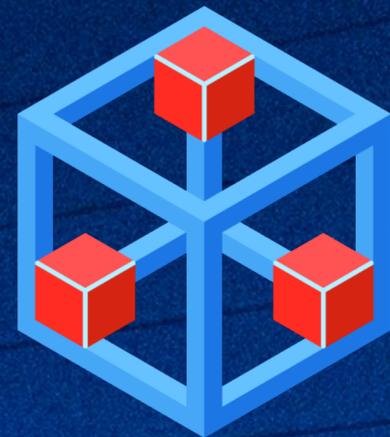


LA SOBRIÉTÉ AU CENTRE DE CALCUL RÉGIONAL



calmip



LE CONTEXTE

- Augmentation importante du coût de l'électricité en 2022.
- La circulaire du MESR de septembre 2022 rappelle l'objectif de réduction de nos consommations d'énergie de 10% d'ici 2024 par rapport à 2019.
- Les tutelles de CALMIP lui demandent, fin 2022, de réaliser 10% d'économie d'énergie pour l'année 2023.
- Fait suite à d'autres mesures de sobriété mises en place à l'Espace Clément Ader, comme la récupération de chaleur pour chauffer le bâtiment, le refroidissement du supercalculateur à l'eau chaude.
- D'autres sont à venir comme la contribution à la boucle de chauffage de la métropole du secteur de Rangeuil/Montaudran



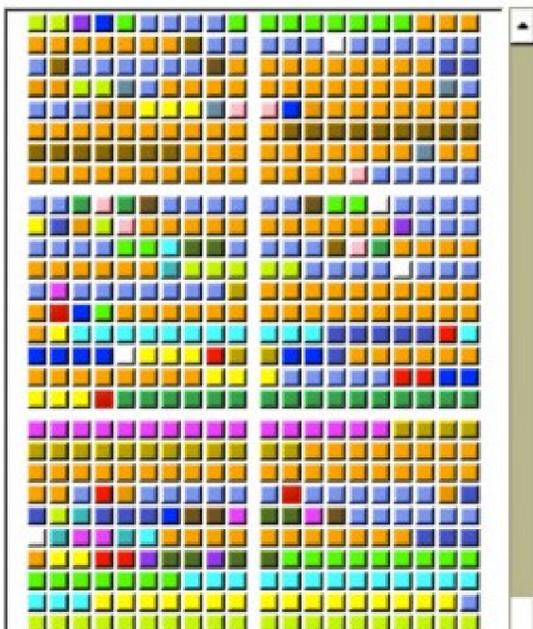
LE SUPERCALCULATEUR OLYMPE L'ARCHITECTURE

- ✓ Cluster de calcul Atos-Bull Sequana (1,365 PFlop/s)
- ✓ 13 464 cœurs - 76 To de RAM
- ✓ Processeurs Intel® Skylake à 2,3 Ghz 18-cores
- ✓ 48 GPU Nvidia Volta V100
- ✓ Réseau d'interconnexion rapide (Infiniband EDR)
- ✓ Espace de stockage temporaire de 1,5 Po (Lustre)
- ✓ Hébergé à l'Espace Clément Ader
- ✓ Refroidissement à eau (Régime Température 38°- 43°) - Mutualisé avec Météo-France
- ✓ 200 KW Linpack



LE SUPERCALCULATEUR OLYMPE L'ORDONNANCEUR

- ✓ Des traitements (jobs) soumis par les utilisateurs (x jobs par an, x faiblement parallele nœud + walltime + x fortement parallele nœud 20 + x job sur GPU)
- ✓ Un ordonnanceur de travaux (slurm)
- ✓ Une exécution très majoritairement asynchrone, et de type « gruyère »



Jobs	Partitions	Reservations	Burst Buffers	Visible Tabs
▷ 628390	standard	dvoidhyn	compress	RUNNING 02:06:59 100 r1i0n0,r1i6n[27-28],r1i7n[3-5,10-11],r2i1n[5-11,27-28],r2i2r
▷ 628408	standard	ckaul	swift_n_w1	RUNNING 03:19:30 3 r1i4n31,r5i3n[27-28]
▷ 628413	long	rischoeppner	02395_fine	RUNNING 03:15:55 1 r1i7n18
▷ 628415	standard	agoyal	bulk_hse	RUNNING 03:10:47 12 r102u[06-09,11-14,16-19]
▷ 628417	standard	rtrottie	Y32_al32_o96_pc.1	RUNNING 03:04:40 2 r1i6n[23-24]
▷ 628424	standard	eyoung	submit_wrf.sh	RUNNING 02:47:46 10 r1i7n[22-23],r3i1n[34-35],r3i2n[0-1],r6i3n[33-35],r6i4n0
▷ 628431	standard	eyoung	submit_wrf.sh	RUNNING 02:47:04 10 r1i4n[4-6,33-35],r2i1n[32-34],r2i6n15
▷ 628441	standard	bpollard	pb.rc2.1.560	RUNNING 02:36:09 1 r1i1n3
▷ 628443	standard	bpollard	pb.rc1.4.860	RUNNING 02:31:08 1 r1i4n9
▷ 628457	gpu	jrood	sh	RUNNING 02:02:52 1 r103u05
▷ 628481	long	hsitaram	alm	RUNNING 01:42:57 25 r1i7n[24-25],r6i2n[16-35],r6i3n[0-2]
▷ 628484	standard	worotni	T1M_mc12_m0Hb	RUNNING 01:40:37 16 r2i4n[29-35],r2i5n[32-33],r3i0n[9-15]





LES PISTES

- Réduire la consommation des nœuds lorsqu'ils sont *idle* : réduction de la fréquence du processeur, désactivation des cœurs,
- Arrêt de 10 % des nœuds, mais on réduit mécaniquement de 10% les ressources de calcul disponibles pour la recherche
- On plafonne la consommation, en bloquant les jobs si la consommation dépasse 90% de la consommation de 2019
- On arrête les nœuds de calculs inutilisés (on espère qu'ils le sont au moins 10% du temps ~ 2h/jour/noeud)





LES PISTES

- Réduction de la fréquence des processeur : oui, mais insuffisant
- La désactivation des cœurs est compliquée : problèmes au niveau vérification hardware

- Réduire la consommation de cœurs lorsqu'ils sont *idle* : réduction de la fréquence du processeur, désactivation des cœurs,
- Arrêt de 10 % des nœuds, mais on réduit mécaniquement de 10% les ressources de calcul disponibles pour la recherche
- On plafonne la consommation, en bloquant les jobs si la consommation dépasse 90% de la consommation de 2019
- On arrête les nœuds de calculs inutilisés (on espère qu'ils le sont au moins 10% du temps ~ 2h/jour/noeud)





LES PISTES

- Choix stratégique au niveau de l'unité : volonté de ne pas pénaliser les utilisateurs

- Réduire la consommation des nœuds qui sont *idle* : réduction de la fréquence du processeur, désactivation des c...
- Arrêt de 10 % des nœuds, mais on réduit mécaniquement de 10% les ressources de calcul disponibles pour la recherche
- On plafonne la consommation, en bloquant les jobs si la consommation dépasse 90% de la consommation de 2019
- On arrête les nœuds de calculs inutilisés (on espère qu'ils le sont au moins 10% du temps ~ 2h/jour/noeud)





LES PISTES

- On ne sait pas le faire simplement

- Réduire la consommation des nœuds en arrêtant *idle* : réduction de la fréquence du processeur, désactivation des cœurs
- Arrêt de 10 % des nœuds, mais cela réduit mécaniquement de 10% les ressources de calcul disponibles pour la recherche
- On plafonne la consommation, en bloquant les jobs si la consommation dépasse 90% de la consommation de 2019
- On arrête les nœuds de calcul inutilisés (on espère qu'ils le sont au moins 10% du temps ~ 2h/jour/nœud)





LES PISTES

- Réduire la consommation des nœuds lorsqu'ils sont *idle* . réduction de la fréquence du processeur, désactivation des cœurs,
- Arrêt de 10 % des nœuds, mais on réduit mécaniquement de 10% les ressources de calcul disponibles pour la recherche
- On plafonne la consommation, en bloquant les jobs si la consommation dépasse 90% de la consommation de 2019
- On arrête les nœuds de calculs inutilisés (on espère qu'ils le sont au moins 10% du temps ~ 2h/jour/noeud)





LA MISE EN PLACE

- Piloter le matériel du constructeur de la solution (EVIDEN – ex ATOS)
- On utilise les outils du constructeur de la solution pour la gestion de l'arrêt/démarrage du matériel
- On utilise les fonctionnalités de l'ordonnanceur (slurm)
- Quelques scripts maison pour gérer les cas d'erreurs

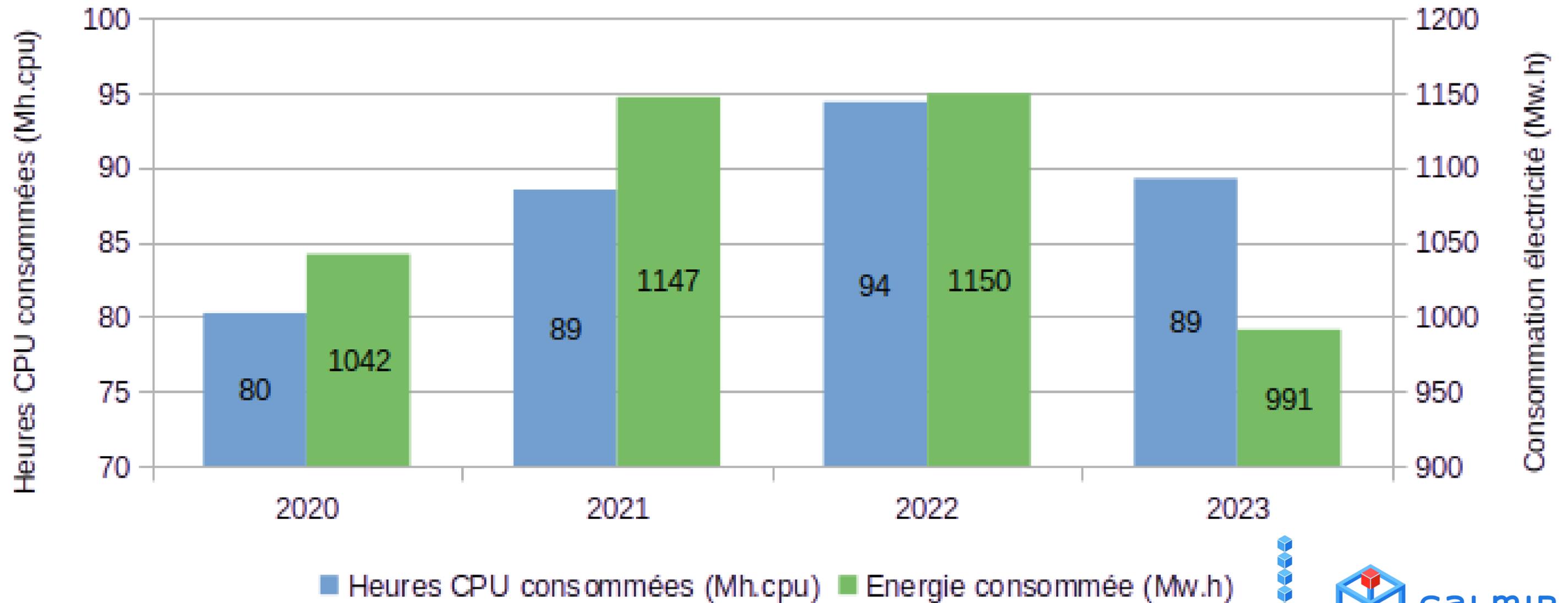
Premier retour d'expérience :

- Pas faire trop compliqué
- Attention aux bugs de l'ordonnanceur (slurm)



LE BILAN

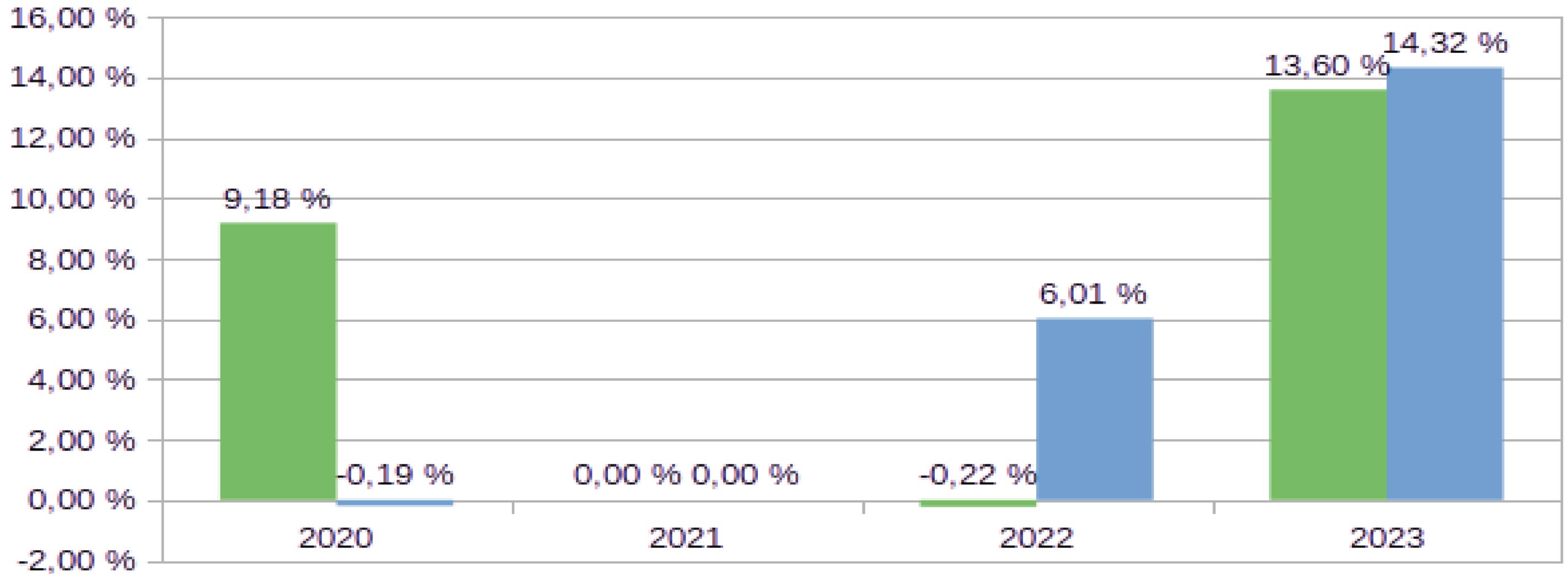
Bilan de la consommation électrique du supercalculateur Olympe au cours des 4 dernières années





LE BILAN

Economie d'énergie réalisée par rapport à 2021



■ Economies réalisées (brute) ■ Economies réalisées (usage)





LE BILAN

- Une mise au point assez longue (3 mois) avec une mise en production réelle en juin 2023
- Un coût de mise en place assez important au niveau RH
- Un léger temps d'attente lors du démarrage des gros jobs le lundi matin, mais bonne acceptation de la part des utilisateurs
- Pas d'augmentation significative des pannes sur les nœuds
- Objectif d'économie de 10 % réussi sur l'année 2023 et un impact finalement très faible sur l'expérience utilisateur





QUESTIONS ?

MERCI

