N° 232 — Décembre 2024 / Janvier 2025

www.linformaticien.com

L'INFORMATICIEN

Réseau
Automatiser avec Python

DevOps
Container lab

InfoCR
Bilan 2024

Palmarès 2024

Le retour des grands comptes

Conférences
AWS, Opentext, Hashiconf

DOSSIER

Datacenters

Au cœur de la digitalisation

L 14614 - 232 - F: 8,50 € - RD

WAICF

WORLD ARTIFICIAL INTELLIGENCE CANNES FESTIVAL

FEBRUARY
13-15, 2025
CANNES, FRANCE

WHERE AI CHANGE MAKERS MEET INDUSTRY LEADERS

12 000 business attendees · **250** exhibitors · **320** speakers

www.**worldaicannes**.com

An event of INSTITUT EUROPIA COMPRENDRE POUR AGIR · PALAIS DES FESTIVALS ET DES CONGRÈS CANNES · CANNES CÔTE D'AZUR FRANCE · DÉPARTEMENT DES ALPES-MARITIMES

Organized by RX In the business of building businesses

www.linformaticien.com
L'INFORMATICIEN

RÉDACTION

88 boulevard de la Villette, 75019 Paris, France.
Tél. : +33 (0)1 74 70 16 30 — *contact@linformaticien.com*

RÉDACTION : Bertrand Garé (rédacteur en chef)
et Victor Miget (rédacteur en chef adjoint)
avec : Olivier Bouzereau, Vincent Bussière, François Cointe, Jérôme Cartegini, Michel Chotard, Alain Clapaud, Guillaume Renouard et Thierry Thaureaux

SECRÉTAIRE DE RÉDACTION : Amélie Ermenault Martin

MAQUETTE ET RÉALISATION : Franck Soulier (chef de studio)

PUBLICITÉ
Tél. : +33 (0)1 74 70 16 30 — *pub@linformaticien.com*

VENTE AU NUMÉRO

France métropolitaine 8,50 € TTC (TVA 5,5 %)

ABONNEMENTS

France métropolitaine 72 € TTC (TVA 5,5 %)
magazine + numérique

Toutes les offres :
www.linformaticien.com/abonnement

Pour toute commande d'abonnement d'entreprise ou d'administration avec règlement par mandat administratif, adressez votre bon de commande à :

L'Informaticien, service abonnements,
88 boulevard de la Villette, 75019 Paris, France.
ou à abonnements@linformaticien.com

IMPRESSION

Imprimé en France par Imprimerie Chirat (42)
Dépôt légal : 1er trimestre 2025

L'INFORMATICIEN est publié par PC PRESSE, S. A. S.
au capital de 130 000 euros.
Siège social : 88 boulevard de la Villette, 75019 Paris, France.

ISSN 1637-5491
Une publication PCpresse,

FICADE

PRÉSIDENT, DIRECTEUR DE LA PUBLICATION :
Gaël Chervet

Une année chargée

2025 ne va pas être de tout repos pour les services informatiques des entreprises. Entre les nouvelles réglementations à mettre en œuvre (NIS 2, DORA, CSRD...), la continuation de la transformation numérique, l'évidente continuation des attaques de tous poils, il va y avoir du pain sur la planche et dès ce début d'année. Comme depuis deux décennies, l'Informaticien essaiera de vous aider dans ces nombreuses tâches, en vous apportant des informations sur les nouveautés dans les produits, les nouvelles tendances, la stratégie des acteurs importants, la cybersécurité, les grandes conférences des industriels de l'informatique... Bref, tout ce qui fait de l'Informaticien un titre majeur comme observateur de l'industrie informatique.

Pour notre part, outre le magazine et le site web, l'équipe est déjà sur la préparation du prochain Top Tech et du Palmarès. Vous trouverez d'ailleurs dans ce numéro les résultats de la dernière édition de ce Palmarès qui apporte cependant peu de surprises, les grands acteurs de l'industrie ayant été plus ou moins plébiscités par les votes. Nous faisons aussi un point sur les centres de données et les axes stratégiques pour 2025 pour ces équipements qui sont le socle de la digitalisation. Outre ces deux rendez-vous, les rubriques habituelles sont présentes avec leur lot de nouveautés et d'observation d'une industrie qui évolue toujours aussi vite.

Il ne nous reste plus qu'à vous souhaiter bon courage, réussite dans vos projets et une superbe année 2025. □

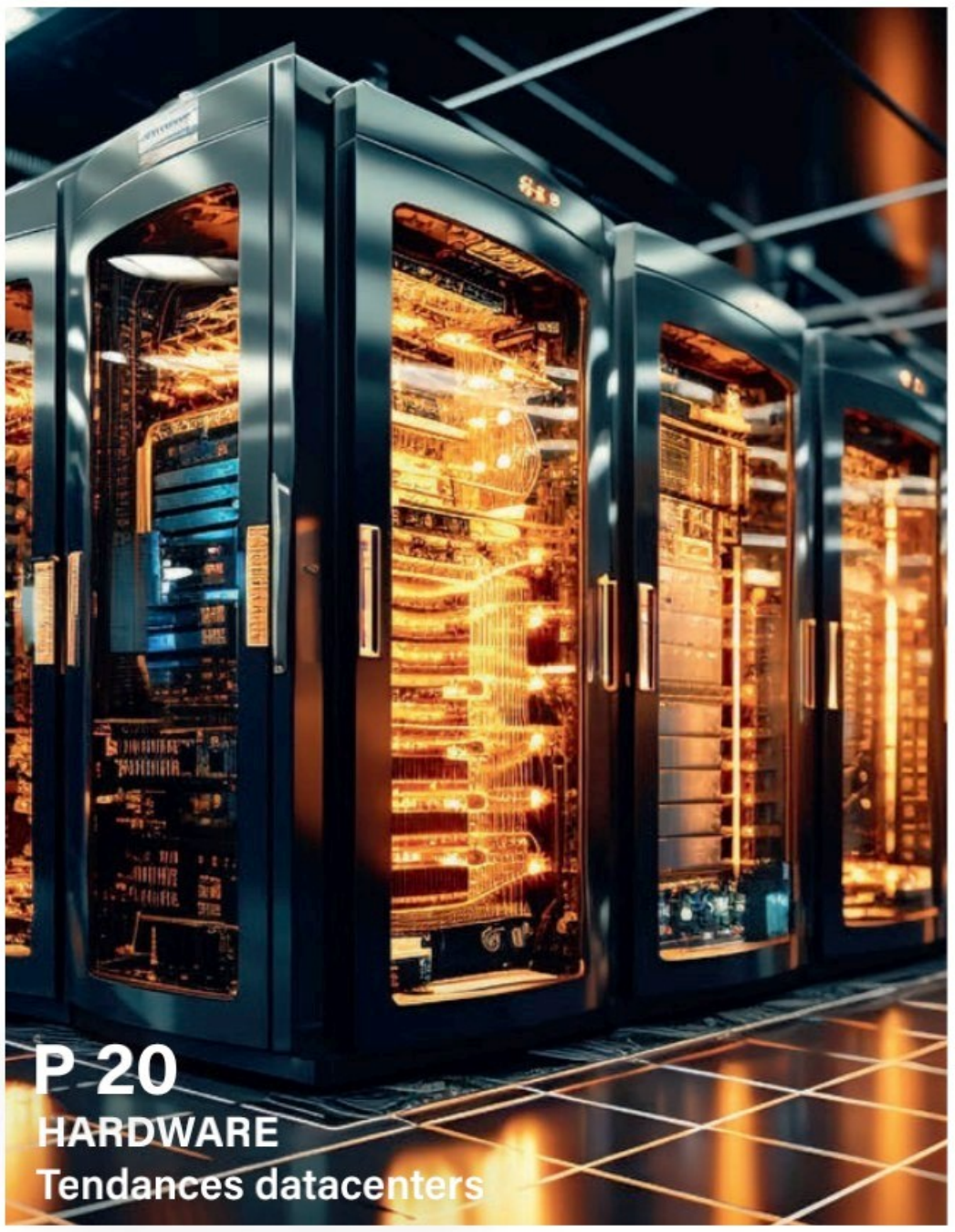
Bertrand Garé
Rédacteur en Chef

PURESTORAGE®

Simplifier le stockage des données pour toujours

Le stockage à la demande (STaaS) vous permet de bénéficier d'une flexibilité financière et d'une simplicité opérationnelle pour répondre durablement aujourd'hui et demain, aux besoins de votre entreprise.
Evergreen//One™ associe l'agilité du stockage dans le cloud public à la sécurité et aux performances d'une infrastructure all-flash. Cette solution STaaS offre une véritable expérience de cloud hybride.

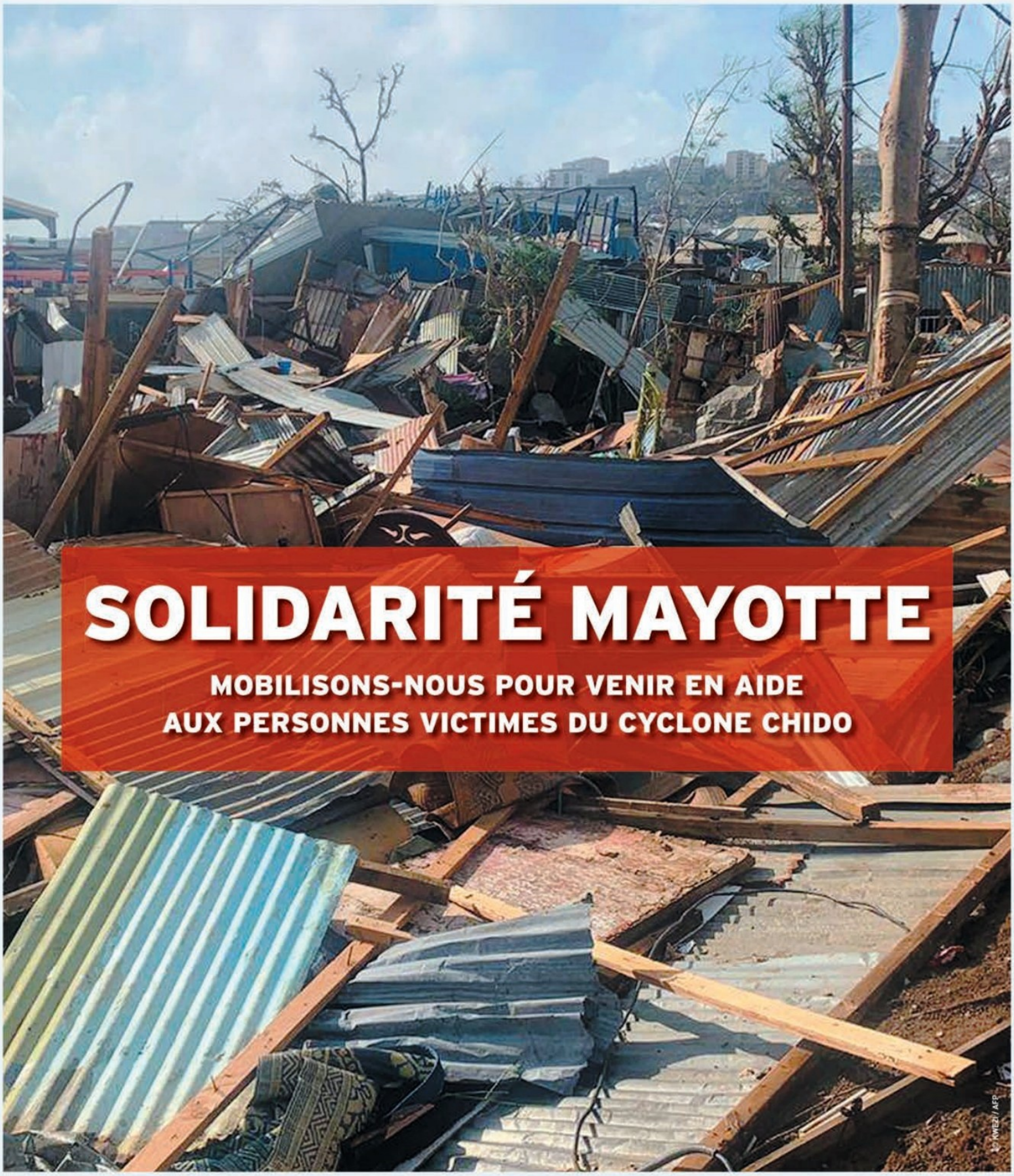
www.purestorage.com/fr/products/staas/evergreen/one.html



P 20
HARDWARE
Tendances datacenters

P 15
DOSSIER
Palmarès L'Informaticien

P 67
INFOCR
Les hauts et les bas de la cybersécurité



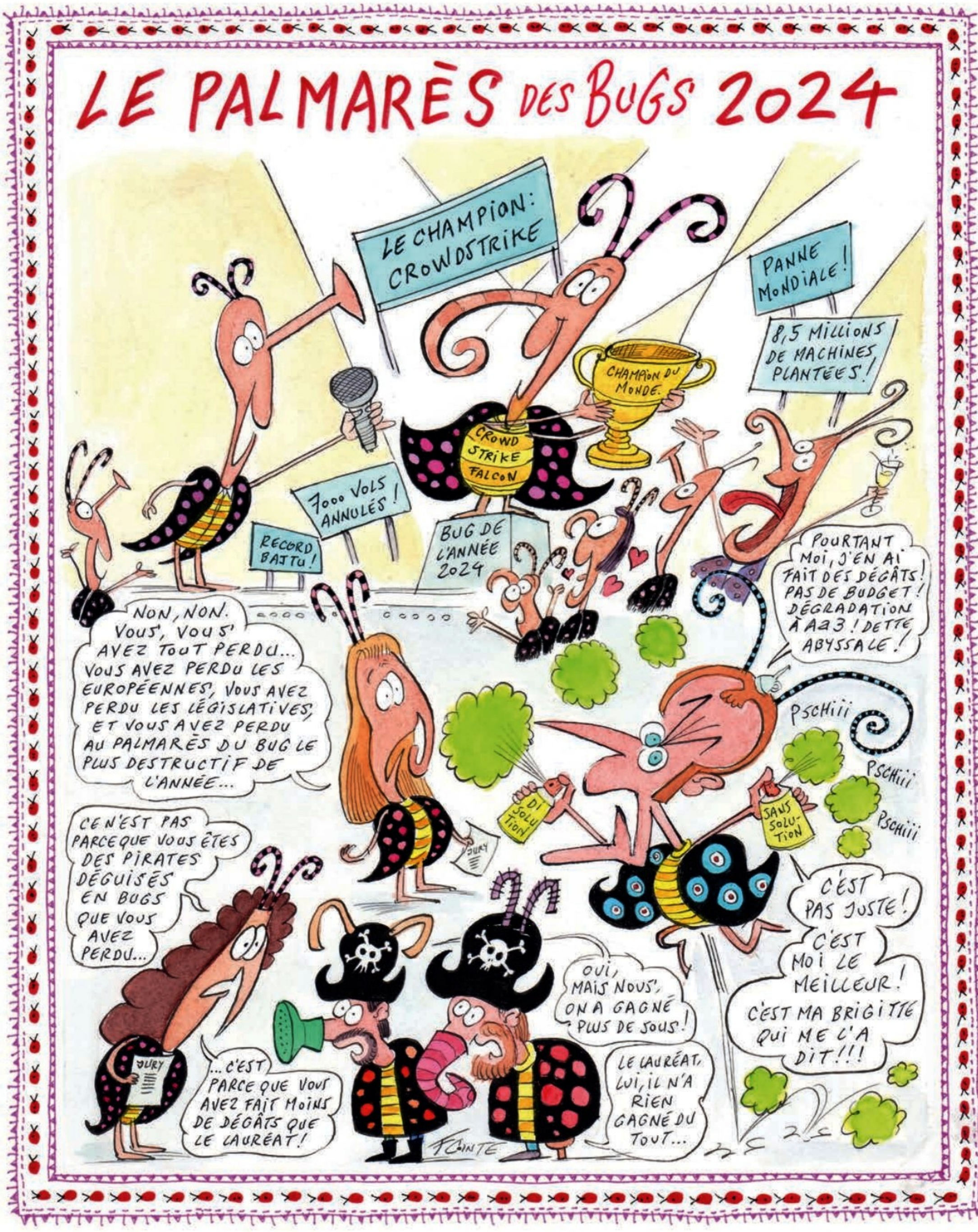
Faites un don sur fondationdefrance.org

ou par chèque à l'ordre de Fondation de France – Solidarité Mayotte -
60509 Chantilly Cedex

Les habitants de Mayotte ont été durement touchés par le cyclone qui a frappé l'archipel le 14 décembre.
La Fondation de France lance un appel à la solidarité nationale pour leur venir en aide.
Merci pour votre mobilisation !

Fondation de France

La Fondation de toutes les causes


LE PALMARÈS DES BUGS 2024
LE CHAMPION: CROWDSTRIKE
CHAMPION DU MONDE.
CROWD STRIKE FALCON
PANNE MONDIALE !
8,5 MILLIONS DE MACHINES PLANTÉES !
7000 VOLS ANNULÉS !
RECORD BATTU !
BUG DE L'ANNÉE 2024
NON, NON. VOUS', VOUS' AVEZ TOUT PERDU... VOUS AVEZ PERDU LES EUROPÉENNES, VOUS AVEZ PERDU LES LÉGISLATIVES, ET VOUS AVEZ PERDU AU PALMARÈS DU BUG LE PLUS DESTRUCTIF DE L'ANNÉE...
POURTANT MOI, J'EN AI FAIT DES DÉGÂTS ! PAS DE BUDGET ! DÉGRADATION À AA3 ! DETTE ABYSSALE !
PSCHIII
PSCHIII
PSCHIII
DI SOLU TION
SANS SOLU-TION
JURY
CE N'EST PAS PARCE QUE VOUS ÊTES DES PIRATES DÉGUISÉS EN BUGS QUE VOUS AVEZ PERDU...
C'EST PAS JUSTE !
C'EST MOI LE MEILLEUR ! C'EST MA BRIGITTE QUI ME L'A DIT !!!
OUI, MAIS NOUS', ON A GAGNÉ PLUS DE SOUS !
LE LAURÉAT, LUI, IL N'A RIEN GAGNÉ DU TOUT...
JURY
... C'EST PARCE QUE VOUS AVEZ FAIT MOINS DE DÉGÂTS QUE LE LAURÉAT !
COINTE

FOCUS

La France, hôte du prochain grand sommet sur l'IA

Le sommet pour l'action sur l'intelligence artificielle se tiendra au Grand Palais, les 10 et 11 février prochains. Le gouvernement espère que l'événement mettra en lumière une troisième voie de l'IA, tournée vers l'innovation certes, mais aussi l'éthique et la confiance.

Après Bletchley Park au (Royaume-Uni) en novembre 2023, et Séoul (Corée du sud) en mai 2024, c'est au tour de Paris de devenir l'épicentre de l'IA. Les 10 et 11 février prochains, c'est sous la verrière du Grand Palais que le sommet pour l'action sur l'intelligence artificielle accueillera plus d'une centaine de chefs d'État, des dirigeants d'organisations internationales, des représentants de petites et grandes entreprises, ainsi que des chercheurs. Sont notamment attendus Donald Trump et Elon Musk, nommé à la tête du futur ministère de l'efficacité gouvernementale, de l'administration Trump.

Trouver un équilibre entre innovation et régulation

L'événement sera articulé autour de cinq axes principaux : l'IA au service de l'intérêt général, l'avenir du travail, l'innovation et la culture, l'IA de confiance, et la gouvernance mondiale de l'IA. Le sommet doit donner un *« élan collectif qui permettra de mettre l'IA au service de chacun, de nos sociétés et de la planète »*, a déclaré Anne Bouverot, envoyée spéciale du président de la République pour le sommet. Le défi est de taille, tant les entreprises spécialisées dans le secteur, notamment américaines, se sont lancées dans une course effrénée à l'innovation, reléguant souvent ces enjeux et questionnements au second plan.

SOMMET POUR L'ACTION SUR L'IA

« Avec le président de la République et le Premier ministre, nous voulons ouvrir une troisième voie pour l'IA. Celle d'une IA éthique, durable et inclusive. Celle d'une IA de confiance. C'est à ces conditions que l'IA sera pleinement adoptée au sein de notre société, de notre économie, de nos administrations et de nos usages communs », a exprimé Clara Chappaz, secrétaire d'État chargée de l'intelligence artificielle et du numérique.

Mettre en avant des projets innovants

Bien sûr, l'enjeu de ce sommet pour le gouvernement est aussi économique, et traduit son ambition de mettre les entreprises tricolores sur le devant de la scène et de placer la France sur la carte mondiale de l'intelligence artificielle.
Dans le cadre du sommet, le gouvernement a annoncé le lancement des défis « Convergence IA », qui ont pour but de mettre en lumière des projets innovants qui s'attaquent à de grands problèmes technologiques ou sociaux, et pour lesquels l'IA peut apporter des solutions bénéfiques. Les projets devront répondre à un ou plusieurs critères, tels qu'impliquer des acteurs de différents pays, proposer des solutions audacieuses et inédites à des problèmes peu explorés (santé, écologie, travail ou mobilité), apporter une véritable valeur ajoutée pour la société et l'économie, avec un focus sur la durabilité et la résilience, rassembler un écosystème diversifié et favoriser l'inclusion.

Un appel à manifestation d'intérêt sur l'IA

Le gouvernement a également annoncé le lancement d'un appel à manifestation d'intérêt (AMI) intitulé *« IA au service de l'efficience »*, visant à favoriser l'intégration de l'IA dans les activités des entreprises, et dont les projets sélectionnés seront mis en avant lors du mommet mondial de l'IA. Cet AMI a pour but d'identifier des cas d'usage concrets de l'IA, contribuant à améliorer la compétitivité, l'innovation et la productivité. Il couvre un large éventail du tissu économique français, incluant les micro-entreprises, les PME, les ETI, les grandes entreprises, ainsi que les administrations publiques et les organisations internationales.
Les projets soumis devront être en cours de déploiement, d'industrialisation ou de passage à l'échelle, et être transférables à d'autres organisations, afin de maximiser leur impact.

FICHES PÉDAGOGIQUES

La direction générale des entreprises (DGE) a publié, sur son site internet, une série de fiches pédagogiques destinées à accompagner les entreprises dans leur utilisation de l'IA. Ces documents abordent plusieurs thèmes pratiques : L'intelligence artificielle générative : à quoi ça sert et comment ça marche ? ; Les cas d'usage de l'IA générative dans les TPE et PME ; Comment choisir parmi les solutions disponibles et les utiliser efficacement ? ; Les précautions pour un usage responsable et sécurisé de l'IA.

La Chine ouvre une enquête sur Nvidia

Nouvel épisode dans la guerre commerciale qui oppose Pékin à Washington, autour de l'industrie des semi-conducteurs. La Chine a annoncé, lundi 9 décembre, avoir ouvert une enquête sur le géant des puces d'IA Nvidia pour des violations présumées des lois antitrust du pays. La société américaine est de très loin le plus gros fournisseur de GPU, des puces indispensables à l'entraînement et à l'inférence des modèles d'IA.

Cette enquête intervient dans un contexte de tensions commerciales qui se cristallisent autour de l'industrie des semi-conducteurs entre Pékin et Washington. Les États-Unis et leurs alliés multiplient les restrictions d'exportation de puces et d'accès aux technologies nécessaires à la production de semi-conducteurs contre leur rival.

De son côté, la Chine n'hésite pas à montrer les crocs. La semaine dernière, le pays a indiqué qu'il ne délivrerait plus de licences d'exportation de terres rares stratégiques pouvant être utilisées à des fins militaires, comme le gallium, le germanium ou l'antimoine, vers les États-Unis.

Cette décision fait suite à l'ajout, lundi 2 décembre, de 140 entreprises chinoises spécialisées dans la fabrication de semi-conducteurs et les composants dédiés au calcul informatique haute performance à la liste noire des exportations soumises à accord préalable. *« Cette action représente l'aboutissement de l'approche ciblée de l'administration Biden-Harris, en concert avec nos alliés et partenaires, pour entraver la capacité de la RPC [République Populaire de Chine, ndlr] à internaliser la production de technologies avancées posant un risque pour notre sécurité nationale »*, a déclaré Gina Raimondo, secrétaire au Commerce des États-Unis, dans un communiqué.

La Chine a réagi dans la foulée, se disant « fermement » opposée à ces nouvelles restrictions. *« L'abus par les États-Unis de mesures de contrôle entrave gravement les échanges commerciaux normaux entre les pays, nuit profondément aux règles du marché et à l'ordre économique et commercial international, et menace sérieusement la stabilité des chaînes industrielles et d'approvisionnement mondiales »*, a notamment déclaré un porte-parole du ministère du Commerce.

Iliad et InfraVia veulent bâtir un leader européen des datacenters

C'est une négociation exclusive qui doit mener à la cession de 50 % du capital de l'opérateur OpCore à InfraVia, filiale de la maison mère de Free. OpCore construit et exploite des datacenters, et est implanté à Paris, Marseille, Lyon et en Pologne.

Iliad et InfraVia veulent mettre sur pied une plateforme de datacenters hyperscale « de référence », et contribuer au développement d'un écosystème « souverain » en Europe, où le marché des datacenters affiche *« une croissance annuelle estimée à plus de 20 % »* dans un contexte d'explosion des usages liés à l'intelligence artificielle et au cloud.

« L'ambition du Groupe Iliad est à la taille de l'Europe. Nous allons investir avec notre partenaire InfraVia 2,5 milliards d'euros dans notre plateforme de datacenters OpCore, pour devenir la première plateforme indépendante européenne », a déclaré Thomas Reynaud, directeur général du Groupe Iliad, cité dans un communiqué.

Via cet accord, OpCore pourra bénéficier de plus de 130 MW de capacité, notamment grâce à un datacenter de 100 MW implanté en région parisienne et dont la construction a déjà démarré. D'autres projets de plusieurs centaines de mégawatts doivent voir le jour en Europe.

OpCore poursuivra en outre le déploiement à grande échelle de ses solutions de refroidissement liquide pour les infrastructures. Une technologie qui fait office de solution d'avenir pour les infrastructures à forte intensité de traitement, telles que celles liées à l'intelligence artificielle.

Derrière cette opération, les deux partenaires entendent aussi doper le développement de Scaleway, la filiale cloud du Groupe Iliad, qui conservera son statut de client privilégié d'OpCore.

RACHATS

Cloud : **Sharp Europe** s'offre le français **Apsia**

Expert en transformation digitale et en intégration cloud, Apsia est tombé dans l'escarcelle de Sharp Europe. Apsia affiche un chiffre d'affaires de 22,5 millions d'euros en 2022 et compte 150 collaborateurs. Spécialisée dans l'implémentation de solutions cloud, elle accompagne également les entreprises de toute taille dans leur transformation digitale.
Son expertise couvre plusieurs domaines, comme le CRM, l'ERP, la cybersécurité, l'exploitation des données et de l'IA, ou encore l'audit cloud. Apsia compte parmi ses partenaires Microsoft ou encore SAP. Elle travaille pour des entreprises de secteurs d'activité divers, allant de la finance à l'industrie, en passant par la banque, la santé ou encore l'éducation.

Via cette acquisition, l'entreprise japonaise dit vouloir se renforcer sur le marché européen des services IT, et poursuit sa stratégie de développement engagée depuis plusieurs années, marquée par des investissements au Royaume-Uni et en Suisse. Apsia conservera son identité de marque en France.

TeamViewer rachète **1E** pour doper ses solutions d'assistance

Spécialisé dans les solutions de connectivité à distance, l'allemand TeamViewer a annoncé, lundi 10 décembre, avoir racheté 1E. Cette entreprise, basée à Londres et valorisée à 720 millions de dollars, développe une plateforme de gestion de l'expérience numérique des employés (DEX), qui permet d'identifier et de résoudre automatiquement des problèmes informatiques dans un environnement informatique d'entreprise.
Ce rachat s'intègre dans la stratégie de TeamViewer, visant à fournir une solution complète d'assistance informatique pour les entreprises. En combinant son expertise d'accès à distance à la plateforme 1E, l'entreprise souhaite améliorer la prévention proactive des problèmes informatiques et le support à distance. Cette acquisition lui permettra en outre d'enrichir son offre pour ses 4 500 clients grands comptes, et lui ouvrira les portes du marché des PME.
TeamViewer compte étendre sa présence en Amérique du nord et ambitionne de déployer ses solutions DEX en Europe, au Moyen-Orient, en Afrique et en Asie-Pacifique. L'entreprise espère boucler le rachat début 2025, sous réserve des approbations réglementaires.

Capgemini finalise le rachat de **Syniti**

Syniti, spécialiste des services sur logiciels de gestion des données d'entreprise, emploie près de 1 200 personnes. Ces effectifs vont désormais renforcer les équipes de Capgemini en matière de transformation digitale fondée sur les données, en particulier pour les projets à grande échelle basés sur SAP, tels que les migrations vers SAP S/4HANA.
La société dispose également d'une forte expertise dans plusieurs secteurs d'activité, notamment les sciences de la vie, l'aéronautique, la défense, l'industrie, les produits de grande consommation, la distribution et l'automobile. Elle accompagne ses clients dans leurs projets stratégiques de transformation, tels que les migrations complexes de données ERP, les activités de consolidation (comme les migrations vers SAP S/4HANA ou les migrations cloud), ainsi que dans le cadre de fusions, acquisitions et cessions, et pour la mise en conformité des données.

Atos vend **Worldgrid** et boucle une augmentation de capital

Atos a finalisé la vente de son activité Worldgrid à Alten. Pour rappel, Worldgrid fournit des services d'ingénierie pour les entreprises du secteur de l'énergie et des services publics, en France, en Allemagne et en Espagne. Elle conçoit notamment des systèmes de pilotage pour les centrales nucléaires. Elle emploie 1 100 personnes et réalise un chiffre d'affaires de 170 millions d'euros. Alten, entreprise d'ingénierie et fournisseur de services IT, a annoncé qu'elle garantirait la continuité des services pour les clients stratégiques, ainsi que pour les collaborateurs de Worldgrid.
Annoncée début novembre, la vente de Worldgrid a été conclue pour 270 millions d'euros et permettra de réduire la dette d'Atos d'une somme à peu près équivalente. En difficulté, le groupe français a enregistré une baisse de 11% de son activité au troisième trimestre et continue de se séparer de ses activités non stratégiques pour tenter de redresser la

barre. Dans le cadre de sa restructuration, le groupe a réussi à boucler son augmentation de capital de 233 millions d'euros, grâce au soutien de ses créanciers.

Stockage moléculaire : **Biomemory** lève **17 millions d'euros**

Spécialisée dans le stockage de données sur ADN, Biomemory a bouclé un tour de table en série A de 17 millions d'euros. L'opération a été menée par Crédit Mutuel Innovation, avec la participation du fonds French Tech Seed de Bpifrance, Paris Business Angels, Sorbonne Venture by Audacia & Aloe Private Equity, Adnexus, Prunay, Next Sequence et Accelerem.
L'entreprise utilisera ces fonds pour soutenir sa R&D et finaliser la première génération de son système, optimisé par des procédés biotechnologiques. La technologie de stockage de Biomemory confère une meilleure densité de données, permettant de stocker *« toutes les données de l'humanité dans l'espace d'une seule baie 19 pouces »*, promet un communiqué de Bpifrance. De quoi réduire considérablement l'occupation spatiale et l'empreinte carbone des centres de données.
L'entreprise souhaite étendre sa technologie de stockage à l'échelle de l'exaoctet d'ici 2030, pour une utilisation en centre de données. Biomemory entend également renforcer ses effectifs et signer de nouveaux partenariats avec des acteurs de l'industrie et des fournisseurs cloud.

Tenstorrent lève **693 millions de dollars** pour ses technologies d'IA

Tenstorrent, société spécialisée dans les ordinateurs et puces pour l'intelligence artificielle, a annoncé avoir bouclé un financement de série D à hauteur de 693 millions de dollars.
L'opération a été dirigée par Samsung Securities et AFW Partners, principaux investisseurs, avec la participation de nombreux autres, dont XTX Markets, Corner Capital, Protagonist, MESH, Exportation et Développement Canada, le Régime de retraite des employés municipaux de l'Ontario, LG Technology Ventures, Hyundai Motor Group, Fidelity Management & Research Company, Innovation Engine, Baillie Gifford, Bezos Expeditions, et bien d'autres.
Tenstorrent se concentre sur le développement d'ordinateurs et de puces d'intelligence artificielle équipés de ses cœurs Tensix. La société propose également des piles logicielles open source pour l'IA, ainsi que des licences d'IP en IA et en RISC-V, permettant à ses clients de personnaliser leurs solutions.

Avec cette nouvelle levée de fonds, l'entreprise prévoit de renforcer ses piles logicielles, de construire des systèmes et des clouds destinés aux développeurs d'IA, de recruter davantage de développeurs et d'élargir son réseau de centres de conception à l'international.

Jotelulu lève **6,8 M€** pour sa plateforme de services cloud

Petit acteur du cloud basé en Espagne, Jotelulu a annoncé avoir bouclé un tour de table de 6,8 millions d'euros. À la manœuvre, Kibo Ventures, un fonds de capital-risque espagnol, avec la participation de ses investisseurs historiques Adara Ventures, Bankinter, G2A et Big Sur Ventures.
Jotelulu développe une plateforme de services cloud pour les PME. Elle fournit des ressources en remote desktop, mais aussi en calcul, en stockage ou encore en plans de reprise après incident. Elle compte des centaines de clients dans les domaines du retail, de la construction, du conseil, de l'énergie, du droit ou encore des services, et revendique quelque 40 000 utilisateurs en Europe.
Actuellement, Jotelulu est présente en Espagne, en France et au Portugal. Avec ces nouveaux fonds, elle ambitionne de soutenir son expansion internationale, en Europe et en Amérique latine.

Liquid AI lève **250 millions de dollars**

Liquid AI a bouclé un tour de table de 250 millions de dollars en série A. Premier partenaire de l'entreprise : le fondeur AMD.
Via cette opération, Liquid AI veut doper le développement et le déploiement de ses petits modèles d'IA dédiés aux tâches spécifiques en entreprise. *« Nous sommes fiers que nos nouveaux partenaires de pointe dans l'industrie croient en notre mission ; ensemble, nous prévoyons de libérer des expériences d'IA souveraines pour les entreprises et les utilisateurs »*, a déclaré le PDG de la startup, Ramin Hasani.
L'entreprise va, en outre, développer son infrastructure de calcul, préparer ses produits pour l'edge et le déploiement sur site, et optimiser les piles d'inférence et d'ajustement de ses modèles.
Partenaire stratégique privilégié de Liquid AI, AMD a participé à cette nouvelle levée de fonds. C'est en outre AMD qui fournit les GPU Instinct sur lesquels les modèles de Liquid AI sont entraînés.

Atempo et Eviden renforcent leur partenariat

L'éditeur de solutions de gestion et de protection des données s'allie à Eviden afin de proposer une solution souveraine pour la protection des données hautement sensibles, particulièrement conçue pour les environnements informatiques hybrides.

Cette collaboration repose sur l'intégration du module de sécurité matérielle HSM (Hardware Security Module) Trustway Proteccio d'Eviden, avec les solutions phares d'Atempo : Tina pour la sauvegarde et la restauration, et Miria pour l'archivage ou la migration de grands volumes de données. Grâce au HSM Trustway Proteccio d'Eviden, seul HSM détenteur de la qualification renforcée de l'ANSSI (Agence nationale de la sécurité des systèmes d'information), ce partenariat permet aux solutions logicielles d'Atempo de se conformer aux normes les plus strictes en matière de sécurité, par l'usage d'une couche avancée de chiffrement pour protéger les données en transit, et au repos.

Mint devient une référence pour Celonis

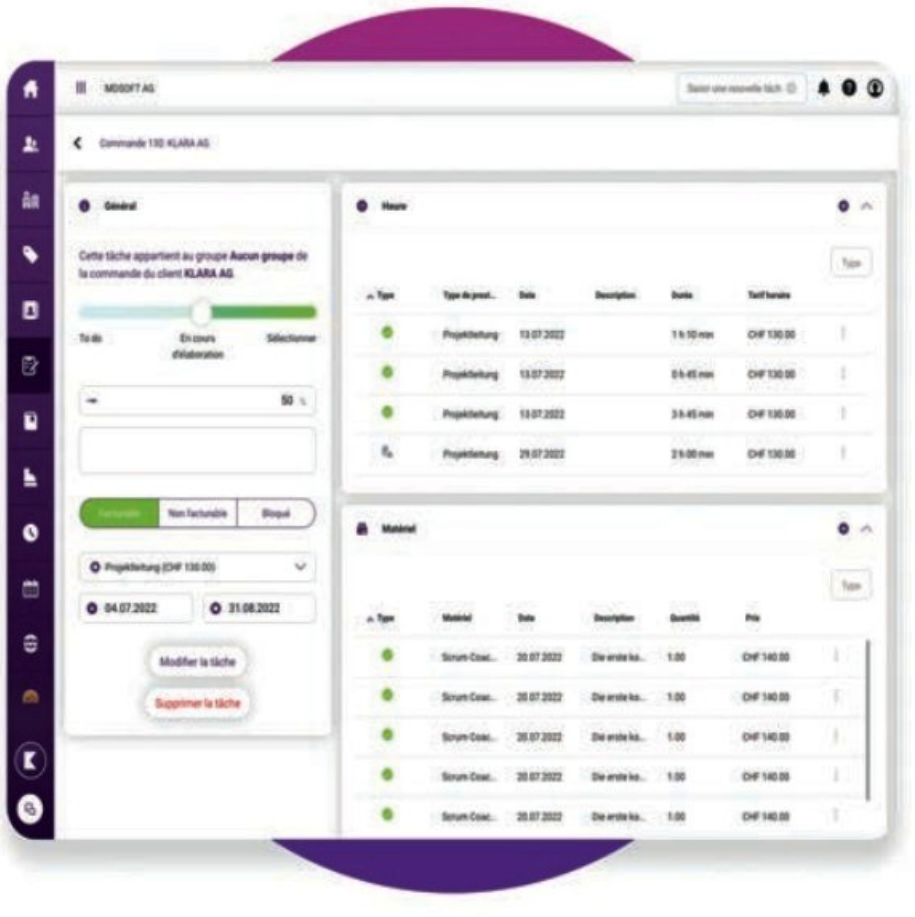
Cette alliance fait de Mint, un éditeur de solutions de gestion des ressources publicitaires, un partenaire privilégié de Celonis pour ce secteur d'activité.

Dans le cadre de cette collaboration, Mint deviendra la plateforme de référence pour les solutions Celonis dans le secteur du marketing et des médias. Celonis a déjà ce type de partenariats avec d'autres éditeurs pour d'autres verticaux. Mint va, pour sa part, intégrer les fonctionnalités de process mining de Celonis dans son logiciel. Avec cette nouvelle solution, les entreprises pourront identifier les inefficacités opérationnelles liées à la planification et à l'achat média, simplifier les processus ainsi que la collaboration entre équipes internes et agences, accélérer les validations et l'activation des campagnes, tout en réduisant les erreurs et les tâches manuelles dans la gestion de leurs campagnes publicitaires, comme les saisies incorrectes de données ou les mauvaises allocations budgétaires.

La plateforme révèle également des opportunités d'optimisation qui génèrent des gains financiers, stratégiques et environnementaux. Ce partenariat permettra à Mint d'aider les organisations à optimiser leurs workflows publicitaires, gérer leurs budgets de manière plus efficace, de maximiser leurs performances médias et d'améliorer leur ROI global.

Klara embarque l'IA de Toucan

Klara est un éditeur qui accompagne les entreprises dans le management du développement des compétences des collaborateurs bureau et terrain, pour booster la performance de l'entreprise via une plateforme en ligne. Pour répondre aux besoins de ses clients grands comptes, Klara intègre Toucan.



Pour répondre aux besoins spécifiques des grands groupes en matière d'insights personnalisés et de tableaux de bord accessibles, Klara a choisi d'intégrer une solution d'analytics performante et facilement déployable. Ce déploiement concerne une soixantaine de grands comptes, dont Carrefour, Safran, Banque Populaire Rives de Paris et CNP Assurances, et vise à renforcer l'efficacité des prises de décision en fonction de l'usage de Klara. La plateforme propose maintenant 40 tableaux de bord personnalisés, adaptés aux différents utilisateurs, des administrateurs aux utilisateurs quotidiens. La plateforme no-code de Toucan a permis de les embarquer dans la plateforme rapidement et sans complexité technique, évitant ainsi à Klara une construction en interne qui serait plus coûteuse et moins performante, pour se concentrer sur son cœur de métier. Cette mise en place a permis une hausse de 30 % des scores de satisfaction client, soulignant l'impact positif de l'analytics intégré de Toucan sur la qualité de service de Klara. Pour Toucan, cette intégration est dans la droite ligne de sa nouvelle stratégie qui se concentre sur la proposition de sa solution en marque blanche pour les éditeurs et les startups.

ESET s'intègre avec **OpenCTI**

ESET et Filigran se rapprochent dans un partenariat qui autorise l'ingestion des flux de Threat Intelligence d'ESET dans OpenCTI, l'outil open source de Filigran.

La solution présente plusieurs avantages. A partir de son vaste réseau de terminaux protégés, les chercheurs ESET produisent à la fois des renseignements (Cyber Threat Intelligence) en temps réel, ainsi que des analyses détaillées, permettant une détection et une réponse précises aux cyber-attaques. Les données fournies par ESET enrichissent les analyses par des éléments de contexte et permettent une détection plus précoce des menaces. Cela soutien les analystes dans leurs taches d'identification et améliore le traitement des menaces. Cette intégration prise en charge par les éditeurs pérennise l'interaction entre les renseignements d'ESET et les outils d'analyse d'OpenCTI. L'utilisation des standards TAXII 2.1 et STIX 2.1 par ESET garantit un partage fluide des données, et optimise les processus de réponse aux menaces. ESET fournit des flux de données structurés qui peuvent être directement exploités dans OpenCTI.

Partenariat **Flaminem** Neo4j

L'éditeur de solutions à destination des responsables de la conformité dans le secteur financier s'associe à la base de graphe de Neo4j.

Flaminem a choisi Neo4j pour remplacer les bases relationnelles classiques, incapables de modéliser efficacement les réseaux complexes d'actionnariat. Grâce à Neo4j, Flaminem propose une solution connaissance des clients qui évalue le risque en temps réel, permet des personnalisations pour s'adapter aux besoins du client, avec des performances élevées par le clustering et la cohérence causale. En 2024, Flaminem a intégré Neo4j Bloom pour offrir des visualisations de données plus poussées, facilitant l'identification des anomalies et des réseaux de fraude. Cette interface graphique intuitive accélère l'analyse et offre une valeur ajoutée considérable à ses clients.

Gcore multiplie les partenariats

Le fournisseur de solutions d'intelligence artificielle pour les environnements de périphérie signent plusieurs partenariats stratégiques.

Avec LightOn, Gcore vise à simplifier le déploiement des projets d'IA générative à grande échelle, quelles que soient les contraintes géographiques, grâce à l'alliance entre la plateforme Paradigm de LightOn et l'infrastructure IA Edge mondiale de Gcore. Paradigm peut être mis en œuvre en quelques instants et les utilisateurs locaux pourront l'utiliser de manière optimale. Les clients s'affranchissent également de la contrainte de déployer l'infrastructure localement et de la manager. Cela permet d'économiser des ressources supplémentaires à la fois au niveau humain mais aussi sur les infrastructures. Le second partenariat avec UltraEdge vise à renforcer la couverture territoriale des services IA cloud et edge computing en France. UltraEdge met à disposition de Gcore son réseau de plus de 250 sites répartis sur le territoire national, dont six NetCenters stratégiques situés à Courbevoie, Lyon, Lille, Bordeaux, Nantes et Strasbourg. Cette infrastructure permettra à Gcore de déployer ses services cloud IA au plus près des utilisateurs finaux. L'objectif est de garantir un temps de latence minimal et une expérience utilisateur optimale. Cette alliance stratégique permet notamment à Gcore de déployer ses solutions d'IA de manière uniforme sur tout le territoire. Ainsi qu'une accessibilité et une performance optimales pour tous les utilisateurs, indépendamment de leur localisation. Le partenariat permettra aux entreprises françaises d'accéder à une infrastructure de nouvelle génération intégrant des serveurs d'IA sophistiqués alimentés par des GPU NVIDIA. Cette alliance stratégique bénéficiera particulièrement aux secteurs des médias, du divertissement, des jeux, de la technologie, des services financiers et du commerce de détail.

AGENDA

CES
7-10 janvier 2025
Différents sites
Las Vegas, USA

World Economic Forum
20-24 janvier 2025
Différents sites
Davos, Suisse

Université des DPO
6-7 février 2025
Maison de la Chimie
Paris

WAICF
13-15 février 2025
Palais des Festivals
Cannes

Web Summit Qatar
23-26 février 2025
Doha Exhibition and Convention Center — Doha, Qatar

MWC
3-6 mars 2025
Fira Gran Via
Barcelone Espagne

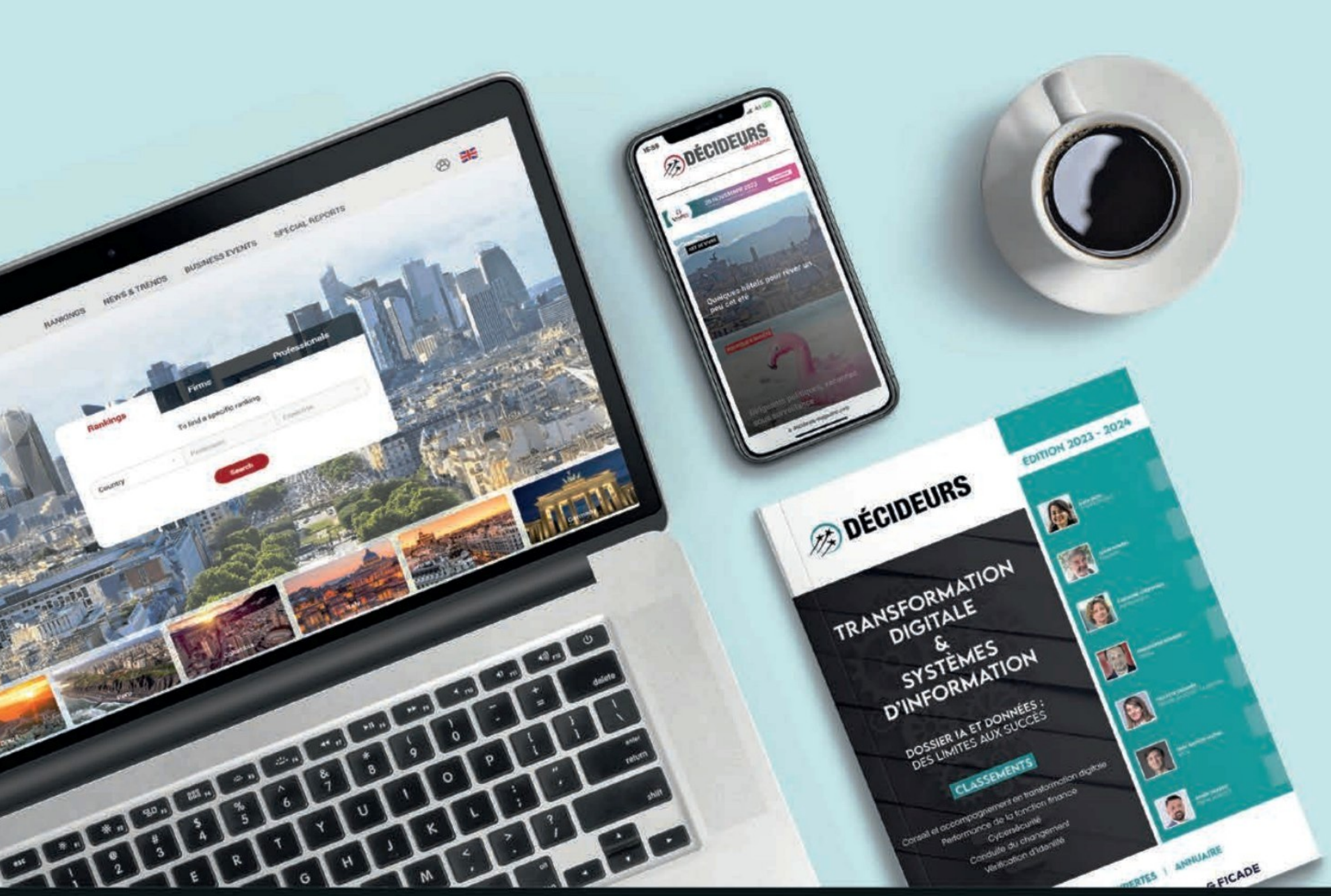
Nvidia GTC
17-20 mars 2025
Convention Center
San Jose, USA

Adobe Summit
18-20 mars 2025
Venetian Convention and Expo Center — Las Vegas, USA

DÉCIDEURS

L'information stratégique pour bien choisir vos partenaires

Conseil et accompagnement en transformation digitale | Performance de la fonction finance | Cybersécurité
Conduite du changement | Vérification d'identité



COMMANDER LE GUIDE

TRANSFORMATION DIGITALE & SYSTÈMES D'INFORMATION

DÉCIDEURS AGENCE DE NOTATION

FICADE

PALMARÈS 2024

UN ÉVÉNEMENT INSTALLÉ COMME UN VÉRITABLE RENDEZ-VOUS DANS L'IT

Cette nouvelle édition du Palmarès confirme l'importance et la place que prend désormais notre enquête annuelle sur les différentes familles de l'IT. Avec plus de 6000 votes comptabilisés, le Palmarès est devenu un véritable baromètre de l'utilisation des différents produits IT dans les entreprises.

Premier enseignement de cette édition 2024, le retour en force des grands noms de l'industrie qui ont clairement fait l'effort de participer pour se retrouver en haut des classements. Autre innovation de cette vague, le prix de la rédaction a aussi permis de nous exprimer sur les matériels et logiciels qui ont marqué notre esprit durant l'année écoulée. Nous avons, de plus, pu mettre en avant nos coups de cœurs.

Bonne année et meilleurs vœux pour 2025 en espérant vous retrouvez encore plus nombreux pour l'édition 2025 de ce Palmarès.

splunk>
a CISCO company

Depuis sa création en 2003, Splunk aide les entreprises à découvrir toute la profondeur de leurs données comme les spéléologues explorent les grottes. C'est d'ailleurs de cette activité, « spelunking » en anglais, que vient le nom de Splunk.

L'objectif de Splunk est de construire un monde numérique plus sûr et plus résilient. Chaque jour, Splunk donne corps à cette vision en aidant les équipes de ses clients à préserver le fonctionnement et la sécurité de leur organisation.

Les grandes entreprises utilisent cette plateforme unifiée de sécurité et d'observabilité pour assurer la sécurité et la fiabilité de leurs systèmes numériques.

En 2024, Splunk a été racheté par Cisco pour poursuivre sa mission : aider les clients à renforcer leur résilience sur l'ensemble de leur empreinte numérique.

Les organisations font confiance à Splunk pour éviter que les incidents d'infrastructure, d'application et de sécurité ne deviennent des problèmes majeurs, pour se protéger des menaces, pour se remettre plus rapidement des perturbations et pour saisir les nouvelles opportunités qui se présentent.

Splunk aide les équipes SecOps, ITOps et d'ingénierie à obtenir ces résultats grâce à une visibilité complète, à des délais de détection et d'investigation rapides, et à des processus de réponse optimisés, le tout soutenu par l'IA et à l'échelle nécessaire pour les plus grandes organisations.

Web : www.splunk.com

Linkedin : www.linkedin.com/company/splunk

NUTANIX

Nutanix est l'un des leaders mondiaux des logiciels de cloud computing, offrant aux organisations une plateforme unique pour faire fonctionner les applications et les données à travers les clouds. Avec Nutanix, les entreprises peuvent réduire la complexité et simplifier les opérations, ce qui leur permet de se concentrer sur les résultats de leur activité. S'appuyant sur son héritage en tant que pionnier de l'infrastructure hyperconvergée, Nutanix a la confiance des entreprises du monde entier pour alimenter les environnements hybrides multicloud de manière cohérente, simple et rentable.

Web : www.nutanix.fr

Linkedin : www.linkedin.com/company/nutanix

Open-Prod

Open-Prod est une solution ERP de dernière génération, conçue pour répondre aux besoins spécifiques des entreprises industrielles. La solution offre une large gamme de fonctionnalités pour gérer l'ensemble des processus de l'entreprise, de la production à la distribution. Open-Prod est flexible, adaptable et évolutif, ce qui en fait la solution idéale pour les PMI-ETI qui souhaitent optimiser leurs opérations et gagner en efficacité.

La nouvelle version majeure d'Open-Prod est prévue pour début 2025. Elle apportera une ergonomie repensée, des fonctionnalités améliorées, et une plateforme technique renouvelée, offrant ainsi une expérience utilisateur et des performances accrues. De plus, la solution est open-source, ce qui constitue un élément essentiel pour une gouvernance des données transparente et souveraine.

Open-Prod s'est imposée comme une solution de GPAO de référence chez les industriels, en particulier dans les secteurs de la maroquinerie, de la plasturgie, et des machines spéciales. Plus de 250 industriels utilisent chaque jour la solution et apprécient sa flexibilité pour optimiser leur processus.

Web : www.open-prod.com

Linkedin : www.linkedin.com/company/erp-open-prod

VERTIV™

Fournisseur mondial de solutions de continuité et d'infrastructures numériques critiques, Vertiv propose un ensemble de matériels, de logiciels et de services conçues pour permettre un fonctionnement optimal et sans interruption des applications informatiques critiques.

En 2024, Vertiv a dévoilé son portefeuille VertivTM 360AI, un ensemble de solutions innovantes pour soutenir l'adoption croissante de l'IA et du HPC, afin de répondre à l'augmentation de la densité informatique et contribuer à améliorer l'efficacité énergétique des infrastructures IT et data centers. Grâce à son portefeuille de solutions d'alimentation et de refroidissement le plus complet du secteur, Vertiv accompagne les entreprises dans le déploiement des infrastructures d'IA et de HPC, aussi bien dans les data centers d'entreprise que les sites de colocation et les sites d'hébergement cloud.

Ses partenariats étroits avec les constructeurs informatiques et les fabricants de puces leaders dans le monde ont permis à Vertiv de concevoir des architectures de références pour les infrastructures informatiques d'IA et HPC, permettant des déploiements rapides, tout en réduisant le temps de conception. Récemment, Vertiv et Nvidia ont collaboré pour le co-développement d'une architecture de référence d'alimentation et de refroidissement complète de 7 MW pour la plate-forme Nvidia GB200 NVL72, permettant de transformer les architectures de data centers en « usines d'IA ». Parmi les lancements marquants, le VertivTM AI Hub, un guide de déploiement pour les clusters d'IA comprenant des conceptions de référence d'infrastructure complètes pour les principaux chipsets GPU.

Web : www.vertiv.com

Linkedin : www.linkedin.com/company/vertiv

Catégorie	Sous-catégorie	1er prix	Prix de la rédaction
Applicatifs	Logiciels ITOM / ITSM	ServiceNow	SolarWinds
Applicatifs	Observabilité et monitoring	Splunk	Centreon
Applicatifs	Logiciel CRM	Salesforce	HubSpot
Applicatifs	Marketing Digital	Plezi	Hootsuite
Applicatifs	Logiciel finance/comptabilité	NetSuite	Finastra
Applicatifs	Logiciel Supply Chain	SAP	Kinaxis
Applicatifs	Logiciel RH	Workday	Lucca
Applicatifs	Logiciel ERP	Open-Prod	IFS Cloud
Applicatifs	IA/Analytics	Mistral AI	Tableau
Applicatifs	Solution de partage de fichiers	Google Drive	Nextcloud
Applicatifs	Gestion de l'Information	Microsoft SharePoint	OpenText
Applicatifs	Solution de gestion de contenus /documentaire	Notion	Obsidian
Applicatifs	Gestion de projet/collaboration	Jamespot	Asana
Applicatifs	Solution de communication unifiée	Zoom	Jitsi
Applicatifs	Virtualisation	Nutanix	Proxmox
Sécurité	SIEM	Splunk	Tenable
Sécurité	EDR	HarfangLab	CrowdStrike
Sécurité	NDR	Custocy	Gatewatcher
Sécurité	XDR	Sekoia.io	Cato Networks
Sécurité	IAM / PAM	Memority	Okta
Sécurité	Gestion de vulnérabilités	Tenable	Qualys
Sécurité	Logiciels de protection des mails	Vade	Cleanmail
Sécurité	Logiciels de protection des données	Rubrik	Zama
Sécurité	Threat intelligence	Cato Networks	Mandiant
Sécurité	Solution IPS / IDS	Gatewatcher	Palo Alto
Sécurité	WAF / WAAP	Ubika	Imperva
Sécurité	Firewall	Fortinet	Stormshield
Sécurité	Matériel de passerelle sécurisée / VPN	Cisco	WireGuard
Sécurité	SOAR	Sekoia.io	IBM
Sécurité	Plateforme de bug bounty	YesWeHack	Yogosha
Sécurité	Logiciels anti DDOS	6Cure	Cloudflare
Sécurité	Sensibilisation	DriveLock	MailinBlack

Catégorie	Sous-catégorie	1er prix	Prix de la rédaction
Hardware	Baies stockage	Pure Storage	Atempo
Hardware	Serveurs	Dell	HPE
Hardware	Postes de travail	Dell	Asus
Hardware	Téléphonie d'entreprise	Zoom	Adista
Hardware	Systèmes de visio-conférence	Zoom	Poly
Hardware	Mobilité (tablette, smartphone)	Samsung	Xiaomi
Hardware	Imprimantes	Epson	HP inc.
Hardware	Périphériques et accessoires	Logitech	Jabra
Hardware	Infrastructures critiques Data Center	Vertiv	Schneider
Réseau	Logiciel de monitoring	EasyVista	Paessler
Réseau	Routeurs / Switch	Extreme Networks	Netgear
Réseau	Bornes Wifi	HPE Aruba	TP-Link
Dev/DevOps	Solution CI/CD	GitLab	Jenkins
Dev/DevOps	Plateforme low-code/no-code	Zapier	Odoo
Dev/DevOps	Infra as code	Ansible	Terraform
Cloud	Cloud public	OVH	Outscale
Cloud	Cloud privé	Nutanix	Canonical
Cloud	Opérateur	S3NS	Bleu
Cloud	Collocation	Data4	Equinix
Cloud	Backup	Rubrik	Commvault
Cloud	Hyperconvergence	Nutanix	Scale Computing
Data	Datalakes / data warehouses	Rivery	Starburst
Data	SGBD relationnels	PostgreSQL	Snowflake
Data	SGBD NoSQL	MongoDB	Redis
Data	Datascience et machine learning	AWS SageMaker	Databricks
Data	Solution d'archivage	Quantum	Docaposte

Marché

Tech Show Paris 2024 : un rendez-vous incontournable pour l'industrie des datacenters

Le Tech Show Paris 2024, qui s'est tenu les 27 et 28 novembre à la Porte de Versailles, a réuni les acteurs majeurs de l'industrie des datacenters autour de cinq salons co-organisés : Cloud Expo Europe, DevOps Live, Cloud & Cyber Security Expo, Data & AI Leaders Summit et Data Centre World. Une occasion unique pour prendre le pouls d'un secteur essentiel en plein essor.

Ces deux jours intenses ont permis aux visiteurs de plonger au cœur d'un secteur stratégique en pleine expansion. Avec des thématiques aussi variées que le cloud computing, le développement logiciel, la cybersécurité, l'intelligence artificielle et les centres de données, l'événement s'est affirmé comme une véritable plateforme d'échanges et d'innovation pour les professionnels de la tech.

Présentée sur le salon, l'Immersion Cooling est une technologie innovante de refroidissement qui consiste à immerger les composants informatiques dans un liquide non conducteur.

Une affluence record

Cette édition 2024 a marqué un tournant avec une affluence impressionnante : plus de 10 700 visiteurs, soit près du double par rapport à l'année précédente. L'enthousiasme autour de l'événement témoigne de l'importance croissante des datacenters dans un monde de plus en plus numérique. Ce succès s'explique également par la qualité et la diversité des intervenants, avec 290 conférenciers de renom venus partager leur expertise, ainsi que 270 exposants présentant leurs dernières solutions technologiques. Longtemps considérés comme de simples infrastructures techniques, les datacenters se redéfinissent aujourd'hui comme des piliers stratégiques de l'économie numérique. Ce changement de paradigmes est porté par plusieurs facteurs : l'explosion des données, les technologies d'intelligence artificielle, la durabilité et l'efficacité énergétique, ou encore l'avènement des datacenters Edge.

L'IA au cœur des transformations des datacenters

L'intelligence artificielle (IA), en particulier, transforme profondément la manière dont les datacenters sont conçus, gérés et exploités. *« L'intelligence artificielle influence déjà certaines de nos pratiques, notamment en matière de modélisation 3D (BIM) et de design technique. Notre cœur de métier demeure toutefois fondamentalement centré sur l'ingénierie et la construction. Là où l'IA exercera un impact véritablement transformateur, c'est dans la réinvention des modèles économiques et l'évolution des technologies associées aux datacenters. L'émergence des supercalculateurs, combinée à l'augmentation exponentielle des capacités de calcul, alimente une croissance fulgurante des besoins énergétiques »*, explique Olivier Piquart, directeur commercial chez Cap Ingelec.

Des algorithmes sophistiqués sont désormais capables de gérer automatiquement la répartition des charges de travail en fonction des fluctuations de la demande, ou de prédire les pannes matérielles avant qu'elles ne surviennent. *« Nous exploitons l'intelligence artificielle pour optimiser la maintenance prédictive. En anticipant les problèmes avant qu'ils ne surviennent, l'IA nous permet de réduire drastiquement les coûts de maintenance tout en limitant de façon significative les pannes. De plus, cette approche contribue à diminuer la consommation énergétique, car elle détecte les équipements en surchauffe, en surrégime ou opérants dans des conditions non conformes aux paramètres définis pour nos clients. Grâce à ces capacités d'analyse avancées, les ajustements nécessaires sont réalisés presque automatiquement »*, souligne Sami Slim, PDG de Telehouse France.

Efficacité énergétique et durabilité

Selon l'Agence internationale de l'énergie, les datacenters ont consommé entre 2 et 3 % de l'électricité mondiale en 2023. Bien que déjà significative, cette part est appelée à croître de manière exponentielle dans les années à venir. De ce fait, la consommation d'énergie des datacenters constitue un enjeu majeur qui pousse les acteurs du secteur à développer des infrastructures moins énergivores et plus respectueuses de l'environnement. *« L'émergence de l'intelligence artificielle et de l'hyperconvergence entraîne une hausse continue des besoins en densité énergétique. Alors qu'une baie consommait 10 kW il y a peu, les modèles actuels atteignent désormais 100 kW. Ces changements imposent une refonte des infrastructures, notamment en ce qui concerne les alimentations électriques et les systèmes de refroidissement »*, ajoute Olivier Piquart. Pour maîtriser la consommation énergétique, les opérateurs de datacenters se tournent massivement vers des technologies de refroidissement innovantes, telles que le Direct Liquid Cooling et l'Immersion Cooling. Ces solutions qui remplacent les systèmes de refroidissement traditionnels à air dissipent la chaleur de manière beaucoup plus efficace, en faisant respectivement circuler un fluide non conducteur à proximité des CPU/GPU pour absorber la chaleur, ou en immergeant complètement les serveurs dans un liquide diélectrique non conducteur.

Pour Sami Slim, CEO de Telehouse France, les nouvelles technologies de refroidissement deviennent incontournables : *« le refroidissement liquide représente une révolution technologique dans la gestion thermique des infrastructures. En substituant l'air par un fluide, cette approche impose une refonte complète des conceptions, l'acquisition de compétences spécifiques et la mise en place d'infrastructures adaptées, notamment pour le traitement et la gestion de l'eau. Bien qu'elle offre des avantages considérables en matière d'efficacité énergétique et de durabilité, cette innovation requiert une transformation profonde des pratiques actuelles »*, poursuit le PDG de Telehouse France.

Thomas Hombert,
PDG d'Etix Everywhere

« L'approche Edge consiste à décentraliser les grands datacenters traditionnels pour s'implanter dans des villes secondaires ou en périphérie du réseau, là où les clients recherchent souvent une plus grande proximité avec leurs infrastructures. Cette stratégie peut s'adresser aux collectivités locales ou aux entreprises désireuses de maintenir leurs serveurs à proximité de leurs activités »

L'avènement de l'Edge computing

Le paradigme centralisé des datacenters traditionnels est en pleine mutation sous l'impulsion de l'Edge Computing, une tendance qui transforme radicalement l'approche de traitement et de stockage des données. Alors que l'Internet des objets (IoT) se développe à une vitesse exponentielle, et que les applications critiques nécessitent des temps de réponse ultras rapides, la gestion centralisée des données dans de gigantesques datacenters éloignés montre ses limites. Les datacenters Edge ou de proximité s'imposent ainsi comme une solution clé pour répondre à ces nouveaux défis.

Avec ce modèle, les données sont traitées au plus près des utilisateurs et des terminaux. Cela permet de réduire considérablement la latence, tout en garantissant des performances optimales. Pour répondre aux exigences croissantes de réactivité dans des secteurs tels que l'IoT ou les applications en temps réel, les micro datacenters, compacts et modulaires, rencontrent un énorme succès. Ils peuvent être déployés très rapidement, que ce soit dans des zones urbaines ou dans des régions éloignées et difficiles d'accès. *« Nos micro datacenters sont fréquemment implantés en dehors des centres-villes pour répondre aux besoins des clients en quête d'une infrastructure locale. Cette localisation permet de réduire la latence et d'assurer une continuité d'activité en cas de rupture de connexion externe. Ces solutions se révèlent particulièrement adaptées aux exigences des industries et des chaînes de production »*, explique Béranger Cadoret, responsable grands comptes chez Grolleau.

Un événement incontournable

Le Tech Show Paris 2024 a confirmé son rôle de rendez-vous incontournable pour les professionnels des datacenters et des technologies associées. Cet événement a non seulement mis en lumière les tendances actuelles du secteur, mais il a également souligné l'importance croissante des datacenters pour l'économie numérique. □

J.C

Marché Des exemples récents

Concrètement, les grandes tendances décrites dans l'article précédent voient leur application dans les offres et annonces des différents acteurs autour des centres de données. En voici quelques exemples récents.

UltraEdge, le nouveau nom des centres de données de SFR cédés à Morgan Stanley.

L'efficacité énergétique et la durabilité sont des questions clés pour ce secteur. Si le refroidissement par eau avance rapidement, les innovations se multiplient dans ce domaine. Ainsi Vertiv, le fournisseur de solutions de protection des actifs critiques des centres de données, travaillent conjointement avec le fournisseur d'espace de centres de données sur un refroidissement combinant liquide et air. La solution autorise le passage d'un refroidissement par air à un refroidissement liquide afin de prendre en charge l'informatique à haute densité. Les ingénieurs de Vertiv et de Compass ont collaboré à cette vision pour une solution de refroidissement d'avenir, avec Vertiv développant et fabriquant la solution technologique. Les premières unités seront déployées sur un site de Compass au premier trimestre 2025 dans le cadre d'un accord de fourniture pluriannuel de plusieurs milliards de dollars.

Appelé Vertiv CoolPhase Flex, la nouvelle génération du système intégré Vertiv Liebert DSE à haut rendement, intègre des capacités de refroidissement liquide avec des technologies de refroidissement par air à base de réfrigérant et de rejet de chaleur dans un seul système compact. Lorsque les clients des data centers planifient leurs stratégies de croissance, Vertiv CoolPhase Flex peut être utilisé initialement comme système de refroidissement à détente directe (DX) avec économiseur intégré pour soutenir un refroidissement par air efficace, puis, à mesure que davantage d'informatique à haute densité est déployée, les capacités de refroidissement liquide peuvent être engagées simplement et rapidement pour prendre en charge les applications de refroidissement liquide.

Compass a également déployé la plateforme de service Vertiv Next Predict, en s'appuyant sur l'analyse prédictive pour permettre à l'entreprise de passer d'une maintenance réactive à une maintenance proactive sur certaines unités de refroidissement. La plateforme offre une visibilité sur les opérations de l'équipement et utilise les données historiques et d'exploitation pour déterminer avec précision quand la maintenance est nécessaire.

Prêts à soutenir les charges IA

De nombreux datacenters se sont aussi adaptés pour soutenir les charges de travail que demande l'intelligence artificielle et pas seulement générative. Le plus spectaculaire est sans doute le hub de puces graphiques de Sesterce. Dans le cadre d'un partenariat avec Dell, Sesterce a annoncé le lancement d'un cluster de calcul haute performance (HPC) basé sur le Dell PowerEdge XE9680 intégrant la plateforme NVIDIA HGX H200 et équipé avec des GPU NVIDIA H200 Tensor Core offrant 141 gigaoctets (Go) de mémoire HBM3e à 4,8 téraoctets par seconde (To/s) — soit près du double de la capacité du NVIDIA H100 Tensor Core GPU avec 1,4 fois plus de bande passante mémoire. Le NVIDIA H200, avec sa mémoire plus importante et plus rapide, accélère l'IA générative et les grands modèles de langage,

tout en faisant progresser le calcul scientifique pour les charges de travail HPC avec une meilleure efficacité énergétique et un coût total de possession réduit. Associée à la plateforme NVIDIA Quantum-2 InfiniBand, cette infrastructure repose sur une architecture puissante qui optimise la connectivité E/S, tout en garantissant scalabilité et performances optimales pour les projets d'IA les plus exigeants.

Le coolflex de Vertiv

Outre cet exemple, Digital Realty a réalisé un autre projet au Danemark. Le supercalculateur NVIDIA DGX SuperPOD est équipé de 191 systèmes NVIDIA DGX H100 intégrant plus de 1 500 GPU NVIDIA H100 Tensor Core et interconnectés à l'aide de la plateforme réseau NVIDIA Quantum-2 InfiniBand. Le DGX SuperPOD comprend également le logiciel NVIDIA AI Enterprise pour les modèles pré-entraînés, les frameworks optimisés et les bibliothèques logicielles accélérées pour la science des données. L'infrastructure avancée de Digital Realty, prête pour l'IA, servira de base à ce puissant système, permettant au Centre National d'innovation en IA — formé grâce à cette collaboration — au Danemark d'accélérer la recherche et l'innovation dans divers domaines, notamment les soins de santé, les sciences de la vie et la recherche sur le climat.

La montée en puissance de l'Edge Computing

Unitel Cloud Services, fournisseur de solutions cloud souveraines, a annoncé lors des « Scality Days Paris » l'intégration de Artesca, la solution de stockage objet développée par Scality. Cette collaboration renforce l'offre d'Unitel en matière de gestion des données, avec une externalisation native et optimisée entre les usages Edge et le Cloud souverain, une réponse aux besoins croissants en performance, sécurité et proximité des entreprises françaises. L'intégration d'Artesca permet à Unitel Cloud Services de proposer à ses clients une synchronisation native entre les applications gérées à l'Edge (à la périphérie du réseau) et leur Cloud souverain, garantissant ainsi une gestion fluide des données en temps réel, où qu'elles soient localisées. Cette solution assure une continuité des opérations pour les entreprises, en permettant une réplication et une protection des données critiques dès leur création, tout en respectant les exigences réglementaires strictes en matière de souveraineté et de localisation des données.


Une vue du Nexus hub de Sesterce à Marseille

Autre exemple intéressant, UltraEdge met à disposition de Gcore son réseau de plus de 250 sites répartis sur le territoire national, dont six NetCenters stratégiques situés à Courbevoie, Lyon, Lille, Bordeaux, Nantes et Strasbourg. Cette infrastructure permettra à Gcore de déployer ses services au plus près des utilisateurs finaux. L'objectif est de garantir un temps de latence minimal et une expérience utilisateur optimale. Cette alliance stratégique permet notamment à Gcore de déployer ses solutions d'IA de manière uniforme sur tout le territoire. Ainsi qu'une accessibilité et une performance optimales pour tous les utilisateurs, indépendamment de leur localisation. Le partenariat permettra aux entreprises françaises d'accéder à une infrastructure de nouvelle génération intégrant des serveurs d'IA sophistiqués, alimentés par des GPU NVIDIA. Cette technologie de pointe permettra l'entraînement de grands modèles de langage (LLM). L'objectif est aussi de garantir le déploiement d'applications d'IA en périphérie. La combinaison de l'infrastructure décentralisée d'UltraEdge et du réseau mondial de Gcore offre des avantages significatifs pour le marché français. Avec plus de 180 points de présence à travers six continents et une capacité réseau totale dépassant 200 Tbps, ce partenariat garantit une latence ultra-faible pour les applications critiques. □

B.G

Contrôleur Fadu, le révolutionnaire du SSD

Fadu, société coréenne, se distingue par son concept fabless et ses innovations sur l'architecture et les contrôleurs SSD pour obtenir des performances supérieures à ses concurrents.

Côté depuis 2023 sur la bourse de Seoul, avec une récolte de 150 M$ de capitaux frais, Fadu est quasiment méconnu en Europe, mais fait partie des acteurs de premier plan en Asie et aux USA avec des partenariats avec Western Digital et Meta. Le constructeur a d'ailleurs la volonté de développer plus largement ce type de partenariat technologique et commercial. L'entreprise a actuellement trois générations de contrôleurs SSD à son catalogue. Cela permet au client de concevoir son propre SSD autour du contrôleur de Fadu ou, plus simplement, de choisir une solution clé en main du constructeur ou d'utiliser les disques SSD en OEM.

Ainsi, Western Digital intègre un de ces trois contrôleurs dans ses SN861 PCIe 5.0 NVM, SSD certifiés pour supporter les systèmes GB200 NVL 75 de Nvidia pour des traitements massifs de données dans des applications d'intelligence artificielle.

Des différences qui comptent

Pour faire face à la charge, il est souvent nécessaire de multiplier les cœurs Arm, l'architecture du fabricant permet de contourner ce souci avec un plan de contrôle programmable, ce qui permet en plus d'optimiser la consommation énergétique et la qualité de service des disques.

Précisément, l'architecture, pour soutenir de larges charges, intègre le « deep pipeline ». Les commandes NVMe sont exécutées dans un processus à plusieurs étapes où chaque étape est réalisée par des puces dédiées (PPU ou Packet Processing Unit). Plusieurs pipelines peuvent être exécutés simultanément.

L'autre innovation du constructeur autorise le placement des données sur les SSD. Le logiciel optimise le placement des données améliorant la performance des applications, et étend l'endurance des disques tout en réduisant l'amplification des écritures sur le disque. Autre avantage de la solution, l'espace disque est optimisé. Le placement flexible est particulièrement indiqué pour les environnements nativement cloud, l'optimisation des applications de cache, les applications avec un nombre d'écriture intensive ou des tâches mixtes. Avec cette fonction de placement, il est même possible de créer différents espaces de nommage isolés qui ont la possibilité de bénéficier d'un ramasse-miette optimisé. Cette technologie est un standard proposé par l'Open Compute Project et a été approuvée récemment comme spécification NVMe qui avait été initiée par Samsung, Meta et Google.

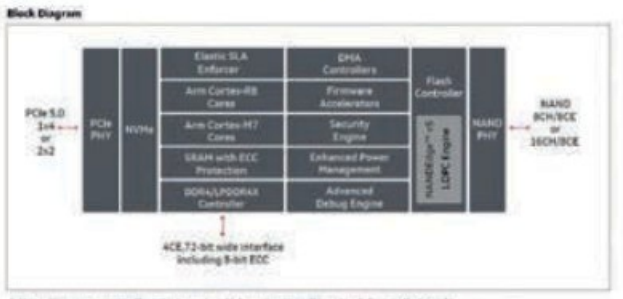
Une roadmap publique

2024 a vu l'extension de la gamme Echo de disques SSD PCIe 5.0 de Fadu. Il faudra cependant attendre 2026 pour une extension vers PCIe 6.0, et l'introduction de l'architecture CXL (mémoire et swithches) qui permet, entre autres, le partage des disques SSD NVMe. Sur le catalogue actuel, cela sera la gamme Sierra qui va bénéficier de ce passage à la génération future de PCIe. L'équipement devrait doubler l'efficacité énergétique comparativement à la génération précédente, et profiter pleinement des performances de PCIe 6.0, soit des flux supérieurs à 28 GB/s. Des fonctions avancées et de sécurité pour les charges dans le Cloud sont aussi annoncées. □

B.G

La différence entre l'architecture d'un cntrôleur classique et celui de Fadu.

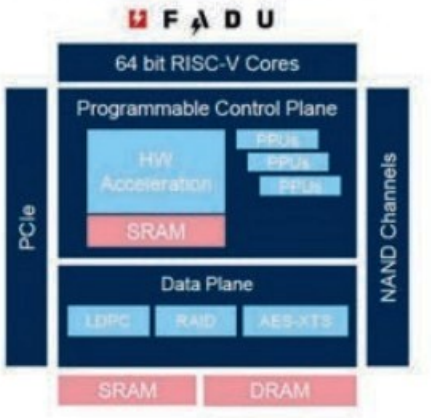
Typical SSD Controller Design



https://www.marvell.com/content/dam/marvell/en/public-collateral/storage/marvell-ssd-mv-ss1331-1333-product-brief.pdf

Lots of Arm cores needed to scale performance

FADU Controller Architecture



High performance scalability from programmable control plane ➡ power efficiency ➡ great QoS

Stratégie

Intel : un ambitieux changement de cap désormais remis en question

L'entreprise cherche, depuis quelques années, à développer son activité de fonderie et à retrouver davantage d'autonomie en contrôlant une plus grande partie de sa chaîne de valeur. Une stratégie que la retraite anticipée de Pat Gelsinger, qui en était le grand architecte, pourrait toutefois remettre en cause.

Les prochains processeurs Panther Lake d'Intel, qui devraient sortir courant 2025, ne seront plus intégralement fabriqués par TSMC, mais en grande partie à domicile. C'est ce qu'a annoncé Pat Gelsinger, le patron d'Intel, lors du dévoilement des résultats de l'entreprise pour le troisième trimestre 2024. Panther Lake constituera le prochain CPU phare d'Intel pour PC, succédant à Arrow Lake et Lunar Lake.

Pour réussir son pari, le géant américain des semi-conducteurs mise sur sa nouvelle technique de gravure, baptisée 18A, qui vise à combler le retard technologique d'Intel sur TSMC dans les gravures les plus fines. Le développement de cette technologie fait partie d'un plan mis en œuvre par Pat Gelsinger, lors de son arrivée à la tête de la société en février 2021, visant à rapatrier une partie de la production de puces, sous-traitée au fondeur taïwanais, en interne.

Retrouver une excellence technologique perdue

Une stratégie qui a pour but de rétablir les marges de l'entreprise, en baisse depuis plusieurs années, alors que ses clients se tournent vers des puces moins chères, et qu'Intel accuse un retard technique de plus en plus significatif sur ses concurrents directs. Au niveau de la conception, la société américaine est notamment supplantée dans les puces pour smartphone par l'architecture Arm, une société implantée au Royaume-Uni et détenue par le conglomérat japonais Softbank. Quant à la fabrication des puces en elle-même, Intel accuse du retard sur le taïwanais TSMC pour les gravures les plus fines, nécessaires pour concevoir les semi-conducteurs de pointe, utilisés dans l'IA. *« Intel a clairement loupé le train de l'IA. Leur GPU Ponte Vecchio, par exemple, n'est pas du tout utilisé pour faire de l'IA, mais plutôt du calcul intensif »*, commente Antoine Chkaiban, analyste chez New Street Research, un cabinet d'intelligence de marché.

La fabrication de puces en interne, grâce au procédé 18A, doit aussi servir de vitrine pour permettre à Intel de développer son activité de fonderie (fabrication de puces pour des sociétés tierces, comme le fait TSMC), en particulier sur le sol américain et européen, afin de répondre à un désir de souveraineté sur les semi-conducteurs au sein des deux blocs. Au total, Intel a investi plus de 100 milliards de dollars dans la construction d'infrastructures visant à rapatrier une partie de la production de semi-conducteurs, aujourd'hui principalement réalisée en Asie du sud-est, sur le sol américain.

Intel a de grandes ambitions… qui tardent à se concrétiser

Une stratégie qui pourrait s'avérer payante, selon Mike Demler, expert indépendant spécialisé dans l'industrie des puces. *« Si Intel atteint ses objectifs de production avec le procédé 18A, cela lui permettra effectivement d'accroitre*

Les Panther Lake vont succéder aux Arrow Lake avec une fabrication maison chez Intel.

ses marges et d'attirer davantage de clients pour son activité de fonderie. Le procédé 18A va sans doute permettre à Intel de rattraper TSMC. Amazon a déjà signé pour produire une puce réseau grâce au procédé 18A, ainsi qu'un microprocesseur Xeon ».

Toutefois, l'ambitieuse stratégie d'Intel autour du procédé 18A a également accumulé retards et problèmes techniques.

En septembre dernier, l'entreprise a notamment subi un gros "bad buzz" après que Broadcom a testé la technique 18A pour la fabrication de ses wafers et affirmé qu'à peine 20 % des puces ainsi réalisées passaient les premiers tests de conception. Autrement dit, le procédé 18A ne semble pas du tout prêt pour la production industrielle, alors même qu'Intel entend le déployer dés courant 2025. Apple et Qualcomm auraient aussi été approchés par Intel, et auraient décliné l'usage du processus 18A pour raisons techniques.

Pat Gelsinger a été débarqué par le Board d'Intel. Quelle stratégie le fondeur va t-il suivre ?

Il faut cependant faire preuve d'un peu de patience avant de juger le succès de la stratégie du géant américain des semi-conducteurs, selon Mike Demler. *« Intel est une grande entreprise dans une phase de transition majeure, qui requiert des années. Intel œuvre en outre pour devenir la seule société américaine capable de produire des semi-conducteurs de pointe aux États-Unis, ce qui accroît encore largement la difficulté et le coût de sa tâche, alors que l'entreprise n'a pas encore touché les fonds du Chips Act »*.

Nouveau dirigeant, nouveau cap ?

La retraite anticipée de Pat Gelsinger, annoncée début décembre, pourrait toutefois conduire à une remise en cause du tournant stratégique amorcé avec son arrivée à la tête de l'entreprise quatre ans plus tôt. Difficile, en effet, de ne pas voir dans ce départ abrupt un désaveu du cap voulu par celui qui avait promis de sauver l'entreprise en rétablissant son activité de fonderie. La démission de Pat Gelsinger aurait été poussée par le conseil d'administration de l'entreprise, qui se serait impatienté face aux retards accumulés par Intel dans sa nouvelle feuille de route, et sur ses rivaux comme TSMC et AMD, qui ont précipité la chute de son action en bourse.

Intel n'a pas été aidé par la stagnation du marché des PC et smartphone, peu reluisant sur les deux dernières années après que les consommateurs ont profité du surplus de cash dont ils ont bénéficié pendant la pandémie pour s'acheter de nouveaux appareils. Confrontée à une baisse de ces revenus, ainsi qu'à des investissements massifs visant à entamer sa mue stratégique, l'entreprise a accusé la plus lourde perte de son histoire (16,6 milliards de dollars) au troisième trimestre 2024. Plus tôt dans l'année, le groupe avait tenté d'arrêter l'hémorragie en licenciant 15 % de son personnel début août. Début novembre, l'entreprise a également reporté la construction de sa méga-usine prévue à Magdebourg, en Allemagne, afin de répondre à la volonté de souveraineté européenne sur les semi-conducteurs.

Des mesures qui n'ont pas suffi à arrêter les pertes, poussant Pat Gelsinger vers la sortie. Une erreur, selon Mike Demler. *« Le conseil d'administration s'est impatienté à cause de la baisse du cours de l'action et du coût de la stratégie de fonderie, qu'il a pourtant lui-même approuvée en nommant Pat Gelsinger. Avec les licenciements qui ont lieu chez Intel, virer Pat Gelsinger va plomber encore plus le moral des employés sans rien faire pour améliorer la situation. D'autant que la plupart des problèmes qui ont causé la baisse de la valeur de marché d'Intel sont le fait de son prédécesseur, ce qui montre que toute stratégie en matière de technologie prend des années avant que son impact soit visible »*.

Dans l'attente d'un remplaçant, qui n'avait au moment de la rédaction de ces lignes pas été annoncé, l'intérim est assuré par un duo constitué du directeur financier et de la responsable de la division PC. Une lourde tâche attendra le successeur de Pat Gelsinger, dont le conseil d'administration attendra sans doute un changement de cap. *« Il est probable que cela accélère des décisions qui étaient déjà en considération, comme la vente d'Altera et de Mobileye »*, estime Mike Demler.

Une scission de l'entreprise en deux, avec d'un côté les activités de fonderie, et de l'autre la partie design, pourrait également être à l'ordre du jour. L'entreprise a déjà flirté avec cette idée par le passé, et plusieurs anciens cadres ont récemment cosigné une tribune dans le magazine américain Fortune prenant parti pour cette solution, affirmant qu'il s'agirait du seul moyen de sauver la société. □

G.R

En 2025, les chefs d'entreprise adoptent enfin l'observabilité

Par Stéphane Estevez, EMEA Observability Market Advisor chez Splunk.

Les temps d'arrêt et les ralentissements ont des conséquences importantes sur les entreprises, telles que l'atteinte à la réputation ou une baisse de la fidélité des consommateurs, qui engendrent des pertes financières conséquentes sur le long terme. Pour tenter de les neutraliser au mieux, des investissements en matière d'observabilité sont nécessaires. Et cela passe par une sensibilisation des dirigeants d'entreprise et du comité exécutif. Pour ce faire, les équipes d'ingénieurs et ITOps devront, dès 2025, établir le lien entre les problèmes de performance des systèmes et les indicateurs commerciaux essentiels, par exemple, le taux de croissance, la conversion des clients et le chiffre d'affaires.

Maintenir la continuité d'activité des entreprises et assurer la résilience numérique des systèmes sont des enjeux de taille. Prouver la valeur de ces enjeux à travers des visualisations compréhensibles par les dirigeants sera primordial en 2025, et va de pair avec l'amélioration de l'expérience utilisateur.

Mesurer le ROI différemment

Il y a quelques années, la plupart des discussions sur l'observabilité étaient centrées sur le besoin de logs, de métriques et de traces pour permettre d'alimenter l'APM et la supervision de l'infrastructure. Aujourd'hui, le monitoring de l'expérience utilisateur revient au cœur des débats. Il s'agit non seulement de superviser les systèmes back-end, mais aussi d'évaluer l'expérience réelle de l'utilisateur de bout en bout, qu'il s'agisse d'un collaborateur, d'un administré ou d'un client, y compris la connectivité sur les réseaux appartenant ou non à l'entreprise.

En 2025, les équipes d'ingénieurs et ITOps commenceront à mesurer le retour sur investissement différemment. Par exemple, l'impact de l'expérience numérique sur les applications critiques en contact avec les clients sera désormais mesuré par la perte de revenus, la satisfaction des clients, les scores NPS et les taux de fidélisation de la clientèle.

Les initiatives d'observabilité conçues dès le départ en pensant aux clients permettent aux équipes d'ingénieurs et ITOps d'examiner des relations telles que l'impact de la vitesse du site sur les taux de conversion et l'existence d'une corrélation entre les performances et l'abandon par les utilisateurs. Ainsi, les entreprises exploiteront les données d'observabilité dès le début du cycle de développement logiciel, optimisant alors le code et les fonctionnalités des applications. Au-delà de l'amélioration de l'expérience, cette approche permettra de hiérarchiser les problèmes, d'aligner les équipes, d'encourager la proactivité et de faciliter la mesure du retour sur investissement des initiatives d'observabilité.

Le rôle de l'IA dans cette transformation

Les entreprises ne manquent pas de données pour prendre des décisions cruciales, mais elles rencontrent toujours des difficultés pour les exploiter. L'essor rapide de l'intelligence artificielle (IA) change la donne grâce à sa capacité à corréler et à résumer des données sous différents formats, favorisant une meilleure compréhension des relations entre les données.

En 2025, l'IA aidera donc les équipes techniques à relier plus facilement les données d'observabilité et les risques métiers. En outre, cela permettra aux dirigeants d'entreprises et aux responsables technologiques d'avoir une vue d'ensemble sur ce qui se passe réellement dans leur organisation.

Adopter une approche d'observabilité pour l'IA permettra également d'alimenter une sous-catégorie de l'IA : l'AIOps. L'usage de l'IA et du machine learning aidera les équipes IT à absorber le volume d'alertes engendré par cette visibilité globale. Ces technologies vont rendre ce volume de données consommable en les filtrant et soulignant celles qui sont prioritaires et requièrent une intervention humaine, tout en apportant du contexte.

L'objectif final de l'observabilité est d'offrir une expérience positive à l'utilisateur final. Pour ce faire, les équipes doivent analyser le coût des temps d'arrêt ou des ralentissements dans l'ensemble de l'entreprise, à n'importe quel niveau de granularité et en temps réel, puis transmettre ces informations dans des termes compréhensibles par toute l'entreprise. D'autant que les données collectées de manière exhaustive pour l'observabilité permettront d'alimenter d'autres branches du département informatique, notamment les équipes chargées de la cybersécurité. ■

Feuille de route

Cognizant veut se faire une place importante en France

L'Informaticien a récemment pu avoir un entretien avec Long Le Xuan, directeur général pour Cognizant France. Il nous a fait part de sa stratégie de développement de l'ESN internationale dans l'Hexagone.

Dans notre pays, le DG de Cognizant indique : *« sur le périmètre France, nous gérons les grands clients français avec un modèle de delivery et de production qui est hybride, c'est-à-dire avec des ressources en France, mais également des ressources délocalisées, en offshore, notamment en Inde, pour servir nos plus grands clients français puisqu'on est très orienté grands comptes »*.

Un positionnement premium

Après ses passages chez Atos et Econocom, Long Le Xuan a été intéressé par la culture internationale de Cognizant, mais aussi son positionnement différent comparativement aux ESN françaises. Il précise : *« Cognizant, c'est une entreprise internationale, au sens où elle est américaine par son siège, qui est dans New Jersey, côté New York, et cotée au Nasdaq, avec un modèle de delivery très centré sur l'Inde, avec 250 000 personnes en Inde sur les 350 000 du groupe, et un modèle de business, de relation client, qui est très local. C'est cette hybridation qui m'a intéressé, en plus d'avoir cette dimension américano-indienne européenne, et évidemment, son positionnement de marché. J'étais chez Atos et Econocom, sur une dimension ESN, avec des métiers très drivés par les services managés, par le digital workplace, par le workspace, par l'infrastructure, par le Cloud. Les marchés sur lesquels est positionnée Cognizant sont l'intelligence artificielle, la data, le cloud, la cybersécurité, l'IoT, le digital engineering et le monde applicatif au sens large. On est donc sur un positionnement ESN un peu plus premium, dans le même marché que les modèles qui étaient soit tirés vers l'infogérance, soit tirés vers les couches basses de l'IT »*.

Think, design, build, run

Ce classement premium se caractérise par *« le think design build run, c'est-à-dire le cycle complet, avec les volumes drivés par le run, mais de grandes expertises en build, des prestations de conseil qui sont établies, et même une activité qui fait de l'advisory, du conseil auprès des clients, ce qu'on appelle de l'AMOA, c'est pour quoi je rajoute le think au design build run, mais on est sur le cycle complet d'une ESN sur le marché français, et c'est précisément ce qui était intéressant en termes de positionnement de compte ; là où, par exemple chez Econocom, j'étais à la fois sur des grands comptes et du mid market, des ETI, et du public. Le modèle de Cognizant en France va surtout chercher les grands comptes, on ne va pas sur le mid market, assez peu sur le public pur, pas sur le public ministère et collectivité, mais on va plutôt sur du para-public »*, ajoute le DG de Cognizant.

Une force de frappe globale

Sur le marché des services peu concentré dans notre pays, avec de nombreux acteurs centrés commercialement sur la France ou sur l'Europe, Cognizant peut s'appuyer sur une force de frappe équivalente à celle de Capgemini, avec ses équipes et son chiffre

COGNIZANT EN BREF

Cognizant s'appuie sur ses compétences globales : en France, en nearshore et offshore, selon les besoins de ses clients. Ses solutions et services sont structurés sur le modèle design, build & run, et s'articulent autour de plusieurs grands axes :

- **Transformation digitale**
- **Gestion de data**
- **Cybersécurité**
- **Cloud**
- **IA**
- **Application /developpement/maintenance**

Exemples de solutions phares

- **Bluebolt**
- **Cognizant Ocean**
- **Cognizant Flowsource**

Secteurs cibles

- **Pharmaceutique**
- **Banque, services financiers**
- **Défense**
- **Industrie**
- **Biens de consommation**

Cognizant France cible exclusivement les grands comptes.

d'affaires de près de 20 milliards de dollars. Notre interlocuteur y voit ainsi une forte opportunité : *« si on le prend basiquement, 70 % du business de Cognizant est aux États-Unis, sur le modèle américain, c'est toute la plateforme ISIT qui est délivrée, ça c'est la puissance globale, l'offre, elle, est globalisée via le marché américain. Nous sommes un acteur moins connu en France, mais nous avons cette capacité, puisque le groupe fait quasiment 20 milliards de dollars à l'année, c'est à peu près l'ordre de grandeur que fait Capgemini, donc en termes d'offres, nous réalisons ce que font Capgemini et Accenture à l'échelle globale. La différence en France, c'est qu'on n'a pas la même notoriété, par l'historique, les relations, l'intimité... En rencontrant les clients, j'ai fait un constat, les grands comptes sont tous en train de se massifier, c'est-à-dire que la demande, notamment dans l'industrie, et même de plus en plus dans la banque, qui était à l'époque localisée, c'est-à-dire que vous pouviez répondre à des appels d'offres nationaux, sur périmètre territoire français, le Covid a induit une globalisation, déjà des nécessités de piste d'économie, parce qu'on parle beaucoup d'économie en France. Pourquoi il y a la concurrence, c'est parce qu'il y a beaucoup de demandes de réduction de coût informatique, on voit de plus en plus de global procurement, donc des achats globaux, et des organisations IT centrales, qui s'organisent et qui font en sorte, que les donneurs d'ordre français, qui à l'époque géraient des grands périmètres fonctionnels et techniques, mais sur un territoire national, sur le périmètre France. Par la massification, les scopes géographiques se sont élargis. Aujourd'hui, c'est souvent soit européen, soit mondial ».* Il ajoute : *« nous sommes un challenger, mais nous avons une capacité globale et nous sommes très bien placés pour être complémentaire dans un écosystème où les entreprises de Tiers 2 dans le service peuvent avoir besoin de nous pour répondre à des besoins hors de France. Nous les aidons ainsi à défendre leur position sur leur marché local ».* Il continue : *« l'union des deux se fait sur une trajectoire de transformation, qu'on co-construit pour apporter une haute proposition de valeur à nos clients qui voient d'un bon œil, justement, cette notion de consortium, qui vient un peu dépoussiérer les modèles jugés parfois un peu trop protectionnistes, ou défensifs par d'autres acteurs du marché ».*

DONNÉES CLÉS

- Fondé il y a 30 ans en Inde, Chennai
- Fondé il y a 20 ans en France
- + de 336 000 employés dans le monde
- Chiffre d'affaires global : 19,4 Md$

Long Le Xuan, DG de Cognizant France

Des partenariats technologiques de premier plan

Le cadre de l'ESN met aussi en avant les nombreux partenariats de premier plan de Cognizant avec les principaux acteurs de l'industrie informatique que cela soit avec des fournisseurs d'infrastructures ou des éditeurs de logiciels.

Un rétroplanning d'actions

En s'appuyant sur ces différentes observations et forces de Cognizant, L'ESN a développé une feuille de route d'actions pour monter en puissance dans notre pays, en s'appuyant donc sur la force de son écosystème, et en choisissant des partenaires de qualité et pas forcément en nombre pour s'immiscer dans les projets. La stratégie se déroule donc sur trois points : croissance, confiance pour les partenaires et les clients, et notoriété ou travail sur la marque Cognizant pour se faire mieux connaître. Tout cela se déroule avec un plan de recrutement précis et une affirmation dans le rôle de conseil joué par l'ESN. ☐

B.G

Conformité

Valiantys lance un nouveau service de conformité à DORA

Le fournisseur de services sur la plateforme d'Atlassian introduit un service de GRC (gestion des risques et de la conformité), afin de répondre aux exigences de DORA (Digital Operational Resilience Act).

La solution repose sur l'apport de différents partenariats. Le premier est celui avec HYCU, un éditeur de gestion et de protection des données, qui propose une protection automatisée des données et récupération des données hors site pour plus de 80 actifs informatiques. L'éditeur fournit, de plus, des fonctions de reprise après sinistre. Landsweeper est le second partenaire impliqué dans la solution qui apporte sa technologie de découverte automatique des actifs et de visualisation des applications SaaS. Appfire complète l'ensemble avec des logiciels qui améliorent, étendent et connectent les principales plateformes mondiales de collaboration d'entreprise.

Explorer DORA

L'offre de l'ESN est conçue pour aider les organisations à répondre aux exigences de conformité complexes des réglementations émergentes, avec un accent initial porté à la loi sur la résilience opérationnelle numérique (DORA) de l'Union européenne. Elle vise à proposer une réponse complète aux défis posés par cette réglementation, en particulier dans les nouveaux domaines de la gestion des risques des tiers et de la résilience opérationnelle. Alors que les organisations s'appuient de plus en plus sur des applications cloud et des solutions as-a-service, la conformité avec DORA monte un défi complexe et chronophage. L'émergence de DORA, dont l'entrée en vigueur est prévue pour début 2025, représente un changement significatif dans la création d'une approche uniforme de la gestion des risques des tiers dans le secteur financier de l'UE. Valiantys a intégré ces différentes solutions dans Jira Service Management permettant aux organisations d'adopter une approche plus résiliente et rationalisée des défis réglementaires. Cette solution est un élément clé de la stratégie globale d'expansion du développement de solutions sur JSM et Confluence d'Atlassian par Valiantys, permettant aux entreprises de réaliser une transformation à l'échelle de l'organisation, et de s'assurer qu'elles sont équipées pour répondre aux demandes actuelles et futures.

Une solution globale

La solution GRC de Valiantys propose un système de gestion intégré englobant la découverte des actifs, la sauvegarde, les tests et la gestion documentaire. Fournissant une source unique et centralisée pour la préparation à la conformité DORA, la plateforme inclut des fonctionnalités de tests de résilience, des tableaux de bords, des notifications et des informations axées sur DORA. Elle a pour but d'optimiser le processus de conformité avec DORA, en réduisant drastiquement le temps et les ressources requis. Elle supprime la nécessité de consacrer des centaines d'heures au catalogage manuel des actifs, d'utiliser plusieurs outils de sauvegarde et de gérer une documentation volumineuse. □

B.G

DORA DEVIENT EFFECTIVE LE 17 JANVIER PROCHAIN

La loi sur la résilience opérationnelle numérique (DORA) définit les normes techniques que les institutions financières et leurs fournisseurs de services technologiques critiques doivent mettre en œuvre dans leurs systèmes TIC (technologies de l'information et de la communication) au début de l'année 2025. D'ici le 17 janvier 2025, un cadre contraignant et complet sera mis en place pour la gestion des risques liés aux technologies de l'information et de la communication (TIC) dans le secteur financier de l'UE. Mais aussi celle des tiers essentiels qui leur fournissent des services connexes, tels que les plateformes en nuage ou les services d'analyse de données. En d'autres termes, leurs principaux fournisseurs. Cette nouvelle réglementation implique que certains contrats, qui n'ont pas été analysés sous l'égide de l'Autorité bancaire européenne, devront être revus. Cela implique un travail important d'analyse des contrats en cours : d'une part, pour identifier les nouvelles clauses qui étaient déjà incluses dans les anciens contrats, et d'autre part, pour identifier les contrats qui devront être renégociés afin d'y inclure de nouvelles clauses. Cette nouvelle réglementation DORA est une avancée pour la sécurité des entreprises. Les cyberattaques en provenance des fournisseurs, notamment lorsqu'ils sont de petite taille, représentent un danger qui n'était pas encore contrôlé.

2025 sera... comme 2024

par Bertrand Garé

L'actualité de l'année technologique en 2024 a été riche en événements mais aussi en nouveautés. Tout en haut de l'affiche, évidemment, l'intelligence artificielle générative a été de toutes les annonces des acteurs du marché, de tous les communiqués de presse et de beaucoup de conversations, salons et autres réunions, tenant à nous montrer que l'innovation est le nouveau totem de ce monde. Nous étions donc tenus d'en parler longuement et d'analyser cette rupture dans nos modes de vies et de travail, balançant entre humain augmenté et scénarii futuristes sur un monde fait de robots, d'intelligence artificielle, de drones plus ou moins meurtriers. Entre espoir et angoisse. Vu la rapidité du développement de cette technologie et les investissements phénoménaux qu'elle génère, il faut donc s'attendre, en 2025, à revivre la même chose avec tout de même quelques surprises, puisque selon une étude récente, les ETI et PME utilisent plus cette technologie que nos grandes entreprises. Il faut bien pallier le manque de ressources pour de nombreux secteurs d'activité puisque les salaires et l'emploi proposés n'attirent personne. Vu les sommes mises en jeu, il ne faudrait pas que la technologie déçoive. Il serait nécessaire aussi que la technologie passe réellement au stade industriel pour faire baisser son prix. Les entreprises, que ce soit en interne, ou sur des plateformes en ligne, ont du mal à trouver d'autre avantages que des automatisations apportant des gains de temps mais de peu de valeur pour l'ensemble de l'entreprise. Des agents, on parle d'agentique maintenant, ont pris le relai des grands modèles. Plus petits et plus spécialisés, ils correspondent mieux aux besoins actuels des entreprises, mais sont encore loin du compte pour offrir des gains de valeur globaux.

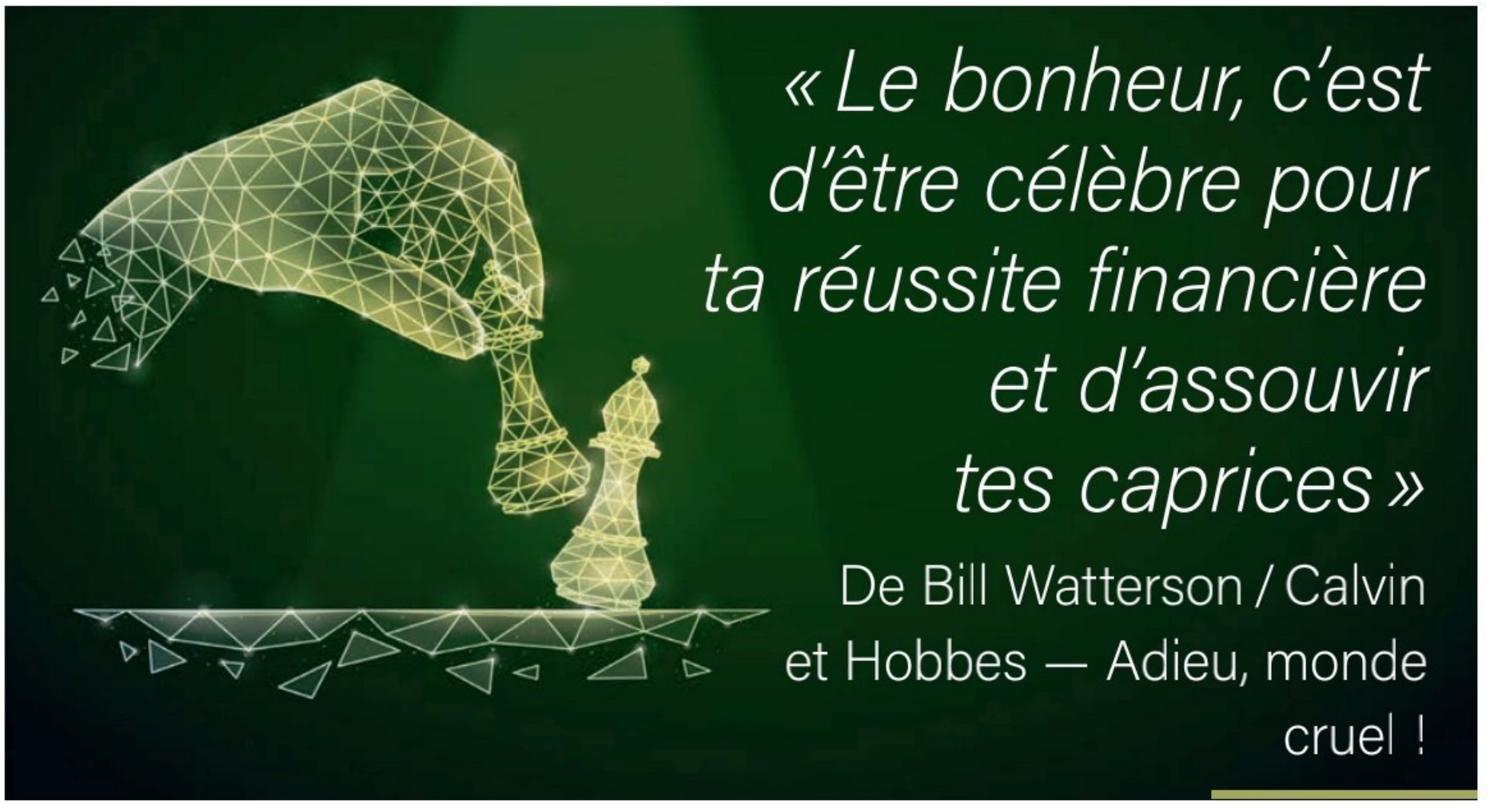
Si 2025 devrait voir des déploiements plus larges sur cette technologie, il n'est pas sûr que l'intelligence artificielle générative dépasse l'esprit humain. Sans compter que la technologie est gourmande en énergie et demande toujours plus d'électricité pour autoriser juste la puissance de calcul nécessaire aux opérations. On voit apparaître des centres de données aux dimensions impressionnantes. Le PDG de Ciena, une entreprise fournissant de la connectivité en fibre optique, déclarait dans une interview relayée par un confrère de ZDNet.fr : *« Vous avez des centres de données qui font plus de deux kilomètres »*. Un autre expert dans le même article notait : *« D'ici 2026, on s'attend à ce que le traitement de l'IA dans le monde nécessite 40 gigawatts d'énergie spécifiquement pour les centres de données d'IA, soit huit villes de New York »*. Cette ville consomme 5 gigawatts par jour en moyenne. En clair, la deuxième tendance de l'année 2024, le Green IT, a pris un sérieux coup au menton alors que Gartner prédit ainsi dans une nouvelle étude que, d'ici 2027, les pénuries d'électricité pourraient restreindre les opérations de 40 % des centres de données dédiés à l'IA. Le même cabinet constate que l'énergie requise pour faire fonctionner les serveurs optimisés pour l'IA atteindra 500 térawattheures (TWh) par an en 2027, soit 2,6 fois le niveau enregistré en 2023.

Si toutes les entreprises du secteur informatique se transforment en parangon du bien de la planète, cela ne résiste pas au mur de la réalité. Google a enregistré en 2023 une hausse de 13 % de ses émissions de CO_2, principalement attribuée à l'expansion de l'IA et à la consommation électrique de ses centres de données. Les émissions globales de Microsoft en 2023 ont augmenté de 29,1 % par rapport à 2022. Et la consommation d'eau du créateur d'Azure a

également augmenté, passant de 6,4 millions de mètres cubes en 2022, à 7,8 millions de mètres cubes en 2023, due aux besoins de refroidissement des centres de données Cloud et IA. Les hyperscalers font feu de tout bois pour utiliser des énergies décarbonées, ce terme est d'ailleurs plus adapté que « énergies propres » puisque, entre autres Google et Microsoft, pensent installer des minicentrales nucléaires dans leurs centres de données ou rachètent carrément un site nucléaire comme celui du hélas trop connu Three Miles Island. Chaque technologie a donc ses avantages et ses défauts, mais être un humain augmenté sans avoir la lumière partout n'est pas forcément le choix à suivre. Malgré tout cela, et les parodies de conférence sur le climat, on continue de nous donner bonne conscience avec des similis droits à polluer comme celui de racheter un quota carbone quand on prend l'avion ou à chaque requête sur internet en utilisant tel ou tel navigateur ou application va permettre de sauver les rhinocéros blancs ou de planter des arbres on ne sait où. On oublie parfois que la meilleure économie d'énergie est celle que l'on n'utilise pas. Si vous voulez mon avis, plutôt que de planter des centres de données partout, il serait judicieux de mettre le paquet d'investissement sur des technologies comme la fusion nucléaire ou d'autres permettant de fournir une énergie durable, à la fois dans le temps pour répondre aux besoins à venir. Il faut donc hélas s'attendre à une flopée de nouveaux centres de données toujours plus grands juste pour permettre à l'IA de fonctionner pour traiter les masses de données nécessaires.

Autre célébrité en 2024, mais déjà au hit-parade des geeks depuis des années, le Bitcoin ! Celui-ci est revenu sur le devant la scène avec les déclarations des futurs dirigeants des USA. Que ce soit Elon Musk ou Donald Trump, tous deux sont des aficionados des cryptomonnaies. Bon Elon Musk, comme toujours, préfère la sienne, le DOGE Coin. Toutes leurs déclarations sur le sujet ont alimenté une spéculation comme jamais, avec des hauts et des bas.

Etonnamment, ces deux personnes qui veulent les USA « Great Again », s'attaquent en fait à un des piliers de la puissance du pays : le dollar. Celui-ci est souvent employé comme une arme stratégique par ce pays qui en contrôle la valeur au gré de sa situation économique et de ses besoins commerciaux. Le bitcoin est très différent et peut être miné par tout le monde. Ils délaisseraient donc une arme importante pour une monnaie qu'en fait presque tout le monde peut détenir. Cette monnaie est en fait un cartel de milliardaires comme Elon Musk, le patron de Microstrategy ou encore le PDG de Galaxy Digital, Mike Novogratz et les patrons de Coinbase ou Binance. Ce dernier a été condamné pour blanchiment d'argent et violations des sanctions. Comme les arbres ne montent pas jusqu'au ciel comme le dit le proverbe boursier, on peut s'attendre cependant à la possible vente massive de bitcoins par un de ces milliardaires, faisant replonger la cryptomonnaie alimentant ainsi un nouveau cycle de spéculation sur sa future hausse. Au passage, les Nord-Coréens apprécient aussi les bitcoins mais préfèrent les escamoter pour financer leur programme de missiles qui finissent sur le tête des Ukrainiens. L'année dernière, ils ont signé une année record avec 1,3 milliards de dollars de bitcoins volés. Voilà tout ce qui faisait que la cryptomonnaie méritait de figurer sur notre podium de ce qui a été en 2024 et sera encore en 2025. Pas besoin donc de boule de cristal et d'études savantes pour comprendre que les tendances de 2024 vont continuer l'année prochaine. Il ne faut donc pas s'attendre à une autre révolution l'année prochaine... À moins que ! □



Intégration

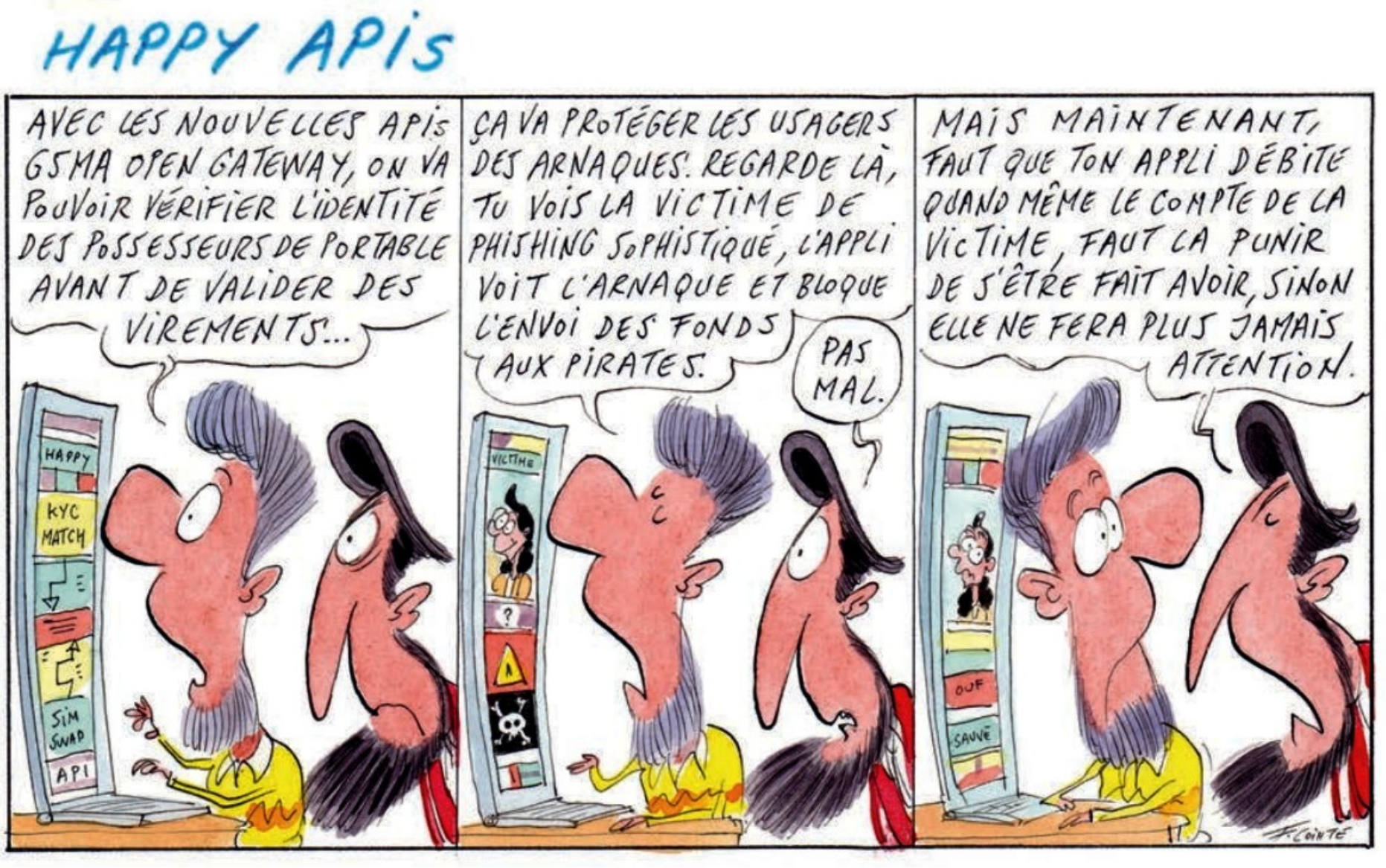
Les opérateurs français se rallient à Open Gateway de la GSMA

Bouygues Telecom, Free, Orange et SFR fournissent deux API de réseau pour aider les entreprises en ligne à lutter contre la fraude et l'usurpation d'identité numérique.

Dans le cadre de l'initiative mondiale GSMA Open Gateway, les quatre principaux opérateurs mobiles français — Bouygues Telecom, Free, Orange et SFR — annoncent qu'ils uniront leurs forces pour fournir des services conçus pour aider les développeurs d'applications et les entreprises à lutter contre la fraude en ligne et à protéger les identités numériques des clients mobiles.

Ainsi, ils lancent deux interfaces programmables d'application (API) pour le marché français — KYC Match et SIM Swap — développées selon la nouvelle norme CAMARA qui vise à harmoniser les spécifications entre les opérateurs de téléphonie mobile. La France est le premier pays au monde où les quatre principaux opérateurs lancent l'API appelée KYC Match, qui renforce la façon dont les entreprises en ligne vérifient l'identité de leurs nouveaux clients. Les API ont été développées et rigoureusement testées sur le marché français avec un certain nombre d'institutions financières, telles que BforBank (la branche en ligne du Crédit Agricole) ou Fortuneo (filiale du Crédit Mutuel Arkéa), qui utilisent KYC Match pour sélectionner les nouveaux clients en partenariat avec DQE Software. Plus d'une vingtaine d'entreprises en France utilisent déjà les API des opérateurs télécoms pour prévenir les fraudes par prise de contrôle de compte, et vérifier l'identité des clients lors des transactions et des processus d'accueil.

Le contexte explique parfaitement cette union, alors que les transactions en ligne explosent et que les tentatives de fraude se multiplient. 80 % des entreprises françaises déclarent avoir subi des tentatives de fraude en ligne, 45 % d'entre elles affirmant que la fraude en ligne a augmenté au cours des 12 derniers mois. Le nombre de cas d'usurpation d'identité numérique a augmenté de 40 % au cours des quatre dernières années, selon les chiffres du ministère français de l'Intérieur. Ces nouvelles API contribueront à protéger les identités mobiles des consommateurs, en ajoutant une couche supplémentaire de défense contre les fraudeurs. □ **B.G**



Code

Automatiser la gestion d'un réseau avec Python

Python propose plusieurs bibliothèques puissantes pour automatiser la gestion des réseaux. Cela va de la configuration des équipements réseau au suivi des performances, en passant par la gestion des paramètres de sécurité ou la surveillance. Nous allons voir dans cet article quels sont les principaux outils disponibles pour automatiser un réseau avec Python.

L'automatisation d'un réseau avec Python implique l'utilisation de bibliothèques adaptées à certaines tâches. Python dispose de nombreuses bibliothèques dédiées au pilotage des équipements réseau. Elles apportent généralement une couche d'abstraction particulière afin de simplifier la communication avec les équipements. Netmiko propose des fonctionnalités de gestion de matériels réseau bien spécifiques. Napalm ouvre les portes de l'abstraction multifournisseurs. Scapy est dédiée à l'analyse du réseau. Psutil se consacre à la surveillance des ressources réseau. L'association de ces différents outils permet de créer des scripts Python qui vont simplifier et automatiser de nombreuses opérations sur les réseaux.

Exécution de commandes en mode sécurisé avec Paramiko

La bibliothèque Paramiko permet d'envoyer des commandes aux appareils réseau via le protocole SSH, afin qu'elles soient exécutées dans leur console CLI. Le résultat est récupéré par le script Python qui pourra les traiter comme bon lui semble. Le script ci-dessous utilise Paramiko pour demander à un routeur sa table ARP en vue d'identifier sur quel port d'un switch est connectée une machine du réseau.

```
# import des bibliothèques et fonctions nécessaires
import sys
import paramiko
from time import sleep
# définition de l'adresse IP du routeur à interroger
router="192.168.1.113"
# création d'une connexion SSH vers le routeur
connection = paramiko. SSHClient()
connection.set_missing_host_key_policy(paramiko. AutoAddPolicy())
connection.connect(router, username=" admin", password="password")
router_connection = connection.invoke_shell()
print('Connexion réussie au routeur %s' % router)
# désactivation de l'affichage dans la console
router_connection.send('terminal length 0\n')
sleep(1)  # pause d'une seconde
# exécution de la commande show arp
router_connection.send(« show arp\n")
sleep(2)
# Affichage de la sortie après décodage au format texte UTF-8 (ASCII)
print(router_connection.recv(5000).decode(« utf-8"))
# Fermeture de la connexion SSH
connection.close()
```

Gestion des appareils réseau avec Netmiko

Netmiko est une librairie Python multi-vendeurs destinée à simplifier l'accès SSH aux équipements via la librairie Paramiko. Elle est très pratique pour automatiser la configuration des appareils réseau tels que les routeurs, les switches, les firewalls ou autre. Elle permet de se connecter à des appareils via le protocole sécurisé SSH et d'exécuter des commandes comme lors d'une connexion directe. Netmiko apporte une couche d'abstraction vers un grand nombre d'équipements réseau. Il n'est plus nécessaire de connaître par avance les commandes propres à l'appareil à piloter. Son installation, comme pour la plupart des packages Python, se fait via pip avec la commande suivante : pip install netmiko

Le script qui suit permet de se connecter à un routeur Cisco et d'exécuter des commandes depuis son interface CLI :

```
# import de la commande nécessaire au script
from netmiko import ConnectHandler
# Création d'un dictionnaire regroupant les paramètres de connexion
device = {
'device_type': 'cisco_ios',
"host": '192.168.113.1',
"username": "admin",
"password": "password",
"secret": 'enable_password'
}
# Connexion à l'appareil via la fonction ConnectHandler qui reçoit en
# paramètre le dictionnaire avec les informations de connexion
connection = ConnectHandler(**device)
# Passage en mode privilégié
connection.enable()
# Exécution de commandes depuis l'espace CLI du routeur :
output = connection.send_command('show ip interface brief')
print(output)
# Déconnexion
connection.disconnect()
```

Netmiko utilise la définition device_type qui puise dans sa base de connaissances, afin de savoir comment gérer les communications avec l'équipement cible. Elle permet une connexion directe à un nombre donné d'appareils, sans avoir à connaître les commandes spécifiques au matériel. En revanche, elle n'est pas capable d'automatiser la gestion de configuration. Pour cette partie-là, il va falloir employer Napalm ou ansible-runner.

NAPALM

La bibliothèque Napalm est conçue pour manipuler la configuration des équipements réseau. Alors que Netmiko dispose d'une base de connaissance très riche, Napalm ne sait pour sa part travailler qu'avec un nombre assez restreint d'équipements. Elle reconnaît ceux de marques Arista, Juniper et Cisco employant un système d'exploitation IOS, IOS XR, NX-OS ou NX-OS SSH. Son installation (toujours via pip, comme souvent en Python) se fait via la commande suivante : pip install napalm

Le script qui suit permet de récupérer des informations sur un périphérique réseau :

```
# import de la commande nécessaire au script
from napalm import get_network_driver
# sélection du pilote de l'équipement réseau
driver = get_network_driver("ios")     # pilote Cisco
# Connexion à l'équipement via son adresse IP
device = driver('192.168.113.1', "admin", "password")
device.open()
# Récupération d'informations générales sur le matériel
informations = device.get_facts()
print(informations)        # Affichage des informations
device.close()                          # Fermeture de la connexion
```

Ce code, très basique, ne teste pas le succès des opérations. Dans la *« vraie vie du code »*, il faut bien entendu le faire pour éviter un plantage brutal.

Surveillance réseau avec Scapy

Le leitmotiv de Scapy est la surveillance via la manipulation de paquets réseau. Vous pouvez l'utiliser pour effectuer des tests de connectivité (ping, traceroute, analyse de paquets) et surveiller l'état du réseau. Le script qui suit envoie un ping à un serveur (ici le dns de CloudFare) et analyse la réponse reçue :

```
# importation des fonctions employées par le script
from scapy.all import ICMP, IP, sr1
# Définition de l'adresse IP cible
target_ip = « 9.9.9.9 »
# Création d'un paquet ICMP, le protocole employé par le ping
packet = IP(dst=target_ip)/ICMP()
# Envoi du paquet et attente de réponse
response = sr1(packet, timeout=1)
if response :
print(f"Réponse reçue de {target_ip} : {response.summary()}")
else :
print(f"Aucune réponse reçue de {target_ip}, serveur inaccessible")
```

Netmiko est une librairie Python multi-vendeurs, dont le but est de simplifier l'accès SSH aux équipements à l'aide de la librairie Paramiko. Elle supporte de nombreux constructeurs.

Gestion de configuration avec Ansible

Ansible est un outil d'automatisation très populaire de gestion de configuration des systèmes Linux. Il a recours à des fichiers YAML pour décrire l'état souhaité (Desired State) des systèmes. Les configurations souhaitées sont contrôlées et, si cela s'avère nécessaire, corrigées automatiquement. Ansible peut être automatisé grâce à la bibliothèque ansible-runner qui permet de l'intégrer dans des scripts Python. Voici un exemple de script exécutant un playbook Ansible :

```
import ansible_runner
# Exécution d'un playbook Ansible
playbook_to_run = ansible_runner.run(private_data_dir='/chemin_vers_le_
playbook', playbook='configuration.yml')
print(f"Status : { playbook_to_run.status}")
```

La difficulté ne tient pas dans l'exécution du playbook Ansible, mais dans son écriture. La bibliothèque ansible-runner se charge du reste.

Planification des tâches avec APScheduler

Si vous souhaitez automatiser des tâches récurrentes, telles que des sauvegardes ou des vérifications de l'état du réseau, la bibliothèque APScheduler est faite pour vous. Son rôle est de planifier des scripts Python quels qu'ils soient. Cela vous permet de faire la même chose qu'avec la Crontab de Linux ou le planificateur des tâches de Windows, mais en vous plaçant un cran au-dessus et avec du code standard quel que soit l'OS (système d'exploitation). Le script qui suit planifie une tâche de vérification du réseau toutes les 30 minutes :

```
from apscheduler.schedulers.blocking import BlockingScheduler
from scapy.all import ICMP, IP, sr1
# définition de la fonction de contrôle du réseau
def check_network() :
target_ip = « 1.1.1.1"                  # Adresse du serveur DNS Cloudfare
packet = IP(dst=target_ip)/ICMP()
response = sr1(packet, timeout=1)
if response :
print(f"Réponse reçue de {target_ip} : {response.summary()}")
else :
print(f"Aucune réponse reçue de {target_ip}. Serveur inaccessible !")
scheduler = BlockingScheduler()
scheduler.add_job(check_network, "interval", minutes=30)
scheduler.start()
Psutil
```

Nous pouvons utiliser la bibliothèque Psutil pour automatiser la surveillance de la bande passante, des connexions réseau et d'autres paramètres système. Voici un petit exemple de script pour surveiller l'utilisation du réseau :

```
import psutil
# Récupération des informations sur le réseau
net_stats = psutil.net_if_stats()
# Affichage des statistiques
for interface, stats in net_stats.items() :
print(f"Interface {interface} :")
print(f" Up : {stats.isup}")
print(f" Vitesse : {stats.speed} Mbps")
```

Th.T

Package

Une déferlante d'offres Cyber pour les PME attendue en 2025

Si les textes d'application du règlement européen se font encore attendre en France, NIS 2 finira bien par s'imposer à tous. Un éveil aux bonnes pratiques Cyber pour bon nombre de PME et collectivités locales jusqu'à maintenant très vulnérables. Le texte pourrait bien en toucher plus de 100 000 !

En fin d'année 2024, Bouygues Telecom Entreprises lançait le bSecure Firewall Pack, un bundle réunissant un firewall Fortinet installé chez le client avec un support téléphonique. La supervision de l'installation depuis le SOC de Bouygues Telecom Entreprises, en option, est sans nul doute l'atout numéro 1 de cette offre. Le pack est disponible à partir de 99 € HT par mois et ne demande pas d'investissement initial. L'option SOC est accessible sur devis, en fonction du nombre de postes à surveiller.

Bouygues Telecom Entreprises déploie des firewalls Fortinet auprès des entreprises clientes de son bSecure Firewall Pack, mais aussi surtout une surveillance en mode managé via ses SOC.

Un effort de démocratisation du marché Cyber

Des offres de SOC managées accessibles aux entreprises de taille moyenne ont émergé ces dernières années, mais cette offre Bouygues Telecom Entreprises est assez représentative des efforts des grands acteurs du marché démocratisation auprès des PME, voire des particuliers.

On se souvient de l'acquisition d'ITrust par Free Pro en 2023, l'idée était alors de muscler l'offre Free Pro, alors essentiellement de l'accès fixe et mobile et de l'hébergement. S'appuyant sur le SIEM Reveelium et le SOC d'ITrust, cette expertise à rejoint le catalogue Free Pro sous la bannière Cyber XPR.

L'automatisation gagne le monde Cyber

L'émergence de ces offres réellement accessibles n'est pas le fruit du hasard. Outre la fenêtre d'opportunité commerciale qui s'ouvre avec le spectre que NIS 2 fait peser sur les entreprises, ces offres témoignent surtout du niveau croissant d'industrialisation de la Cyber. Impossible pour un MSSP ou un opérateur de lancer un service SOC managé à un tarif PME. Le modèle économique est bien entendu d'automatiser et d'industrialiser au maximum le traitement des alertes pour réduire les interventions humaines. Les SIEM « Next-Gen » et les XDR amènent leurs capacités d'automatisation des traitements et les prestataires misent sur le SOAR pour traiter l'immense majorité des faux positifs générés par les infrastructures de leurs clients.

L'acteur français qui va sans doute le plus loin dans cette logique est Orange Cyberdefense, puisque l'acteur historique est solidement implanté auprès des grandes entreprises avec ses CyberSOC et qu'il adresse le marché PME avec l'offre micro-soc depuis 2022, en droite ligne avec son service d'EDR managé. L'opérateur revendique 2,5 millions d'assets sous surveillance de cette offre. A cette échelle, l'enjeu est de traiter les alertes au moindre coût, mais il est aussi de négocier au plus juste les licences avec les éditeurs d'EDR. En effet, Orange Cyberdefense a poussé la logique jusqu'au bout avec le lancement d'Orange Cybersecure en juin 2024.

Si on ajoute à ces poids lourds, un SFR ou encore de Docapost sous l'influence de Guillaume Poupard, les ESN régionales et petits acteurs de la Cyber vont devoir innover pour exister sur ce marché de la Cyber de masse. Ainsi, Alexis Missoffe, CEO de Winthorpe Company, a choisi de proposer aux PME un abonnement couvrant leur sécurité à 360°, avec la Cyber et une couverture d'assurance. *« Nous externalisons toute la problématique Cyber pour le compte de nos clients : organisation de tests de vulnérabilité, protection Cyber en H24 en mode supervisé. En cas d'alerte, nous effectuons la levée de doute ou nous déclenchons le traitement de la crise. S'il y a besoin, nous faisons intervenir l'assurance. Le but est vraiment de proposer une offre tout en un ».* □

A.C

Intelligence

Celonis joue sur l'écosystème et les agents IA

Lors de Celosphere, l'événement clients et partenaires de Celonis, l'éditeur a mis en avant différentes nouveautés, dont la possibilité d'utiliser des agents d'intelligence artificielle et d'étendre son concept de Process Intelligence dans l'écosystème de l'entreprise utilisatrice de sa plateforme.

Sans surprise, la principale annonce de l'éditeur a été l'intégration de l'intelligence artificielle dans sa plateforme et l'annonce d'AgentC. Avec les outils fournis, les entreprises peuvent développer leurs propres agents d'intelligence des processus. Certains agents préconstruits sont disponibles auprès de partenaires de l'éditeur. Tous ces agents d'IA sont alimentés par Celonis Process Intelligence, ce qui leur permet d'analyser et de comprendre l'organisation et le fonctionnement de l'entreprise, et de proposer des actions d'amélioration et d'optimisation. Celonis Process Intelligence fournit des données essentielles mises en perspective par le contexte commercial.

Les premières intégrations à la plateforme incluent Microsoft Copilot Studio, IBM watsonx Orchestrate, Amazon Bedrock Agents, et des environnements de développement open-source comme CrewAI. Les premiers agents d'intelligence artificielle préconstruits sont disponibles auprès des partenaires de Celonis : ISV, Rollio et Hypatos. De plus, lorsque les clients développent leurs agents, ils peuvent bénéficier du soutien de partenaires de conseil spécialisés tels qu'Accenture, IBM Consulting et EY. Un module, AI Assistant Builder, permet aussi aux clients de créer des assistants d'IA et des copilotes au sein de Celonis. Les assistants IA peuvent classer les données et fournir des recommandations, en raisonnant sur des tâches telles que la priorisation des tickets de service client ou l'évaluation du niveau de risque lié à la suppression d'un blocage de crédit qui bloque une commande.

Une logique d'écosystème

Networks étend la Process Intelligence en connectant les entreprises entre elles. Networks permet de créer un nouveau niveau de transparence des processus partagés au-delà des frontières d'une seule organisation, de sorte que chaque processus, au sein des entreprises et entre elles, puisse être optimisé. Un des principaux cas d'utilisation est la supply chain, où une collaboration efficace entre les fournisseurs et les clients est cruciale pour un processus fluide de bout en bout. La solution permet des échanges de données à l'échelle plus efficacement que par des solutions EDI classiques ou les simples échanges entre applicatifs par des API.

Alex Rinke, CEO de Celonis, dans sa session plénière d'ouverture lors de Celosphere 2024.

Outre ces nouveautés, la manifestation a fait une large place aux retours d'expérience des clients sur la solution. De nombreux grands comptes français étaient à la fois dans des ateliers ou des présentations, ou dans les salles pour suivre les débats. Cela marque clairement l'empreinte sur le marché qu'a pris Celonis dans notre pays en peu de temps, comparativement aux outils classiques de BPM (Business Process Management) par sa vision autour de l'optimisation des processus dans les grandes entreprises. □

B.G

AWS re:Invent L'IA et le cloud au service de la transformation numérique

AWS a présenté des outils majeurs lors de re:Invent (du 2 au 6 décembre à Las Vegas), dont Amazon Q Developer pour moderniser les systèmes hérités, Aurora DSQL pour des bases de données distribuées plus performantes, et les instances EC2 Trn2 pour l'IA avancée. Avec Amazon Nova, AWS a également élargi son offre en IA générative et renforcé Bedrock pour une personnalisation plus fiable des modèles. L'entreprise mise aussi sur des centres de données moins énergivores, consolidant son rôle dans la transformation numérique et durable.

Lors de la conférence, Amazon Web Services (AWS) a levé le voile sur une série d'innovations destinées à redéfinir les normes technologiques des entreprises. Ces annonces couvrent les domaines de l'intelligence artificielle générative, la modernisation des systèmes hérités, les bases de données distribuées et les infrastructures cloud. Avec des outils comme Amazon Q Developer, Aurora DSQL, les puces Trainium2 (et bientôt Trainium3), et Amazon Nova, AWS a voulu montrer que l'entreprise est un acteur clé de la transformation numérique, aidant les entreprises à réduire leurs coûts, à augmenter leur efficacité et à exploiter les opportunités offertes par la technologie de pointe.

Amazon Q Developer : les défis des systèmes hérités

La modernisation des systèmes hérités reste un défi majeur pour les entreprises. Ces systèmes, bien qu'essentiels, sont souvent obsolètes, coûteux à maintenir et difficiles à moderniser. Des milliers d'entreprises dans des secteurs tels que la finance, la santé et la logistique utilisent encore des technologies comme Windows. NET, VMware et les mainframes. Bien que solides et fiables, ces systèmes limitent souvent l'agilité et augmentent les coûts opérationnels. Par exemple, migrer des applications Windows. NET vers Linux peut prendre des mois, voire des années, en raison de la complexité des dépendances et des configurations spécifiques.

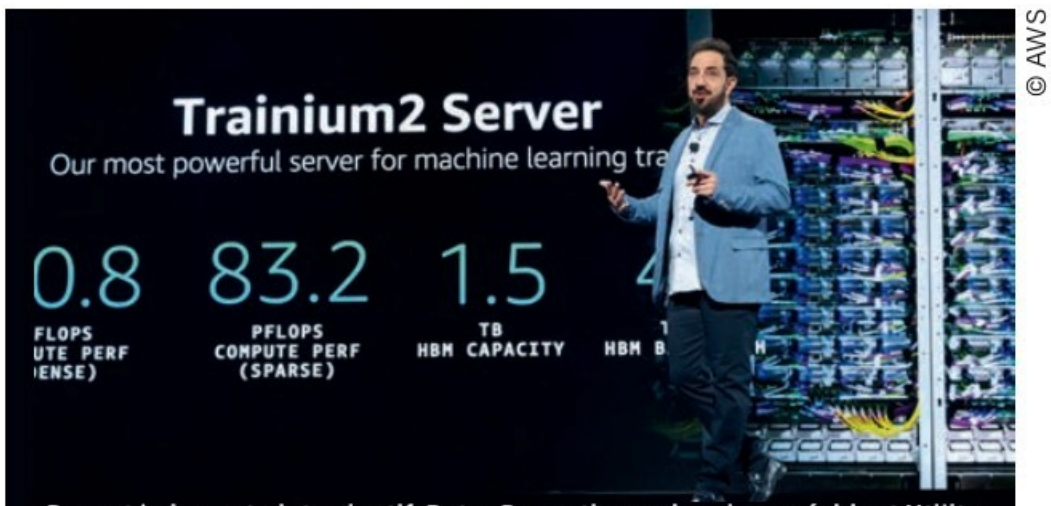
Avec Amazon Q Developer, AWS propose une solution qui combine l'intelligence artificielle générative et l'automatisation pour réduire les délais et les coûts de migration. *« Avec Amazon Q Developer, nous simplifions la transformation des systèmes hérités en éliminant la complexité et en accélérant les délais de modernisation »*, a déclaré Mai-Lan Tomsen Bukovec, vice-présidente de la technologie chez AWS.

Ainsi, Amazon Q Developer est conçu pour réduire drastiquement les délais de transformation des applications Windows. NET en environnements Linux. Ce changement, en plus de simplifier la gestion des systèmes, permet une réduction des coûts de licences de 40 % et ouvre la voie à de nouvelles pratiques de développement. Le processus de transformation est simple : le développeur sélectionne une application via l'interface d'Amazon Q dans son environnement de développement. Ensuite, des agents intelligents prennent en charge l'analyse des composants, génèrent des codes de correction, exécutent des tests et appliquent les changements nécessaires.

Conversion des charges VMware

Par ailleurs, Amazon Q Developer offre une solution rapide et automatisée pour convertir les charges VMware en architectures natives cloud. La technologie identifie automatiquement les dépendances et convertit les configurations réseau complexes en équivalents AWS.

Les mainframes, bien qu'utilisés dans des secteurs critiques comme la finance et les assurances, sont difficiles à moderniser en raison de leur complexité. Amazon Q Developer surmonte cet obstacle en automatisant des tâches telles que la documentation des programmes COBOL, l'analyse des connexions logiques et la décomposition des applications en composants cloud-natifs. Ces outils permettent aux entreprises de moderniser leurs systèmes en quelques semaines, au lieu de plusieurs années.



© AWS

Durant le keynote introductif, Peter Desantis, senior vice-président Utility Computing d'ASW, a dévoilé les UltraServers, des solutions pour répondre aux besoins croissants des modèles d'intelligence artificielle.

Aurora DSQL, une base de données SQL plus rapide

AWS a aussi annoncé une évolution majeure dans le domaine des bases de données avec l'introduction d'Aurora DSQL, une solution SQL distribuée qui offre une cohérence forte et une faible latence. Conçue pour répondre aux besoins des entreprises, Aurora DSQL redéfinit les normes en matière de bases de données distribuées. Selon AWS, Aurora DSQL garantit une cohérence forte pour toutes les transactions, quelles que soient les régions AWS concernées. Son architecture permet aux applications de lire et d'écrire des données depuis n'importe quelle région AWS, avec une synchronisation en temps réel entre toutes les copies des données.

© AWS

Aurora DSQL offre également des performances de lecture et d'écriture quatre fois supérieures à celles des bases de données distribuées traditionnelles. Son architecture serverless élimine le besoin de partitionner ou de provisionner manuellement les bases de données, permettant une scalabilité infinie et une gestion simplifiée. La clé de la performance d'Aurora DSQL réside dans sa technologie de synchronisation avancée, basée sur l'Amazon Time Sync Service. Ce service utilise des horloges atomiques et des satellites pour synchroniser les serveurs à travers le monde avec une précision de l'ordre de la microseconde, garantissant donc une faible latence. *« Aurora DSQL est un outil indispensable pour les entreprises cherchant à évoluer à l'échelle mondiale sans compromis sur la performance et la cohérence des données »*, a déclaré Ganapathy Krishnamoorthy, vice-président des services de bases de données chez AWS.

Instances EC2 Trn2 et nouveaux UltraServers

Durant re: Invent, AWS a présenté la dernière nouveauté de sa collaboration avec Nvidia. Matt Garman, directeur général d'AWS, a annoncé l'arrivée de la famille d'instances P6 qui sera dotée des derniers GPU Blackwell de Nvidia et qui sera disponible en 2025. AWS annonce une puissance de calcul jusqu'à 2,5 fois supérieure à celle de la génération actuelle de GPU. *« AWS et Nvidia collaborent depuis 14 ans pour s'assurer que nous excellons dans l'exploitation et le fonctionnement des charges de travail sur GPU »*, a déclaré Matt Garman à propos de ce partenariat.

Dans le même registre, AWS a introduit les nouvelles instances EC2 Trn2, alimentées par les puces Trainium2. Les instances Amazon EC2 Trn2 offrent une amélioration de 30 à 40 % des performances/prix par rapport aux GPU actuels (P5e et P5en). Chaque instance regroupe 16 puces Trainium2 et délivre une puissance de calcul de 20,8 pétaflops, rendant possible l'entraînement et le déploiement de modèles IA comportant des milliards de paramètres. *« Trainium2 est conçu pour gérer les charges de travail d'IA générative les plus vastes et les plus complexes, en offrant les meilleures performances au meilleur coût. Avec Trn2, nos clients peuvent entraîner et déployer les plus grands modèles au monde plus rapidement et à moindre coût »*, a expliqué David Brown, vice-président Compute et Networking d'AWS.

UltraServers, une solution pour les extrêmes

Pour les besoins les plus extrêmes, AWS introduit les Trn2 UltraServers, une nouvelle architecture reliant 64 puces Trainium2 via l'interconnexion ultra-rapide NeuronLink. Ces serveurs fournissent une capacité de calcul de 83,2 pétaflops, quadruplant les ressources d'une instance standard. Cette innovation permet aux entreprises d'entraîner et de déployer des modèles encore plus volumineux, tels que les modèles de langage de nouvelle génération contenant des trillions de paramètres, tout en réduisant le temps d'entraînement et en accélérant la mise sur le marché. On notera aussi que pour exploiter pleinement les puces Trainium2, AWS propose le SDK Neuron, compatible avec les frameworks populaires tels que JAX et PyTorch. Ce logiciel simplifie l'optimisation des modèles IA tout en minimisant les changements de code, facilitant ainsi l'adoption des instances Trn2.

Collaboration stratégique avec Anthropic

En collaboration avec Anthropic, AWS a construit un UltraCluster EC2 composé de centaines de milliers de puces Trainium2. Ce projet, baptisé Project Rainier, multipliera par cinq la puissance de calcul utilisée par Anthropic pour entraîner ses modèles actuels, comme Claude. Par ailleurs, AWS a renforcé sa collaboration avec plusieurs partenaires majeurs pour démocratiser l'accès aux modèles IA. Ainsi, Databricks utilisera Trn2 pour sa plateforme Mosaic AI, permettant à ses clients de personnaliser leurs modèles avec des données d'entreprise. De son côté, Hugging Face intègrera Trainium2 via la bibliothèque Optimum Neuron, offrant des outils améliorés pour développer et déployer rapidement des modèles. Enfin, Poolside utilisera les Trn2 UltraServers pour réduire ses coûts d'entraînement de 40 % tout en augmentant la vitesse d'exécution. Les instances Trn2 sont d'ores et déjà disponibles dans la région AWS US East (Ohio) et les Trn2 UltraServers seront déployés progressivement.

AWS a également évoqué Trainium3, une nouvelle puce de 3 nm. Prévue pour fin 2025, cette technologie sera quatre fois plus performante que les Trn2 UltraServers actuels, offrant une puissance et une efficacité énergétique sans précédent. Avec Trainium3, les clients pourront

construire des modèles encore plus grands et les déployer en temps réel à une vitesse très rapide.

Amazon Nova : Une IA générative accessible et puissante

Par ailleurs, AWS enrichit son offre de produits avec Amazon Nova, une gamme de modèles de fondation optimisés pour la rapidité, la précision et la polyvalence. Ces modèles, disponibles sur Amazon Bedrock, sont conçus pour répondre à un large éventail de besoins. Nova Micro propose ainsi un modèle texte ultra-rapide et économique, tandis que Nova Lite et Nova Pro sont des modèles multimodaux capables de traiter et de générer texte, images et vidéos. Enfin, Nova Canvas et Nova Reel ont été conçus pour produire des images et vidéos. Pour donner un exemple, Nova Reel permet aux entreprises de générer des vidéos, de six secondes dans un premier temps, à partir de simples descriptions textuelles, avec des options de personnalisation avancées pour le style et la mise en page. AWS prévoit d'étendre cette capacité à des vidéos de deux minutes, élargissant les cas d'utilisation dans les domaines de la publicité, du marketing et de la formation.

Distillation de modèles et personnalisation

Amazon Nova prend en charge la « distillation de modèles », une technique permettant de transférer les connaissances d'un modèle plus grand vers un modèle plus petit et plus efficace. Cela réduit les coûts et améliore la vitesse sans compromettre la précision. Cette capacité est intéressante pour les entreprises cherchant à optimiser leurs modèles pour des cas spécifiques tout en minimisant les coûts d'exploitation.

En outre, les modèles Nova sont entièrement intégrés à Amazon Bedrock, ce qui permet aux entreprises de les personnaliser facilement avec leurs propres données. Grâce à des outils comme Retrieval Augmented Generation (RAG), Nova peut s'appuyer sur les informations internes des entreprises pour produire des réponses précises et contextualisées. *« Nova met à la disposition des entreprises des capacités d'IA générative de pointe, tout en maintenant un coût et une complexité opérationnelle faibles »*, a déclaré Rohit Prasad, vice-président de l'intelligence artificielle chez Amazon.

Amazon Bedrock, une plateforme pour l'IA générative

Amazon Bedrock, une plateforme permettant de construire et de déployer des applications basées sur des modèles de fondation, s'est enrichie de fonctionnalités avancées pour faciliter le développement et l'adoption de l'IA générative.

© AWS

Pour cette édition 2024 de re: Invent, plus de 60 000 personnes étaient présentes.

AWS a introduit des outils de vérification automatique pour éviter les « hallucinations » des modèles d'IA, un problème où les réponses générées peuvent être incorrectes ou non factuelles. Ces vérifications permettent aux entreprises de s'assurer que les modèles restent conformes à leurs règles internes et produisent des résultats fiables, même pour des cas d'utilisation critiques comme la finance ou la santé.

Bedrock facilite désormais la création de systèmes multi-agents capables de collaborer sur des tâches complexes. Un agent peut, par exemple, analyser des données financières, tandis qu'un autre effectue une évaluation concurrentielle. Ces agents travaillent ensemble sous la supervision d'un agent principal, coordonnant les actions et consolidant les résultats pour produire une vue complète et cohérente.

Soutenir l'innovation avec des centres de données

Outre ses avancées dans l'intelligence artificielle et en modernisation, AWS a également présenté des innovations dans la conception de ses centres de données, pour répondre aux besoins croissants en puissance de calcul tout en réduisant leur impact environnemental. Pour soutenir les charges d'IA haute densité, AWS a intégré des systèmes de refroidissement liquide innovants, capables de gérer efficacement la chaleur générée par des puces comme Trainium2 et les serveurs de calcul. Ces systèmes combinent refroidissement à air et refroidissement liquide pour maximiser les performances tout en minimisant la consommation d'énergie.

AWS s'est engagé à réduire son impact environnemental en adoptant des matériaux durables et en optimisant la conception structurelle de ses centres de données. Les nouvelles infrastructures réduisent la consommation énergétique mécanique de 46 % et l'empreinte carbone incorporée dans le béton de 35 % par rapport aux normes industrielles. □

M.C

OpenText World

Une nouvelle version et Titanium X

Lors d'OpenText World 2024 (du 18 au 21 novembre à Las Vegas), OpenText a dévoilé Cloud Editions 24.4, et annoncé l'achèvement de son projet Titanium X, combinant cloud, intelligence artificielle et cybersécurité pour répondre aux besoins croissants des entreprises. La nouvelle plateforme se concentre sur le multi-cloud, l'automatisation, l'expérience client et une sécurité renforcée.

Durant OpenText World 2024, la société canadienne a dévoilé une mise à jour majeure de sa plateforme, Cloud Editions 24.4, et annoncé l'achèvement de son ambitieux projet Titanium X, un ensemble intégré de suites de gestion de l'information aux environnements cloud publics et privés. Avec une combinaison de technologies avancées en cloud, en intelligence artificielle et en cybersécurité, OpenText entend s'imposer comme un acteur incontournable dans la transformation numérique des entreprises modernes.

Dans un contexte où les environnements numériques deviennent de plus en plus complexes et interdépendants, OpenText mise ainsi sur le multi-cloud et l'IA pour offrir des solutions intégrées qui facilitent la gestion des données, automatisent les flux de travail et libèrent le potentiel humain. Mark Barrenechea, PDG et directeur technique d'OpenText, a résumé cette vision lors de son discours d'ouverture durant un premier keynote. *« Les organisations possèdent deux atouts uniques — le talent et les données. En réimaginant la manière dont nous gérons et utilisons ces ressources grâce à des solutions intégrées d'IA et de cloud, nous pouvons libérer un potentiel humain sans précédent »*.

Surmonter les obstacles du multi-cloud

Adopté par une majorité d'entreprises, le multi-cloud est devenu un pilier stratégique pour réduire la dépendance à un seul fournisseur, accroître la flexibilité et renforcer la résilience. Cependant, cette multiplicité de plateformes pose des défis de connectivité, de sécurité et de conformité. Avec Cloud Editions 24.4, OpenText cherche à surmonter ces obstacles en offrant une intégration fluide entre les différents environnements cloud, sans déplacer les données. Muhi Majzoub, vice-président exécutif et directeur produit chez OpenText, a souligné l'importance de cette approche. *« OpenText Cloud Editions 24.4 est conçu pour autonomiser les employés en alliant IA et connectivité sécurisée des données. Alors que l'IA devient un collaborateur clé au sein des organisations, cette version offre aux dirigeants les capacités nécessaires pour faire évoluer leurs opérations durablement, gagner un avantage compétitif et obtenir des résultats dans le futur »*.

Améliorer l'expérience des clients

L'un des axes principaux de Cloud Editions 24.4 est l'amélioration de l'expérience client à travers OpenText Experience Cloud, un portefeuille d'outils intégré qui repense les interactions entre les entreprises et leurs clients. OpenText Communications (Exstream), par exemple, facilite les transitions vers le cloud en offrant des accélérateurs de conception et des capacités d'orchestration des communications à grande échelle. Les entreprises peuvent ainsi optimiser leurs campagnes publicitaires et leurs échanges avec les clients via des emails, SMS et autres supports numériques.

Avec OpenText Core Messaging, les interactions clients deviennent encore plus riches grâce à la prise en charge des Rich Communication Services (RCS), notamment via WhatsApp ou SMS enrichis, s'adaptant aux préférences des utilisateurs finaux. Les secteurs sensibles comme la santé, la finance ou les administrations publiques bénéficieront également de Core Fax, une solution qui migre les serveurs de fax locaux vers un modèle cloud sécurisé. Cette évolution garantit la continuité des services critiques tout en respectant les normes de sécurité et de confidentialité.

© Open Text

Mark Barrenechea, PDG et directeur technique d'OpenText, durant le keynote introductif d'OpenText World 2024 à Las Vegas, fin novembre.

Automatisation et amélioration de la productivité

Au cœur de Cloud Editions 24.4, l'intelligence artificielle joue un rôle central avec le déploiement de plus de 100 agents intelligents intégrés dans l'écosystème OpenText Aviator. Ces outils permettent d'automatiser les processus et d'améliorer la productivité dans tous les domaines de l'entreprise. Parmi les innovations notables, la société canadienne a développé OpenText Content Aviator, un outil qui facilite l'accès aux données archivées grâce à des interfaces conversationnelles intuitives basées sur des modèles de langage avancés. On peut aussi évoquer OpenText Intelligence Aviator, dont la mission est de simplifier la prise de décision grâce à une interaction en langage naturel avec les données analytiques, rendant les analyses accessibles à tous. Enfin, OpenText propose aussi OpenText DevOps Aviator, un outil qui utilise l'IA pour convertir des séquences vidéo en tests manuels, accélérant ainsi les délais de développement tout en améliorant la qualité.

Cybersécurité et conformité renforcées

Dans un contexte où les cybermenaces sont en constante évolution, Cloud Editions 24.4 introduit des fonctionnalités de sécurité avancées pour protéger les données sensibles et garantir la conformité des entreprises. Ainsi, la solution Threat Detection and Response apporte des outils puissants pour identifier et neutraliser les menaces en temps réel. Par ailleurs, l'API Trust d'OpenText permet de classifier les données sensibles et de vérifier leur conformité avec les réglementations en vigueur.

En matière de gouvernance, OpenText a franchi deux étapes majeures. OpenText Content Management a obtenu la certification Protected B du gouvernement canadien, autorisant son utilisation pour des données sensibles dans le cloud public AWS. De son côté, la plateforme IT Management (ITMX) d'OpenText a été certifiée FedRAMP, permettant son adoption par les agences fédérales américaines. Pour finir, Mark Barrenechea et les équipes en charge de la sécurité, dont Stephan Jou, le senior directeur de la sécurité analytics, ont aussi évoqué l'évolution des moyens de sécurité. Sans trop entrer dans les détails, la quasi-fin du mot de passe est un sujet qu'OpenText a validé avec l'intégration de la biométrie pour se substituer aux mots de passe et assurer une meilleure protection des données et des infrastructures des clients.

Aider les développeurs au quotidien

OpenText ne s'adresse pas uniquement aux DSI et aux responsables IT. Avec Cloud Editions 24.4, les développeurs bénéficient de nouvelles fonctionnalités qui rationalisent les flux de travail et améliorent la productivité. Parmi les nouveaux outils, AI-Assisted User Stories génère automatiquement des récits utilisateur basés sur les besoins des projets, réduisant le temps consacré à la planification.

L'intelligence artificielle joue un rôle central dans Cloud Editions 24.4, avec le déploiement de plus de 100 agents intelligents intégrés dans l'écosystème OpenText Aviator. Ces outils permettent d'automatiser les processus et d'améliorer la productivité dans tous les domaines de l'entreprise.

Video-to-Defect Translation, une autre innovation importante, transforme des séquences vidéo en étapes de test exploitables, simplifiant le travail des équipes. En outre, l'OpenText AI Data Cloud permet aux développeurs de créer des applications robustes et conformes, exploitant pleinement la puissance de l'IA tout en respectant les normes de sécurité. Il faut souligner l'importance d'OpenText Core Application Observability, qui vient améliorer l'observabilité des applications cloud. Cela permet aux développeurs de détecter et résoudre efficacement les problèmes liés à l'infrastructure, aux applications et au réseau, tout en optimisant les environnements cloud.

La GenAI en soutien des décisions

L'intelligence artificielle générative (GenAI) est l'une des pierres angulaires de Cloud Editions 24.4. OpenText Intelligence Aviator exploite cette technologie pour permettre aux utilisateurs d'interagir directement avec les données analytiques en utilisant un langage naturel. Les équipes peuvent ainsi poser des questions complexes sans avoir besoin de maîtriser des outils d'informatique décisionnelle traditionnels, obtenant des insights et des visualisations en quelques clics. Mark Barrenechea a résumé cette ambition en ces termes : *« Grâce à l'IA, nous donnons à la prochaine génération de travailleurs les moyens de réaliser leur plein potentiel, leur permettant de se concentrer sur les tâches stratégiques et à fort impact. Il est temps de laisser la technologie gérer les détails, afin que vos équipes puissent se concentrer sur ce qui compte »*.

Avec CE 24.4, OpenText ne se contente pas d'apporter des solutions immédiates aux entreprises. La société poursuit une stratégie à long terme, avec des mises à jour régulières, tous les 90 jours, pour intégrer de nouvelles capacités, notamment dans le domaine de l'intelligence artificielle. Ainsi, lors d'un second keynote, les équipes techniques ont présenté la feuille de route des prochaines mises à jour avec les introductions progressives de Cloud Editions 25.1, 25.2, 25.3 et enfin 25.4. Ces mises à jour seront disponibles progressivement en janvier, avril, juillet et octobre. □

M.C

Infrastructure as Code

HashiCorp repense Terraform

Présenté par HashiCorp, lors de la HashiConf de Boston du 14 au 16 octobre 2024, Terraform Stacks doit simplifier la gestion et le déploiement des configurations Terraform interdépendantes dans les charges de travail Infrastructure as Code (IaC).

Traditionnellement concentré sur la gestion des cycles de vie de l'infrastructure (ILM) et de la sécurité (SLM), HashiCorp est resté fidèle à lui-même lors de sa conférence annuelle, HashiConf, qui s'est tenue à Boston (États-Unis). L'occasion pour le spécialiste de l'automatisation de l'infrastructure de présenter une flopée de nouveaux produits et fonctionnalités sur ces deux volets, avec un gros focus sur son outil phare, HashiCorp Terraform.

Mieux gérer les configurations interdépendantes

Dans le domaine de l'ILM, l'entreprise a annoncé la bêta publique de Terraform Stacks, une évolution qui *« apporte une incroyable richesse, à commencer par la capacité à modéliser des environnements complexes »*, s'est enthousiasmé Armon Dadgar, le cofondateur et CTO de HashiCorp. Concrètement, ses principales capacités consistent à pouvoir coordonner, déployer et gérer efficacement les configurations Terraform interdépendantes dans des flux de travail IaC.

Armon Dadgar a illustré ces capacités avec un exemple concret : dans un scénario classique de déploiement, la création d'un cluster EKS (Elastic Kubernetes Service) sur Amazon, suivie de l'installation d'une application sur ce même cluster nécessitait deux configurations distinctes. Avec Terraform classique, le déploiement de l'application échouait fréquemment si le cluster EKS n'était pas créé avant. *« Désormais, vous pouvez modéliser ces deux composants ; Terraform Stacks comprend automatiquement qu'ils sont interdépendants »* — à savoir que l'application dépend du cluster Kubernetes — *« et [Terraform Stacks] comprend alors qu'il doit créer le premier composant avant le deuxième »*, poursuit Armon Dadgar. Grâce à cette avancée, les équipes n'ont plus à gérer manuellement les relations complexes entre configurations, réduisant ainsi le risque d'erreurs humaines.

Armon Dadgar,
le cofondateur et CTO de HashiCorp

« [Terraform Stacks] apporte une incroyable richesse, à commencer par la capacité à modéliser des environnements complexes »

Automatiser la migration des configurations Terraform

Les utilisateurs peuvent, en outre, déployer plusieurs fois la même infrastructure avec des variations selon les exigences — par exemple dans le cas de déploiements dans plusieurs régions (États-Unis, Europe), zones de disponibilités ou comptes de fournisseurs de cloud. L'application des variations sur ces infrastructures répétées peut être automatisée grâce à des règles d'orchestration.

Toujours par souci de simplification, HashiCorp a également rendu disponible en bêta publique Terraform Migrate. Cet outil automatise la migration des configurations de Terraform vers d'autres instances de Terraform, comme HCP Terraform ou Terraform Enterprise. Par la même occasion, il encourage la migration vers des versions plus robustes et évolutives de Terraform. □ **V.M**

UNE **GESTION RENFORCÉE** DES SECRETS

Second volet pour HashiCorp : la sécurité a aussi eu droit à son florilège de nouveautés, notamment pour ce qui concerne la gestion des secrets au sens des certificats numériques, des identifiants de bases de données, des mots de passe et des clés de chiffrement d'API. L'entreprise a présenté des fonctionnalités axées sur l'identité pour renforcer la sécurité de l'accès aux ressources, telles que la disponibilité générale d'une fonctionnalité de HCP Vault Secrets : la rotation automatique, qui permet le renouvellement des secrets (identifiants) après un certain temps ou selon une règle définie, pour HCP Terraform. Les secrets dynamiques et les informations d'identification dynamiques sont, quant à eux, accessibles en bêta. Ils génèrent des secrets et informations d'identification temporaires et à la demande. La bêta publique de HCP Vault Radar automatise, pour sa part, l'analyse des secrets non gérés, et fournit des conseils et mesures de remédiation aux équipes de sécurité pour éviter leur prolifération.

Stratégie

Broadcom accélère sur le cloud privé avec VMWare

Lors de VMware Explore 2024 à Barcelone, Broadcom a présenté les fonctionnalités de VMware Cloud Foundation (VCF), visant à renforcer son offre de cloud privé, alors même qu'une part non négligeable de ses clients envisage de se passer de ses solutions. L'entreprise a insisté sur l'importance d'un cloud privé unifié, pour offrir aux entreprises une meilleure maîtrise de leurs données et de leurs coûts, tout en répondant aux exigences de conformité.

Le mot d'ordre de cette édition de VMware Explore 2024 à Barcelone était le cloud privé unifié, afin d'optimiser les coûts et l'exploitation des données privées des entreprises, tout en se démarquant clairement du cloud public. *« Aujourd'hui, vous êtes confrontés aux conséquences des 3C que j'associe au cloud public. À savoir : le coût — le cloud public finit toujours par coûter plus cher que prévu ; la complexité — c'est une plateforme supplémentaire à gérer, ce qui n'est pas toujours simple ; et la conformité — répondre aux exigences réglementaires devient de plus en plus difficile avec l'utilisation du cloud public »*, a déclaré Hock E. Tan, directeur général de Broadcom, lors de la session d'ouverture de l'événement. En conséquence, 83 % des CIO envisagent de rapatrier leurs environnements du cloud public vers leurs propres infrastructures, avance Broadcom, à la lumière d'une enquête récente menée auprès de CIO par une grande banque mondiale. Selon Hock E. Tan : *« L'avenir de l'entreprise est privé : un cloud privé, une IA privée à l'échelle. Il s'agit de rester sur place, de garder le contrôle et, bien sûr, de continuer à utiliser le cloud public si nécessaire pour l'élasticité et la surcharge des charges de travail »*.

Reste à savoir si ces départs vers le cloud privé profiteront à VMware ou à la concurrence, Microsoft en tête. Broadcom fait, en effet, l'objet de vives critiques depuis l'acquisition de VMware. Forrester Research, prédisait qu'en 2023, 20 % des clients de VMware quitteraient l'écosystème de l'entreprise. La raison ? Des augmentations de prix significatives depuis le rachat, et un support de moins en moins efficace selon Forrester. *« Par conséquent, et malgré la domination de l'entreprise dans ces technologies, de nombreux clients entreprises de VMware explorent des alternatives à ses produits de virtualisation, de gestion du cloud, de calcul pour utilisateurs finaux et d'infrastructure hyperconvergée »*, écrivait Michele Pelino, analyste principal du cabinet de recherche et de conseil.

Pour ceux qui resteront, Broadcom renforce son offre, afin qu'ils puissent bénéficier d'une plateforme complète, intégrant des briques de conformité, de sécurité et d'intelligence artificielle. À Barcelone, la multinationale américaine a ainsi présenté, ou représenté, plusieurs nouveautés intégrées à sa plateforme VCF,

LA SOUVERAINETÉ POUR SÉDUIRE LES EUROPÉENS

La souveraineté numérique et le cloud souverain représentent des arguments commerciaux de poids pour les entreprises qui souhaitent séduire la clientèle européenne, ou du moins ne pas se mettre à dos leurs partenaires locaux soucieux de savoir ce qu'il advient de leurs données une fois dans le cloud. Le cloud privé est présenté par certains acteurs comme une solution souveraine, car il permet aux entreprises de gérer leurs données en interne (on-premise) sans dépendre des régulations externes ni craindre le droit extraterritorial de certains pays. Broadcom, qui a récemment acquis VMware, fait partie de ceux-là lorsqu'il promeut VCF 9. *« VCF fournit la solution ultime de cloud souverain. Nous vous offrons un fournisseur de cloud national, localisé dans votre pays spécifique, délivrant des services de cloud souverain qui respectent les lois nationales »*, a assuré Hock E. Tan, lors du VMware Explore 2024 à Barcelone.
L'entreprise a expliqué que cinquante de ses fournisseurs de services cloud VMware (VCSP) proposent des services de cloud souverains, dont trente pour la seule région Europe, Moyen-Orient et Afrique. Il a assuré que ses partenaires VCSP répondaient à l'exigence de traitement local des données via une entité juridique. Broadcom collabore, en outre, avec des partenaires européens comme OVHcloud, pour offrir des solutions de cloud privé souverain, tout en permettant une certaine flexibilité d'hébergement, y compris avec des hyperscalers, mais sous certaines conditions. Rappelons que ce sont ces derniers qui cristallisent l'essentiel des craintes autour la confidentialité des données dans le cloud. Si bien qu'AWS ou encore Microsoft travaillent eux-mêmes à développer leurs clouds souverains.

déployable sur site dans les data centers, dans des clouds hyperscale ou à la périphérie.

Réduire les coûts d'exploitation

Pour rappel, VMware a combiné plusieurs composants de sa plateforme (vSphere, vSAN, NSX, etc.) pour créer une plateforme unifiée, baptisée VMware Cloud Foundation (VCF) 9. Cette version, qui n'a pas encore été lancée sur le marché, doit permettre aux clients de déployer un cloud privé complet, avec tous les composants nécessaires à l'infrastructure, à l'automatisation et aux opérations.

Broadcom a renforcé son offre de cloud privé, dévoilant, lors de VMware Explore 2024, de nouvelles solutions pour améliorer la gestion des données, la sécurité et l'intelligence artificielle.

VCF9 embarque l'extension des capacités de gestion des données avec VMware Tanzu Data Services, intégrée nativement à VCF. Cette fonctionnalité doit faciliter l'harmonisation, le déploiement, la gestion et la consommation des données critiques, tout en renforçant la sécurité et la gouvernance des données. Broadcom a d'ailleurs annoncé le lancement prochain de Tanzu Platform 10 pour les environnements autogérés et isolés — cette solution a déjà été présentée lors de VMware Explore, à Las Vegas, en août dernier. Bien qu'elle ait été optimisée pour les clouds privés estampillés VCF, elle peut également être déployée dans un cloud public.

Autre nouveauté : la résilience après une cyberattaque et la récupération d'urgence, avec l'extension de la prise en charge de VMware Live Recovery à Google Cloud VMware Engine (GCVE). Ce dernier peut être utilisé comme environnement de reprise isolée (IRE) pour les charges de travail VCF. Il s'appuie sur la protection VMware Live Recovery des sites GCVE, et vient s'ajouter aux solutions VMC on AWS et IRE sur site. Broadcom a également étendu son programme de modernisation en intégrant un outil de maturité et d'optimisation du cloud privé dans VCF, ainsi qu'une nouvelle certification VMware Cloud Foundation.

Gérer les charges de travail d'IA

Broadcom a, en outre, dévoilé VeloRAIN, une nouvelle solution dans son portefeuille VeloCloud. Celle-ci vise à accélérer et optimiser les charges de travail de mise en réseau IA. Concrètement, elle améliore la gestion des réseaux pilotés par l'IA, en priorisant les applications et le trafic, promet Broadcom. Et ce, afin de résoudre les problèmes d'encombrement des réseaux entre sites qui freinent l'adoption des solutions de périphérie et des charges de travail IA dans les infrastructures edge distribuées.

Deux nouvelles appliances VeloCloud Edge 4100 et 5100, conçues pour les grandes organisations et les centres de données, optimiseront, quant à elles, les performances et la résilience pour les charges de travail à grande échelle.

Broadcom a aussi annoncé le lancement du Broadcom Advantage Partner Program. Destiné aux fournisseurs de services gérés (MSP) et aux partenaires VeloCloud Titan, ce programme aidera les MSP à fournir à leurs clients le portefeuille VeloCloud SD-WAN en services gérés.

Sécurité et IA

Quelques autres annonces ont été faites en matière de sécurité. Broadcom a introduit une nouveauté dans VMware vDefend avec vDefend Intelligent Assist, un outil propulsé par l'IA générative, qui vise à améliorer la détection, l'analyse et la neutralisation des menaces. Il intègre également un copilote pour réduire les faux positifs et le nombre d'alertes. Parallèlement, VMware Avi Load Balancer a été amélioré pour optimiser l'équilibrage de charge dans les environnements VCF et Kubernetes.

Broadcom progresse aussi dans sa stratégie autour de l'intelligence artificielle, visant à offrir plus de flexibilité aux clients dans le choix du matériel, des modèles et des logiciels utilisés pour déployer des plateformes d'IA. Dans cette optique, l'entreprise a annoncé la prise en charge du service Azure AI Video Indexer sur les plateformes VMware Private AI. Pour rappel, Azure AI Video Indexer est un service de Microsoft qui analyse les vidéos et l'audio, et gère les charges de travail d'IA générative, soit dans des centres de données, soit en périphérie. □

V.M

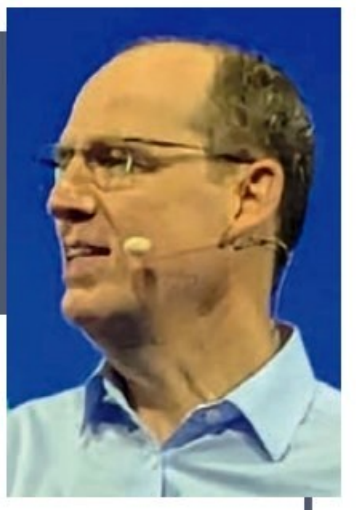
OVHcloud Summit 2024

Le nordiste se rêve en hyperscaler

OVH a 25 ans cette année. Le petit hébergeur du nord, spécialiste de l'hébergement Web et des serveurs Bare Metal, se rêve aujourd'hui en hyperscaler. Rebaptisé OVHcloud, il mise sur ses partenariats avec AMD, Broadcom, Nvidia et Nutanix pour conquérir les entreprises.

Si Michel Paulin a cédé la place de directeur général à Benjamin Revcolevschi, c'est bien Octave Klaba, l'emblématique président d'OVHcloud, qui a lancé l'édition 2024 de l'OVHcloud Summit. Incorrigible fan de technologie, il n'a pu céder au plaisir de présenter son dernier joujou, la On-Prem Cloud Platform, tout OVHcloud dans un demi-rack. Au-delà de ce tour de force d'intégration, OVHcloud entame un nouveau chapitre de son histoire.

Benjamin Revcolevschi a souligné qu'OVHcloud réalise un chiffre d'affaires mondial d'1 milliard de dollars, qu'il est présent dans 140 pays et que la moitié de ses 43 datacenters sont localisés hors de France. *« Aujourd'hui, plus de la moitié de notre chiffre d'affaires est réalisé à l'international, et l'un des secrets les mieux gardés est que notre deuxième source de CA après la France sont les USA ! »*



Benjamin Revcolevschi, Directeur général d'OVHcloud

« À la question comment faites-vous pour lutter contre les GAFAM et proposer les solutions dont nous avons besoin, ma réponse est simple : construire une proposition de valeur alternative qui plait à nos clients. Elle plait, car vous trouvez chez nous l'essentiel des services de Cloud privé, de Cloud public et de Web Cloud »

Le DG estime qu'OVHcloud propose déjà 90 % des services cloud les plus consommés.

OVHcloud en pointe sur les puces AMD

En termes d'annonces produit, si OVHcloud a pris le virage du logiciel, on n'oublie pas que pour faire tourner les 40 produits cloud au catalogue du Roubaisien, il faut du silicium, beaucoup de silicium. OVHCloud a été le premier dans le monde à lancer des AMD Epyc 4000 sur son offre Bare Metal et les nouveaux Ryzen 9000 seront disponibles d'ici quelques semaines dans la gamme RISE. Les AMD Epyc de cinquième génération « Turin » et leurs 192 cœurs/384 threads arrivent, soit potentiellement 36 000 vCPU dans un seul rack ! Côté GPU, OVHcloud lorgne sur les Nvidia H200NLV et Blackwell, et les AMD Instinct MI325X d'ici quelques semaines (ou quelques mois pour certains).

Sur l'offre Hosted Private Cloud, OVHcloud s'appuie sur ses partenariats avec Broadcom / VMware et avec Nutanix. Le Français est maintenant partenaire Pinnacle de Broadcom et une nouvelle gamme de solutions ex-VMware arrive. L'offre Public VCF (VMware Cloud foundation) arrivera en mode « as a Service » pour les clients existants d'OVHcloud dès le mois de janvier 2025, puis pour tous en mai 2025. Côté Nutanix, ça bouge aussi avec une offre Nutanix sur la gamme serveur Scale depuis le mois de novembre 2024.

Côté logiciels, la Data Platform, déjà accessible en bêta, sera officiellement lancée en février prochain. S'appuyant notamment sur Apache Iceberg, Spark et Trino, elle offre déjà une cinquantaine de connecteurs. OVHcloud a aussi annoncé la disponibilité officielle de son offre AI endpoint, un outil destiné aux développeurs et destiné à consommer facilement des inférences de modèles. Une quarantaine de modèles et LLM sont d'ores et déjà accessibles via l'outil. Par ailleurs, Octave Klaba a présenté Omisimo, une nouvelle approche vis-à-vis des LLM. Cette API joue le rôle de routeur de requêtes entre les LLM Open Source. *« Une voie alternative à la course aux LLM de plus en plus gros »*, estime Octave Klaba.

Enfin, Ranger rejoint la liste des solutions managées au catalogue d'OVHcloud. Managed Rancher Service by Suse permet de provisionner une infrastructure Rancher en quelques clics. Celle-ci va pouvoir gérer des clusters Kubernetes où qu'ils se trouvent au moyen d'un simple dashboard. □

A.C

PRA

L'informatique des radiologues Imasud se relève d'un incendie

Après l'incendie de la clinique Sainte-Marguerite de Hyères, Florian Martin, le DSI d'Imasud, a soigné la sécurité physique des serveurs et la redondance des liens de 15 sites de radiologie dans le Var.

Quelles sont les applications critiques de vos utilisateurs ?
Florian Martin : Nos deux applications principales sont le RIS — le système d'information radiologique qui inclut le dossier patient — et notre PACS (Picture Archiving and Communication System). Sans elles, l'activité du groupe devient très difficile. Plusieurs services de post-traitement se greffent dessus pour améliorer les clichés et conforter les diagnostics par l'IA.

En mai 2024, la clinique Sainte-Marguerite a subi un incendie important. Qu'est-ce qui a été endommagé ?
FM : L'incendie s'est produit le vendredi 25 mai vers 3 h du matin. Il n'y a pas eu de blessé, mais le rez-de-chaussée a subi de gros dégâts matériels, en particulier le service radiologie et une partie de l'IT, qui pour des raisons historiques, était encore hébergée sur ce site. Avec la chaleur, plusieurs matériels ont fondu (dont les façades de certains de nos serveurs), mais heureusement les onduleurs positionnés en bas n'ont pas explosé. L'absence de combustion au niveau des serveurs nous a permis de récupérer 80 % des disques immédiatement et d'envoyer en récupération les 20 % endommagés.

Comment s'est déroulée la reprise d'activités ?
FM : Vers 9 h nous avons réalisé un état des lieux de la situation, mis en lumière les priorités et établi une stratégie de remise en marche de l'activité. Nous avions un plan de reprise d'activités reposant sur d'anciens serveurs, mais on ne l'avait jamais testé faute de temps. Suite à l'incendie, il nous fallait rebâtir des services d'infrastructure, adapter notre routage réseau, et surtout remonter notre PACS principal devenu inopérant pour l'ensemble des sites. Nous avons décidé de repartir sur une base saine. Nous avions commandé de nouveaux serveurs qui sont arrivés par chance quelques jours plus tôt. Grâce à leur disponibilité et performances, nous avons pu rapidement restaurer les sauvegardes déportées qui s'étaient correctement effectuées avant l'incendie. L'incendie a eu lieu un weekend. Nos deux seuls sites recevant du public le samedi matin ont travaillé en mode dégradé, à l'ancienne. Radiologues et hôtesses d'accueil se sont adaptés à cette situation, tandis que l'équipe IT préparait le redémarrage d'activités sur les nouveaux serveurs.

FICHE D'IDENTITÉ

Raison sociale : Imasud (groupe Vidi)
Activités : Radiologie, imageries médicales
Applications : RIS, PACS, dépistages assistés par l'IA, portail de diffusion de comptes-rendus
Chiffres d'affaires : 30 millions d'euros
Effectif : 250 utilisateurs, dont 55 radiologues
Equipe IT : 5 personnes

Finalement, vous avez modernisé une grande partie de l'infrastructure en un week-end ?
FM : Notre système d'information était déjà réparti sur 16 sites. Sur chacun, nous avons un serveur d'authentification et un serveur cache du système PACS, grâce auquel nous avons pu repartir et restituer immédiatement les données d'imagerie les plus récentes des patients, les plus utiles à court terme. La consolidation s'effectue au fil de l'eau entre chaque site, grâce aux liens en fibre optique. Plutôt que de redémarrer notre PACS principal sur son site initial, il nous a fallu trouver une solution dans l'urgence. Nous avons décidé d'héberger les nouveaux serveurs virtualisés sous VMware, dans un premier temps au siège d'Imasud. En novembre, ils ont rejoint le datacenter XL360 certifié HDS au centre de Toulon, où nos liens intersites sont gérés. Les équipes d'astreinte du datacenter XL360, de Koesio et de notre PACS Infinitt ont fait preuve d'un professionnalisme exemplaire en étant disponibles et performantes lors de la relocalisation et du redémarrage de notre infrastructure.

La reprise complète des données a mis plusieurs mois. Pourquoi ?
FM : Nous avons pu accéder au site de la clinique incendiée fin août seulement, pour des raisons d'enquête administrative. C'est seulement à ce moment-là que nous avons pu constater l'état de la baie informatique. Les disques durs ont dû être nettoyés et décontaminés par un prestataire mandaté par l'assurance, puis envoyés pour récupération dans une autre entreprise. En novembre 2024, nous étions toujours en attente de certaines données.

Quelle leçon retenez-vous de cette expérience ?
FM : Plusieurs en réalité. J'avais davantage anticipé l'inondation que l'incendie. C'est la raison pour laquelle j'avais surélevé les serveurs. Mais, avec l'incendie, on se rend compte que le danger vient aussi d'en haut, à cause de la chaleur. Une bonne pratique consiste selon moi à placer

ses serveurs les plus critiques au milieu des racks. Dans un datacenter de colocation, les protections en place font qu'il y a peu de chance qu'il se passe quelque chose de grave. Les sauvegardes déportées nous ont sauvés ; encore faut-il s'assurer qu'elles se déroulent bien. La souscription d'une assurance multirisque demeure importante. Les personnes impliquées et les prestataires externes doivent être disponibles. Nous le disons toujours, mais le négligeons souvent, il faut passer à la pratique. C'est la seule façon concrète de vérifier le PRA et de l'améliorer. En l'occurrence, après une heure de restauration de sauvegarde à distance et malgré des liens gigabits entre les sites, j'ai pris la décision d'aller chercher physiquement les sauvegardes pour les ramener en local. Ce pari nous a fait perdre 2 heures, puis gagner 3 jours. Rien ne vaut des équipements connectés sur un seul et même switch !

La restauration d'infrastructure au lendemain de l'incendie a-t-elle restitué l'intégralité des usages métiers ?

FM : Nous avons réussi la migration en un jour et demi. Les utilisateurs n'ont pas vu la différence le lundi matin et ont pu travailler normalement, ce qui prouve qu'elle a été efficace. Certaines fonctionnalités non-essentielles n'ont néanmoins pas pu être restaurées immédiatement, car elles nécessitaient un matériel spécifique qui est arrivé quelques semaines plus tard.

Comment s'est déroulée la migration vers le datacenter ?

FM : Initialement, nous avions prévu de préparer notre infrastructure durant l'été pour migrer à l'automne. Avec l'incendie et la mise en production éclair des serveurs, il a fallu réorganiser la migration sur une sorte de jeu de chaises musicales car une bonne partie des ressources était en production. Nous avons profité du weekend du 11 novembre pour finaliser ce projet et tester toutes les redondances mises en place. Le matin de la reprise d'activité, nous avons eu quelques effets de bord mineurs durant 30 minutes car nous en avons profité pour renforcer la sécurité, le filtrage intersites et normaliser les filtrages de chaque site. En cas de nouvelle panne, les redondances mises en place au niveau de l'infrastructure s'avèrent rassurantes.

Avez-vous souffert de matériels obsolètes ?

FM : Certains de nos serveurs avaient jusqu'à 9 ans, mais deux fusions successives ont retardé leur renouvellement. L'obsolescence des serveurs commençait à nous freiner pour accompagner la croissance du groupe. Petite anecdote à ce propos : la carte-mère d'un ancien serveur du service RH nous a lâchés un matin, plutôt que de restaurer la VM sauvegardée la veille et perdre quelques heures de données, notre fournisseur de matériel local, CMS informatique, avait par chance en stock le même modèle dans ses locaux à quelques kilomètres de là, ce qui nous a permis de redémarrer en un temps record et sans aucune perte de données. En effet, il a suffi de remonter nos disques dans cet ancien serveur.

Un projet de télétravail a-t-il été mis en place ?

FM : Plus de la moitié de nos 55 radiologues étaient demandeurs initialement et le nombre a grandi par la suite. Lorsqu'ils sont d'astreinte le week-end, au lieu de se déplacer, ils peuvent se connecter de chez eux, depuis un poste d'interprétation médicale, via un VPN avec double authentification. Ils peuvent ainsi interpréter en quelques minutes les images et éviter un temps de transport qui peut être très long selon la saison. Le bénéfice pour le patient est indéniable. Ce projet a duré deux mois en tout, grâce à une solution MFA clé en main et du matériel médical. En cas de nouvelle pandémie, cela sera précieux.

Pour les logiciels d'IA, quelles précautions spécifiques recommandez-vous ?

FM : Nous utilisons pour le moment trois logiciels d'IA, Lunit et Icad pour le dépistage de cancer du sein, et MilVue pour la détection de fractures ; nous allons poursuivre dans cette voie pour aider à déceler d'autres pathologies comme des cancers du poumon ou de la prostate. Pour choisir une solution d'IA, nous comparons les résultats de plusieurs logiciels concurrents en leur soumettant des radios de plus en plus anciennes de cas cliniques connus. L'IA que nous choisissons est celle qui donne la meilleure satisfaction en termes de sensibilité. L'IA ne doit pas laisser passer une densité suspecte sans être trop sensible et attirer l'attention du radiologue sur de trop nombreux faux positifs évidents. L'anonymisation des clichés est importante, en particulier si les images doivent sortir de l'infrastructure où elles sont produites. Nous privilégions l'installation sur site aux solutions SaaS pour cette raison. En règle générale, nous nous assurons de travailler avec des prestataires et des éditeurs de confiance, compétents et disponibles.

Quels sont vos projets en 2025 ?

FM : Nous allons travailler sur trois axes. La cybersécurité, car nous devons savoir comment rebondir lorsque l'on sera attaqué. Nous devons également faciliter l'expansion du groupe et le moderniser avec des solutions d'IA. Enfin, avec le projet Drim box, nous allons nous rallier au réseau national de partage et de consultation d'images des praticiens. □

Propos recueillis par Olivier Bouzereau

Requêteur Airmod choisit Tuito

Airmod développe des produits et des équipements pour la TV payante, l'IoT ou la vidéosurveillance... Récemment, l'entreprise a choisi Tuito, un outil de QueryX, pour donner accès aux données à ses salariés sans avoir recours aux équipes IT.

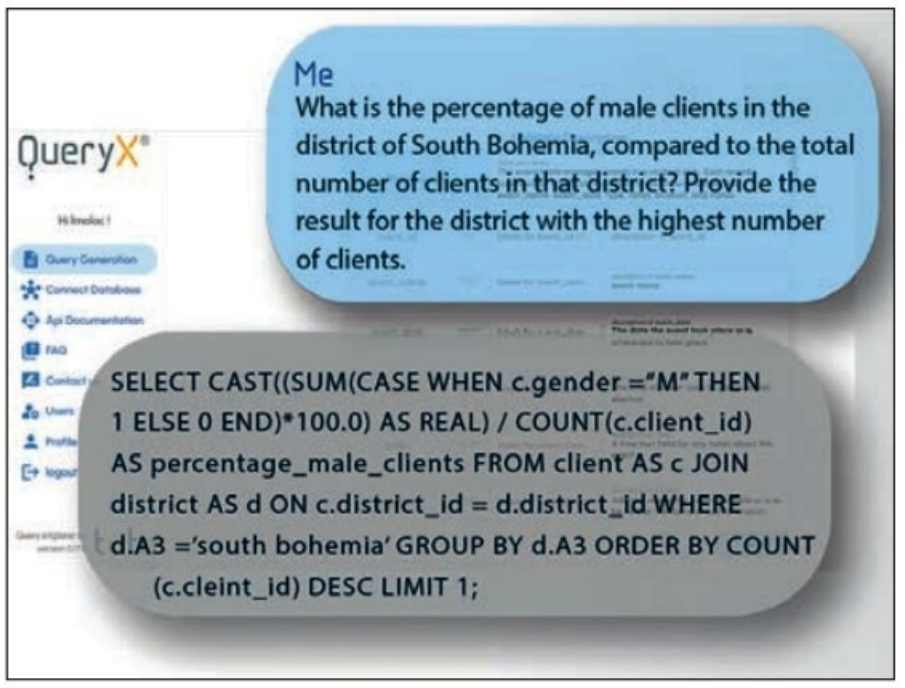
Même s'ils travaillent dans l'électronique, tous les salariés d'Airmod n'ont pas de compétences spécifiques sur les bases de données. Ludovic Delapeyrière, chef des opérations chez Airmod explique : *« Depuis 20 ans, nous avons construit une base de données riche et complexe sur laquelle nous avons branché nos outils maison. Ceux-ci nous permettent de traiter 90 % de nos besoins. Restent les 10 %, les fameux* « corner cases » *pour lesquels nous, les non-techniciens, étions démunis. Et c'est là que QueryX s'est avéré un atout de choix pour nos équipes pour qui exploiter les données est clé ».*

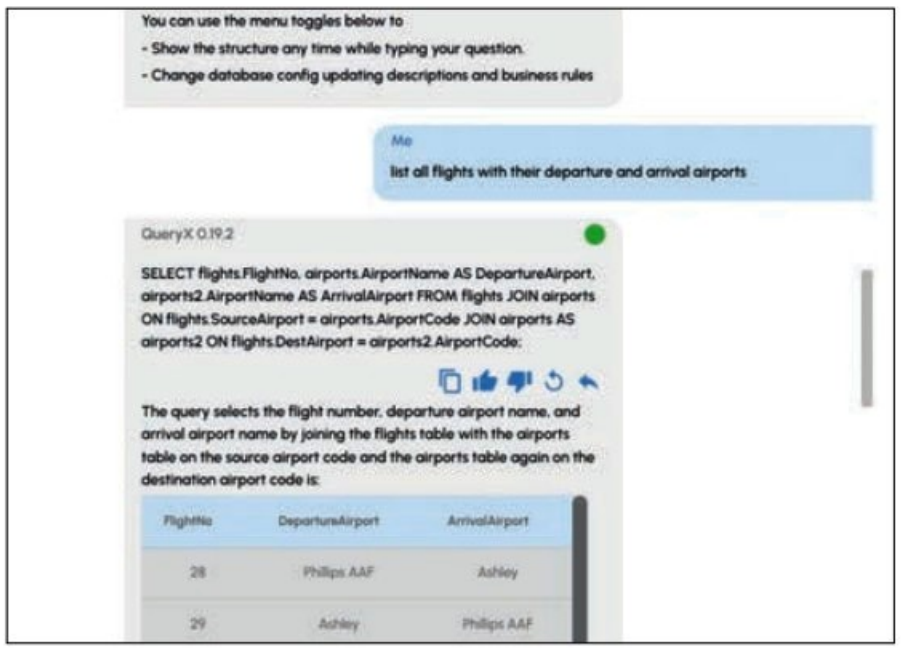
L'objectif était somme toute pragmatique : permettre aux équipes métiers de gagner en autonomie et en productivité grâce à QueryX en leur facilitant l'accès aux données, sans compétences techniques et sans recourir aux équipes IT d'Airmod.

Un pionnier du traitement du langage

Laurent Molac, DG de Tuito, indique : *« depuis 2018, nous travaillons sur l'intelligence artificielle et nous avons développé un produit QueryX sur l'intelligence artificielle générative, qui permet l'accès aux données en grand nombre pour les exploitants métiers qui n'ont pas de connaissances particulières en informatique. Il réalise les requêtes par la voix ou le texte. Le logiciel réalise la requête en SQL. Le requêteur est multilingue, dont la plupart des langues couramment parlées dans le monde ».* Il permet d'interroger les bases de données, même complexes, par de simples questions en langage naturel, par écrit (Text2SQL) ou à l'oral (Voice2SQL).

Si, au début, la solution a utilisé ChatGPT, *« la solution générique la moins chère et la plus efficace »*, le logiciel a été conçu sur différents LLM (Large Language Models) pour répondre à différentes tâches. La solution ne demande pas de « Fine tuning » mais une configuration,

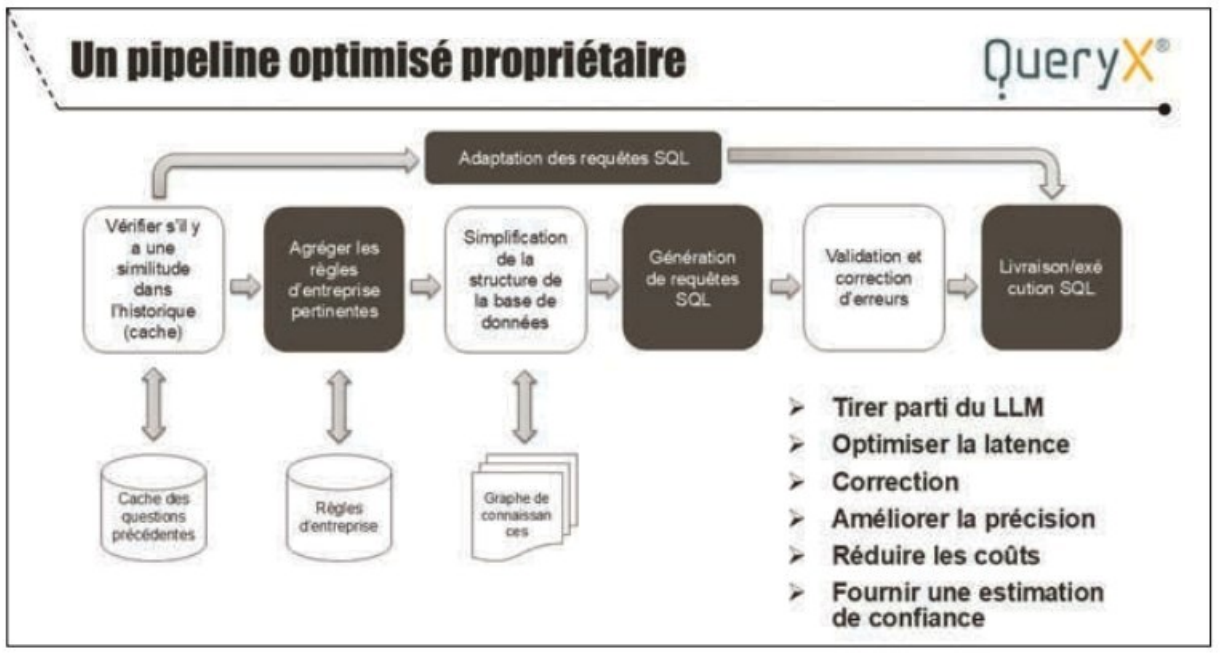




ce qui est différent de l'entraînement mais rend la solution plus facile à déployer. Tuito, l'éditeur de QueryX, fournit aussi du conseil pour cette étape. Un différenciateur de la technologie de Tuito reste la capacité du logiciel à conserver les informations contextuelles. En gardant en mémoire les questions précédentes, la solution permet aux utilisateurs d'affiner leurs recherches au fil de la conversation. Cette fonctionnalité améliore considérablement l'expérience utilisateur, en permettant une exploration plus fluide et plus naturelle des données. En plus de cela, la possibilité de déployer QueryX sur site, ce qui garantit la sécurité et la confidentialité des données, cruciales pour les entreprises opérant dans des secteurs sensibles.

AIRMOD EN BREF

Airmod est un fabricant et un bureau d'études de confiance, spécialisé dans la conception et le développement d'équipements électroniques numériques sécurisés pour un large éventail d'industries. Avec une forte concentration sur l'innovation. Basée à La Ciotat (13), Airmod accompagne plus de 200 clients dans le monde, pour qui elle gère des flux de 400 à 500 000 produits par an. La société emploie une dizaine d'experts.



Un outil pour les exploitants métiers

L'adoption de l'outil a été l'opportunité d'aller chercher des informations présentent dans ses bases de données mais jusqu'ici difficiles d'accès sans faire appel à un IT expert SQL. Véritable solution « no code », elle permet aux équipes Airmod d'opérer en totale autonomie et de gagner en productivité. *« Concrètement, déployer QueryX nous permet de solutionner simplement et rapidement certains points bloquants, de répondre à des questionnements ponctuels ou nouveaux sur tous les aspects de notre business, de nos processus, de nos clients, etc. C'est l'opportunité d'investiguer plus librement en accédant à une information difficile d'accès jusque-là »*, insiste Ludovic Delapeyrière.

Le constructeur a choisi la version sur site du logiciel. *« Cela signifie que la solution n'a pas d'accès à la base de données. Pour nous lancer, il nous a suffi de communiquer le détail de l'architecture de notre base à QueryX »*, précise Ludovic Delapeyrière. Il ajoute : *« Chez Airmod, nous avons fait le choix d'utiliser QueryX en mode hors connexion, ce qui signifie que QueryX n'a aucun accès à nos bases de données. Lorsqu'un besoin de données spécifiques se présente, l'équipe exprime sa demande à QueryX. En réponse, l'outil nous fournit la requête SQL appropriée, qui est ensuite intégrée dans la base de données d'Airmod pour obtenir les résultats souhaités »*. Les données de l'entreprise ne sortent pas de son périmètre de sécurité. Seule la structure de la base de données est partagée. Lors de cette étape préalable, Tuito regarde la mise en forme de la base de données et sa structure, et peut ainsi accompagner son client en expliquant quelles données peuvent être utilisées et à quoi elles sont utiles. Il existe aussi une version SaaS. QueryX, en retour, fournit la traduction du besoin sous la forme d'une requête SQL. L'utilisateur n'a plus qu'à la copier/coller dans l'outil maison pour obtenir sa réponse. *« Un processus qui ne dure le plus souvent que 3 à 5 minutes, alors que cela pouvait prendre des heures voire des jours précédemment »*, conclut Ludovic Lapeyrière. Laurent Molac indique que le logiciel est 10 fois plus rapide que la procédure d'écrire une requête SQL par un SQL Coder Copilot qui permet d'accélérer l'écriture des requêtes SQL, et en fait un compagnon de route des gestionnaires de bases de données, des analystes des données ou des développeurs SQL.

Une solution flexible

L'éditeur fournit de plus des API ou des exemples de code open source pour une intégration simple avec les principaux outils d'interaction ou de visualisation. □

B.G

TUITO EN BREF

L'entreprise est à peu près de la même taille que son client Airmod et emploie une dizaine de salariés. Elle aussi est sise à La Ciotat. Pour les deux entreprises, il n'y a pas de déclarations officielles de chiffre d'affaires. L'éditeur développe des solutions dans le domaine de la voix et du son, ainsi que le texte. Elles s'appliquent dans différents secteurs d'activité. La solution possède des clients en France, en Italie et aux USA. Récemment, l'éditeur a conclu un partenariat avec Thales qui vise à déployer la solution propriétaire de Tuito, QueryX, au sein d'une division du secteur aérien de Thales. Ce projet pilote se concentrera sur l'amélioration des processus de maintenance, en permettant aux agents de terrain d'accéder rapidement et facilement aux informations cruciales directement depuis les bases de données structurées.

Sensibilisation
L'humain premier rempart

L'adage est hélas trop bien connu, en cybersécurité, le problème se trouve entre la chaise et l'écran. En clair l'utilisateur est voué aux gémonies et responsable de tout. L'ouvrage de Michel Gérard vient a contrario de cette idée reçue, et propose des stratégies pour faire de l'utilisateur le premier rempart de la sécurité informatique. Loin du Yakafocon habituel, l'auteur explique le pourquoi, indique les méthodes qui font de la sensibilisation à la cyber une véritable aide pour élever le niveau de sécurité de l'entreprise en s'appuyant sur les humains qui la composent. Bien sûr, tout ce qui est décrit ne se fait pas en un jour, mais c'est aussi le chemin pour faire de l'utilisateur le héros du moment et d'améliorer à la fois la sécurité mais aussi l'engagement de vos utilisateurs. À la suite de la lecture, vous saurez tout sur la méthode Okispri qui met en place les différents éléments pour y parvenir. Alors redonnons à l'utilisateur le premier rôle et non celui du bouc émissaire.

L'utilisateur, ce héros

« Les cyberattaques, c'est le seul truc qui vous fout la boîte en l'air en 24 heures ». Ainsi s'exprimait Olivier Piquet, directeur général de Lise Charmel, groupe de lingerie fine française.

Un salarié qui ouvre sa messagerie personnelle, clique sur un e-mail vérolé et en trois heures de temps dans la nuit, 98 % des machines sont cryptées. L'entreprise est à l'arrêt et mettra 10 mois à retrouver un fonctionnement normal. Un passage en redressement judiciaire sera nécessaire pour surmonter l'épreuve.

En 2021, les cyberattaques ont coûté environ 6 000 milliards de dollars à l'économie mondiale, soit 190 000 dollars par seconde. La cybercriminalité est devenue 5 fois plus rentable que l'ensemble des crimes transnationaux combinés.

Si la cybercriminalité était un pays, ce serait la 3ème économie mondiale derrière les États-Unis et la Chine.

Pratiquement pas une semaine ne passe sans que l'actualité ne se fasse l'écho d'un incident de sécurité majeur. Fuite de données, systèmes bloqués par un ransomware, détournement de fonds, usurpation d'identité, sont autant de types d'attaques fréquemment rencontrés.

Dans la grande majorité des cas, un défaut de comportements d'un ou plusieurs utilisateurs est en cause.

Il peut s'agir de celui qui clique sur un e-mail de phishing permettant l'introduction d'un code malveillant. Ou encore de celui qui utilise son mot de passe et son e-mail professionnels sur un site qui se fait pirater, offrant ainsi à des pirates les informations nécessaires pour se connecter au système d'information. Cela peut aussi être le cas de celui qui, trop naïf, se fait duper sur un réseau social et divulgue des informations sensibles. On peut aussi citer le voyageur imprudent qui utilise un réseau wifi public. Les exemples se multiplient et se déclinent dans des variantes infinies.

Face à des infrastructures de mieux en mieux sécurisées, les cybercriminels utilisent la vulnérabilité principale d'un système d'information : son utilisateur.

Pourquoi chercher à crocheter une serrure, quand il suffit de trouver quelqu'un à qui demander d'ouvrir la porte ?

Notre utilisateur, héros, est bien celui qui est ici au cœur de l'aventure cyber.

Notre époque ne manque d'ailleurs pas de défis à relever pour toutes les organisations dans lesquelles leurs collaborateurs sont étroitement impliqués. Le climat, le numérique, l'intelligence artificielle, l'économie, la sécurité et la cybersécurité, autant de sujets de préoccupations et d'enjeux pour l'ensemble des entreprises et des administrations.

Ma conviction ici est qu'aucun de ces défis ne pourra être relevé sans embarquer les collaborateurs.

Pour Thucydide, stratège et historien athénien, la capacité d'une cité à se défendre tient plus à la volonté de ses citoyens à la défendre que de l'épaisseur de ses remparts.

Rapportée à notre domaine, cette citation renvoie les remparts à tous les dispositifs de cybersécurité déployés ; les citoyens étant, bien sûr, les utilisateurs du système d'information.

Dans notre monde, la notion même de rempart a disparu. Les systèmes d'information se sont tellement ouverts qu'ils tiennent davantage de la fête foraine où tout le monde peut entrer, charge à chaque activité d'effectuer ses propres contrôles, que du parc d'attractions dans lequel on accède librement à toutes les activités après un contrôle unique à l'entrée.

C'est la notion même du zéro trust.

[...]

Le piratage informatique ne passera pas par moi

« Le piratage informatique ne passera pas par moi » est l'engagement que chaque utilisateur doit prendre.

L'objectif de toute stratégie de sensibilisation est bien que les collaborateurs s'engagent dans cette voie. Avant d'investir dans telle ou telle opération de sensibilisation, se pose la question de mesurer l'impact qu'elle aura sur la progression vers cet objectif. Il peut être en effet parfois tentant de se laisser séduire par telle ou telle nouvelle modalité.

Cependant, dans un monde de temps et de budget contraints, il faut privilégier les actions les plus efficaces et dont les effets sont mesurables.

Dans la durée, de plus en plus d'utilisateurs s'engagent dans cette voie et acceptent d'afficher cet engagement, notamment en publiant notre badge.

Les techniques de sensibilisation impliquantes, décrites ci-après, renforcent cet élan.

La sensibilisation impliquante

Trop souvent les organisations sont déçues des résultats de leurs opérations de sensibilisation, tant du point de vue de la participation que du résultat sur l'adoption des bonnes pratiques.

Nos équipes n'ont de cesse de rechercher les meilleures pratiques et méthodes favorisant l'adhésion et l'engagement vers un comportement cyber sécurisé. Nous cherchons ainsi à guider chaque utilisateur vers un engagement dans lequel il fait sienne la maxime *« le piratage informatique ne passera pas par moi »* et qu'il ne soit pas une source d'incident de sécurité.

C'est ainsi que nous nous appuyons sur différents travaux de Robert-Vincent Joule et de Fabien Girandola, tous deux de Aix-Marseille Université, en matière de psychologie de la persuasion et de l'engagement.

La combinaison de ces travaux nous amène à la mise en œuvre d'une démarche qui s'appuie sur quatre piliers :

1. Expliquer les enjeux et développer la sensibilité au sujet.

En effet, avant toute chose, il est indispensable que chacun puisse considérer le sujet comme important.

2. Inculquer un minimum de connaissance.

Sans chercher l'expertise, la connaissance d'un minimum d'éléments sur le sujet est nécessaire.

3. Savoir se poser les bonnes questions.

Dans cette étape, nous cherchons à faire en sorte qu'une personne, face aux différentes situations auxquelles elle se trouve confrontée, prenne un temps minimum pour se poser les bonnes questions et y apporter les bonnes réponses.

4. Adopter les nouveaux comportements.

C'est par la mise en pratique régulière de l'étape 3, qui s'appuie elle-même sur les deux premières étapes, que de nouveaux réflexes et comportements finissent par être adoptés et deviennent naturels. Des actes engageants viennent renforcer l'adoption des comportements attendus dans la durée.

La mise en oeuvre de ces piliers sera complétée par :

→ Le ciblage de façon appropriée d'une population segmentée ;

→ L'identification des freins et des objections ;

→ Un mix pédagogie et tests ;

→ La déclinaison d'une approche sectorielle permettant la mise en exergue de situations au plus près du vécu des utilisateurs (santé, collectivités territoriales, industrie, retail...)

→ L'utilisation des différents leviers possibles pour maximiser taux de participation et engagement ;

→ La définition d'une stratégie de sensibilisation et de programmes annuels en fonction des objectifs définis ;

→ L'obtention d'actes engageants, orientant les utilisateurs vers une adoption durable des bons comportements.

D'un point de vue pratique, l'application des travaux de recherche sur la psychologie de la persuasion et de l'engagement, nous a conduits à la mise en œuvre de différentes techniques appliquées à la mise en œuvre d'une stratégie de sensibilisation en ligne.

Nous en présentons ici 12 qui nous semblent les plus utiles, afin de maximiser l'impact des opérations de sensibilisation.

D'autres techniques peuvent être imaginées et mises en œuvre une fois compris les principes de la psychologie de la persuasion et de l'engagement.

Technique N°1
Du bon usage de la peur

La peur est très souvent utilisée quand il s'agit de cybersécurité.

On mettra ainsi en avant les profils menaçants des pirates et des organisations criminelles qui les utilisent, les conséquences désastreuses des attaques sur les organisations et sur les collaborateurs qui en sont victimes.

Si l'usage de la peur peut se contester, il n'en reste pas moins qu'à partir du moment où l'on parle des risques et des enjeux liés à la sécurité, il est plus que probable qu'un sentiment de peur naisse dans l'esprit de beaucoup d'utilisateurs.

La question est donc de savoir ce que l'on fait de ce sentiment de peur.

La recherche montre que lorsqu'on utilise la peur pour persuader une personne du bien-fondé du comportement qu'on lui demande, on peut provoquer deux types de réactions : la gestion de la peur ou la gestion du danger.

La première réaction est contreproductive, car il s'agit d'éviter le problème en déniant ou minimisant le sujet, en le contournant, en transférant la responsabilité à un tiers ou encore tout simplement en faisant preuve de fatalisme.

La deuxième réaction correspond à ce qu'on cherche à obtenir, à savoir la mise en œuvre des bonnes pratiques permettant d'éviter que la menace se concrétise.

Ici aussi la recherche montre qu'on obtient cette deuxième réaction à condition d'expliquer, en même temps que l'on suscite la peur, les mesures à prendre pour faire face au danger, leur efficacité et leur simplicité de mise en œuvre.

En matière de sensibilisation à la cybersécurité, il faut ainsi, au même moment qu'on met en avant la menace, les risques et les conséquences des incidents, expliquer aux utilisateurs que l'application de quelques règles et comportements simples permettront de les éviter.

En conclusion les appels à la peur sont particulièrement efficaces quand :

→ Ils décrivent une menace en accentuant la sévérité et la vulnérabilité ;

→ Ils font état de l'efficacité des recommandations et de leur facilité d'exécution.

Technique N°2
Traitement direct du sujet

Le message délivré a tendance à convaincre davantage sans avertissement préalable.

En étant averti avant, il existe une tendance à imaginer les résistances.

Même si ce phénomène est moins marqué sur les sujets techniques ou l'avertissement a moins d'impact négatif, il peut être intéressant d'apporter directement l'information sans annonce.

Cela peut être le cas de campagnes très ciblées, dans lesquelles le message d'invitation commence directement à traiter du sujet choisi et propose de suivre immédiatement le support fourni. Le message peut aussi présenter un risque et les dangers associés, tout en proposant de suivre le support qui donne simplement les moyens de s'en protéger.

Technique N°3
Développement de l'implication

Plus le sujet sera perçu comme important et plus les utilisateurs auront tendance à se sentir impliqués. Cela peut se traduire notamment par une communication appropriée mettant en avant que le sujet est essentiel pour l'organisation et, que la direction et tout le management sont alignés pour en faire une priorité.

L'implication sera également d'autant plus forte que les situations présentées colleront au plus près du vécu des utilisateurs. C'est là tout l'intérêt de disposer de contenus et de modules qui mettent en avant des situations et des messages au plus près du contexte vécu par l'utilisateur : secteur, métier, pays...

La possibilité de personnalisation de ces modules est également un bon moyen pour chaque organisation d'adapter les messages à son propre contexte, son vécu et sa population.

[...]

La démarche en mouvement

Éléments de contexte

Dans certaines organisations, la population concernée est de taille relativement faible. Dans un tel contexte, la mise en œuvre d'une stratégie de sensibilisation est plutôt simple. Une seule langue est nécessaire et de nombreux moyens de sensibilisation peuvent être utilisés, incluant la mise à disposition de documents, des campagnes en ligne, des ateliers, jusqu'à la sensibilisation one to one.

En revanche, dans les organisations plus conséquentes en taille de population, et qui constituent la grande majorité des organisations avec lesquelles il nous a été donné de travailler, les contraintes sont toutes autres.

En effet, si on doit toucher plusieurs centaines, milliers, dizaines de milliers, voire largement plus de cent mille personnes, les moyens employés devront répondre à certains critères précis :

→ Les langues : les supports proposés doivent être disponibles dans les différentes langues des utilisateurs. En effet, les utilisateurs auront en général peu de goût à suivre une sensibilisation qui n'est pas proposée dans leur langue.

→ La simplicité de mise en œuvre des supports par les utilisateurs : ils doivent en effet être conçus pour être facilement utilisables par le plus grand nombre. Il faut que l'apprenant accède au message de façon quasi immédiate sans avoir à se poser de questions sur la façon de mettre en œuvre du support.

→ Le fait de s'adresser à des types de populations très diverses en âge, métier, éducation numérique, pays, culture, etc.

→ La facilité, la rapidité et le coût de déploiement : déployer une sensibilisation auprès d'une large population doit en effet être accessible financièrement, pouvoir se faire dans un laps de temps court et pour tout le monde en même temps.

→ L'intégration au système d'information : l'adhérence au système d'information existant doit être la plus faible possible, afin d'avoir les coûts d'intégration les plus bas possibles. C'est l'intérêt des solutions Saas (Software as a service). Il n'en reste pas moins qu'il faudra prévoir d'automatiser au maximum les imports et mises à jour des bases d'utilisateurs.

→ La capacité à s'adapter à l'organisation : chaque entreprise ou administration connait ses contraintes d'organisations. Elles sont plus ou moins centralisées ou décentralisées. Il peut ainsi être nécessaire de tenir compte des répartitions par pays, par filiale, par branche, par région, par métier, etc.

→ La capacité à tenir la charge : envoyer une campagne de sensibilisation vers dix ou cent mille personnes en même temps, nécessite de la puissance, notamment quand une fraction non négligeable de cette population se connecte et démarre le support en même temps.

→ La sécurité et le respect des données personnelles : même si les règles et pratiques peuvent différer d'un pays à l'autre, les grands principes restent les mêmes. La solution mise en œuvre devra notamment être conforme au RGPD.

Incarner la cybersécurité

Le RSSI ou à tout le moins, une personne de son équipe, doit incarner le message de cybersécurité.

En effet, plusieurs études et recherches montrent que l'incarnation d'un message par un leader ou une figure d'autorité, comme un RSSI, peut considérablement augmenter l'efficacité de la communication, en particulier dans des domaines complexes comme la cybersécurité.

En voici quelques extraits : une étude publiée dans le Journal of Business Ethics a révélé que la communication éthique incarnée par les dirigeants augmentait l'engagement des employés et leur adhésion aux politiques de l'entreprise.

→ Une étude de McKinsey a démontré que les employés sont cinq fois plus susceptibles d'agir sur une communication, lorsqu'elle est délivrée par un leader visible et engagé, plutôt que par un canal impersonnel ou automatisé.

Cela souligne l'importance de l'incarnation du message pour stimuler l'action et l'engagement.

→ Une recherche de Harvard Business Review a montré que les messages de sensibilisation, notamment en matière de sécurité, sont perçus comme plus pertinents et urgents lorsque communiqués directement par des experts internes.

Cela conduit à une augmentation de 20 à 30 % de la conformité aux politiques de sécurité, selon les entreprises étudiées. □

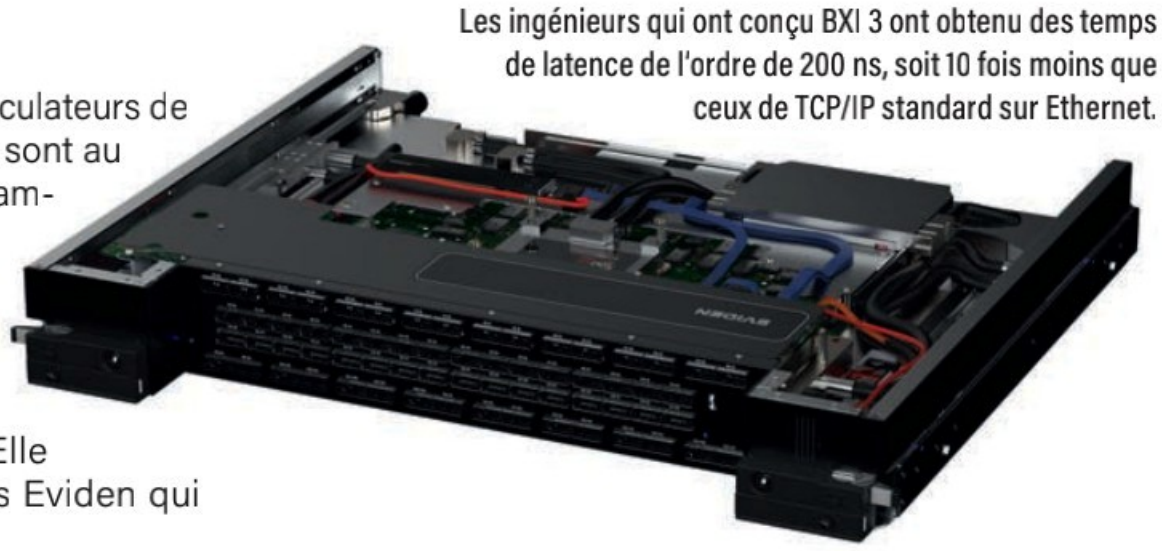
▪ ISBN : 978-2-9589186-2-0
▪ Prix : 17,50 € TTC ▪ 74 pages
▪ Ouvrage auto-édité par l'auteur : Michel Gérard

Eviden BXI 3

La pépite d'Eviden

La nouvelle a fait la une des quotidiens économiques : L'État français veut nationaliser l'activité « Advanced Computing » d'Atos, le fabricant des supercalculateurs, notamment utilisés par la dissuasion nucléaire. Une expertise dans le HPC unique qui se matérialise sous la forme d'une ASIC : BXI 3.

Eviden sait construire des supercalculateurs de tout premier plan, dont plusieurs sont au TOP 500 mondial. Il maîtrise notamment le refroidissement liquide des lames de calcul à très large échelle, mais aussi une technologie d'interconnexion qui lui est propre, BXI. La version 3 de ce composant d'interconnexion clé vient d'être dévoilée. Elle équipera les prochains supercalculateurs Eviden qui seront vendus en 2025.



Les ingénieurs qui ont conçu BXI 3 ont obtenu des temps de latence de l'ordre de 200 ns, soit 10 fois moins que ceux de TCP/IP standard sur Ethernet.

Il s'agit de la nouvelle évolution d'une technologie dont l'origine remonte aux systèmes eXascale de Bull de 2015, BXI signifiant Bull eXascale Interconnect. La version 3 permettra de supporter des clusters de 64 000 nœuds, un nœud, le NIC pouvant être un serveur avec 2 CPU AMD, Intel ou ARM par exemple ou un GPU Nvidia. Si l'interface PCIe Gen 5 dicte la bande passante pour tous les fournisseurs de solutions HPC, l'un des réels différenciateurs du composant BXI 3 réside dans sa faible latence : *« Avec du TCP/IP classique sur Ethernet, on ne peut faire de basse latence »*, explique Eric Eppe, vice-président d'Eviden Group en charge de l'HPC, l'IA et le quantique, et membre du board de l'Ultra Ethernet Consortium (UEC). *« Le switch doit faire remonter les paquets dans toute la stack TCP/IP pour les routers, ce qui est impossible à faire en moins de 200 ns. En modifiant la stack, on peut faire beaucoup mieux, et c'est une optimisation qui sera intégrée à l'Ultra Ethernet à l'avenir »*.

Autre fonctionnalité implémentée dans l'ASIC BXI 3, le « in order delivery » des paquets, ce que TCP/IP ne permet pas, et qui s'avère très pénalisant dans les supercalculateurs. Autre moyen de booster les échanges, le « packet spreading » pour exploiter au maximum les capacités des switchs en répartissant les flux de données sur tous les ports disponibles.

Enfin, contrairement à l'Ethernet classique et à Infiniband, BXI 3 pratique l'Offloading Collective Operations. La gestion des calculs distribués des applications HPC ou IA repose sur l'API MPI (Message Passing Interface) dans le monde du HPC et sur NCCL (NVIDIA Collective Communication Library) ou RCCL (ROCm Communication Collectives Library) d'AMD dans le monde de l'IA. Ces API sont habituellement exécutées au niveau du serveur, mais Eviden a câblé ces instructions dans le hardware de son ASIC, ce qui lui permet d'accroître de 35 % la rapidité d'exécution de ces API. Un gain particulièrement appréciable, lorsqu'on sait que selon un White Paper publié par Meta, en phase d'apprentissage, les GPU sont en pause jusqu'à 70 % du temps, en attente de réception des données... □

A.C

ERIC EPPE, VICE-PRÉSIDENT D'EVIDEN GROUP EN CHARGE DE L'HPC, L'IA ET LE QUANTIQUE

« Les clusters HPC fonctionnent tous sur un même modèle, avec des dizaines de milliers de nœuds interconnectés sur lesquels on va répartir les calculs. Cela suppose beaucoup d'échanges de données entre ces nœuds, car le calcul va être distribué entre tous ces derniers. Selon le calcul envisagé, chacun des 10 000 nœuds d'un cluster va devoir communiquer des données aux 10 000 autres pour mener le calcul à son terme. Cela nécessite une bande passante extrêmement élevée, de l'ordre de 400 ou 800 Gbit/s par nœud, soit le maximum de ce qu'il est aujourd'hui possible de faire sur un réseau scale out. Ce qui est en train de changer aujourd'hui avec l'IA, c'est que le besoin de bande passante est comparable à celui du HPC, car les volumes de données mis en œuvre lors du training des IA sont encore plus importants »

Médecine

Insilico brille et affiche ses ambitions au GITEX 2024

La Scale Up spécialiste du développement de nouvelles molécules par l'IA séduit de nombreux laboratoires et accélère son développement.

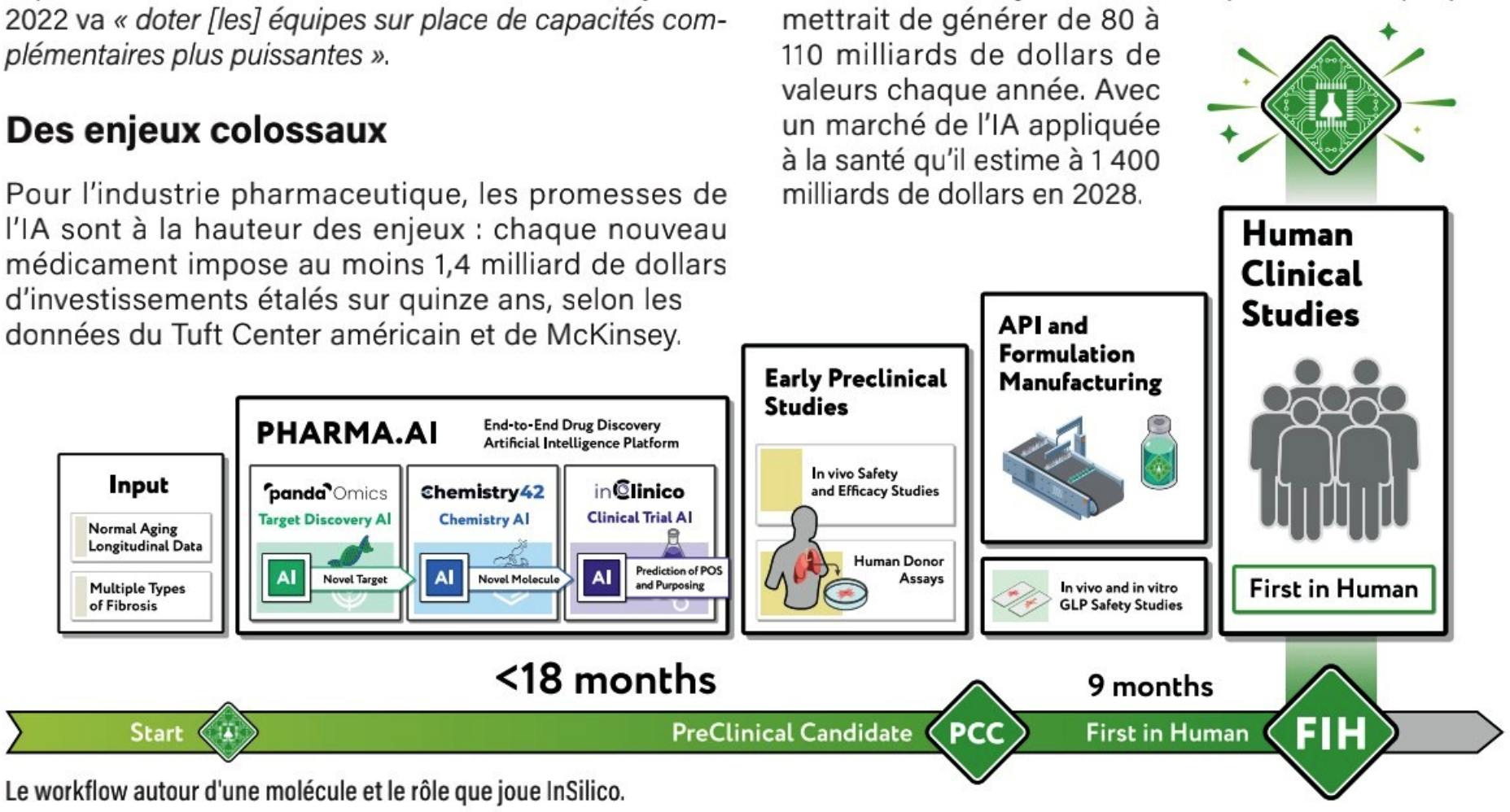
L'IA est l'avenir de 'Big Pharma' et Insilico Medicine son prophète, en particulier au GITEX 2024 de Dubaï. Créée en 2014, la scale-up américaine a élaboré pharma.ai, une plateforme de découverte de molécules et de conduite des essais cliniques, pilotée de bout en bout par l'intelligence artificielle. Sa vocation se résume dans la jolie formule de son fondateur Alex Zhavoronkov, docteur en biologie et ingénieur en data sciences : *« capter et déstabiliser les protéines criminelles des cancers »*. Après avoir eu quelques difficultés pour convaincre l'écosystème de financer son embryon de jeune pousse, il a levé près de 400 millions de dollars depuis 2017. Il a depuis noué des partenariats avec les plus grands labos comme Fosun, Novo Nordisk, Pfizer, etc., ce qui lui permet de délivrer un message sans ambages aux géants de la pharmacie mondiale : *« soit ils renoncent à l'IA et sont dans l'obligation de licencier, soit ils redoublent d'efforts en s'associant à des biotechs »*. Comme pour lui faire écho, Changchun Xia, le patron de la R&D de Sanofi en Chine, indiquait que le partenariat de 21,5 millions de dollars conclu en janvier 2022 va *« doter [les] équipes sur place de capacités complémentaires plus puissantes »*.

Alex Zhavoronkov, docteur en biologie et ingénieur en data sciences

Des enjeux colossaux

Pour l'industrie pharmaceutique, les promesses de l'IA sont à la hauteur des enjeux : chaque nouveau médicament impose au moins 1,4 milliard de dollars d'investissements étalés sur quinze ans, selon les données du Tuft Center américain et de McKinsey.

Sans garantie de succès, puisque d'après Deloitte, 80 % des recherches menées entre 2000 et 2015 ont débouché sur des impasses. Or, d'après McKinsey, l'incorporation de l'IA dans les tuyaux de la R&D pharmaceutique permettrait de générer de 80 à 110 milliards de dollars de valeurs chaque année. Avec un marché de l'IA appliquée à la santé qu'il estime à 1 400 milliards de dollars en 2028.



Le workflow autour d'une molécule et le rôle que joue InSilico.

D'où l'intérêt des labos pour les travaux de cette combinaison de spécialistes en bio technologies et en data sciences. Selon David Del Bourgo, le CEO du français White_Lab Genomics, l'IA permet *« de diviser par trois le temps nécessaire à la découverte de nouvelles molécules, avec l'objectif de le ramener à cinq ans, contre quinze années aujourd'hui »*.

Ainsi, avec ses outils comme PandaOmics pour la génomique et Chemistry42, son moteur de chimie générative pour la conception de composés zéro, le 'Pipeline' ou portefeuille de traitements en gestation, d'Insilico comptait en juillet dernier dix-huit molécules en oncologie, contre les maladies du système nerveux central (SNC), immunothérapie, etc. Sur ce total, cinq étaient en essais cliniques, dont quatre en phase I et une en phase 2. Parmi elles, un nouvel inhibiteur des facteurs de croissance des fibroblastes, ces fameuses 'protéines criminelles' qui jouent un rôle crucial dans la prolifération des maladies. Sans oublier, en janvier dernier, la livraison d'une molécule KAT6A à l'italien Menarini Group et à sa filiale américaine Steamliner, qui poursuivront la mise au point et commercialiseront cet inhibiteur des cancers du sein. Ou, en juillet dernier, un second candidat contre les tumeurs solides développé avec le chinois Fosun.

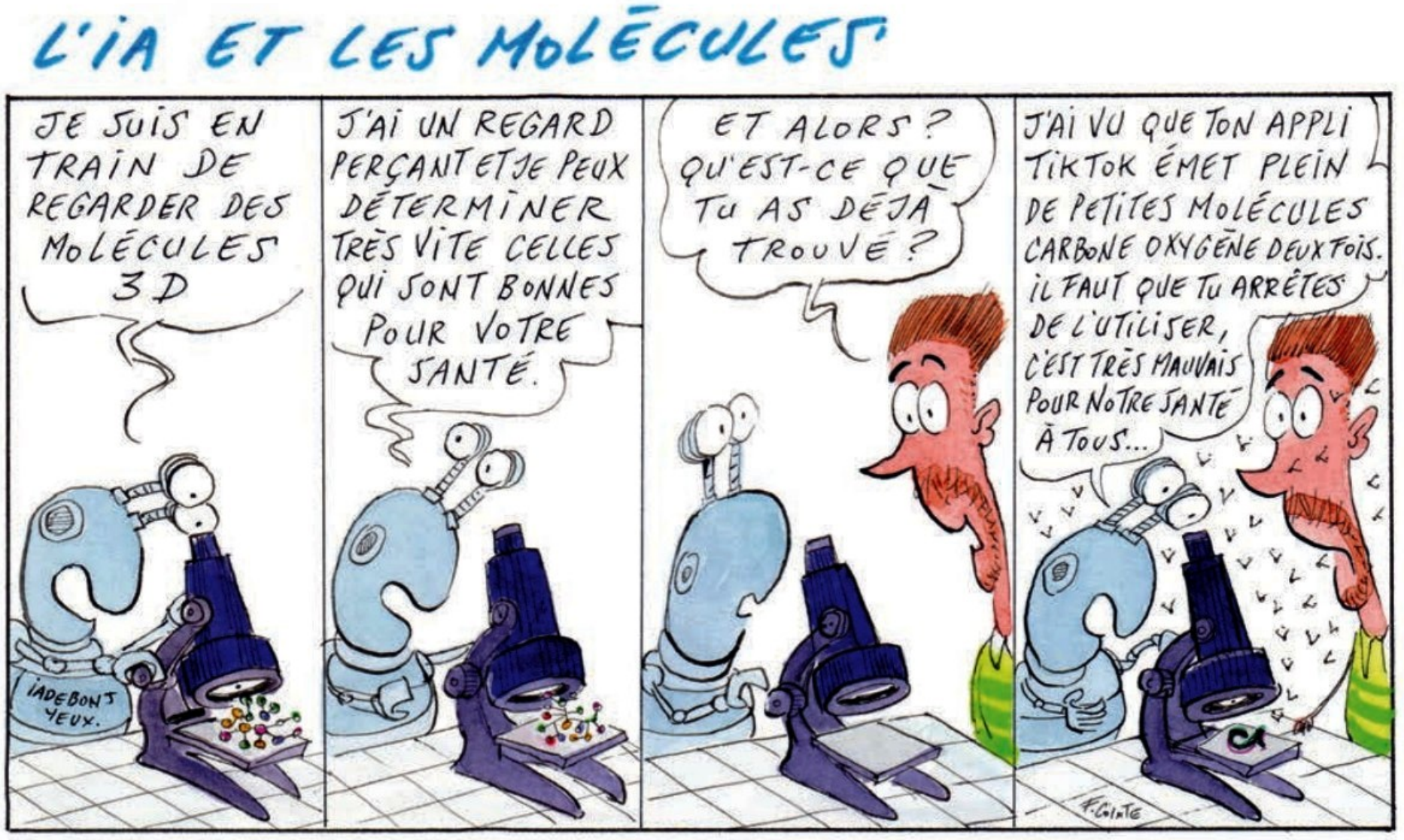
Surtout, le champ d'action d'Insilico ne se limite plus à la recherche médicale, depuis son deal conclu avec Syngenta, le leader suisse des produits phytosanitaires. Prosperity Ventures 7, le fond de capital risque d'Aramco lui a apporté trente-cinq millions de dollars. Car la compagnie pétrolière saoudienne — qui affiche une capitalisation boursière de 700 milliards de dollars ! — compte appliquer ses algorithmes à l'énergie, à la chimie verte, voire aux nouveaux matériaux.

Les promesses de l'IA appliquée à la recherche médicale sont conséquentes. Mais n'effacent pas totalement le risque d'échec. La britannique Benevolent AI peut en témoigner : l'échec en octobre de l'une de ses molécules se solde par cent-quatre-vingts suppressions d'emplois. Sans compter un préjudice de réputation à la hauteur des espoirs déçus...

En France, Whitelab Genomic n'est pas en reste

Créé en 2019 par son CEO David Del Bourgo, et le généticien Julien Cottineau, WhiteLab Genetics emprunte lui aussi la voie de l'IA pour, selon le premier *« développer de nouvelles grandes molécules ou médicaments à base d'ARN et d'ADN en machine, avant de les livrer aux laboratoires pour les essais cliniques »*. La Deep Tech, qui a levé dix millions d'euros en 2022 auprès d'Omnes Capital et du suisse Debiopharm, se focalise sur l'ophtalmologie et l'oncologie. Associée à Sanofi, au laboratoire TaRGeT (Université de Nantes et Inserm) et à l'Institut Imagine, elle a déjà obtenu des résultats contre la DMLA et les podocythopaties (lésions de cellules filtrantes dans les reins). Le tout depuis une très forte volumétrie de données alimentant sa plateforme de simulation et de machine learning. Comme Insilico, White Labs a dû convaincre un capital risque *« qui repose sur la validation des hypothèses par le marché, qui veut savoir si le marché du développement de molécules par l'IA existe »*, conclut-il. □

V.B

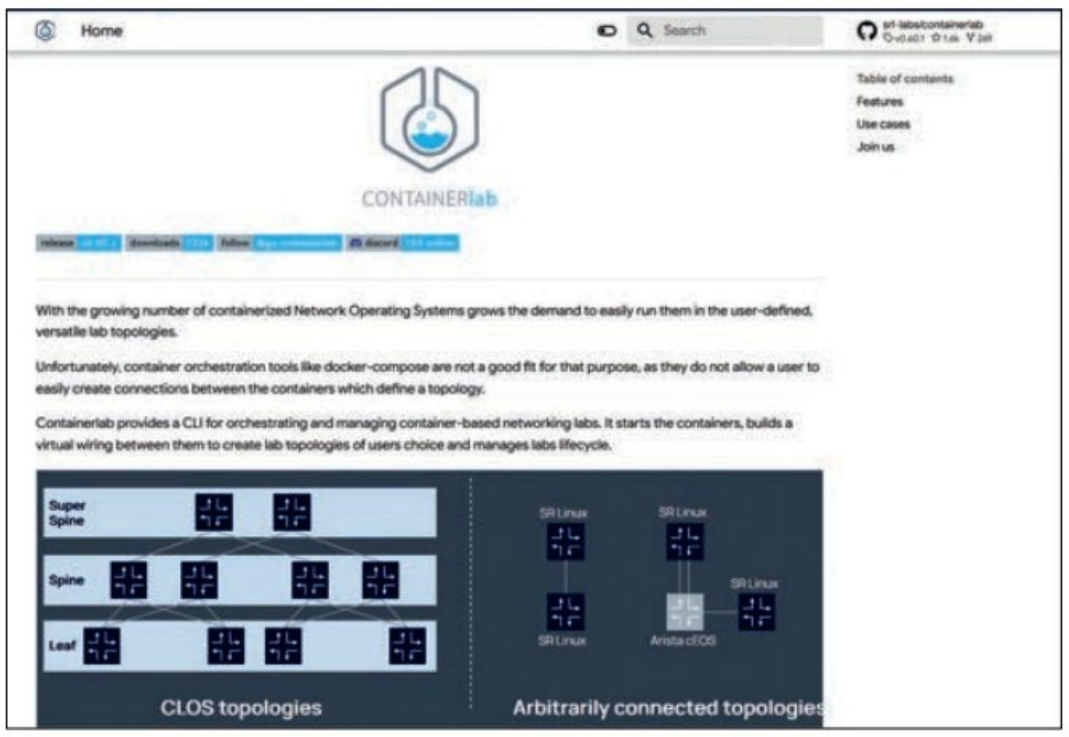


Simulateur

Créer des réseaux virtuels avec containerlab

Containerlab est un outil puissant permettant de créer et de gérer des réseaux virtuels en utilisant des conteneurs Docker. Il est largement utilisé pour la simulation de topologies réseau. Nous allons voir, dans cet article, comment créer des réseaux et y accéder avec cette « Rolls Royce » des simulateurs réseaux.

Containerlab a donc pour vocation la création de réseaux virtuels en vue d'effectuer des simulations complexes avec des conteneurs. Il se base sur des fichiers YAML pour définir les différentes topologies qui vont conduire à l'automatisation et à la gestion d'environnements de test pour des réseaux, des équipements ou des configurations spécifiques. Ce processus permet également de reproduire des réseaux complexes dans un environnement isolé, sans avoir à acheter le moindre matériel.



Si vous voulez absolument tout savoir sur containerlab, rendez — vous sur le site du projet à l'adresse https://containerlab.dev/

Prérequis

Outre containerlab lui-même, vous allez devoir installer Docker sur votre machine. Containerlab s'appuie dessus pour exécuter des conteneurs. S'il n'est pas déjà présent, vous pouvez l'installer à partir de www.docker.com/get-started. Docker est gratuit et multiplateformes et son installation est très simple. Vous n'avez qu'à suivre les instructions indiquées sur son site. Containerlab est disponible pour Linux, Mac et Windows. Sous Linux et Mac OS X, la procédure est la même. Pour installer container lab, vous pouvez employer ces instructions depuis un shell « puissant » :

```
echo « deb [trusted=yes] https://netdevops.fury.site/apt/ /"|
sudo tee – a /etc/apt/sources.list.d/netdevops.list
```

Vous pouvez aussi employer le package curl ou wget. Vous pouvez, par exemple, l'installer avec cette instruction — après avoir installé curl au préalable :

```
curl – sL https://containerlab.dev/setup | sudo – E bash – s "all"
```

Vérifiez l'installation en tapant simplement : containerlab version

Si cela ne fonctionne pas, c'est qu'il y a eu un problème durant l'installation et que vous devez recommencer. Vous pouvez employer une autre procédure. Les différentes méthodes d'installation de containerlab sont toutes répertoriées à l'adresse : https://containerlab.dev/install/. Une fois que tout est installé, vous allez pouvoir charger et décharger directement les labos avec la commande containerlab ou, si vous préférez, son alias clab. Containerlab ne fonctionne pas à proprement parler sous Windows. Néanmoins, vous pouvez l'installer et l'utiliser sous WSL (Windows Subsystem for Linux). Si WSL n'est pas encore sur votre système, allez sur le lien https://docs.microsoft.com/en-us/windows/wsl/install et suivez les instructions. Il est possible, selon votre version et votre installation de Windows, que vous ayez à mettre à jour le composant WSL avant de procéder. Après avoir activé WSL, il vous faudra choisir une distribution Linux. Nous vous conseillons de prendre une Debian ou une Ubuntu récupérées depuis le Windows Store. Installez ensuite Docker Desktop pour Windows, toujours depuis le site de Docker. La procédure d'installation de containerlab sous WSL est sensiblement la même que sous un Linux classique. Bien qu'il soit possible d'utiliser WSL, il vaut mieux installer une vraie machine virtuelle Linux. La seule bonne raison d'employer WSL est le manque de ressources, notamment en termes de RAM sur votre machine. Il vaut mieux directement partir du principe que la simulation de réseaux est un processus qui nécessite d'importantes ressources en termes de processeur et surtout de mémoire vive. Avec des configurations trop légères, vous marcherez sur « trois pattes » et perdrez beaucoup de temps.

Créer une topologie réseau avec containerlab

Containerlab utilise des fichiers de topologie au format YAML pour décrire les réseaux. Ce fichier spécifie les conteneurs à utiliser, leur configuration réseau et les connexions

existantes entre eux. Voici un exemple de fichier de topologie pour un réseau virtuel simple avec trois nœuds réseau (deux routeurs et un switch) :

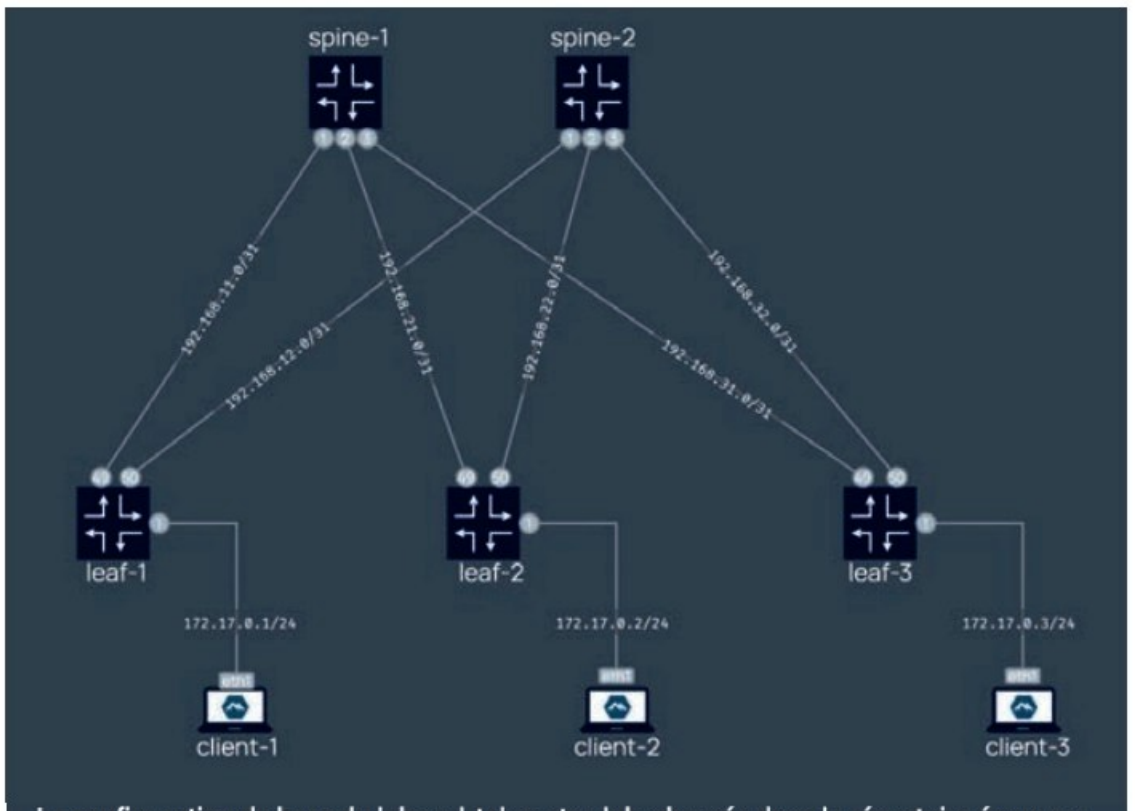
name : my_network
topology :
nodes :
router1 :
kind : vr-1
image : vrnetlab/vr-srx
interfaces :
- name : eth1
address : 192.168.1.1/24
- name : eth2
address : 192.168.2.1/24
router2 :
kind : vr-1
image : vrnetlab/vr-srx
interfaces :
- name : eth1
address : 192.168.2.2/24
- name : eth2
address : 192.168.3.1/24
switch1 :
kind : linuxbridge
image : linuxbridge
interfaces :
- name : eth1
address : 192.168.1.2/24
- name : eth2
address : 192.168.3.2/24
links :
- endpoints : ["router1: eth2 », « router2: eth1 »]
- endpoints : ["router1: eth1 », « switch1: eth1 »]
- endpoints : ["router2: eth2 », « switch1: eth2 »]

Génération du réseau virtuel

Placez votre fichier topology.yaml dans un répertoire de votre choix. Utilisez ensuite la commande containerlab (ou son alias clab) pour créer et démarrer la topologie réseau : *containerlab deploy --topo topology.yaml*

CONFIGURATION DE BASE DU LABO

La configuration de base du labo srl-telemetry-lab (https://github.com/srl-labs/srl-telemetry-lab) chargée dans le répertoire éponyme est composée de trois feuilles (leaves), deux épines dorsales (spines) et trois clients. Feuilles et épines dorsales utilisent respectivement les chassis Nokia SR Linux IXR-D2 et IXR-D3L. Chaque élément réseau de cette topologie est équipé d'un fichier de configuration de démarrage qui est appliqué au lancement de la création du nœud. Une fois démarré, les nœuds réseau sont disponibles avec leurs interfaces, leurs protocoles sous-jacents et leurs services de superpositions configurés. Le réseau généré exécute le service Layer 2 EVPN entre les feuilles. Vous pouvez retrouver ces fichiers dans les sous-répertoires du labo, ainsi qu'à l'adresse https://github.com/srl-labs/srl-telemetry-lab/blob/main/configs/fabric.



La configuration de base du labo srl-telemetry-lab, chargée dans le répertoire éponyme, est composée de trois feuilles (leaves), deux épines dorsales (spines) et trois clients. Source https://github.com/srl-labs/srl-telemetry-lab

Vous pouvez vérifier si tout fonctionne correctement avec la commande suivante : *containerlab inspect*

Accès aux nœuds et test de la connectivité

Vous pouvez vous connecter à un conteneur individuel via Docker. Si, par exemple, vous souhaitez accéder au conteneur router1, tapez : *docker exec – it router1 /bin/bash*

Vous vous retrouverez alors dans le shell bash du routeur. Une fois à l'intérieur des conteneurs, vous pouvez tester la connectivité entre eux en utilisant la commande ping. Pour tester la connexion avec router2, lorsque vous êtes dans router1, tapez : *ping 192.168.2.2*

Arrêt et suppression du réseau virtuel

Lorsque vous avez terminé de travailler avec votre simulation, vous pouvez arrêter et supprimer de la mémoire vive les conteneurs que vous avez lancés en exécutant la commande suivante : *containerlab destroy --topo topology.yaml*

Un labo de télémétrie offert par Nokia

Nokia propose sur github un labo très intéressant prêt à l'emploi, le Nokia SR Linux Streaming Telemetry Lab (https://github.com/srl-labs/srl-telemetry-lab/tree/main?-tab=readme-ov-filepour). Vous pouvez partir de ce labo de télémétrie qui représente une base solide pour en créer d'autres afin de tester différentes configurations. Les variations peuvent concerner pour l'essentiel la taille du réseau (nombre de nœuds/switches virtuels et nombre d'hôtes), mais aussi

Les outils de télémétrie fournis dans le labo de Nokia (gnmic, prometheus et grafana) permettent d'obtenir des visualisations variées et plutôt intéressantes des réseaux virtuels générés.

les types de nœuds. Vous pouvez utiliser comme « matériel virtuel » au moins tout ce que propose containerlab. Il est également possible de récupérer sur le site de docker ou ceux de constructeurs, d'autres images virtuelles de switches et de routeurs. Il faut cependant faire attention à un point important : les nœuds de réseau virtuels n'implémentent pas tous les protocoles. Si, par exemple, vous souhaitez récupérer des informations via le protocole SNMP (Simple Network Management Protocol), il faut impérativement choisir les hôtes réseaux compatibles parmi ceux disponibles. Le Nokia SR Linux est proposé en priorité dans le projet containerlab car il s'y intègre efficacement et implémente SNMP. Les configurations sont proposées, selon le process containerlab, sous la forme de fichiers YAML. Elles peuvent être chargées dans une vm Linux « classique » Debian (version 12 ou supérieure), Ubuntu (version 22.04 ou supérieure) ou avec une autre distribution ou bien encore directement dans CodeSpaces (https://github.com/codespaces/new/srl-labs/srl-telemetry-lab?quickstart=1. Pour installer le labo, il faut au préalable créer et se positionner dans le répertoire racine dans lequel vous souhaitez installer vos configurations de test. Si, par exemple, vous souhaitez créer un répertoire labs depuis votre répertoire utilisateur dans lequel seront stockées toutes vos configurations, vous devrez exécuter les commandes suivantes dans votre shell Linux :

cd ~
mkdir labs
cd labs
git clone https://github.com/srl-labs/srl-telemetry-lab.git

Si tout s'est bien passé, le labo Nokia est normalement installé sur votre machine (virtuelle ou physique). Ce labo de télémétrie s'appuie sur la norme de couverture réseau YANG des données d'état et de configuration (https://learn.srlinux.dev/yang/). La topologie du labo consiste en une topologie Clos avec des switches Nokia SR Linux s'exécutant sous forme de containers avec, en plus, une pile de télémétrie de streaming composée des applications gnmic, Prometheus et Grafana. Gnmic joue le rôle de collecteur de télémétrie, Prometheus celui de base de données Time-Series, et Grafana assure la visualisation des données. Vous trouverez la description de ces différents outils à l'adresse https://github.com/srl-labs/srl-telemetry-lab.

Charger un labo

Pour charger un labo, déplacez-vous dans son répertoire (avec la commande cd) et servez-vous du programme containerlab. Si le répertoire du labo est srl-telemetry-lab, lancez les instructions qui suivent. Sinon, remplacez simplement le nom du répertoire par celui qui vous intéresse :

cd srl-telemetry-lab/
 # si le fichier YAML de description du labo est st.clab.yml
sudo clab deploy – t st.clab.yml ou sudo containerlab deploy – t st.clab.yml

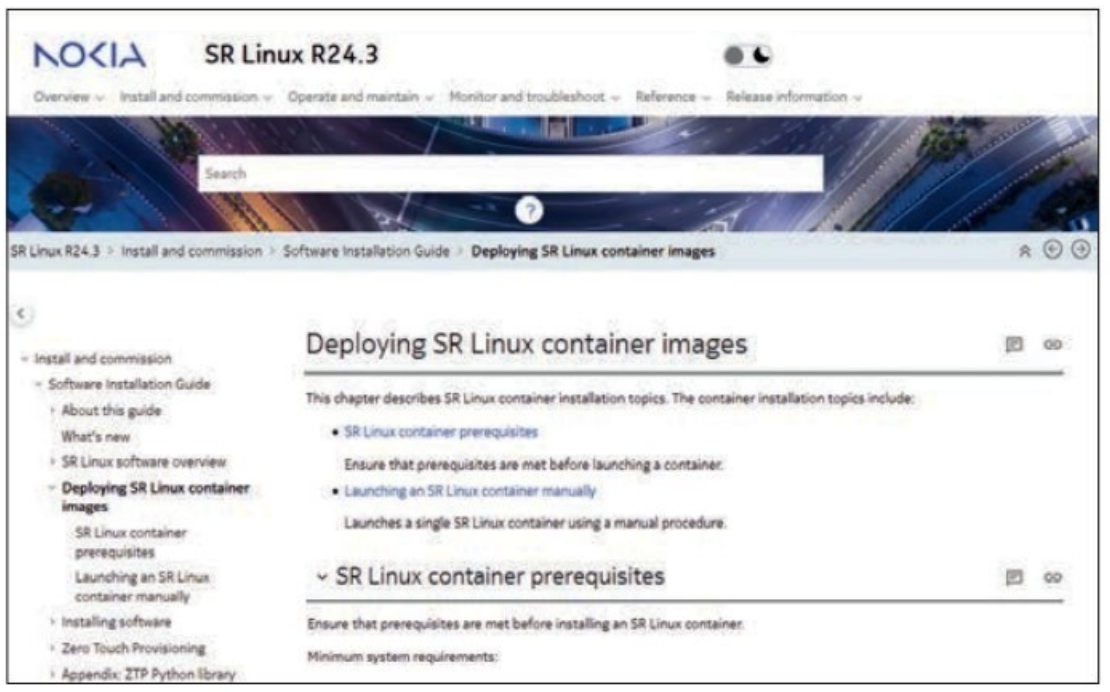
Pour réafficher les données de description du réseau et l'état de ses nœuds, employez cette instruction : *sudo clab inspect – t st.clab.yml*

Pour reconstruire et relancer un réseau virtuel : *sudo clab deploy --reconfigure – t st.clab.yml*

Cela s'avérera nécessaire si, par exemple, un des nœuds n'est plus en cours d'exécution.

Si tout se passe bien, vous devriez obtenir quelque chose de ce genre :

```
+---+------------+-------------------+-----------------------
--------------------+--------------+----------+-------------
---+--------------+
| #| Name| Container ID| Image| Kind| State| IPv4 Address
|IPv6 Address|
```



Vous pouvez aussi récupérer des images de conteneurs docker de nœuds réseau directement. Le site de Nokia propose — évidemment — des images de Nokia SR Linux comme ici à l'adresse https://documentation.nokia.com/srlinux/24-3/books/software-install/install-containers.html

```
+----+------------+--------------------+------------------------
--------------------+--------------+----------+-----------------
---+---------------+
| 1| client1| bacd21474868| ghcr.io/hellt/network-multitool|
linux| running| 172.80.80.31/24| N/A|
| 2| client2| f33b782b85ca| ghcr.io/hellt/network-multitool|
linux| running| 172.80.80.32/24| N/A|
| 3| client3| 75c7c2c3015c| ghcr.io/hellt/network-multitool|
linux| running| 172.80.80.33/24| N/A|
| 4| gnmic| 59d31645bd43| ghcr.io/openconfig/gnmic:0.33.0
| linux| running| 172.80.80.41/24| N/A|
| 5| grafana| ef8b89dddaae| grafana/grafana:10.2.1| linux|
running| 172.80.80.43/24| N/A|
| 6| leaf1| 2dabd15c0f40| ghcr.io/nokia/srlinux:24.3.2| nokia_
srlinux| running| 172.80.80.11/24| N/A|
| 7| leaf2| d9abf5aa6ca8| ghcr.io/nokia/srlinux:24.3.2|
nokia_srlinux| running| 172.80.80.12/24| N/A|
| 8| leaf3| 00bdf292f84e| ghcr.io/nokia/srlinux:24.3.2|
nokia_srlinux| running| 172.80.80.13/24| N/A|
| 9| loki| 04c2e8359fc7| grafana/loki:2.9.2| linux| running| 172.80.80.46/24|
N/A||10| prometheus| ea1ea6ae0326| quay.io/prometheus/prometheus: v2.47.2
| linux|running|172.80.80.42/24|N/A|
| 11| promtail| ed8fce78c2e6| grafana/promtail:2.9.2| linux| running|
172.80.80.45/24| N/A|
| 12| spine1| f07c6e2d814f| ghcr.io/nokia/srlinux:24.3.2| nokia_srlinux| running
| 172.80.80.21/24| N/A|
| 13| spine2| bf442fe24cc4| ghcr.io/nokia/srlinux:24.3.2| nokia_srlinux| run-
ning| 172.80.80.22/24| N/A|
| 14| syslog| 3da409e6cfe3| linuxserver/syslog-ng:4.1.1| linux| running|
172.80.80.44/24| N/A|
+-----+------------+-----------------+-----------------------------------------+-
-------------+------------+-----------------------+-----------+
```

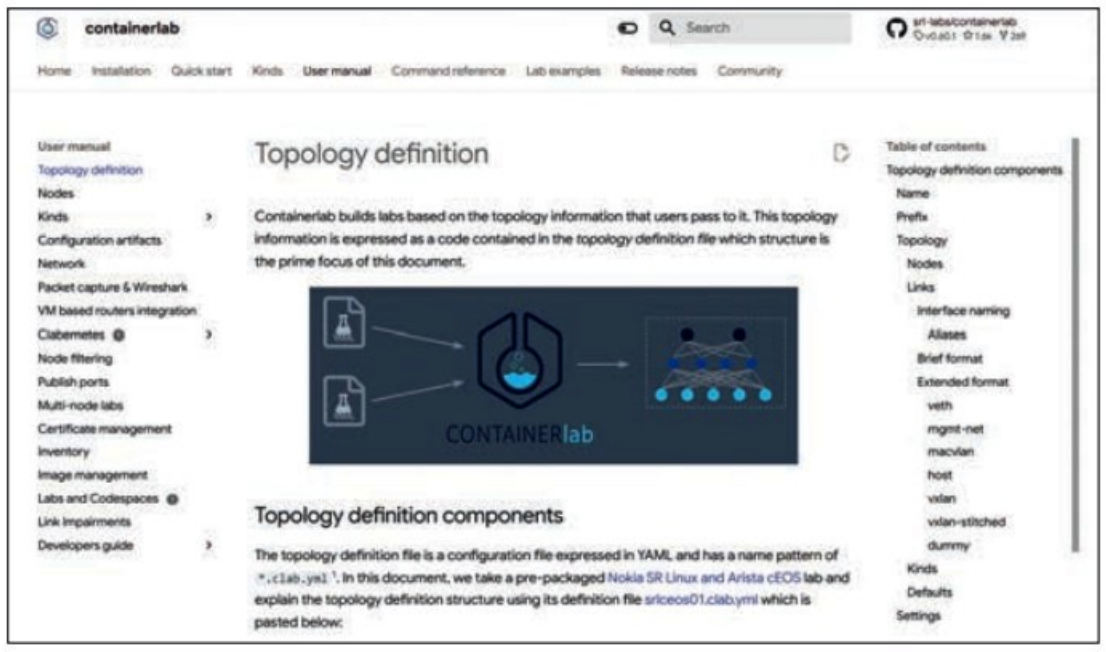
Le labo de Nokia fournit également un script de génération de trafic. Voici comment le lancer : ***sudo./traffic.sh start all***

Il faut bien entendu avoir lancé auparavant le réseau virtuel, comme ci-haut. Pour arrêter le script de génération de trafic, tapez : ***sudo./traffic.sh stop all***

Si vous devez décharger le labo, lancez la commande suivante : ***sudo clab destroy – t st.clab.yml***

ÉLÉMENTS DU **FICHIER YAML**

Le fichier YAML a une structure bien particulière. Voyons de quels éléments il est composé.
name : le nom de la topologie. Cela peut être n'importe quoi. Il doit néanmoins être unique au cas où vous voudriez en associer plusieurs.
topology : définit la structure du réseau. Contient des sous-éléments nodes et links.
nodes : chaque nœud (routeur ou switch) est défini avec son nom, son type (kind), l'image docker à utiliser (image), et les interfaces réseau (interfaces) avec pour chacune son nom (name) et son adresse IP (address).
links : définit les connexions entre les nœuds. Chaque lien (endpoints) associe deux interfaces de nœuds différents.
Vous trouverez tous les détails de la topologie d'un fichier YAML containerlab à l'adresse https://containerlab.dev/manual/topo-def-file/.



Vous trouverez tous les détails de la topologie d'un fichier YAML containerlab à la page du manuel qui lui est dédiée à l'adresse https://containerlab.dev/manual/topo-def-file/

N'oubliez pas que, sur le même principe que celui des vm, chaque labo chargé consomme une certaine quantité de RAM et que tant que vous ne l'avez pas « détruit » (déchargé avec la commande destroy), il continue à la monopoliser. Ces labos consommant pas mal de RAM (16 Go pour celui-ci), il faut impérativement en décharger un avant d'en lancer un autre pour ne pas risquer de saturer la mémoire vive. Au passage, si vous employez une vm Debian, Ubuntu ou autre sur un OS physique (Windows, Mac OSX ou Linux, peu importe), il faudra au moins 4 Go de RAM pour l'OS physique et, pour la vm Linux, 2 Go pour son OS plus ce que consomme votre labo. En clair, pour un réseau virtuel de 16 Go, mieux vaut prévoir un minimum de 24 Go en tout sur votre poste. Ce sera un peu plus « souple » avec 32 Go. Et si vous pouvez monter à 48 ou 64 Go, ne vous gênez surtout pas.

Accéder aux éléments du réseau

Une fois qu'un labo a été déployé, les différents nœuds SR Linux sont accessibles via SSH, à travers leur adresse IP de gestion. Celle-ci est affichée dans le résumé sous forme de tableau affiché après l'exécution des commandes deploy et inspect. Ces nœuds sont également accessibles via leur nom d'hôte affiché dans le même tableau dans la colonne Name, et défini dans le fichier de topologie en YAML. Les clients Linux « simples » ne peuvent, eux, être accédés directement via SSH, mais ils peuvent l'être avec une commande docker exec.

accès à une feuille (leaf) ou une épine dorsale (spine) SR Linux via SSH

```
ssh admin@leaf1
ssh admin@spine1
# accès à un client Linux via Docker
docker exec – it client1 bash
```

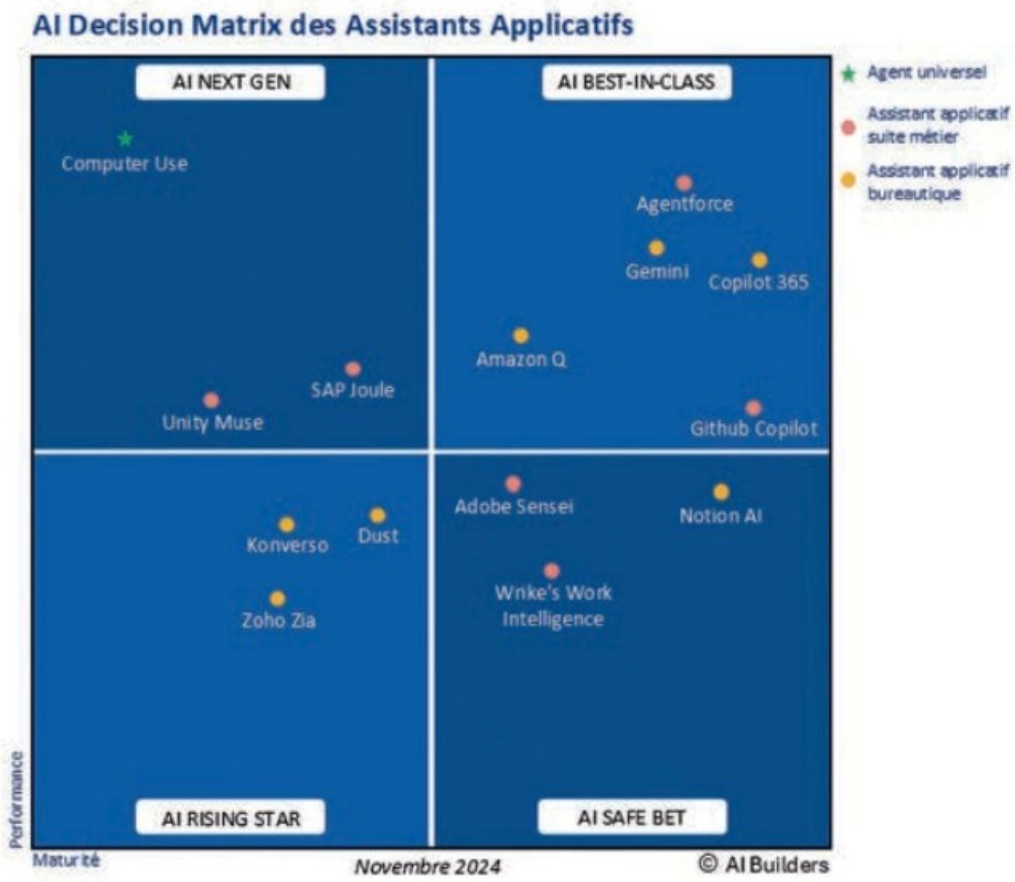
T.T

IA AI Builder passe au crible les assistant applicatifs

Le cabinet de conseil AI Builders Research a présenté son benchmark AI Decision Matrix. Il vise à comparer les performances et la maturité des assistants applicatifs qui s'invitent dans le monde de l'entreprise pour automatiser toujours plus de tâches. Explications.

Gartner estime que d'ici à 2027, plus de la moitié des modèles de GenAI utilisés en entreprise, seront spécialisés dans un domaine, un secteur spécifique ou une unité opérationnelle, contre 1% en 2023.

L'IA générative et les LLM se dirigent de plus en plus vers des solutions d'automatisation pour la réalisation d'une tâche spécifique, qui remodéliseront profondément les modes de travail. *« La vraie révolution, c'est qu'on passe d'un monde descriptif à un monde actionnable »* ; explique Stéphane Roder, président d'AI Builders. Cette tendance de fonds est poussée par les éditeurs, qui développent des assistants applicatifs intégrés à un logiciel ou à une suite logicielle, afin d'automatiser des tâches répétitives au sein de l'entreprise. Ces agents sont capables de raisonner et d'analyser des problématiques, d'attribuer des outils à des tâches complexes et de les exécuter.



Les résultats de l'IA Decision Matrix des assistants applicatifs d'AI Builders.

Face à cette mutation rapide et aux nombreuses solutions déjà disponibles sur le marché, AI Builders, un cabinet de conseil indépendant spécialisé dans la transformation Data IA des entreprises, a présenté AI Decision Matrix, un benchmark qui passe en revue les principales solutions d'agents applicatifs du marché, afin d'aider les entreprises à sélectionner la solution la plus adaptée à leurs besoins.

Focus sur la performance et la maturité

Pour réaliser son benchmark, le cabinet a divisé les assistants applicatifs en deux grandes familles : les suites bureautiques type Copilot de Microsoft et Gemini de Google, et les suites métiers telles que Agentforce de Salesforce et Joule de SAP. Ils ont été évalués selon deux critères : la performance et la maturité. La performance mesure la qualité de la réponse obtenue, la personnalisation de l'assistant — *« Un point très important pour nos clients et pour tout utilisateur qui souhaitent pouvoir réaliser une tâche propre à ses besoins »*, explique Pauline Delavalade, analyste en chef chez AI Builders — ensuite, tout ce qui concerne la sécurité et la gestion des données et la complexité des tâches réalisées.

La maturité s'attarde, quant à elle, sur la pénétration du marché par la solution, l'étendue des intégrations internes et externes, la facilité de mise en place et d'utilisation : *« on veut que ça soit facile, intuitif, que l'on puisse facilement décrire une tâche en langage naturel. Et enfin, est-ce une solution qui va répondre à un petit nombre de personnes, ou est-il possible de l'étendre au sein de l'entreprise »*, décrypte Pauline Delavalade. La maturité analyse également la scalabilité et les intégrations interne et externe. *« Aujourd'hui, les besoins des métiers, c'est de pouvoir avoir accès à tous les outils que j'utilise, et donc pour cela, il faut que l'assistant applicatif puisse aller chercher des données dans les outils externes »*, développe l'analyste.

Construire ses propres agents

Selon les résultats obtenus, les solutions ont été classées dans quatre familles suivant ces deux critères. Les Best in Class regroupent les solutions leader du marché, à la fois robustes et performantes, et dont l'intégration est jugée fiable dans les systèmes des entreprises. Deux suites métier y ont été intégrées, Now Assist de Service Now et Agentforce de Salesforce. De l'analyse d'AI Builders, Agentforce est la plus performante. Au-delà des capacités de génération automatique d'emails ou la création de chatbot pour la relation clients, c'est son outil de fabrication d'agent personnalisé, Agent Builder, qui fait la différence. Agentforce met aussi à disposition des agents prêts à l'emploi.

Sur les assistants applicatifs bureautiques, Amazon Q, Gemini et Copilot 365 ont été placés dans la catégorie AI Best-in-class. La suite de Microsoft se démarque essentiellement grâce à Microsoft Copilot Studio, qui permet

aussi de personnaliser ses propres agents. Il est également possible d'intégrer les agents d'autres plateformes, comme ceux de SAP ou de Service Now.

Vient ensuite l'assistant applicatif Gemini. Plus exclusif, il est axé sur l'écosystème Google et peut ainsi être intégré dans BigQuerry, GoogleWorkspace, GoogleColab entre autres. *« C'est sa grande force et cela lui permet de fournir un large éventail de fonctionnalités »*, adapté aux outils dans lesquels il est déployé. Gemini peut, par exemple, fournir des traductions en direct sur Google Meet, ou résumer des discussions dans Google Chat. Il est également possible de créer ses propres gems, des Gemini personnalisés qui peuvent être déployés sur des tâches spécifiques.

Des agents qui s'entraident

La seconde catégorie, dite Next Gen, correspond aux solutions performantes, mais dont la maturité doit encore progresser pour être pleinement exploitables. C'est dans cette catégorie qu'ont atterri Github Copilot et Joule, l'assistant applicatif de SAP. *« Nous ne l'avons pas mis [Joule ndlr] dans la première catégorie car elle manque encore de maturité »*. Toutefois, cette solution pourrait bien s'y retrouver sous peu. En effet, *« des fonctionnalités très performantes ont été annoncées pour 2025, comme la possibilité de créer des agents et des multi-agents qui s'entraideront pour répondre à une tâche spécifique »*, décrit Dimitri Calmand, analyste Data/IA chez AI Builders Research.

Stéphane Roder,
président de AI Builder

« SAP est un vieux dinosaure avec une grosse inertie qui, normalement, ne va pas vite et qui, pour le coup, est allé vite et a proposé un agent d'IA comme les autres, certainement aussi par peur de se faire distancer. C'est assez important pour le souligner. Si SAP le fait, c'est qu'on assiste sous nos yeux, à l'émergence d'un standard »

L'agent applicatif métier Amazon Q a, quant à lui, été classé dans la catégorie AI Safe BET aux côtés de Notion AI ou encore Adobe Sensei. Cette catégorie regroupe les solutions solides et matures, adaptables à de nombreux cas d'usage. Amazon Q peut être intégré à deux suites métiers, business et développeurs, et propose, pour ces derniers, des fonctionnalités de complétion, de suggestion et de modernisation de code. Amazon annonce un gain de rapidité de 25 % pour les développeurs. L'outil peut être intégré à des sources de données externes, type Google Drive ou Microsoft 365.

Les agents applicatifs assignés à la famille des Rising Stars sont les solutions prometteuses qui ont le potentiel pour devenir des leaders. Parmi elles, on trouve la plateforme d'IA conversationnelle Konverso, Muse, Zia et Wrike's Work Intelligence de Wrike, qui visent entre autres à automatiser des tâches manuelles, et Zia, un assistant d'IA pour les entreprises développé par Zoho.

Le benchmark est amené à être mis à jour très régulièrement. Stéphane Roder insiste : *« les solutions évoluent toutes les semaines, tous les mois. Il va y avoir des acteurs qui vont se repositionner, rebrander leurs assistants applicatifs, donc il faut garder en tête que cette étude est valable aujourd'hui »*. Une prochaine version est attendue au printemps prochain.

De l'IA oui, mais à quel prix ?

Ce que le benchmark n'explore toutefois pas, c'est la dimension coûts et modèles tarifaires appliqués par ces acteurs. Lesquels sont pourtant essentiels pour les entreprises qui espèrent tirer de la valeur de l'IA générative. Dans une étude Gartner, menée entre juin et juillet 2024, auprès de plus de 300 DSI, 90 % d'entre eux ont déclaré que la gestion des coûts limitait leur capacité à tirer de la valeur pour leur entreprise. Les DSI rencontrent également des difficultés à comprendre les coûts réels de la GenAI, et comment ceux-ci évoluent, ce qui pourrait les amener à commettre des erreurs de calcul des coûts de 500 % à 1 000 %, estime Gartner.

« Vous devez comprendre les composants des coûts et les options des modèles de tarification, et savoir comment réduire ces coûts et négocier avec les fournisseurs. Les DSI devraient créer des preuves de concept pour tester non seulement le fonctionnement de la technologie, mais aussi comment les coûts évolueront », previent Hung LeHong, vice-président analyste chez Gartner.

A cela s'ajoute que la valeur commerciale de l'IA générative ne se matérialise pas toujours forcément comme on s'y attend. Dans une enquête Gartner, réalisée au deuxième trimestre de 2024 auprès de 5 000 DSI dans plusieurs pays, les répondants ont déclaré gagner en moyenne 3,6 heures par semaine avec la genAI. Toutefois, tous n'en tirent pas les mêmes avantages. *« Les gains de productivité générés par l'interface genAI ne sont pas répartis de manière égale. Ils varient selon les employés, non seulement en raison de leur intérêt personnel et de leur niveau d'adoption, mais aussi en fonction de la complexité de leur travail et de leur niveau d'expérience »*, prévient Hung LeHong. □

V.M

Télétravail

À la recherche de la flexibilité

Remote, un éditeur de solutions RH en ligne, a réalisé une étude sur les différents modes de travail auprès de 4 000 dirigeants internationaux dans 10 pays dont la France. Le principal enseignement de ce sondage indique que les organisations offrant davantage de flexibilité en termes d'horaires et de lieux de travail ont gagné des recrutements aux dépens d'entreprises plus rigides.

Les réponses dans ce sondage sont sans équivoques : 71 % des responsables du recrutement ont vu leur entreprise perdre des employés au profit d'organisations offrant davantage de flexibilité en termes d'horaires et de lieux de travail, sur les six derniers mois. De plus, 85 % des recruteurs constatent une demande croissante de flexibilité de la part des salariés. 79 % des entreprises françaises qui ont des collaborateurs dans d'autres pays ont constaté une augmentation de leurs effectifs au cours de l'année écoulée, avec une croissance particulièrement marquée en Allemagne (86 %), au Royaume-Uni (81 %), et aux États-Unis (72 %). Ces entreprises sont aussi celles qui connaissent la plus forte croissance et 84 % d'entre elles recrutent actuellement. Particularité des USA et de la France, le recours à des ressources temporaires comme des freelances ou des prestataires de services est plus marqué. 34 % des entreprises françaises utilisent ces capacités extérieures. Il n'empêche qu'une large majorité d'entreprises dans le monde se sont converties ou ont élargi le travail à distance (70 %). Cette possibilité de travail à distance favorise aussi l'inclusion des femmes dans le monde du travail, dans une proportion désormais proche des 40 %. Autre remarque, le recours aux ressources extérieures provient aussi de la nécessité de s'adapter rapidement aux changements de la conjoncture.

PAS UNE PRIORITÉ PREMIÈRE POUR LES SALARIÉS FRANÇAIS

En juillet dernier, ADP a réalisé une étude sur le télétravail en France. Les résultats sont parfois surprenants. La flexibilité des horaires de travail est importante pour 26 % des salariés français (contre 31 % en 2023). Quel que soit leur âge, ils classent la flexibilité après le salaire (63 %), le plaisir au travail (44 %) et la sécurité de l'emploi (34 %). De plus, seuls 12 % des répondants (contre 15 % en 2023) estiment que la flexibilité du lieu de travail est un critère important dans leur emploi, soit une proportion légèrement inférieure à la moyenne en Europe (14 %), ainsi qu'à celles observées en Asie-Pacifique (15 %), en Amérique latine (15 %) et en Amérique du nord (17 %).

LES BÉNÉFICES CONSTATÉS DE LA FLEXIBILITÉ

JUST OVER **40%**

of hiring leaders surveyed say their company's flexibility initiatives have resulted in **enhanced work-life balance** for employees.

ALMOST **37%**

of hiring leaders surveyed say their company's flexibility initiatives have resulted in **higher productivity levels** for employees.

ALMOST **34%**

of hiring leaders surveyed say their company's flexibility initiatives have resulted in **increased employee engagement.**

ALMOST **32%**

of hiring leaders surveyed say their company's flexibility initiatives have resulted in reduced absences among employees.

ALMOST **26.5%** of hiring leaders surveyed say their company's flexibility initiatives have resulted in a **higher quality of job candidates** when hiring.

Une différence d'attractivité

47 % des entreprises françaises, avec un mode d'organisation du travail exclusivement en présentiel, ont du mal à trouver des candidats aux compétences recherchées, contre seulement 35 % de celles qui pratiquent le télétravail. Parmi les principaux avantages du travail à distance, les dirigeants français sondés citent également un meilleur équilibre travail-vie personnelle (44 %) et une meilleure satisfaction des employés (38 %). Les modèles de travail flexibles permettent notamment aux entreprises d'élargir leurs options en matière de recrutement. L'étude montre que l'augmentation de la productivité (35 %) et l'amélioration des performances commerciales (34 %) sont deux des principaux bénéfices attribués au travail à distance. Cela contraste avec la stratégie de certaines grandes entreprises, qui ont introduit des politiques de retour au bureau (RTO), invoquant la productivité comme justification de cette décision. Point intéressant, l'étude pointe que les entreprises flexibles attirent des candidats plus qualifiés

LES DIFFÉRENTS NIVEAUX D'EMBAUCHE

JUST OVER

85%

of those who make hiring decisions said they had expanded their global workforce with **senior-level** (e.g. director or VP) roles within the last year.

ALMOST

84%

of those who make hiring decisions said they had expanded their global workforce with **executive-level** (e.g. CEO) roles within the last year.

JUST OVER

80%

of those who make hiring decisions said they had expanded their global workforce with **mid-level** (e.g. manager) roles within the last year.

ALMOST

75%

of those who make hiring decisions said they had expanded their global workforce with **entry-level or junior** (e.g. coordinator or associate) roles within the last year.

ou de meilleur qualité (26,5 %) et connaissent un meilleur taux d'acceptation des postes proposés (24 %). De plus, la flexibilité assure une meilleure rétention des ressources internes en évitant de perdre des employés de valeur.

L'IA est un problème

L'utilisation de l'intelligence artificielle générative dans la rédaction des curriculums vitae devient un réel problème pour les entreprises avec des documents qui réhaussent sur le papier les qualités d'un candidat alors qu'ils ne sont pas forcément qualifiés pour le poste (65 % des responsables RH). Sur ces 65 %, 74 % indiquent que cela devient un vrai problème lors du recrutement, alors que les équipes RH sont déjà surchargées.

Les fausses annonces

Autre problème rencontré, près de 4 entreprises sur 10 ont posté de fausses annonces durant l'année écoulée. 34 % des répondants considèrent que 20 à 30 % des annonces sont frauduleuses ou du spam. Ceci devient une vraie question alors que le recours à ce type de plateforme comme Linkedin ou Indeed devient monnaie courante dans les entreprises. Cependant, le recrutement par ces plateformes connaît des fortunes diverses. Près d'un tiers des répondants indiquent avoir un taux entre 20 et 30 % de succès. Seulement 15 % assurent avoir des taux d'embauche par ces moyens proches de 40 %. Ces chiffres doivent être de plus pondérés par secteur d'activité.

Au global, 70 % des répondants pensent que la flexibilité est importante ou très importante quand ils proposent un travail. En ce sens les entreprises s'alignent plutôt sur les désirs des salariés et voient les possibilités proposées comme un avantage offert. Ainsi, si 60 % des entreprises proposent des horaires variables, seuls 41 % des employés utilisent cette formule. Il en est de même pour la plupart des possibilités offertes que ce soit le temps partiel, les durées de congés payés illimités...

Des bénéfices indéniables

La flexibilité comporte cependant des bénéfices importants comme l'augmentation de la satisfaction des employés (40,3 % des répondants). Quasiment dans la même proportion, elle apporte un meilleur équilibre entre vie personnelle et professionnelle. À plus de 30 %, elle augmente la performance des salariés et réduit les coûts fixes comme ceux de la location de bureaux ou de fournitures bureautique. Elle permet d'accéder à un pool de ressources plus larges, réduit le turn-over et contribue sur plusieurs aspects à la politique RSE de l'entreprise.

Des adaptations nécessaires

Reste qu'il faut s'adapter aux différentes règles locales, ce qui peut être un problème pour les entreprises globales. De plus, la gestion d'équipes distantes ne se réalise pas comme des équipes en présentiel, du fait d'emploi du temps asynchrone. Les répondants pointent aussi la difficulté à maintenir la culture de l'entreprise dans ces conditions. Moins de 30 % de ceux-ci pensent que ce n'est pas un problème.

Les coûts grimpent aussi. Plus de la moitié des sondés indiquent avoir constaté des hausses de coûts du fait de la mise en œuvre de solutions pour soutenir le travail à distance, ce qui n'est pas toujours compensé par la baisse des coûts sur les locaux. ☐

B.G

LES ÉLÉMENTS VUS COME UN BÉNÉFICE POUR LES SALARIÉS

JUST OVER **60%**

of hiring leaders surveyed said their company offers **flexible working hours** as an employee benefit.

ALMOST **49%**

of hiring leaders surveyed said their company offers **remote work** as an employee benefit.

ALMOST **33%**

of hiring leaders surveyed said their company offers **part-time work options** as an employee benefit.

ALMOST **28%**

of hiring leaders surveyed said their company offered a **"work from anywhere"** policy (a flexible location policy) as an employee benefit.

JUST OVER **26%**

of hiring leaders surveyed said their company offered enhanced **flexibility and support for new parents** (such as extended parental leave, ramp-up policies, flexible scheduling, job sharing, or child care stipends) as an employee benefit.

Face à l'urgence, pour être prêts à tout, tout de suite, tout le temps.

Faites un don à notre Fonds d'urgence.

ALLIANCEURGENCES.ORG

1 CLIC, 1 DON, 6 ONG EN ACTION

ALLIANCE URGENCES · ACTION CONTRE LA FAIM · care · handicap international · MÉDECINS DU MONDE · PLAN INTERNATIONAL Jusqu'à l'égalité · solidarités international

Regard sur 2024

Les hauts et les bas de la cybersécurité

Sommaire

C'est la fin d'année, et comme toujours, l'heure est venue de jeter un oeil dans le rétroviseur et de dresser le bilan de l'année écoulée : qu'est-ce qui a bien fonctionné, où aurais-je pu mieux faire, et où me suis-je complètement trompé ? C'est aussi l'occasion de préparer ses bonnes résolutions pour l'année à venir, tout en faisant preuve d'un peu d'indulgence envers soi-même.

Eh bien, pour le domaine de la cybersécurité, c'est un peu la même chose. Si 2024 a été marquée, une fois encore, par une explosion des menaces et de l'élargissement des surfaces d'attaque, révélant une fois encore les insuffisances de nombreux acteurs à y faire face efficacement, tout n'est pas à jeter pour autant.

Les Jeux olympiques de Paris, par exemple, ont démontré qu'il est possible de contrer efficacement la cybercriminalité grâce à une coordination sans précédent. Par ailleurs, les éditeurs de solutions se sont appropriés les avancées en intelligence artificielle pour améliorer l'efficacité de leurs outils — à voir toutefois ce que cela donnera concrètement à l'usage et avec suffisamment de recul. Enfin, malgré des retards initiaux, les directives NIS2, DORA ou encore le Cyber Resilience Act devraient prochainement renforcer la résilience de nombreuses entités, bien que leur mise en œuvre soulève encore des interrogations parmi les principaux concernés. Mais allez, pour une fois, en cette fin d'année, sans sombrer dans l'autosatisfaction ou sabrer le champagne, essayons de voir le verre à moitié plein.

En 2024, les cybercriminels n'ont pas fait de cadeaux

Malgré la défense réussie des Jeux olympiques, l'année 2024 a été marquée par une explosion des cyberattaques, une augmentation de la surface d'attaque et, forcément, une hausse importante des coûts de la cybermenace. Si les observateurs ont bien constaté une appropriation de l'intelligence artificielle par les cybercriminels, les entreprises de cybersécurité ne sont pas en reste.

Annoncés comme un cataclysme cyber, les Jeux olympiques et paralympiques de Paris ont été, au contraire, un franc succès, comme l'indiquent les autorités. Publié quelques jours après la clôture des Jeux, le bilan de l'Agence nationale de la sécurité des systèmes d'information (Anssi) a recensé 141 incidents entre le 26 juillet et le 11 août, dont 22 ont été menés à bien, mais avec un impact relativement faible. *« La clé de ce succès : la mobilisation d'une coalition d'acteurs (Anssi, COM Cyber) et la nomination d'un délégué dédié, qui a permis une coordination efficace autour des entreprises cruciales pour le bon déroulement des Jeux »*, a salué l'éditeur Sesame IT, dans son bilan 2024 sur la cybersécurité.

Les centaines d'organisations impliquées dans les JO étaient pour l'essentiel des fédérations sportives, des PME et des collectivités. Réputées vulnérables, elles ont mis en place des mesures d'hygiène numérique et des bonnes pratiques de base. Sesame IT estime que *« la protection numérique à l'œuvre lors des JO doit devenir la nouvelle norme en matière de cybersécurité »*. Tout l'enjeu consiste désormais à passer à l'échelle supérieure et à entraîner des milliers d'entités à mieux se protéger.

Les ransomwares toujours au top

Dans l'ensemble, la situation s'est aggravée et le paysage cyber de 2024 est plus que jamais moribond. Selon les dernières estimations issues des Technology Market Insights de Statista, le coût total des cyberattaques et autres actes malveillants en ligne devrait atteindre 129 milliards de dollars cette année, contre 94 milliards en 2023.

« Une fois encore, l'année a été riche en nombre d'attaques ayant mené à des vols de données et à leur mise en vente sur des marchés criminels », fait remarquer David Grout, chief technical officer (CTO) chez Mandiant pour la zone EMEA. Si les données divergent d'une étude à l'autre, des tendances se dégagent. Toutes ont effectivement constaté une hausse des incidents de sécurité. Le rançongiciel (ransomware) constitue, une fois de plus, le type de menace le plus répandu. Sur la période de juin 2023 à juin 2024, le Microsoft Digital Defense Report 2024 fait état, quant à lui, d'un doublement des cyberattaques d'une année sur l'autre, avec 78 000 milliards de signaux de sécurité analysés, soit 20 % de plus qu'en 2023. Les rançongiciels comptent parmi les attaques les plus répandues. La firme de Redmond a également observé que les attaques sur les mots de passe sont passées de 4 000 chaque seconde en 2023, à 7 000 par seconde un an plus tard.

Selon le rapport annuel de Check Point Software, les rançongiciels représentent, à eux seuls, 46 % des cas. Le nombre de victimes par ransomware a explosé de 90 % par rapport à 2023, et celles exposées publiquement ont doublé pour atteindre 5 000. Ce dernier chiffre n'est pas anodin et traduit une stratégie de plus en plus adoptée par les pirates, qui consiste à mettre au pilori les entreprises compromises — une façon de leur mettre la pression exercée pour les forcer à payer.

Les dernières estimations issues des Technology Market Insights de Statista

Le coût total des cyberattaques et autres actes malveillants en ligne atteindra 129 milliards de dollars en 2024, contre près de 94 milliards de dollars en 2023. À l'échelle mondiale, toujours selon Statista, cette proportion devrait atteindre 9,22 milliers de milliards de dollars, contre 8,15 en 2023.

> *« Lorsqu'un attaquant veut faire de l'exfiltration de données rapidement, ça ne sert à rien de tout exfiltrer. Or, il peut aller bien plus vite grâce à l'IA pour interroger les bases de données [scripting, ndlr], classifier les informations en fonction de leur nature (financières, personnelles…) et exfiltrer uniquement celles qui sont sensibles »*
>
> Raphaël Marichez, directeur sécurité Europe du sud chez Palo Alto Networks

Des stratégies de défense qui posent question

Autre exemple, une étude de Cohesity, réalisée auprès de 3 000 décideurs IT dans huit pays, montre une augmentation des cyberattaques et, en particulier – roulement de tambour – des ransomwares. Les chiffres sont éloquents : 86 % des entreprises françaises interrogées ont confirmé avoir été victimes d'une attaque par rançongiciel en 2024, contre 53 % en 2023. En outre, 92 % des organisations françaises interrogées, et 69 % à l'échelle mondiale, reconnaissent avoir versé une rançon à des cybercriminels au cours de l'année écoulée, alors même que la plupart affiche une politique de non-paiement des rançons.

Cette hausse s'explique de plusieurs manières, déjà bien connues et documentées. Parmi elles : l'augmentation de la surface d'attaque, la complexité croissante des infrastructures, les attaques de la supply chain, la sophistication et l'efficacité croissante des cybercriminels, la stratégie de Ransomware-as-a-Service qui consiste à vendre des services clés en main à de plus petits acteurs, multipliant ainsi le nombre d'acteurs malveillants et, dans une moindre mesure, l'adoption de l'IA générative. Ces chiffres toujours plus alarmants d'une année sur l'autre soulèvent de nombreuses questions quant à l'efficacité des stratégies actuelles adoptées par les entreprises, alors même que le cadre réglementaire s'est durci ces dernières années. Car c'est aussi, encore et toujours, la combinaison de plusieurs facteurs, allant des insuffisances persistantes en matière de protection au sein des organisations aux comportements à risque des salariés en interne, en passant par le manque de ressources et de budgets alloués à la cybersécurité, qui mène à ces compromissions. Le simple fait de payer la rançon est considéré comme un risque en soi par Cohesity, qui rappelle que cela conforte les criminels dans leurs actions et les pousse souvent à attaquer les bons payeurs à plusieurs reprises.

Petite lumière au bout du tunnel : les conseils d'administration et comités exécutifs prennent, chaque année, un peu plus conscience de l'importance de la cybersécurité, et 2024 ne fait pas exception. Le tissu économique prend aussi toujours un peu plus la mesure du risque. *« Il y a tellement d'attaques réussies depuis plusieurs années, qu'il y aura toujours un membre, au sein d'un conseil d'administration ou d'un comité exécutif (Comex), qui aura déjà vécu une attaque dans une autre entreprise où il siège »*, fait remarquer Raphaël Marichez, CSO régional pour l'Europe du sud chez Palo Alto Networks.

Dans ce contexte, le RSSI est davantage connecté aux organes de direction, que ce soit comme membre permanent ou en tant que consultant régulier, pour garantir que la cybersécurité est prise en compte dans les décisions stratégiques. Il reste toutefois un long chemin à parcourir. Une étude de Trend Micro, portant sur 2 600 RSSI, publiée en avril 2024, a dépeint un manque d'adhésion des conseils d'administration aux enjeux cyber : 79 % des répondants ont déclaré avoir ressenti une pression de la part du conseil d'administration pour minimiser la gravité des risques auxquels leur organisation est confrontée. Autre enseignement : 42 % ont estimé qu'ils étaient perçus comme trop négatifs, et seulement 56 % que leur dirigeant comprenait les risques.

La question de la cybersécurité est encore trop regardée sous le prisme de la mise en conformité. Les entreprises sont soumises à des exigences réglementaires de plus en plus strictes. Une mise en conformité qui coûte cher et qui doit donc faire l'objet d'arbitrages. Les conseils d'administration préfèrent des réponses stratégiques aux questions clés sur la cybersécurité, comme son impact sur les objectifs commerciaux ou le ROI (retour sur investissement), plutôt que des détails techniques ou des métriques complexes. Les RSSI qui alignent la cybersécurité sur la stratégie commerciale gagnent en crédibilité, tandis que les autres sont marginalisés. Près de 46 % des RSSI ayant mesuré la valeur commerciale de leur stratégie ont vu leur influence renforcée.

L'année 2024 a effectivement été marquée par un renforcement des exigences de conformité pour certaines entités, avec notamment l'entrée en application de la directive NIS 2, le 17 octobre dernier. Elle établit un ensemble de mesures juridiques, techniques et organisationnelles que les entités fournissant des services essentiels devront adopter en fonction du niveau de risque, pour renforcer leur défense et leur résilience opérationnelle. Quelque 15 000 structures sont concernées, rien que pour la France. Mais les effets de cette loi ne se feront pas sentir immédiatement.

Son parcours législatif a été bousculé par la dissolution de l'Assemblée nationale en mai 2024. Sa transposition en droit national est passée en Conseil des ministres et doit maintenant être discutée au Parlement. Une fois la loi de transposition votée, les entités concernées disposeront de trois ans pour mettre en œuvre les mesures de sécurité attendues.

Des attaques sur mesure

C'était la grande interrogation de l'année : quel serait l'impact de l'intelligence artificielle et de son pendant génératif dans le paysage cyber en 2024 ? *« Avec l'IA, une seule personne pourrait effectuer, en quelques secondes, ce qu'une équipe de hackers réalise en quelques jours aujourd'hui. Comme pour le Ransomware-as-a-Service (RaaS) en 2014, l'intérêt pour les grands modèles linguistiques (LLM) va se développer sur le darknet »*, prédisait l'éditeur de cybersécurité Quarkslab en 2023. Même son de cloche du côté de Palo Alto Networks, qui prévoyait que les attaquants allaient exploiter les LLM pour améliorer leurs emails d'hameçonnage ou lancer des attaques améliorées afin d'accroître leur taux de succès. Qu'en est-il un an après ?

« La moitié des attaques réussies sur lesquelles Unit42 [la division de recherche et de services en cybersécurité de Palo Alto Networks, ndlr] est intervenue, révèle un intervalle de temps de plus en plus court entre la première compromission et l'accomplissement de l'objectif final de l'attaquant », développe Raphaël Marichez. Cet intervalle est désormais en train de passer sous la barre des 24 heures. *« C'est la première fois que nous observons cela »*. Ce qui suggère, selon l'expert, un recours à des capacités d'IA et d'automatisation.

De son côté, Mandiant a observé que, pour l'essentiel, l'IA générative est utilisée par les attaquants pour réaliser des campagnes de phishing et de spear-phishing (hameçonnage personnalisé) plus convaincantes. *« Notamment en exploitant de nouvelles langues »*, précise David Grout, comme le japonais, qui sont plus difficilement exploitables par les attaquants qui mobilisent des LLM publics et des LLM malveillants comme Worm GPT pour produire des campagnes parfaitement traduites et adressables à de nouvelles cibles. Mandiant a également observé des tentatives d'utilisation de modèles dans l'accompagnement au code pour la création ou l'adaptation de malware.

Pour autant, *« 2024 n'aura pas été l'année de la généralisation des usages malveillants »*, soutient David Grout, qui prévient toutefois que le recours par les attaquants à l'IA générative et à l'IA plus globalement *« s'intensifiera en 2025 »*.

L'IA en 2024, d'abord une affaire de défenseurs

« Aujourd'hui, l'avantage est encore du côté de la défense », fait remarquer David Grout. Aucun doute, selon l'expert, que les acteurs malveillants n'ont pas la main sur la compréhension et le savoir-faire autour de l'IA. Développer, exploiter et maintenir cette technologie mobilisent des compétences techniques avancées et engendrent des coûts et des investissements importants, liés aux exigences de puissance de calcul pour l'entraînement des modèles et aux coûts d'usage notamment. Les cybercriminels n'ont ainsi pas les ressources nécessaires pour exploiter pleinement les avantages de l'IA, à la différence de certaines entreprises.

Car 2024 a été marquée par les très nombreuses annonces d'éditeurs qui ont ajouté de nouvelles capacités d'IA et de GenAI dans leurs solutions et services, afin de gagner en productivité. EDR, SOAR, SOC… tout y passe. Dernier en date, l'intégrateur réseau et sécurité Nomios qui a annoncé, en décembre 2024, intégrer la technologie du français Qevlar AI pour améliorer l'efficacité des analystes de ses centres opérationnels de sécurité (SOC). Qevlar AI va permettre d'investiguer les alertes, analyser et structurer les données de façon autonome, puis générer des rapports détaillés et des actions correctives à mener. Un rapport d'investigation pourra être généré en trois minutes par l'intelligence artificielle, contre 30 minutes auparavant. Éric Bohec, CTO de Nomios, affirme que les analystes pourront, dès lors, *« se concentrer davantage sur des missions critiques et à forte valeur ajoutée pour nos clients »*, comme le traitement des alertes prioritaires, la création de nouveaux playbooks, l'évaluation de nouvelles solutions de sécurité, entre autres.

Un rapport du Capgemini Research Institute, publié en novembre 2024 et intitulé *« Nouvelles défenses, nouvelles menaces : ce que l'IA et l'IA générative apportent à la cybersécurité »*, révèle que plus de 60 % des organisations qui utilisent ces technologies dans leurs SOC ont signalé une réduction de 5 % du temps de détection des incidents, 40 % ont déclaré une diminution de 5 % du temps de remédiation. Des gains a priori relativement faibles qui ne tempèrent toutefois pas l'enthousiasme des entreprises. *« 61 % considèrent l'IA comme critique pour répondre efficacement aux menaces, et permettre la mise en œuvre de stratégies de sécurité proactives contre des attaques de plus en plus sophistiquées »*, note effectivement le rapport. Ces sociétés en sont pour la plupart encore au premier niveau d'adoption, qui consiste essentiellement à

« Ce que nous avons observé est globalement cohérent avec ce que nous avions prédit, avec des tendances qui étaient orientées sur l'utilisation de l'intelligence artificielle par les attaquants »

David Grout, chief technical officer (CTO) chez Mandiant

Des standards de **cryptographie postquantiques** disponibles

En 2024, le Nist (National Institute of Standards and Technology) a publié les premiers standards pour des algorithmes de chiffrement postquantiques, visant à sécuriser les systèmes face à la menace potentielle des ordinateurs quantiques. Ces algorithmes sont prêts à être intégrés et devraient être déployés entre 2025 et 2035, selon les directives américaines. Toutefois, le Nist enjoint les experts en cybersécurité à ne pas perdre de temps et à les intégrer dans leurs systèmes le plus rapidement possible.
Le processus s'annonce complexe. En effet, la transition nécessite de réaliser un inventaire complet des systèmes utilisant la cryptographie actuelle, d'identifier les données sensibles et de déployer une mise à jour progressive des systèmes. À cela s'ajoute que la cryptographie est souvent implémentée de manière isolée par les développeurs, rendant difficile l'inventaire des systèmes à risque. Les équipes de développement devront ainsi être formées à l'intégration des nouveaux algorithmes.

automatiser des tâches fastidieuses et à faible valeur ajoutée. D'autres utilisent déjà l'intelligence artificielle à un niveau plus avancé. Google a annoncé, en fin d'année, que son outil d'intelligence artificielle Bip Sleep, développé par Project Zero et DeepMind, avait détecté une faille zero-day dans SQLite, un moteur de base de données open source. Une première.

La menace du Shadow AI

L'IA apporte aussi son lot de vulnérabilités qui donnent des sueurs froides aux RSSI et qui ont été largement documentées cette année. L'avènement de l'IA générative a fait craindre une explosion de la génération de codes malveillants par des cybercriminels, ou de codes de mauvaise qualité par les développeurs. C'est finalement un autre risque qui leur a volé la vedette. On connaissait le Shadow IT, désormais les organisations doivent composer avec le Shadow AI. Il consiste en l'utilisation non autorisée par des collaborateurs d'applications et d'outils d'IA en dehors des cadres approuvés par l'entreprise.
Ce phénomène inquiète particulièrement les responsables de la sécurité des systèmes d'information (RSSI), car il expose à des risques importants, notamment des fuites de données. Par exemple, un chatbot pourrait analyser des informations financières, IT et RH sans distinction, exposant ces données dans des conversations inappropriées. À cela s'ajoute un manque, voire une absence de segmentation ou d'étiquetage clair des données en interne, ce qui pose un problème de gouvernance. Par exemple, un salarié pourrait très bien interroger un chatbot sur sa fiche de paie, et des informations sensibles pourraient ensuite apparaître dans un échange sur les augmentations salariales, créant des situations délicates.
Face à cela, les entreprises peuvent adopter une posture restrictive, autorisant un nombre limité de solutions et bloquant les autres par défaut. « *Malgré tout, des outils d'IA et de productivité se glissent dans les réseaux sans supervision* », note Raphaël Marichez, et collectent des données, les traitent puis les restituent via des agents conversationnels. D'après le rapport du Capgemini Research Institute, 39 % des organisations ayant mis en place des restrictions d'usage constatent des violations fréquentes desdites règles.
Pour 2025, « *le chantier restera le même : bloquer les accès aux technologies qui ne seraient pas autorisées pour certains cas d'usage, sensibiliser aux bonnes pratiques et bons usages, et rediriger les salariés vers des outils internes ou ayant fait l'objet d'une contractualisation. Nous en sommes encore là* », confie Raphaël Marichez.

Deepfakes : la menace qui monte, qui monte

Cette année qui s'achève a marqué un tournant concernant les deepfakes. De nombreuses campagnes d'arnaques ont été recensées et documentées. Début septembre, les chercheurs de l'Unit 42 de Palo Alto ont publié une étude recensant des dizaines de campagnes d'escroquerie utilisant des vidéos deepfake, et mettant en scène des personnalités publiques, comme des PDG ou des responsables gouvernementaux. Ces technologies ouvrent la voie à un nouveau genre d'arnaques financières très convaincantes, en offrant aux escrocs de nouveaux moyens sophistiqués pour manipuler et piéger les salariés, en se faisant passer pour des dirigeants ou des figures d'autorité. Exemple le plus frappant, ce salarié d'une entreprise financière basée à Hong Kong qui a transféré 25 millions de dollars à des criminels après avoir été trompé par des deepfakes de ses collègues et de son supérieur. À l'avenir, la détection reposera davantage sur le recours à des outils utilisant eux-mêmes l'intelligence artificielle pour identifier des extraits générés ou modifiés par une IA. Comme souvent en cybersécurité, les organisations et leurs collaborateurs ne peuvent se contenter des seules technologies de protection, mais devront redoubler de vigilance. Il incombe également à chaque salarié de développer, plus que jamais, une forme de scepticisme à l'égard de ce qu'il voit et entend. **V.M**

L'analyse comportementale
pour détecter les attaques complexes

Le spécialiste en cybersécurité Netscout lance une nouvelle fonctionnalité d'analytique comportementale pour sa solution Omnis Cyber Intelligence, afin de détecter les attaques complexes. La société aide les entreprises à lutter notamment contre les ransomwares et les attaques zero-day, tout en améliorant la remédiation.

Netscout a réalisé une importante mise à jour de sa plateforme NDR (Network Detection and Response) Omnis Cyber Intelligence (OCI) avec l'intégration de fonctionnalités d'analyse comportementale alignées sur le référentiel Mitre Att&ck. Philippe Alcoy, spécialiste de la sécurité chez Netscout, nous explique comment ces nouvelles capacités renforcent la sécurité des entreprises face aux cyberattaques de plus en plus sophistiquées.

Pouvez-vous nous dire ce qu'apporte cette mise à jour en matière de détection et de réponse aux menaces réseau ?

En matière de détection, Netscout OCI a enrichi son moteur d'analyse comportementale avec des protocoles et des modèles supplémentaires. Elle a amélioré sa capacité de détection des fichiers malveillants pour une vue plus complète des menaces détectées. Les utilisateurs peuvent désormais accéder aux fichiers suspects et les télécharger en toute sécurité et consulter les événements associés à leurs détections. Concernant la réponse, Netscout a étendu la capacité d'OCI en introduisant une architecture ouverte pour l'intégration de tierces parties. Bien que sa conception unique soit principalement destinée à l'intégration avec des outils de remédiation (pare-feux et EDR), sa fonctionnalité personnalisable peut être étendue pour prendre en charge n'importe quelle plateforme avec une interface API.

Comment l'intégration du référentiel Mitre Att&ck et l'analyse comportementale élargie permettent-elles aux entreprises d'améliorer leur capacité à faire face aux menaces ?

Ce référentiel offre une base de connaissance détaillée des tactiques, techniques et procédures de cybersécurité de l'adversaire bien connues par la communauté SecOps. Dans OCI 6.4.0, nous fournissons une visibilité des menaces basée sur ce cadre, comme une alternative à la vue classique de type détection des menaces : intelligence autour des menaces, analyses de comportements, surface d'attaque, violations des règles de conformité, etc. Avec l'intégration de Mitre Att&ck, les opérateurs peuvent désormais lister les détections basées sur des catégories tactiques telles que la reconnaissance, l'accès initial, la découverte, le mouvement latéral, le commandement et le contrôle, ou encore l'exfiltration. Cela offre un nouveau contexte où la chronologie des menaces individuelles fournit un aperçu des objectifs et des motivations des acteurs malveillants.

Quelles sont les principales améliorations en termes d'interopérabilité avec des solutions tierces ?

Cette nouvelle fonctionnalité offre une intégration plus complète. Elle permet aux SecOps d'agir rapidement sur les détections pendant que l'enquête sur la portée complète et la cause première est effectuée. L'analyste peut bloquer le trafic à l'aide d'un dispositif en ligne ou le mettre en quarantaine à l'aide d'un EDR. L'architecture ouverte offre une fonctionnalité personnalisable pour répondre plus précisément aux besoins des clients. L'intégration et la configuration de n'importe quelle tierce partie avec des API exposées peuvent maintenant être réalisées en quelques jours. En plus de cette architecture ouverte, OCI offre déjà une intégration avec Splunk pour permettre aux opérateurs d'effectuer des analyses de sécurité et de détection des menaces dans un tableau de bord. OCI peut également s'intégrer à n'importe quel SIEM, XDR ou SOAR.

Comment votre solution améliore la détection proactive des menaces inconnues ?

La détection des attaques de type « zero-day » nécessite le diagnostic d'activités réseau inhabituelles et suspectes, associé à un processus d'investigation efficace et complet. Les détections analytiques comportementales d'OCI sont conçues pour découvrir les activités réseau anormales et suspectes, qui sont typiques de ces attaques. Elles peuvent également révéler des activités de menaces internes et des changements bénins dans les opérations commerciales du réseau. Un workflow d'investigation efficace est tout aussi important pour la détection d'activités suspectes et, grâce à la capture de paquets en continu d'OCI, les données nécessaires des communications passées deviennent disponibles pour corréler et confirmer les menaces dans le cadre d'investigations en cours. **J.C**

Swarm Learning

Une méthode de protection décentralisée des données

Le swarm learning s'impose comme une solution innovante pour la protection des données des entreprises. Alliant l'intelligence artificielle et la technologie blockchain, elle permet une collaboration décentralisée entre différents modèles d'apprentissage individuels, tout en garantissant un haut niveau de sécurité et de confidentialité des données.

À l'avant-garde de l'innovation technologique, le Lab Innovation du Lamarck Group vante les mérites du swarm learning auprès de ses clients. Basée sur l'IA et la blockchain, cette technologie vise à répondre à un défi majeur : comment partager des informations utiles entre différents acteurs sans exposer leurs données sensibles ? En alignant son fonctionnement sur le RGPD, le swarm learning offre une solution permettant de tirer parti des données de manière sécurisée, tout en respectant la vie privée. Elle présente, en outre, de nombreux avantages en termes de réduction des coûts, de traçabilité et de confiance numérique.

Thomas Boidot-Dorémieux, directeur du Lab Innovation de Lamarck Group, nous explique les enjeux du swarm learning pour la collaboration entre les entreprises.

Pouvez-vous nous présenter le Lab Innovation de Lamarck Group ?

Le Lab Innovation est le département de R&D du groupe Larmarck que l'on a créé en 2019. L'objectif était d'avoir un département de recherche en interne. En tant que cabinet de conseil spécialisé principalement en finance, on voulait pouvoir développer nos propres recherches sur des sujets innovants, qui nous permettent de déboucher sur des applications concrètes pour nos clients. Au départ, nous avons démarré sur deux sujets très particuliers : les risques climatiques en finance et en assurance, et les solutions pour les entreprises et les banques.

Sur quels types de projets travaillez-vous ?

Dans nos études sur la blockchain, nous avons travaillé sur un projet permettant de collecter, partager et monétiser les données extra-financières des entreprises. C'est un outil pour garantir la confidentialité de certaines données que l'on pourrait exploiter. Lorsqu'on travaille sur les données extra-financières, nous pouvons utiliser des données brutes, mais nous voulons éviter qu'elles soient divulguées et sortent des entreprises. Avec les approches classiques, nous avons nos jeux de données sur lesquels nous entraînons nos modèles. Plus nous avons de données, plus le modèle sera performant. Par contre, cela implique de collecter énormément de données, de les centraliser à un endroit et de les exploiter par une entité spécifique. Cela peut fonctionner pour des données publiques, mais ce n'est pas viable pour des données confidentielles. Lorsque c'est centralisé, le système devient plus vulnérable aux pannes et aux fuites de données. Si le système est attaqué, il peut y avoir des fuites massives, ce qui peut être préjudiciable pour les entreprises.

Comment intégrez-vous le swarm learning dans vos projets de recherche actuels ?

Pour le swarm learning, nous essayons de combiner différentes techniques en utilisant du edge computing et en travaillant avec les données en local. Cela signifie que l'entraînement est décentralisé et fractionné sur l'ensemble des endroits où se trouvent les données. Les modèles utilisés sont principalement des réseaux de neurones. Ils sont incrémentaux : à chaque étape, nous faisons tourner le modèle, et récupérons un ensemble de paramètres, puis nous répétons le processus avec les dernières mises à jour. Il faut un chef d'orchestre pour gérer l'authentification, la sécurité, et l'harmonisation des transactions entre les différents nœuds, ainsi qu'un nœud agrégateur pour collecter les résultats et les redistribuer. Cette partie authentification et sécurisation des transactions est possible grâce au concept du swarm learning.

Quelles sont les applications concrètes dans le cadre d'une collaboration interentreprises ?

Nos principaux clients sont les banques qui collectent un ensemble de données sur leurs clients. Ces données sont confidentielles et ne peuvent pas être partagées. Par contre, elles peuvent servir à entraîner des modèles. Chaque banque réalisant la même chose, nous sommes capables d'avoir des modèles entraînés comme s'ils l'avaient été avec l'ensemble des données des différentes banques. Cela permettrait par exemple de créer des projets interbancaires d'analyse de crédit sur l'ensemble de leurs clients, sans qu'elles aient besoin de partager leurs données. Il y aurait également la possibilité de travailler sur des cas d'usage liés au risque climatique. L'objectif serait d'étudier la sensibilité et la résilience des entreprises face au risque climatique, et notamment le risque climatique physique et les aléas climatiques extrêmes.

J.C

Sitting Ducks : une menace dormante qui gagne du terrain

Les attaques de type « Sitting Ducks » constituent une menace encore trop méconnue par les experts en cybersécurité. De plus en plus sophistiquées, elles exploitent les vulnérabilités des configurations DNS pour détourner des noms de domaine et mener ensuite des campagnes malveillantes sous le couvert de leurs propriétaires légitimes.

Bien que fréquentes et représentant un risque significatif pour les entreprises, les attaques dites « Sitting Ducks » restent peu documentées. Spécialisée dans l'identification des dispositifs connectés au réseau et la gestion des DNS (système de nom de domaine) pour les entreprises, Infoblox Threat Intel tire la sonnette d'alarme sur cette menace en plein essor. Selon son dernier rapport, plus d'un million de domaines enregistrés pourraient être vulnérables.

« Infoblox est une entreprise qui s'occupe principalement des DNS pour les entreprises. Depuis plusieurs années, nous intégrons de plus en plus de sécurité informatique dans nos produits. Pour cela, nous avons toute une équipe dédiée à la cybersécurité qui mène des recherches et travaille pour bloquer les noms de domaine piratés ou spécifiquement créés pour des activités malveillantes », explique Jacques Portal, threat researcher chez Infoblox.

Vaccant Vipper et Vextrio Viper : des serpents au cœur des TDS

La société a découvert de nouvelles preuves selon lesquelles un attaquant peut prendre le contrôle complet d'un domaine en s'appropriant ses configurations DNS. Des dizaines de milliers de noms de domaine de marques connues et d'organisations étatiques sont détournés chaque année. Durant son dernier programme de surveillance des DNS lancé en juillet 2024, Infoblox a détecté pas moins de 800 000 domaines vulnérables, dont 70 000 ayant d'ores et déjà été détournés.

> ***« Les principaux défis sont liés aux systèmes de sécurité qui s'appuient sur la réputation des noms de domaine. Cette réputation est souvent basée sur l'ancienneté et le trafic observé sur un domaine »***
>
> Jacques Portal, threat researcher chez Infoblox

Actif depuis décembre 2019, le groupe Vacant Viper détourne environ 2 500 domaines chaque année pour alimenter son système de distribution de trafic malveillant (TDS : Traffic Distribution System), baptisé 404TDS. Ce système sert de rampe de lancement pour des campagnes de spam, la diffusion de contenus pornographiques, l'installation de serveurs de commande et de contrôle (C2) destinés à des chevaux de Troie d'accès à distance (RAT), ainsi que pour des logiciels malveillants tels que DarkGate et AsyncRAT. Cet acteur privilégie des domaines à haute réputation d'entités connues pour contourner plus facilement les défenses traditionnelles. *« Les attaquants exploitent des noms de domaine enregistrés par des entreprises pour mener des activités criminelles, comme le phishing ou la diffusion de logiciels malveillants. Nous avons par exemple observé des noms de domaine appartenant à McDonald's ou à CBS détournés pour rediriger des victimes vers des sites malveillants. Bien que ces attaques ne visent pas directement les entreprises en tant que victimes finales, elles n'apprécient évidemment pas que leur nom soit utilisé pour de telles activités »*, ajoute l'expert.

Si un employé d'une entreprise est victime de ce type d'arnaque, les attaquants peuvent par exemple voler de petites sommes d'argent (comme 5 euros pour une fausse taxe de livraison), mais aussi obtenir des informations sensibles, comme des adresses e-mail ou des mots de passe professionnels. Ces données peuvent ensuite être revendues à prix élevé sur le marché noir.

Un autre groupe appelé Vextrio Viper dirige, quant à lui, le plus vaste programme d'affiliation cybercriminel. Depuis 2020, il redirige du trafic compromis vers plus de 65 partenaires affiliés, dont certains sont connus pour détourner des domaines via des attaques « Sitting Ducks ».

Horrid Hawk et Hasty Hawk : l'émergence de nouveaux acteurs malveillants

Incarnés par Horrid Hawk et Hasty Hawk, les jeunes groupes « Hawks » exploitent les failles DNS pour leurs campagnes frauduleuses. En activité depuis seulement 2023, Horrid Hawk détourne des noms de domaine pour orchestrer notamment des campagnes d'investissement frauduleuses attractives dans plus de 30 langues et sur plusieurs continents. Ces campagnes sont diffusées dans des publicités éphémères sur Facebook, via des domaines en apparence légitimes. Identifié en 2022, Hasty Hawk se concentre, quant à lui, sur des campagnes de phishing massives en usurpant des marques de confiance comme DHL. Il crée également de faux sites de dons liés à des causes sensibles pour tromper les utilisateurs et voler leurs informations. Grâce à un système TDS, ce groupe redirige habilement les utilisateurs vers des contenus adaptés à leur géolocalisation, tout en multipliant les thèmes de campagne pour brouiller les pistes.

Une menace grandissante qui se diversifie

Les acteurs exploitant les attaques « Sitting Ducks » sont capables d'infiltrer des domaines vulnérables à des fins diverses et variées, allant des fraudes financières aux vols d'informations sensibles, en passant par le déploiement de ransomwares. La montée en puissance de ces groupes malveillants impose une vigilance accrue, non seulement pour sécuriser les infrastructures existantes, mais aussi pour anticiper leurs prochaines manœuvres. Pour Jacques Portal, les hébergeurs de noms de domaine ne sont pas assez vigilants : *« Beaucoup d'hébergeurs ne vérifient pas suffisamment les informations des propriétaires avant d'accorder le contrôle d'un nom de domaine. Bien que des solutions existent, les hébergeurs hésitent souvent à ajouter de la complexité par peur de perdre des clients »*.

Les « Vipers » et les « Hawks » incarnent une menace évolutive et diversifiée qui exige des réponses innovantes et adaptées pour contrer leur expansion. *« Les principaux défis sont liés aux systèmes de sécurité qui s'appuient sur la réputation des noms de domaine. Cette réputation est souvent basée sur l'ancienneté et le trafic observé sur un domaine »*. Les domaines anciens, appartenant à des entreprises de confiance, ne sont pas bloqués facilement, ce qui complique la détection des attaques. Il existe des outils, mais une grande partie du travail repose sur la veille manuelle et l'analyse de données accumulées sur plusieurs années. *« Chez Infoblox, nous avons développé une approche différente : au lieu de nous fier uniquement à la date de création d'un domaine, nous considérons un domaine inactif depuis longtemps comme* ***« nouveau » lorsqu'il redevient actif. Nous examinons aussi l'adresse IP associée au domaine, en particulier si elle est liée à des zones suspectes comme la Russie »***, décrit l'expert. Les attaquants s'adaptent toutefois de mieux en mieux. Ils parviennent à détecter et contourner les systèmes automatisés de détection déployés par les entreprises de cybersécurité. ▮

J.C

www.linformaticien.com

ABONNEZ-VOUS À L'INFORMATICIEN

linformaticien.com/abonnement

MAGAZINE

Recevez chaque mois (10 numéros par an) le magazine «papier» et accédez également aux versions numériques.

1 AN FRANCE : 72 €
2 ANS FRANCE : 135 €
1 AN UE : 90 €
2 ANS UE : 171 €
1 AN HORS UE : 108 €
2 ANS HORS UE : 207 €

NUMÉRIQUE

Accédez chaque mois (10 numéros par an) à la version numérique du magazine et retrouvez également via votre compte en ligne les versions numériques des dernières publications.

1 AN : 49 €
2 ANS : 89 €

ÉTUDIANT / ÉCOLE

Abonnez vos étudiants avec une formule dédiée à 60 % du prix normal de l'abonnement sous forme de PDF (10 numéros par an).
Possibilité abonnements groupés en contactant le service abonnements du magazine à abonnements@linformaticien.com.

ABONNEMENT 1 AN : 43, 20 €

Qu'est-ce qui attend les DPO en 2025 ?

Par Paul-Olivier Gibert, président et Patrick Blum, délégué général

En 2025, les délégués à la protection des données (DPO) devront s'adapter à plusieurs évolutions majeures dans le domaine de la protection des données personnelles. Voici les principales tendances et défis qui les attendent.

Une jungle réglementaire en expansion

En 2025, les DPO feront face à une prolifération de nouvelles réglementations qui complexifieront davantage leur mission. Parmi elles, les premières applications concrètes du règlement sur l'IA, le règlement sur la cyber résilience et la mise en œuvre de la directive NIS2. Si ces textes partagent une philosophie commune avec le RGPD — responsabilisation des acteurs, sanctions en cas de manquement — ils introduisent également des chevauchements et des effets de bord. Les DPO devront ainsi jongler avec ces nouvelles exigences tout en gardant une vue d'ensemble pour garantir une mise en conformité cohérente.

Gouvernance : une collaboration à redéfinir

Dans les organisations, le rôle du DPO dépasse aujourd'hui le simple dialogue avec la DSI. Avec la généralisation des outils en mode SaaS et des services décentralisés, les métiers contournent souvent les services IT traditionnels, introduisant des risques nouveaux pour la conformité. Cette évolution appelle à une collaboration renforcée entre tous les départements (marketing, ressources humaines, etc.) et une meilleure intégration des pratiques de protection des données dans les processus opérationnels.

L'IA : opportunités et défis

L'IA, et particulièrement l'IA générative, représente un bouleversement économique et sociétal comparable à celui de l'Internet dans les années 2000. Cependant, cette technologie pose des défis inédits aux DPO. Comment appliquer le RGPD à des systèmes qui collectent et traitent des masses de données sans finalité précise dès leur collecte ? L'équilibre entre innovation et conformité sera crucial, d'autant que ces technologies s'intègrent progressivement aux systèmes de production des entreprises.

Menaces croissantes sur les données

L'adoption croissante de l'intelligence artificielle, ou de l'Internet des objets notamment, pose de nouveaux défis en matière de protection des données. Ces nouveaux usages et nouvelles tendances impliquent toujours plus de données personnelles, et les cyberattaques continuent de se multiplier. Les DPO devront élaborer et tester régulièrement des plans de réponse aux incidents, tout en assurant une communication efficace avec les autorités de régulation et les personnes concernées en cas de violation de données. L'objectif des DPO est ainsi triple : prévenir ces risques, réagir efficacement, tout en renforçant la formation des collaborateurs pour réduire les failles humaines. Les évolutions rapides du paysage technologique et réglementaire imposent donc aux DPO de se former et d'être en veille en permanence pour maintenir et développer leurs compétences.

Sensibilisation accrue des publics

Un changement notable, ces dernières années, est l'attention croissante du public envers la protection de ses données personnelles. Les individus sont de plus en plus conscients de leurs droits en matière de données personnelles. Cette évolution exerce une pression supplémentaire sur les organisations et leurs DPO, notamment à travers une augmentation des demandes d'exercice des droits d'accès, de rectification ou de suppression des données personnelles par les individus. Les entreprises doivent s'adapter à cette nouvelle réalité, en renforçant leur transparence et leur gestion des traitements de données.

Des facteurs d'incertitude à surveiller

2025 sera marquée par plusieurs incertitudes politiques et économiques. L'impact potentiel de la présidence Trump sur le Data Privacy Framework et, plus largement, sur les échanges de données entre l'Union Européenne et les États-Unis, combiné à une instabilité institutionnelle en France, pourrait créer des perturbations. En premier lieu, le gel de la loi d'application NIS2, déstabilisant pour l'entreprise qui travaille sur sa mise en conformité. Ces contextes doivent inciter les DPO à renforcer leur veille et à s'adapter rapidement aux éventuelles évolutions.

Les DPO se trouvent à la croisée des chemins entre innovation technologique, exigences réglementaires et attentes sociétales. Pour relever ces défis, ils devront cultiver une vigilance permanente, renforcer leur collaboration interne et externe, et préparer leurs organisations à un avenir où la protection des données sera plus que jamais au cœur des priorités.

À gauche, Paul-Olivier Gibert, président ; à droite, Patrick Blum, délégué général

Loi Naegelen : un nouveau tournant dans la lutte contre le spoofing

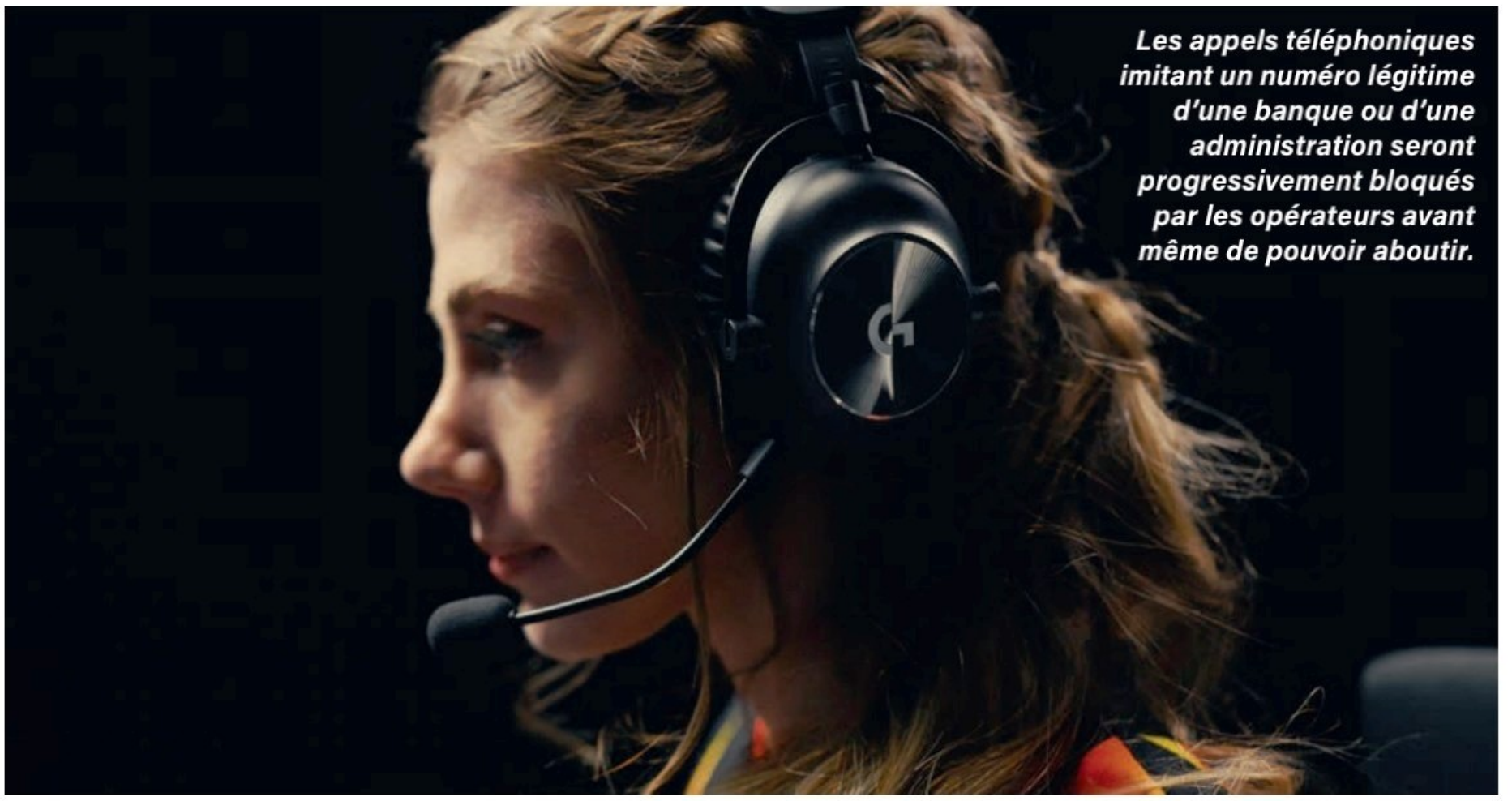
Votée en 2020, la loi Naegelen a été renforcée le 1er octobre dernier pour permettre de lutter plus efficacement contre l'usurpation des numéros de téléphone. Celle-ci impose désormais aux opérateurs de mettre en place un système d'authentification de l'origine des appels, pour mettre notamment un terme à la fraude aux faux conseillers bancaires et autres arnaques téléphoniques.

Les fraudes au « spoofing » – techniques d'usurpation d'identité téléphonique – continuent de faire des ravages dans les entreprises. En affichant un numéro légitime sur le terminal de leurs cibles, les malfrats parviennent à leur inspirer confiance et les manipuler par le biais de scénarios parfaitement rodés. Ils se font généralement passer pour des conseillers bancaires ou des représentants officiels de divers organismes (assurances, supports techniques, services administratifs...), dans le but d'obtenir des informations confidentielles et perpétrer des escroqueries pouvant parfois être très coûteuses pour les entreprises. Programmé à l'origine en 2023 dans le cadre de la loi Naegelen, le nouveau dispositif d'authentification de l'identifiant de l'appelant est finalement entré en vigueur au 1er octobre 2024 avec plus d'un an de retard.

Dans un premier temps, cet outil de contrôle ne fonctionnera pas sur toutes les technologies de réseau. Il sera déployé sur les appels non authentifiés de fixe à fixe (hors lignes de cuivre), puis étendu progressivement au réseau historique en cuivre, aux réseaux mobiles, et enfin aux réseaux téléphoniques d'entreprises. Les arnaques téléphoniques qui ciblent les entreprises devraient donc perdurer encore un certain temps. Même si cette nouvelle réglementation pourrait à terme réduire considérablement les fraudes téléphoniques, la vigilance reste de mise. Comme toujours, les fraudeurs ne manqueront pas de s'adapter pour trouver de nouvelles parades. Dans ce contexte, les entreprises doivent adopter une approche proactive en renforçant leurs propres systèmes de sécurité.

Inquest, une filiale spécialisée dans la prévention et la cybersécurité

La filiale du groupe Stelliant « Inquest » est un acteur majeur en France dans les services d'assurance. Elle possède une forte activité centrée sur l'expertise après sinistre dans divers secteurs tels que la construction, l'aéronautique, l'automobile et la finance. Inquest représente la filiale conseil et prévention du groupe dans les domaines de l'incendie et de la cybersécurité. En cybersécurité, son approche est organisée en trois pôles : la prévention (conseil), la réponse aux incidents et l'investigation numérique. L'ensemble de ses interventions vise à renforcer la résilience des entreprises face aux risques cyber. *« Stelliant assiste les entreprises et les professionnels pour gérer la crise, en*



Les appels téléphoniques imitant un numéro légitime d'une banque ou d'une administration seront progressivement bloqués par les opérateurs avant même de pouvoir aboutir.

Thibaut Carré, directeur cybersécurité d'Inquest, observe une recrudescence des fraudes par spoofing ciblant les entreprises. Pour l'expert, la mise en œuvre du nouveau système d'authentification des numéros d'appel prévu par la loi Naegelen prendra un certain temps avant d'être pleinement opérationnelle.

mettant en place des solutions techniques, des actions de remédiation et des plans de sécurisation pour une reprise d'activité en toute sérénité. L'intervention ne se limite pas aux entreprises déjà affectées ; elle propose également des services de prévention et de conseil aux entreprises souhaitant renforcer leur sécurité, notamment via des audits et des conseils en conformité réglementaire avec les normes RGPD, DORA, et NIS2 », explique Thibaut Carré, directeur cybersécurité d'Inquest.

Le spoofing : un levier pour les fraudeurs

« Dans une fraude par spoofing, le fraudeur utilise des techniques de masquage pour dissimuler son identité et son numéro de téléphone, souvent grâce à des outils spécifiques. Il peut faire afficher un numéro d'appel comme celui d'une banque ou d'un conseiller, et ainsi se faire passer pour une personne de confiance », précise Thibaut Carré. Cette technique est similaire à celle utilisée pour falsifier des adresses email, où des alias légitimes sont créés pour rendre le message crédible. Les fraudeurs emploient des logiciels et services en ligne pour modifier l'identification de l'appelant, masquer leur véritable numéro et tromper leur interlocuteur. *« Le problème actuel réside dans la nécessité de déclarer les numéros de téléphone et de contrôler les outils utilisés par les fraudeurs. Bien que des lois soient en place pour réguler les pratiques commerciales abusives, elles doivent être complétées par des solutions techniques »*. Des dispositifs comme Bloctel, une initiative de l'Arcep pour bloquer le démarchage, permettent par exemple de mettre des numéros sur liste noire. Les technologies actuelles ne sont toutefois pas toujours au niveau requis. *« Les opérateurs et constructeurs de téléphones devraient travailler ensemble pour intégrer des systèmes de sécurité capables de s'aligner sur les lois en vigueur et de renforcer la protection des utilisateurs contre les appels frauduleux. J'ai bon espoir que ce sera mis en place un jour, mais cela va nécessairement prendre un peu de temps »*, ajoute l'expert.

Des indices pour détecter une tentative de spoofing

Dans le domaine bancaire, certains indices doivent alerter : *« dans le cadre de nos activités de réponse aux incidents, nous voyons beaucoup de fraudes en tous genres, dont surtout la fraude aux faux conseillers bancaires, et dans une moindre mesure aux faux supports techniques. Le spoofing constitue, entre guillemets, une technique parfaite pour cela. Elle permet aux fraudeurs de se faire passer pour un agent bancaire ou un support technique légitime, comme celui d'une grande banque ou d'un opérateur »*, rappelle Thibaut Carré. Ces pratiques trompeuses compliquent la tâche des entreprises, car les utilisateurs ont tendance à devenir de plus en plus méfiants, même envers des communications légitimes. *« Nous recevons quotidiennement des signalements de fraudes par téléphone ou par email, y compris des fraudes sophistiquées comme la modification des RIB dans les emails »*. Face à ces menaces, le directeur cybersécurité d'Inquest insiste sur l'importance des mesures organisationnelles : *« pour détecter ces tentatives de fraude, des signes peuvent être observés, surtout dans le contexte bancaire. »* Par exemple, les banques authentiques suivent des processus rigoureux pour confirmer l'identité de leurs appelants, contrairement aux fraudeurs. De plus, les escrocs demandent souvent des informations sensibles, comme des codes de validation de transaction ou des détails de compte, informations que les conseillers bancaires légitimes ne demanderaient jamais. *« Les utilisateurs doivent être attentifs à ces indices pour éviter les pièges de l'ingénierie sociale »*.

Mesures préventives

La mise en place de processus est essentielle pour sécuriser les opérations comptables et financières des entreprises. Thibaut Carré conseille de mettre en place une validation en deux étapes lors, par exemple, des changements de coordonnées bancaires : *« une seule personne ne doit pas pouvoir modifier ces informations et effectuer des transactions sans vérification. Cette approche de gouvernance des paiements est cruciale, notamment pour prévenir les fraudes. Cependant, les petites entreprises ont plus de difficultés à adopter ces mesures, car elles disposent de peu de personnel pour séparer les tâches, contrairement aux grandes entreprises qui peuvent répartir ces responsabilités entre plusieurs collaborateurs »*. En cas de fraude avérée, il faut aller vite : *« la réactivité est cruciale pour tenter de récupérer les fonds. Si les sommes ont été transférées à l'étranger, elles deviennent généralement irrécupérables »*, avertit-il. Thibaut Carré recommande également des assurances contre la fraude ou une assurance cyber pour limiter les impacts financiers de telles attaques.

J.C

Supply chain, le maillon faible de la cybersécurité

À l'heure actuelle, de plus en plus de cybercriminels ciblent la supply chain, car infiltrer une entreprise leur permet d'accéder à d'autres organisations par « rebond ». Il est donc essentiel de comprendre ce qu'est la supply chain, afin que tous ses acteurs puissent travailler de concert à sa protection et à celle de leurs clients.

Le même code a été retrouvé dans plusieurs logiciels compromis, montrant les liens entre diverses attaques.

La supply chain (chaîne d'approvisionnement) englobe toutes les parties tierces avec lesquelles une entreprise peut être amenée à collaborer. Il peut donc s'agir de fournisseurs divers, de partenaires ou bien encore de prestataires. Elle est assez souvent confondue avec la chaîne logistique, alors que celle-ci ne considère que les acteurs ayant un rôle dans la production et la distribution d'un bien. Ces dernières années ont vu une augmentation constante des attaques ciblant la chaîne d'approvisionnement et ayant pour but de distribuer des logiciels malveillants. Les chercheurs d'Eset détaillaient déjà en 2019 des attaques ciblant des éditeurs et des plateformes de jeux vidéo. Ils avaient enquêté de manière approfondie, et analysé l'arsenal utilisé par le groupe Winnti. Celui-ci est encore l'un des groupes les plus efficaces dans l'attaque via la chaîne d'approvisionnement. Non seulement ils avaient compromis plusieurs cibles importantes mais, dans chaque cas, ils avaient pu passer sous les radars pendant de nombreux mois avant d'être débusqués. Cette fois-ci, c'étaient deux jeux et une plateforme de jeu en ligne qui avaient été compromis pour y inclure une petite backdoor (porte dérobée). Ces attaques visaient alors principalement l'Asie et l'industrie du jeu. L'objectif d'Eset était de découvrir comment les chaînes d'approvisionnement numériques des organisations avaient été compromises, jusqu'à finir par intégrer un malware dans leurs applications clientes. Bien que ces logiciels aient utilisé des configurations assez différentes dans chaque cas, ils intégraient tous le même code de porte dérobée et avaient été lancés en utilisant le même mécanisme. L'un des jeux incriminés, produit par l'éditeur thaïlandais Electronics Extreme, s'appelle Infestation. Cela ne s'invente pas.

Une aubaine pour les hackers

Les chercheurs d'Eset avaient déterminé certaines relations entre les incidents survenus sur la chaîne d'approvisionnement et en avaient déduit les conclusions qui suivent :

- l'un des objectifs du groupe était le minage de cryptomonnaie.
- il existait des relations évidentes entre les techniques et les outils employés lors de multiples attaques majeures de la chaîne d'approvisionnement au cours des dernières années. Cela indiquait que ces incidents avaient très certainement été commis, soit par le même groupe, soit par des groupes partageant les mêmes « trousses à outils ». Parmi ces attaques partageant les mêmes technologies, il faut citer celles d'Asus, de CCleaner, de jeux et logiciels ou encore de NetSarang.
- une porte dérobée Windows, appelée PortReuse par ses auteurs, et utilisée par le groupe Winnti, est une espèce d'implant réseau passif qui s'injecte dans un processus déjà en écoute sur un port réseau. Il attend qu'un paquet magique (utilisant donc le WOL, Wake On Lan) entrant déclenche le code malveillant.
- les chercheurs de l'Eset ont, en collaboration avec la société Censys, identifié et prévenu une importante victime asiatique de PortReuse, un fournisseur de matériel et de logiciels mobiles.

Le cloud, grand facilitateur des infections

Un bon tiers des entreprises n'auraient que peu, voire pas du tout, d'informations sur les intrusions de pirates informatiques au sein de leur chaîne d'approvisionnement. La récente faille Log4j a mis en évidence cette situation. Il est par conséquent difficile de découvrir les dégâts avant qu'il ne soit trop tard. Le cloud joue d'évidence un rôle facilitateur dans ces contaminations, particulièrement avec les évolutions de type Saas (Software as a service) et les flux de type API ou EDI. Avec des architectures dans lesquelles les ordinateurs sont reliés entre eux de façon quasi permanente, le risque est démultiplié. Alors qu'à la fin des années 1990, la gestion des vulnérabilités et les tests de pénétration étaient souvent limités à quelques serveurs d'entreprise connectés à Internet, le passage au cloud au cours des dernières décennies a ouvert les systèmes à des milliers d'ingénieurs, de fournisseurs et autres partenaires.

• Le malware ShadowPad est toujours maintenu par son auteur et continue à évoluer. Il conserve la même approche modulaire, mais en ajoutant sans cesse des techniques de furtivité.

L'attaque, fin 2020, du fournisseur de logiciels informatiques SolarWinds a eu une répercussion sur plus de 18 000 clients de la chaîne logistique, parmi lesquels de nombreuses institutions gouvernementales très sensibles. Avec la pandémie du Covid 19, la mise en place du télétravail et une certaine désorganisation au moins temporaire des entreprises, les cybercriminels s'en sont donné à cœur-joie. D'après une étude menée par un spécialiste de la cybersécurité, BlueVoyant, 82 % des grandes organisations ont été victimes d'une violation de données au cours de cette période. Ces organisations comptant en moyenne 1 000 fournisseurs dans leur écosystème, cela fait, sans équivoque aucune, des entreprises de la supply chain les principales sources d'actes de cybercriminalité. Les pirates se tournent de plus en plus vers ces fournisseurs, bien moins sécurisés que leurs clients.

Un domaine bien difficile à contrôler

Une attaque au niveau de la chaîne d'approvisionnement n'a rien de très nouveau, mais elle s'avère souvent difficile à maîtriser. L'organisation d'une entreprise est composée d'un très grand nombre de logiciels, d'outils technologiques et de services. Cela rend souvent sa sécurisation très complexe. Les pirates le savent et visent spécifiquement le maillon au niveau de sécurité le plus faible afin d'atteindre leur véritable cible. Une prise de conscience rapide et une mise à jour complète des stratégies de cybersécurité représentent la seule méthode efficace pour faire face à cette menace et s'assurer que tous les maillons de la supply chain sont sécurisés de manière satisfaisante. L'effort doit être collectif et simultané, car toutes les organisations participant à un écosystème sont aussi vulnérables que leur maillon faible. L'une des raisons pour lesquelles ces prestataires ne sont parfois pas suffisamment sécurisés peut venir du fait qu'ils pensent ne pas détenir directement d'information sensibles sur leur réseau, et ne voient ainsi pas d'enjeu à se sécuriser davantage. Ces petites structures peuvent ainsi se sentir moins concernées par la cybersécurité, alors qu'elles représentent des cibles de choix pour les hackers comme point d'entrée dans des plus grosses structures.

Trois grands types d'attaques

De la demande de rançon à l'espionnage industriel, les cybercriminels traquent les données sensibles. Au travers de la chaîne d'approvisionnement, il existe principalement trois types d'attaques :

• La première a pour objectif de provoquer le crash du système. Cela peut se régler à l'aide d'une simple remontée de sauvegarde, mais pas toujours.

• La seconde consiste à employer un ver. Une fois intégré au réseau, celui-ci va diffuser des informations erronées aux systèmes automatisés d'une entreprise. Dans certains domaines, comme celui de la logistique, cela peut être lourd de conséquences et aller jusqu'à provoquer l'arrêt de toute une chaîne de diffusion ou de production. Le ransomware ayant touché l'entreprise de pipeline pétrolier aux États-Unis en est un exemple criant.

Les données de santé particulièrement visées

Le ministère des Solidarité et de la Santé a décrété, l'année dernière, que la cybersécurité dans le domaine de la santé et du médico-social représentait une priorité nationale. Néanmoins, parler ne suffit pas. Il faut aussi allouer des budgets en conséquence. La numérisation s'est accélérée avec la pandémie, suite notamment au développement de la téléconsultation et de la prise de rendez-vous pour se faire vacciner via des plateformes en ligne. Cela a, par conséquent, agrandi la surface d'attaque pour les cybercriminels. Cette situation ne peut que se développer, constituant encore plus d'opportunités de mettre la main sur des données de santé. Or, ces données représentent une véritable mine d'or pour les pirates qui peuvent gagner beaucoup d'argent rien qu'en les revendant sur le darknet.

• La troisième est une action d'espionnage utilisant un cheval de Troie. Lorsque celui-ci est implanté, il ouvre un canal vers un ordinateur distant grâce à la backdoor qu'il a mise en place. Le pirate a alors accès au système et peut agir à sa guise. Le cheval de Troie peut aussi être doublé d'un keylogger, enregistrer tout ce qui est tapé au clavier et faire des copies d'écran en vue de découvrir des informations sensibles, telles que logins et mots de passe de comptes divers.

Des conséquences multiples

Dans une organisation en cascade, la chaîne d'approvisionnement regroupe une large liste d'acteurs allant du donneur d'ordre aux sous-traitants. Les maillons doivent être irréprochables, afin d'assurer la sécurité informatique de l'ensemble de la chaîne. La moindre faille peut impacter tous les participants et du coup tous leurs clients finaux. Souvent, les techniques de protection telles que la détection d'incidents et l'analyse de comportements anormaux se concentrent sur une seule organisation, alors que leur application devrait s'étendre à l'ensemble de l'écosystème de la chaîne d'approvisionnement. Si les organisations ont davantage conscience des cybermenaces qui les ciblent et se protègent en conséquence, elles n'imposent pas encore cette même rigueur vis-à-vis de leurs parties tierces, bien souvent faute de moyens ou de temps. Évidemment, cette solution est très difficile à mettre en place et nécessite une sélection stricte des sous-traitants et fournisseurs en matière de cybersécurité, mais c'est la seule qui vaille. Selon des rapports récents, seulement 25 % des entreprises ont défini une stratégie de gestion des accès des tiers à leurs systèmes informatiques. L'objectif est de s'assurer qu'aucun acteur ne présente de vulnérabilités et ne pose un danger pour tous les autres maillons de la chaîne. C'est l'enjeu majeur si les entreprises veulent éviter l'effet papillon, avec une surface d'attaque qui ne cesse de s'étendre du fait de la digitalisation croissante, mais aussi de la sophistication toujours accrue des cyberattaques. ▮

T.T

Eviden coordonne le **projet Cyderco**

La ligne d'activité du groupe Atos est en charge de la coordination du projet Cyderco (CYber DEtection, Response and Collaboration).
Ce projet européen vise à développer, tester et valider une plateforme qui soutiendra et améliorera les capacités de détection et de réponse des entités concernées dans les pays ; y compris les SOC (Security Operations Centers) privés et nationaux, pour lutter contre les cybermenaces affectant les réseaux et les systèmes d'information à travers l'Union européenne.

Le projet est coordonné par les équipes d'Eviden en Roumanie et est constitué d'un consortium de quatre partenaires : Eviden en Roumanie, Atos en Espagne, ISEP (Instituto Superior de Engenharia do Porto) et DNSC (Romanian National Cybersecurity Directorate). L'objectif de cette plateforme est de fournir aux SOC des informations essentielles sur les acteurs de la menace, leurs tactiques, techniques et procédures (TTP) et leurs indicateurs de compromission (IoC), afin d'améliorer la collaboration, l'efficacité et la proactivité dans la lutte contre les cyberattaques. Zeina Zakhour, CTO de la cybersécurité chez Eviden, précise : *« L'objectif de ce projet pour nous, c'est de pouvoir, avec nos partenaires, d'améliorer les capacités de détection qui existent aujourd'hui dans les écosystèmes de ce qu'on appelle les Security Operations Center, d'amener des plateformes qui sont plus avancées pour permettre une détection en amont, proactive des menaces. Et l'objectif dans ce cas-là, c'est d'avoir une plateforme qui a des capacités de détection, d'analyser le réseau, d'analyser les terminaux, les machines, d'analyser l'ensemble des données, d'avoir les télémétries dans un environnement technologique, de pouvoir faire cette analyse-là, d'ajouter, bien sûr, l'intelligence artificielle qui permet, en fait, d'analyser rapidement les volumes énormes de données, de pouvoir détecter les signaux, ce qu'on appelle, nous, des signaux faibles d'une potentielle menace pour être proactif ».*
Sur l'aspect intelligence artificielle de la plateforme, elle ajoute : *« dans l'approche qu'on a adoptée, c'est, premièrement, la valeur de la détection, la valeur des algorithmes qu'on est en train de développer, la capacité de pouvoir identifier tous ces patterns et de bénéficier de toutes les données à disposition. Puis, il y a la technologie en elle-même. Nous travaillons par exemple sur l'analyse du trafic réseau. Là, nous sommes en amont dans la détection, et nous voulons aller plus loin sur les différentes couches du modèle OSI et d'en extraire autant d'informations que possible ».*


Zeina Zakhour, CTO Eviden Cybersecurity

Une « actionable threat intelligence »

La plateforme se complète d'une couche de threat intelligence. La CTO de chez Eviden clarifie : *« on appelle cela une actionable threat intelligence, donc, qui nous permet vraiment d'exploiter la donnée, parce que vous savez autant que moi que la valeur de cette donnée-là doit être bien gérée, avec les bonnes sources et bien analysée. Nous souhaitons avoir ces deux mondes sur la plateforme ».*

Un effet communautaire

Si 2024 a été l'année de la mise en œuvre technique de la plateforme, 2025 verra l'extension par souscription du nombre des membres de partenaires. Les inscrits bénéficieront des éléments quasi en temps réel sur les nouvelles attaques. De plus, par un espace de partage, ils pourront échanger sur les bonnes pratiques, des retours d'expérience sur certains sujets, sur certaines menaces, même collaborer ensemble pour analyser un type de menace particulière. Il est aussi envisagé de rendre publics certains post-mortem d'attaques significatives, dont les entreprises de toutes tailles peuvent tirer des enseignements.
La plateforme suit une logique Zero Trust alignée sur les prérequis d'ISO 27001, du NIST Framework et les principales bonnes pratiques et conformités au niveau européen. ▮

B.G

Le Cyderco

Le projet a une durée de 36 mois, à partir du 1er octobre 2023, et un budget total de 2 881 082 euros, avec un taux de financement de 50 % des coûts éligibles de l'action. Le consortium du projet est composé de quatre partenaires issus de trois pays européens, et représente un mélange de grandes entités industrielles (Eviden Technologies Roumanie et Atos Espagne), d'institutions académiques (ISEP), du secteur public et d'institutions de transfert de technologie (DNSC). Ils possèdent une expertise et des rôles complémentaires bien définis qui répondent aux besoins cruciaux pour générer des résultats à haute valeur ajoutée. Les partenaires sont bien placés pour coopérer et collaborer, afin de relever les défis technologiques du Cyderco en combinant le savoir-faire technologique et scientifique, la perspective de l'industrie et de l'utilisateur final, ainsi que la connaissance des affaires et du marché. Les partenaires disposent de toute l'expertise nécessaire pour mener à bien le projet Cyderco.

SMART TECH

DELPHINE SABATTIER

7H30 | 18H30

VOTRE ÉMISSION QUOTIDIENNE DÉDIÉE À L'INNOVATION

Dans l'émission SMART TECH animée par Delphine Sabattier, l'actualité du numérique et de l'innovation prend tout son sens. Chaque jour, des spécialistes décryptent les dernières news, les tendances, et les enjeux soulevés par l'adoption des nouvelles technologies.

N°230 orange™ | N°246 bouygues | N°163 free

B SMART by Change

Libérez votre ambition

Les grandes ambitions produisent de grands résultats. Nos puissantes solutions IA peuvent vous aider à surmonter vos limites.

Rendez-vous sur **HPE.com/fr**

Hewlett Packard Enterprise