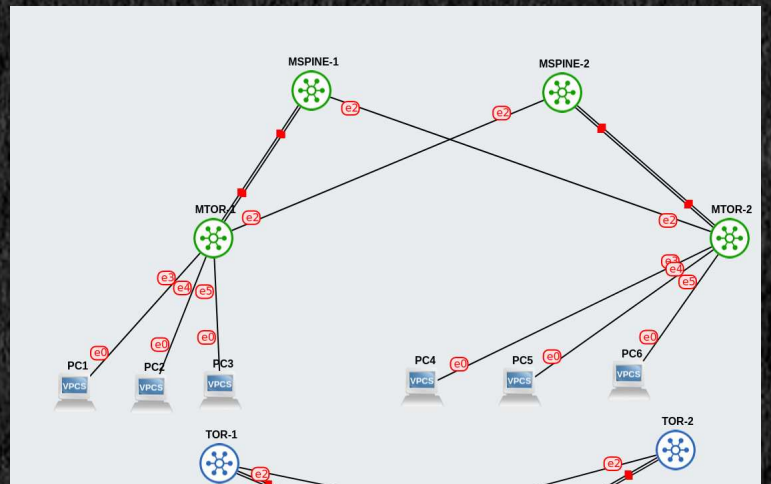
TESTS GNS3 (SIMULATEUR RÉSEAU)

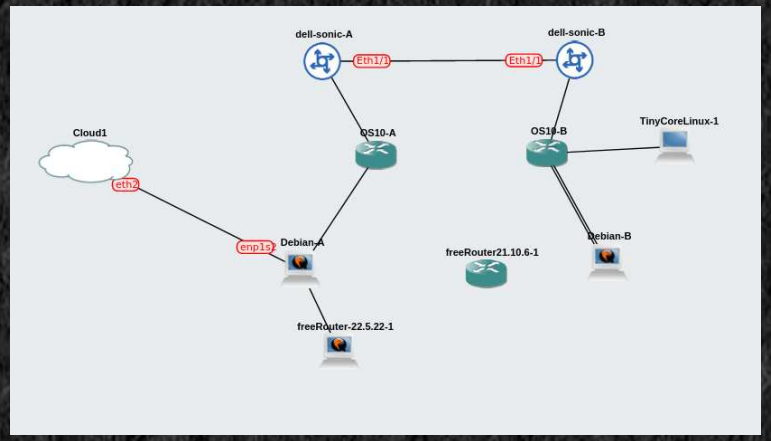




Network OS :

- Dell OS10
- SONiC
  - Entreprise
  - 202205, 202211,202305
- RARE





← Différents tests:

- Tester interopérabilité, topologies & valider
- Se faire la main sur les OS

SONIC

- Software for Open Networking in the Cloud
- Open-source, beaucoup de plate-formes supportées
  - OCP, Microsoft (!) puis Linux Foundation
  - Contributeurs : gros acteurs industriels (Microsoft, Dell, Nvidia/Mellanox, Broadcom, Edge-Core, Marvell, Cisco, Orange…)
- Base Debian Linux + conteneurs Docker
- Prise en main particulière
  - Pas de CLI, commandes shell, config : REDIS
  - Plutôt pilotage (Ansible), que géré à la main

```
Debian GNU/Linux 11 sonic ttyS0

sonic login: admin
Password:
You are on
  ____   ___  _   _ ___ ____
 / ___| / _ \| \ | |_ _/ ___|
 \___ \| | | |  \| || | |
  ___) | |_| | |\  || | |___
 |____/ \___/|_| \_|___\____|

-- Software for Open Networking in the Cloud --
Linux sonic 5.10.0-18-2-amd64 #1 SMP Debian 5.10.140-1 (2022-09-02) x86_64 GNU/Linux

admin@sonic:~$ docker ps
CONTAINER ID   IMAGE                                 NAMES
1ae042b1b220   docker-sonic-telemetry:latest         telemetry
644a5daf2a56   docker-snmp:latest                    snmp
c87838444cce   docker-platform-monitor:latest        pmon
52c0164ce52a   docker-sonic-mgmt-framework:latest    mgmt-framework
d502eee697c7   docker-lldp:latest                    lldp
476d12d4fa0f   309d1d28abfb                          dhcp_relay
562ac05f1288   docker-fpm-frr:latest                 bgp
b15ed2b2d70f   docker-router-advertiser:latest       radv
9328046590de   docker-syncd-bfn:latest               syncd
aa947370b572   docker-teamd:latest                   teamd
39225f382ae1   docker-orchagent:latest               swss
dacc7c570c4d   docker-eventd:latest                  eventd
e0ec03d83a1b   docker-database:latest                database
```

SONIC ENTREPRISE

- Propre à chaque constructeur, basé sur version particulière (parfois ancienne) de SONiC Open-Source
  - Ajouts ou modifications propriétaires, fonctionnalités supplémentaires (non open-source)
    - CLI proche OS10 chez DELL, «SAG» chez EdgeCore
  - Offre commerciale
    - Licences supplémentaires, support et garantie habituels
- Différences significatives entre version open et entreprise
  - Options, commandes subtilement différentes
  - Dump XML de la base : non intéropérable (open/entreprise/constructeurs)

⚠ Avantages/inconvénients proches d'un *Network OS* propriétaire

RARE

- Projet européen Géant
  Renater partie prenante
  Déployé sur Géant
- Petite équipe de contributeurs « académiques »
- Debian + FreeRouter gérée par NIX : petit OS
- CLI complet, fonctionnellement riche
- Uniquement plate-forme physique « Tofino »
  - Pas d'accords avec d'autres fondeurs (royalties…)

```
Debian GNU/Linux 11 localhost ttyS0

localhost login: admin
Password:
Linux localhost 5.10.0-8-amd64 #1 SMP Debian 5.10.46-4 (2021-08-03) x86_64
admin@localhost:~$ telnet localhost  2323
Trying ::1...
Trying 127.0.0.1...
Connected to localhost.
Escape character is '^]'.
welcome
line ready
rare#
show running-config
hostname rare
[…]
!
bridge 1
 mac-learn
 exit
!
vrf definition GLiCID
 exit
!
[…]
!
interface sdn16
 mtu 9200
 macaddr 0063.4c03.7d77
 lldp enable
 lacp 0000.0000.1234 12345 1
 no shutdown
 log-link-change
 exit
!
```

COINCÉS POUR COINCÉS…

- Besoin pressant de ports 100G
- Changement de titulaire du marché
  - Ancien titulaire : délais trop importants, tarif devenu moins favorable, complication achat licences…
  - Nouveau titulaire ou UGAP : **NON** ! (délais, non open, pas d'unification d'OS, interop. existant ?)
- Hors marché : disponibilité **immédiate** de *whiteboxes*
  - Possibilité d'unifier avec notre existant ? (SONiC open source)
- Risque raisonnable et assumé : achat 2 switches, compatibles SONiC, RARE, SONiC entreprise
  - Choix Edgecore 100BF-32X 2x32 ports 100G, chipset «Intel/barefoot Tofino 1»
  - Tarif compétitif, livraison quelques jours

SONIC COMMUNITY = FRANKEINSWITCH

- Expérience décevante
  - Docker : assemblage (hétéroclite ?) de nombreuses fonctions
  - Manque d'intégration, fragilité, lourdeur
- Bugs importants (mieux depuis 202305)
  - Conteneur «swss» (commutation) fragile, diagnostic difficile
    - Non-démarrage si ports 40G (202211)
      ou configuration XML importée incompatible
- Limitations gênantes
  - Routage au plus près pas encore intégré
    BGP à configurer **À PART** pour «FRR»
  - GNS3 : encapsulation Vxlan inactive ⟶ tests incomplets
- Documentation parcellaire, contradictoire

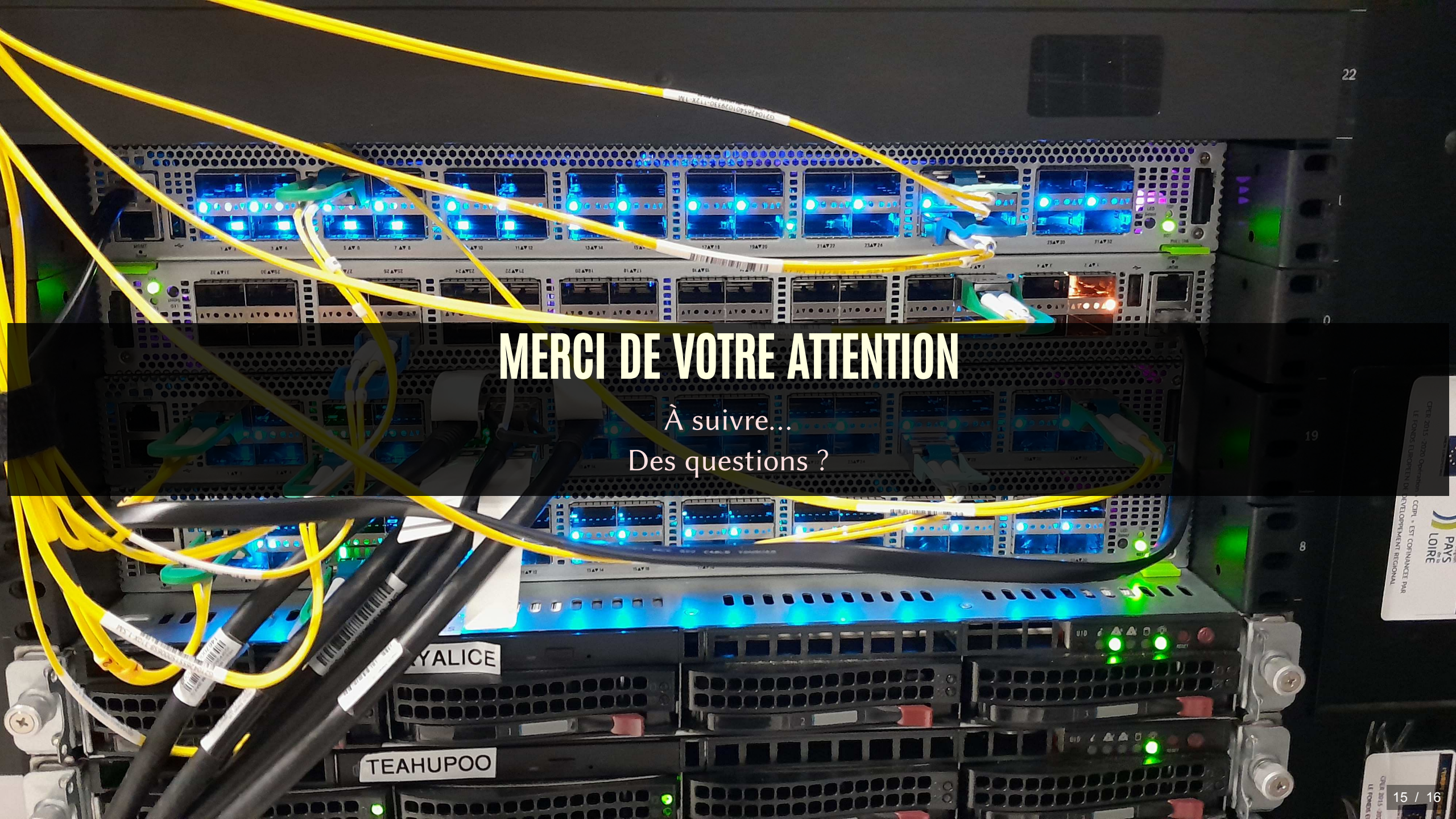
⚠ Cible du produit :
pas l'utilisateur final

RETARDS

- *Fabric Ethernet* pas encore démarrée
  - Raisons multiples
    - «C'est la vie» : beaucoup d'autres soucis extérieurs
    - **MAIS** Sonic Community n'a pas aidé
    - Échange switches «spine» : fenêtre de l'été ratée…
- Switches «leaf» en place et fonctionnels
  - Suffisants pour Nautilus, ceph et l'infrastructure commune
  - Au final, ça fonctionne
- *Fabric Ethernet* peut venir plus tard
  - Nous n'avons pas encore tous les liens physiques
  - Il reste beaucoup de plans B (Rare)

BILAN… MITIGÉ

- *Network OS*
  - « Sonic Community » : difficultés redoutées : pas déçus…
    - Finition perfectible, confiance limitée : irritant, décevant mais pas bloquant
    - Manques fonctionnels **actuels** pour nos besoins avancés
    - Interopérabilité prouvée, fonctionnement stable (après repérage des bugs…)
  - « Sonic Entreprise » : pas ces travers, mais moins intéropérable
  - Rare : prometteur, à tester davantage, avenir dépendant des constructeurs…
- Commutateurs *Open Networking*
  - Un vrai plus indéniable, apporte choix et indépendance
- Commutateurs *WhiteBox*
  - Tarifs compétitifs, disponibilité
  - RoCE, performance et latence : tout OK



MERCI DE VOTRE ATTENTION

À suivre...

Des questions ?

COPYRIGHTS

- Icônes, fonds écrans libres de droits
- Certains fonds proviennent de Wikipedia