

OCP

Open Compute Project

FRnOG #38 – 6 octobre 2023

\$ whoami

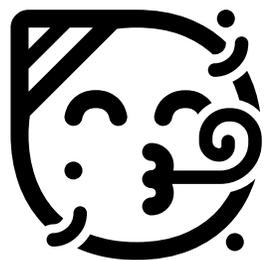
Frédéric Mossmann

CTO @ 2CRSI

- Pas d'AS
- Pas de /24
- Pas full IPv6 à la maison

PGP : C6C2 C97C 99A5 88FC D167 5268 564E 9F05 51BC 5841





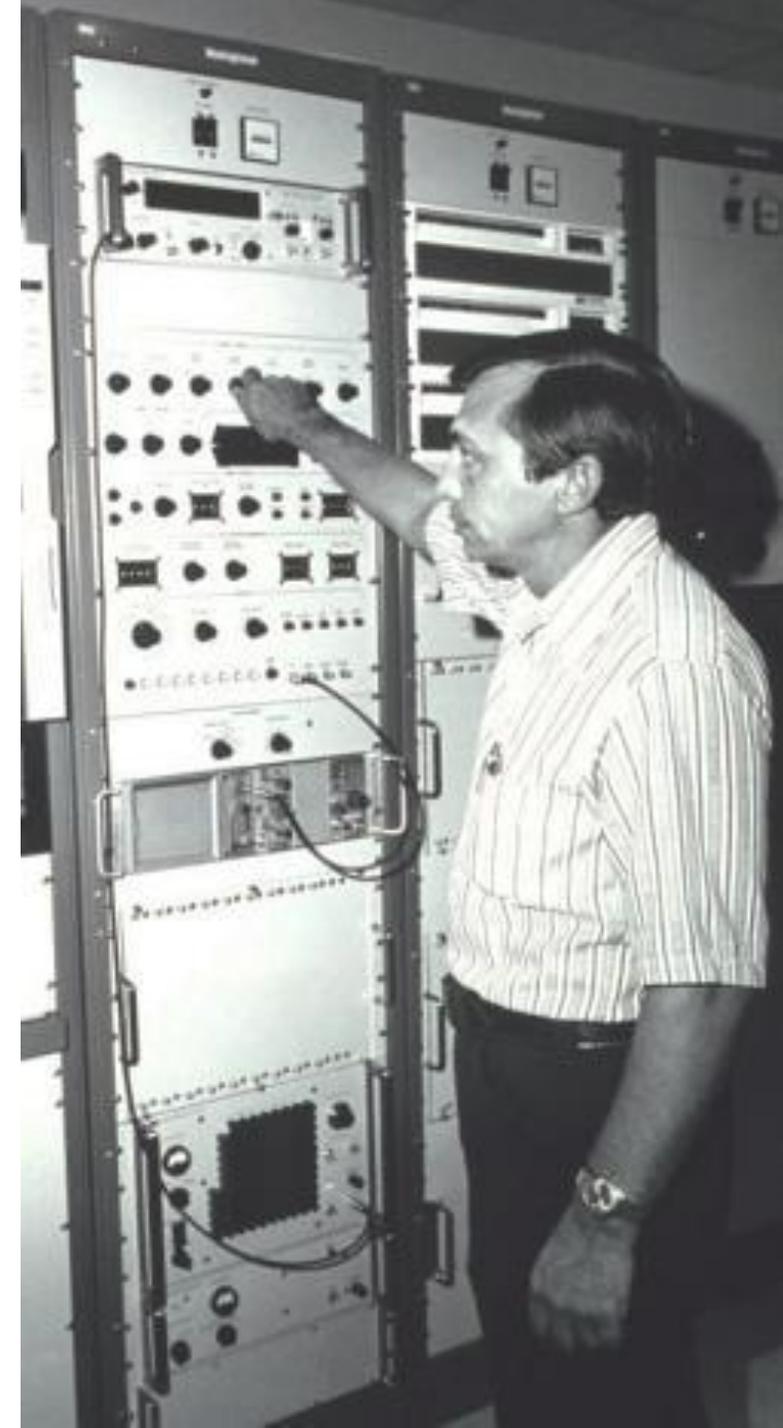
Happy Birthday EIA 310-D !!

Introduction du format en 1922 par AT&T

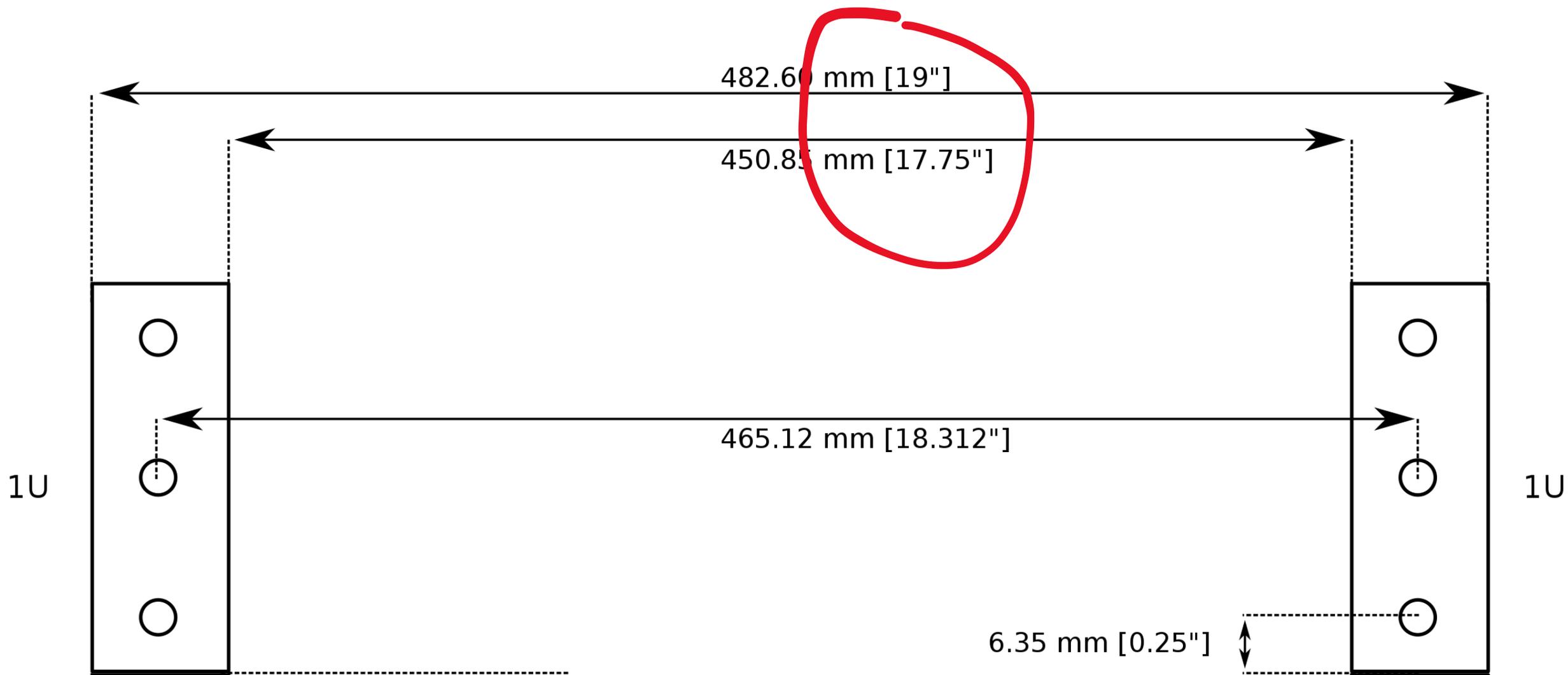
« monter divers modules électroniques »

« stockage de matériel de télécommunications »

— https://fr.wikipedia.org/wiki/Rack_19_pouces



Ça ne fait même pas 19 pouces !



Une brève histoire de l'OCP

Une brève histoire de l'OCP

2011 : Initiative d'un ingénieur de Facebook

(lors de la reconception du Data Center de Prineville, Oregon, USA)

Fondation à but non lucratif + Communauté collaborative

Objectifs principaux :

- Efficacité énergétique
- Simplification de la maintenance
- Réduction des coûts (Opex/Capex) (cf. ci-dessus...)
- Spécifications ouvertes, réunions publiques, vote des membres, etc.

Pensé pour les hyperscalers / CSP

- **Serveur traditionnel** : autonome sans optimisation
- **Serveurs en lames** : regroupement mais dette technique
- **Serveurs OCP** : nouveaux standards

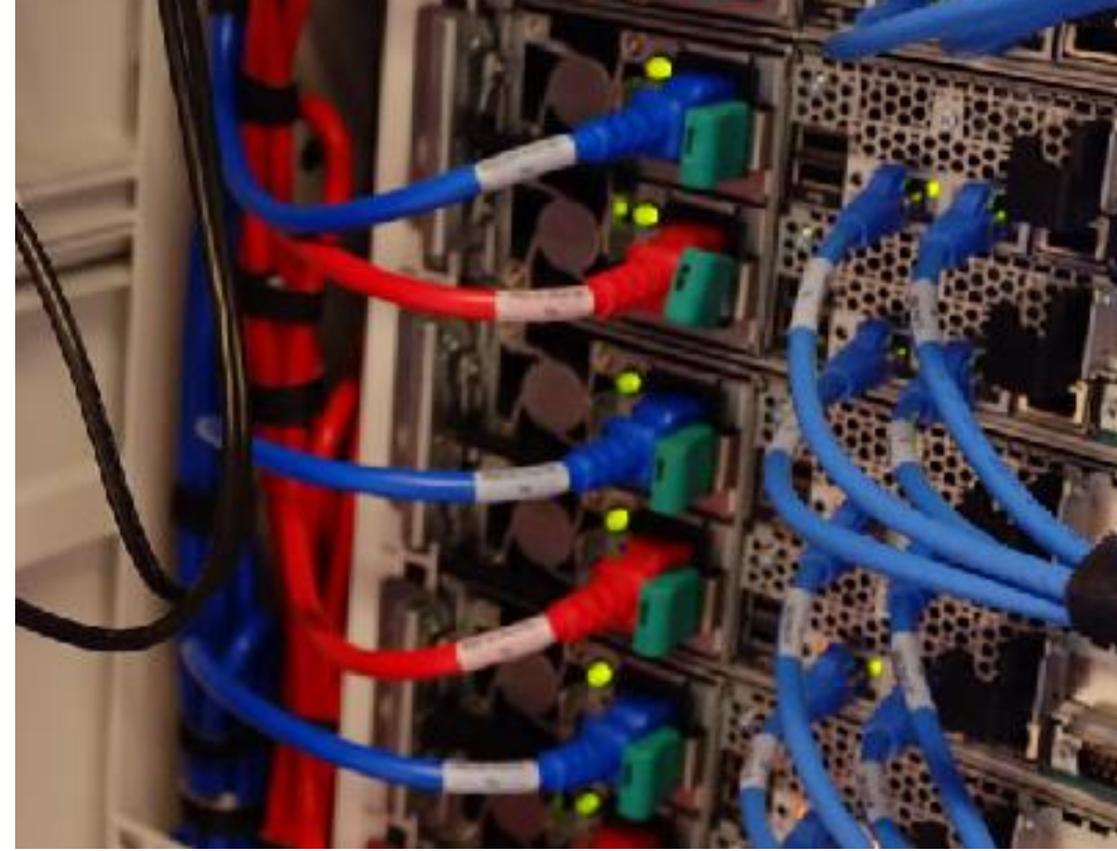


RAID5 sur des alims ?

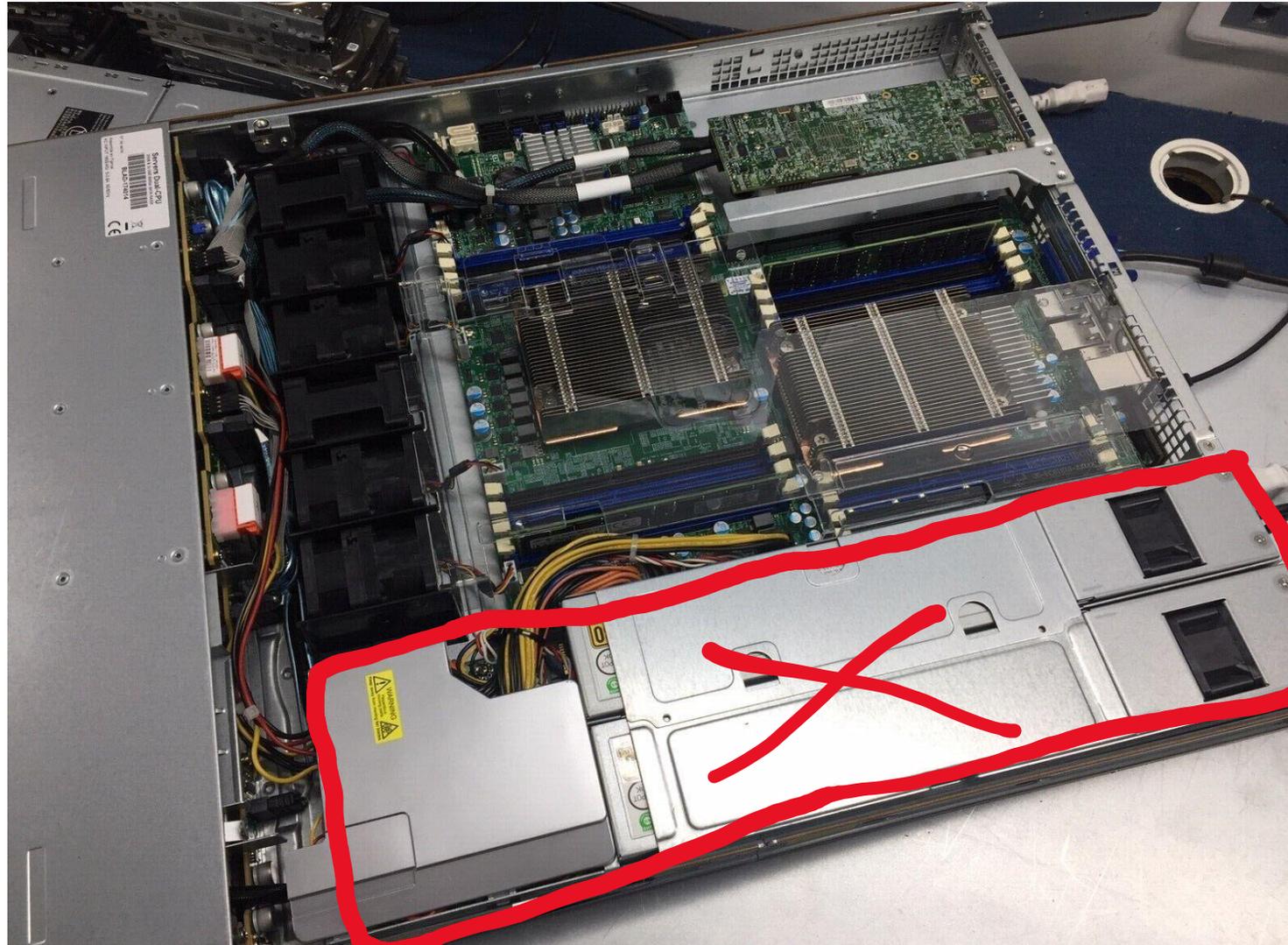
19" : Deux alimentations par serveur

OCP : Alimentations centralisées

- Redondance au niveau du rack
- Alimentations industrielles
- Alimentations plus efficaces
- ...et moins nombreuses



Moins d'alimentation = plus d'I.T. !



Front I/O

Toutes les connexions à l'avant (exemple : réseau)

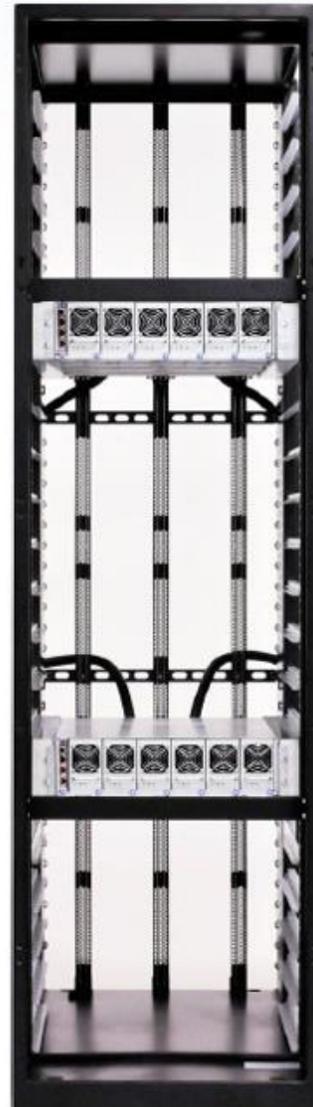
Gain de temps : Pas besoin de passer de l'autre côté du rack !

Et l'alimentation ??

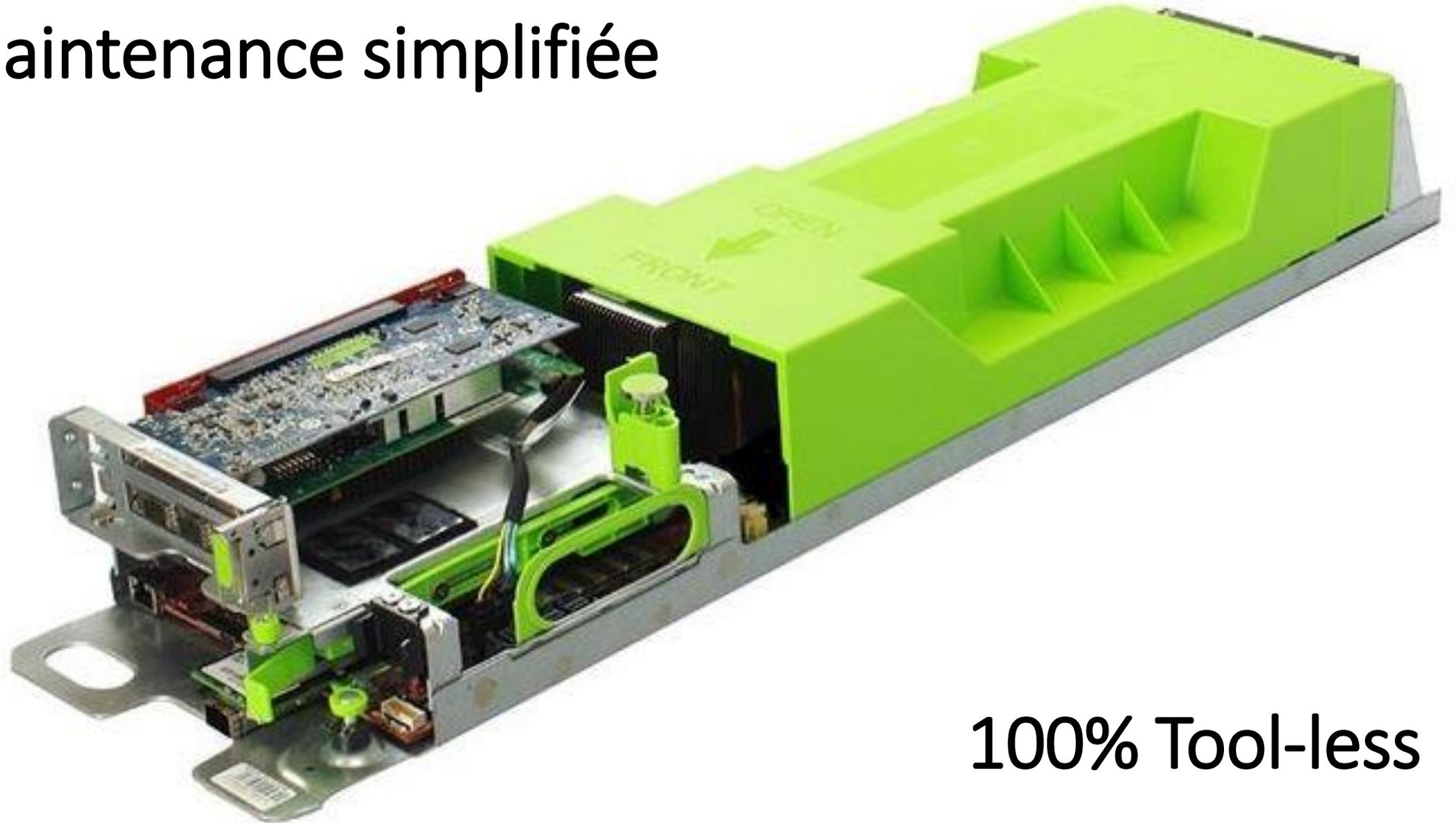
- sans câble
- par barre de cuivre au fond du rack
- connexion par « griffe »



OCP™
ACCEPTED



Maintenance simplifiée



100% Tool-less

(presque) Open Hardware

Faire, c'est bien... partager, c'est mieux !

Créer un « standard » :

- Augmenter la demande → augmenter l'offre de constructeurs
- Augmenter le volume → réduire les coûts

OCP : Pas uniquement des serveurs

OCP est organisé en Projets (et Sub-Projets) :

- Data Center Facility (Modular Data Center)
- Rack & Power
- Cooling Environments (Immersion, Door Heat Exchanger, Heat Reuse)
- Servers (HPC, Accelerators) / Storage / Network

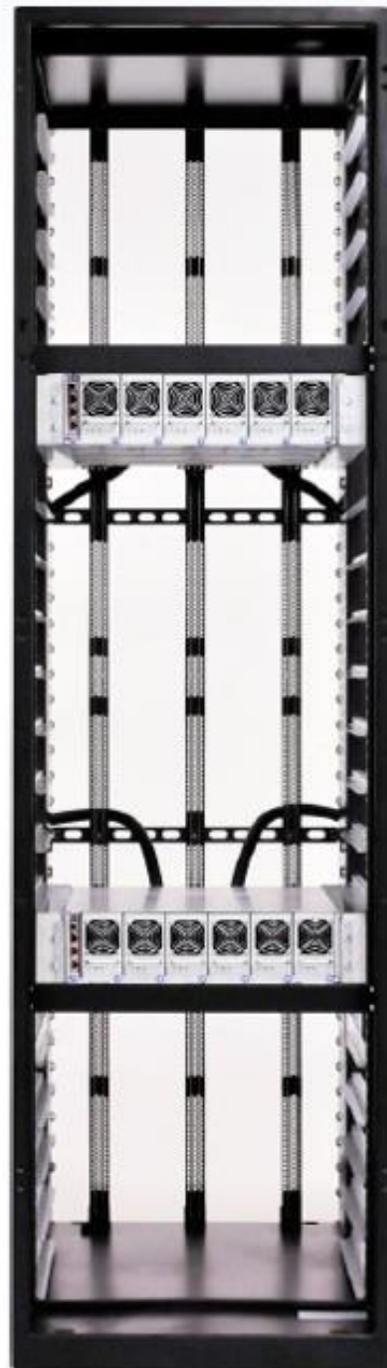
(quelques exemples)

Évolution de l'OCP

Open Rack (généralités)

- Nouvelle largeur : 21" (au lieu de 17,5" +20%)
- Nouvelle hauteur : 1 OpenU = 48mm (au lieu de 44,45mm)
- Même largeur que les racks 19" (600mm)

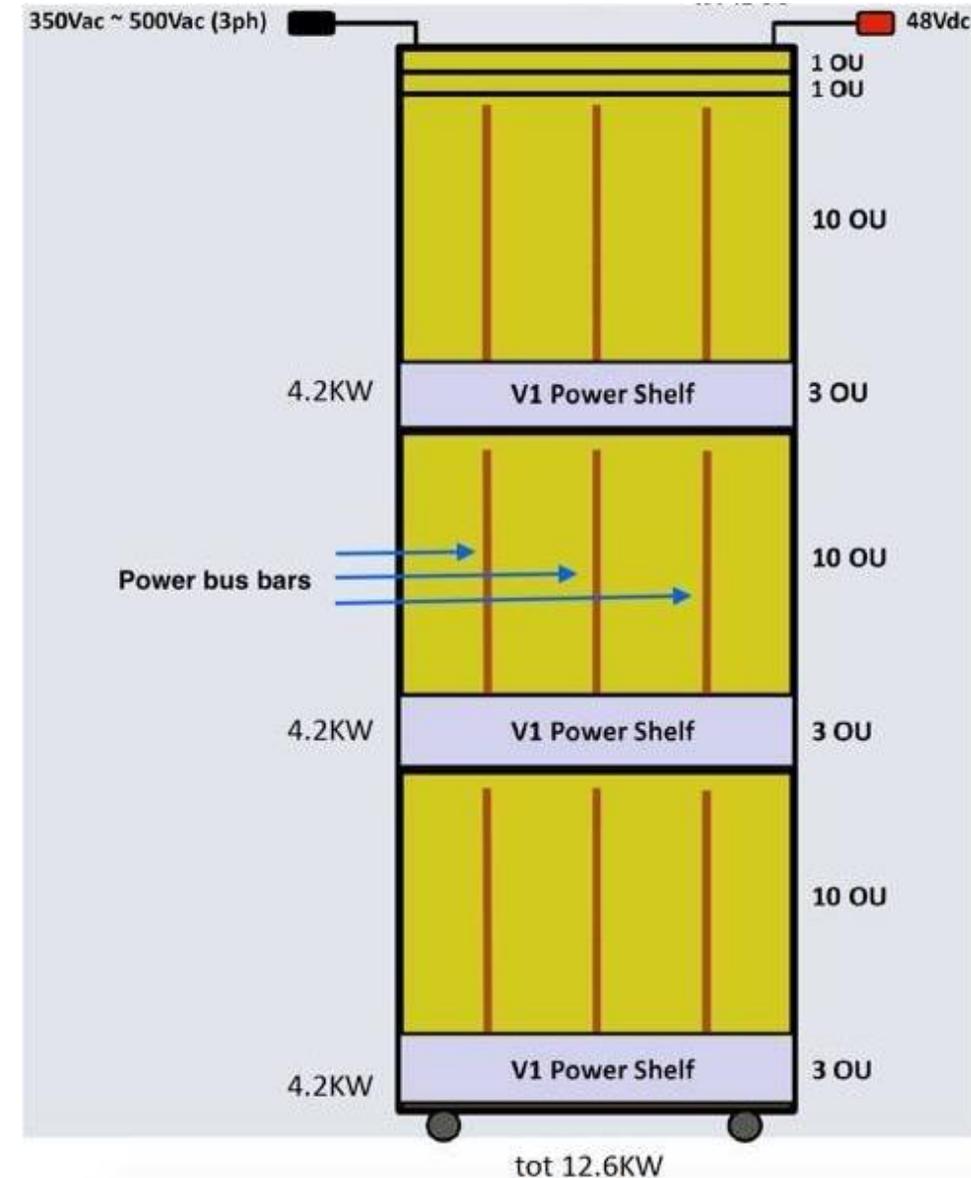
- Arrivée électrique triphasée
- Alimentations regroupées dans un PowerShelf
- Distribution du courant par Power Bus Bars
- Rails standards (équerres)



Open Rack v1 (ORV1)

- 3 zones avec 3 jeux de barres 12V
- Alimentation de secours 48V-54V
- Batteries en rack

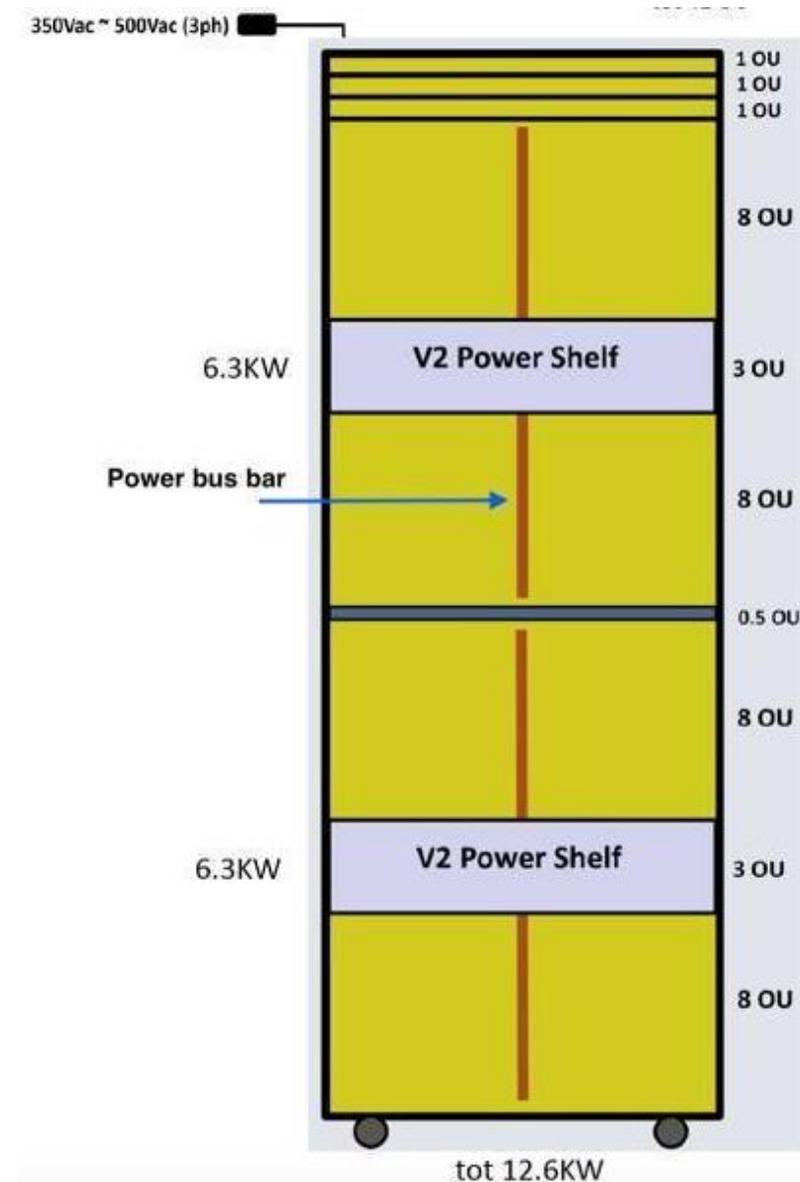
Objectif : Ne plus avoir besoin d'onduleur



Open Rack v2 (ORV2)

- 2 zones avec 1 jeux de barres 12V
- Batteries dans le Power Shelf
- Evolution en 48V envisagée mais non standardisé

Rétro-compatible ORV1



Open Rack v3 (ORV3)

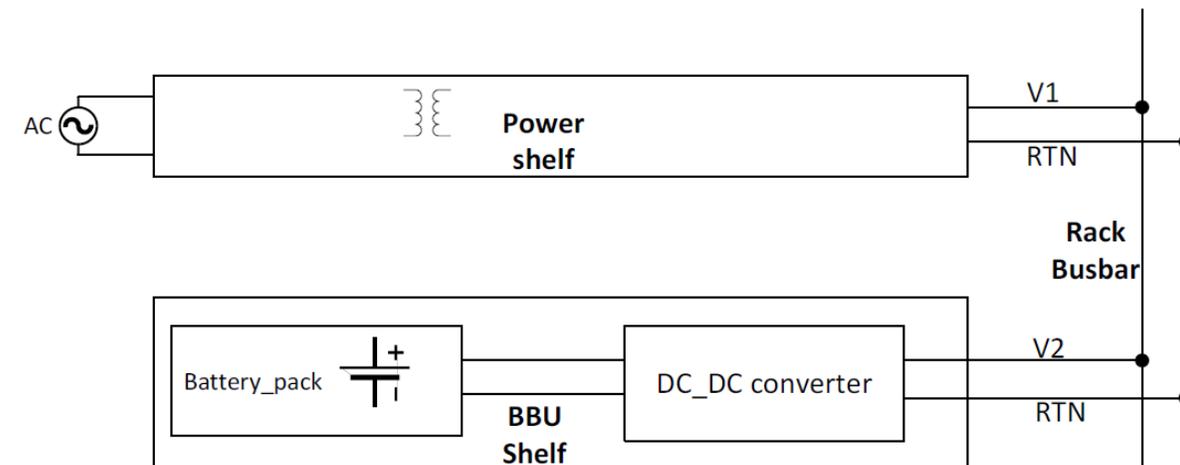
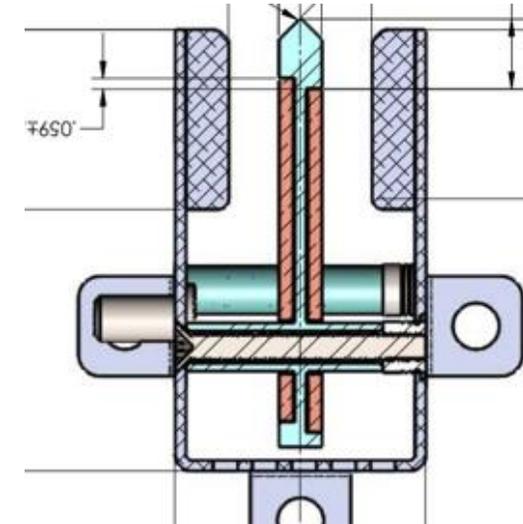
Non rétro-compatible !!

- 1 jeu de barres 48V
 - PowerShelf : nombre et positions libres (avec griffe)
 - Battery Backup Unit : optionnel
 - Nouveau modèle de Power Bus Bar
-
- Convertisseur 12V au besoin



Open Rack v3 (ORV3) – 48V ?

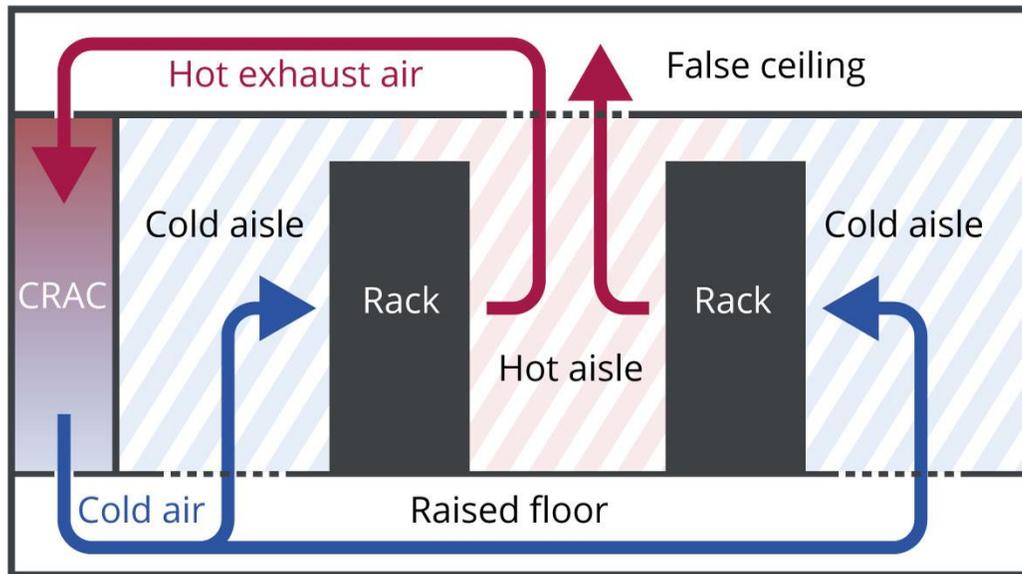
- Distribution 48V plus efficace
(tension plus haute, intensité plus faible, moins de pertes)
- Équipements réseau souvent compatible en DC 48V
- Nouveaux GPU prévus pour 48V
- 48V supporte mieux les pics de charge
- 48V : Power Bus Bars, connecteurs et régulateurs plus petits



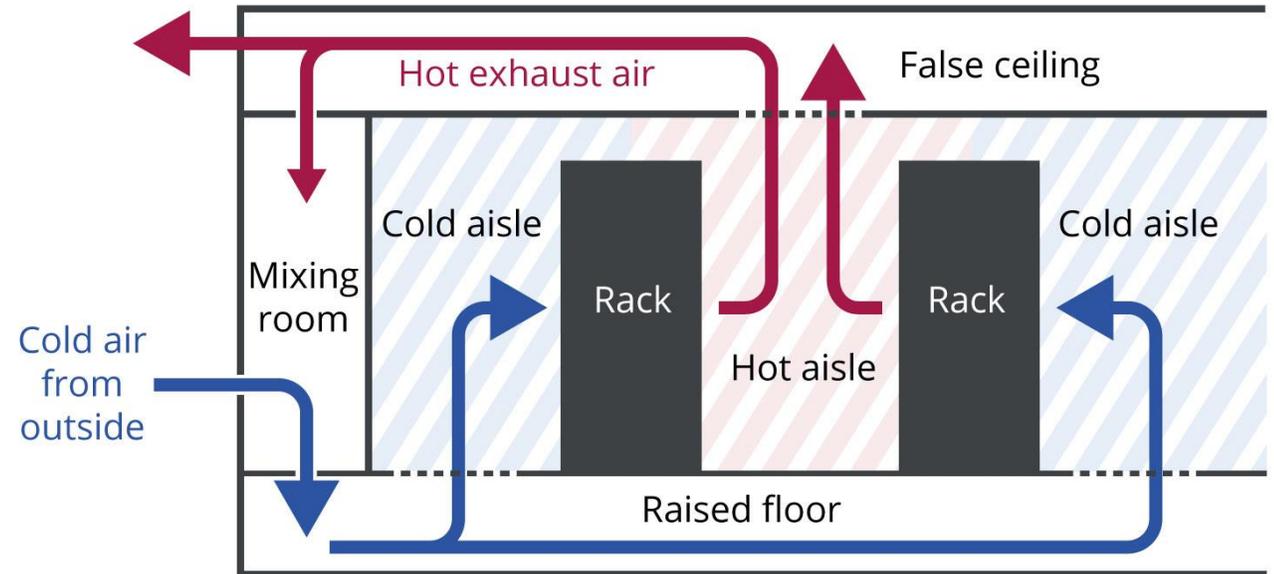
Changement dans les DC

Free Cooling

- Standard OCP : 30-35°C d'air ambiant entrant
- Prévu pour être utilisé sans climatisation



Avec climatisation



Free cooling

Colocation

- **ORV1 et ORV2**
Pas de portes, pas de « sidewall »
- **Open Rack v3**
Standard pour être acceptés en colo

Même footprint : 600mm



Direct Liquid Cooling

- **Open Rack v3 standardise le DLC**
- Manifold prévu dans le rack
- Raccord avec centrage et obturation automatiques



Standards OCP devenus la norme

OCP Mezzanine card 2.0

- Contributeur : Facebook
- Montage sur la carte « à plat »



- Plus compact que le PCIe « perpendiculaire »

OCP NIC 3.0

- Contributeurs : « OCP NIC sub-project »
- Montage coplanaire



- Supermicro SIOM : même idée, non standard

OAM (OCP Accelerator Module)

- Prévu pour les serveurs
- Adapté aux multi-GPU
- Nombre de connexions plus important
- Montage plus industriel



Exemple : Intel Ponte Vecchio, AMD Instinct MI250

Wedge 100 (Edgecore)

- Utilisé massivement par Facebook
- Switch « Bare Metal »
- ONIE (Open Network Install Environment)



<https://www.opencompute.org/documents/facebook-wedge-32x100ge-top-of-rack-switch>
<https://github.com/opencomputeproject/onie>

SONiC (NOS)

« Software for Open Networking in the Cloud »



<https://www.opencompute.org/projects/sonic>
<https://sonicfoundation.dev>

Conclusion

Open Compute Project

- Pas pour tout le monde (exemple : serveurs unitaires)
- MOQ : 1 rack

mais...

- 19" vraiment trop juste pour les système denses (2U4N)
- Transition de plus en plus facile (même footprint, nombreux fabricants)
- Source d'innovation, même en 19"

Merci JJ ! ;-)

Merci ! (et questions ?)

ocp@2crsi.com

<https://www.linkedin.com/in/fmossmann>