

L'INFORMATICIEN

Cloud

Les nouveautés Google

Hardware

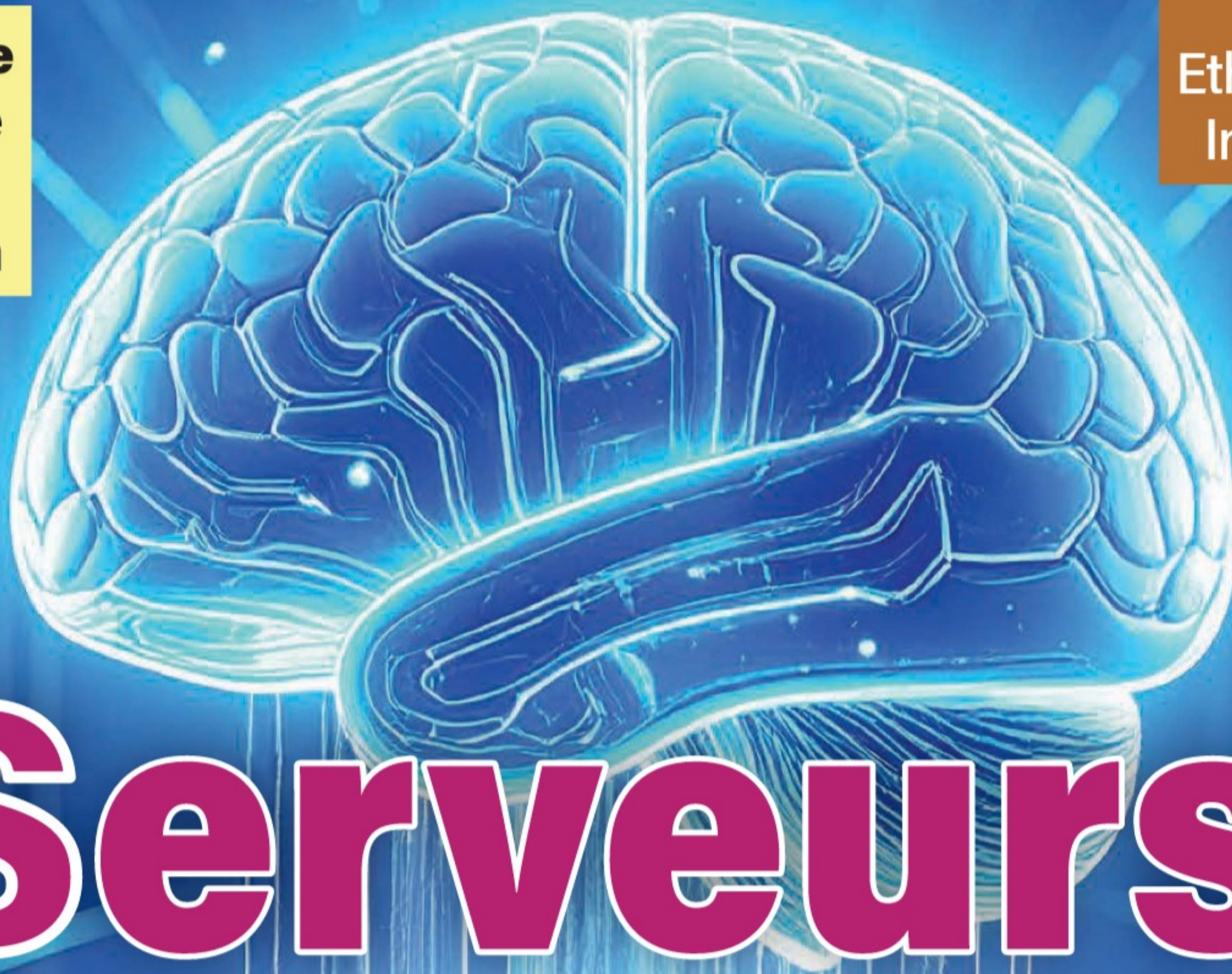
Mainframe
nouvelle
génération

ESN

Vers l'industrie 5.0

Réseau

Ethernet vs
Infiniband



Serveurs

Dédiées à l'IA ?

Retex

« La DSI devient
vraiment légitime
avec l'IA »

Wolide Nagmar,
DSI de Constructa



L 14614 - 236 - F: 8,50 € - RD





Guide les analystes SOC avec l'IA de la détection à la réponse

TOUT COMPRENDRE, RÉAGIR INSTANTANÉMENT, PROTÉGER À GRANDE ÉCHELLE



Analysez

tout votre environnement



Détectez

les menaces complexes



Réduisez

la fatigue liée aux alertes



Automatisez

la gestion des incidents



Accélérez

le temps de réponse



Privilégiez

une solution souveraine



Accélérez vos opérations de sécurité avec la plateforme SOC Sekoia, qui offre une compréhension inégalée de votre environnement et des renseignements exclusifs sur les menaces. Et grâce à l'intelligence Artificielle, bénéficiez d'un véritable co-pilote pour améliorer l'efficacité opérationnelle de votre équipe, de la détection à la réponse.

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 : François Cointe, Patrick Brebion, Olivier Bouzereau,
Vincent Bussière, Christine Calais, Jérôme Cartegini, Michel Chotard,
Alain Clapaud et Thierry Thauereaux.

SECRÉTAIRE DE RÉDACTION : Amélie Ermenault Martin

MAQUETTE ET RÉALISATION : Franck Soulier (chef de studio)

PUBLICITÉ

Antoine Foulon — afoulon@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 : 2^{ème} trimestre 2025

Toute reproduction intégrale, ou partielle, faite sans le consentement de l'auteur
ou de ses ayants droit ou ayants cause, est illicite (article L122-4 du Code de la
propriété intellectuelle). Toute copie doit avoir l'accord du Centre français du droit
de copie (CFC), 20 rue des Grands-Augustins 75006 Paris. Cette publication peut
être exploitée dans le cadre de la formation permanente. Toute utilisation à des
fins commerciales de notre contenu éditorial fera l'objet d'une demande préalable
auprès du directeur de la publication.

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

 **FICADE**

PRÉSIDENT, DIRECTEUR DE LA PUBLICATION :
Gaël Chervet

Le moteur de l'IA

Notre dossier se consacre aux serveurs informatiques. Longtemps considérés comme des « plaques chauffantes » ou des « frigos », selon le format et la chaleur qu'ils dégageaient, ils sont devenus la pierre angulaire, et ce depuis toujours, du calcul et du traitement des données. Avec l'émergence de l'IA, ils évoluent et se dotent de puces graphiques, d'accélérateurs en tous genres. Ils deviennent aussi plus puissants, plus compacts, disparaissent pour revenir sous une forme logicielle ou dans le Cloud. Nous faisons donc un point sur ce composant fondamental de l'informatique, même si, caché dans des centres de données, il ne fait pas souvent la une et reste souvent dédaigné car « c'est de la prod ». Loin des lumières des technologies à la mode ou des applications avec des millions d'utilisateurs, sans eux, pourtant, rien ne serait possible.

Évidemment, ce numéro fait aussi le point sur les applications et les grands événements du printemps comme la Kubecon qui s'est tenue récemment à Londres. Mention spéciale pour la dernière ligne droite afin de déposer votre dossier de participation au Top Tech. Pour rappel, il suffit d'aller sur le site de L'Informaticien, de cliquer sur l'onglet Top Tech. Si vous êtes une entreprise ou une société de services, cliquez sur les formulaires à remplir selon les catégories dans lesquelles vous souhaitez concourir ; au passage, tout est gratuit. Il n'y a pas de frais de dossier ou de candidature. Le formulaire rempli, suivez les instructions et faites-le nous parvenir sur la plateforme mise en place. Voilà ! Vous avez quasiment jusqu'à la fin mai pour candidater, et la remise des prix a lieu le 3 juillet prochain ! ☐

Bertrand Garé
Rédacteur en Chef



SMART TECH

DELPHINE SABATTIER
7H30 | 18H30

VOTRE ÉMISSION QUOTIDIENNE DÉDIÉE À L'INNOVATION

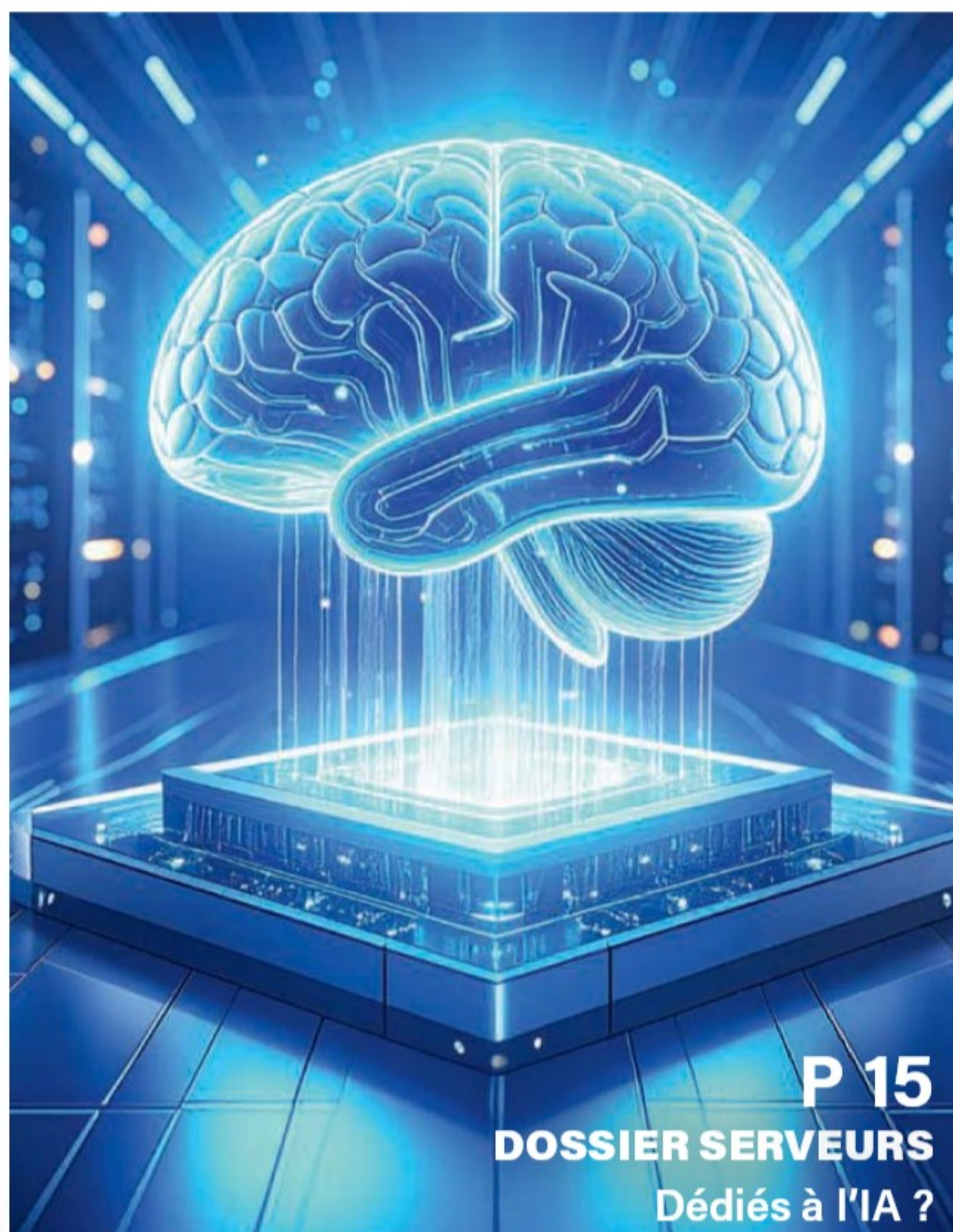
Dans l'émission SMART TECH animée par Delphine Sabbattier, l'actualité du numérique et de l'innovation prend tout son sens. Chaque jour, des spécialistes décryptent les dernières news, les tendances, et les enjeux soulevés par l'adoption des nouvelles technologies.

N°230
orange™

N°246
bouygues
telecom

N°163
free

B SMART
Change

**DOSSIER** **P 15**

Serveurs : Dédiés à l'IA ?

BIZ'IT **P 8****BIZ'IT PARTENARIAT** **P 12****HARDWARE** **P 22**Nouvelle génération Mainframe
NAS Synology
Pure Storage**ESN** **P 28**Partenariat Omron Cognizant
Écritel**TACTIC** **P 31**L'IA manque
de matière première**RÉSEAU** **P 33**Ethernet face
à Infiniband
Test eSIM**LOGICIEL** **P 38**Adobe
Workday
SAP**CLOUD** **P 42**Google
OVHCloud
Kubecon**RETEX** **P 48**HP : contre la fraude documentaire
Constructa**BONNES FEUILLES** **P 51**

Innover ou agoniser

INNOVATION **P 54**Jumeaux numériques
Confiance.ai
Atlassian**DEVOPS** **P 58**

ML pour Python

ÉTUDE **P 62**

Rapport Armis cyberguerre

RH / FORMATION **P 64**Data center école
RH Magellan**INFOCR** **P 67****ABONNEMENTS** **P 76**

Le Grand Saut : comment les DSI des ETI relèvent le défi du digital ?

La transformation numérique est une priorité pour les ETI françaises, et leurs DSI doivent innover

tout en assurant la résilience et la sécurité des systèmes.

L'étude PAC « Le Grand Saut », menée en exclusivité pour Ready For IT, explore les ambitions et les obstacles rencontrés par ces décideurs.

Le numérique, un levier de compétitivité incontournable

Près de 80 % des DSI interrogés considèrent le digital comme le moteur central de leur stratégie de croissance. Comme en référence à un grand saut, l'étude met en évidence que la situation des DSI (spécifiquement en ETI) est en pleine évolution sur l'année 2025.

Et ce, à trois niveaux différents :

1. Un contexte économique et financier qui pousse nos entreprises à une performance globale de la part de toutes ses composantes, avec la DSI en première ligne.
2. L'aspect technologique n'y coupe pas : le déploiement massif des IA et les besoins toujours plus frénétiques de trouver LA nouvelle technologie pour aller encore plus vite, encore plus loin (vs. la concurrence).
3. Le point de vue réglementaire avec les nouvelles directives NIS 2, entre autres.

Face à ces défis, cette 6^{ème} édition de Ready For IT se positionne comme un accélérateur de transformation pour nos entreprises.

Ready For IT 2025 : le rendez-vous stratégique des décideurs IT

Du 20 au 22 mai à Monaco, la 6^{ème} édition de Ready For IT réunira les DSI, CTO, RSSI et Directeurs Innovation des ETI françaises, ainsi que les principaux offreurs de solutions en transition et sécurité numériques. Cet événement s'impose comme une plateforme d'échanges incontournable, permettant aux décideurs de structurer leur transformation numérique dans un contexte marqué par de nouvelles exigences réglementaires et économiques.

En 2025, les DSI doivent composer avec de nouvelles contraintes réglementaires (NIS2), une exigence accrue en matière de RSE et des défis RH toujours plus complexes. La transformation digitale ne se limite plus à un levier d'innovation : elle est désormais un impératif de compétitivité internationale et de maîtrise des risques. Dans ce cadre, les entreprises doivent optimiser leurs



Scannez ce QRcode pour en savoir plus



processus, rationaliser leurs coûts et renforcer leur cybersécurité pour assurer un développement pérenne.

Le programme de Ready For IT 2025 met en lumière ces enjeux à travers des conférences, des tables rondes et des rencontres One to One, offrant aux participants une vision stratégique et des solutions concrètes pour réussir leur transition numérique.

Anticiper l'avenir et structurer la transformation numérique

Face aux évolutions rapides du numérique et aux impératifs de sécurité, nos entreprises doivent adopter une approche proactive et s'appuyer sur des écosystèmes solides pour assurer leur résilience. Ready For IT 2025 sera l'occasion d'échanger sur les meilleures pratiques, d'identifier les leviers d'action et de bâtir des stratégies robustes pour faire de la transformation numérique un moteur de croissance durable.

Ne manquez pas cet événement clé pour anticiper les mutations du numérique et sécuriser la transformation de votre organisation. Rendez-vous à Monaco du 20 au 22 mai !

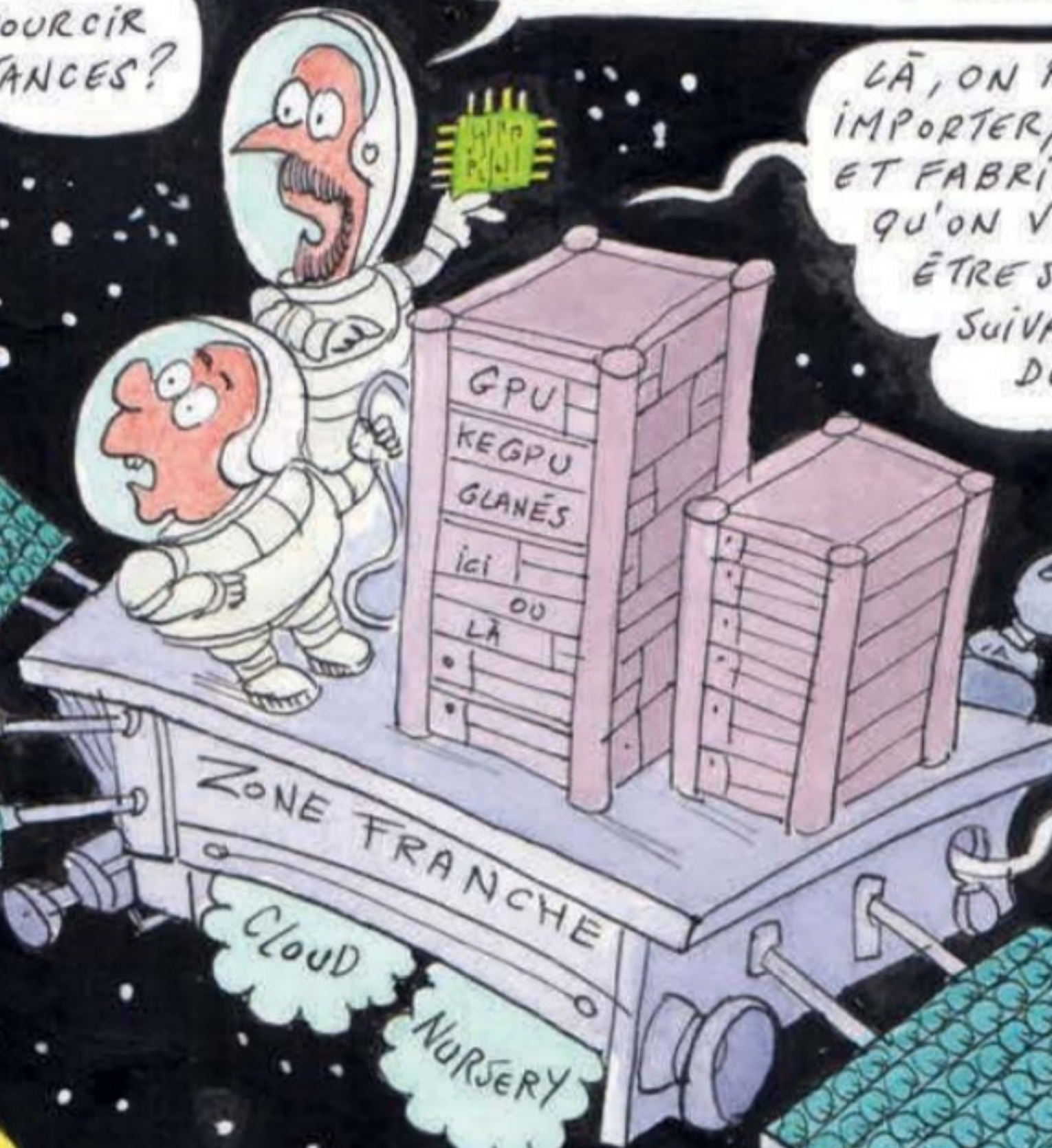
LES SERVEURS DU FUTUR

C'EST POUR LES
REFROIDIR PLUS SIMPLEMENT
QUE VOUS INSTALLEZ VOS
SERVEURS DANS L'ESPACE?

ET POUR ÊTRE
PLUS PRÈS
DES NUAGES
ET RACCOURCIR
LES DISTANCES?

NON. C'EST POUR ÉCHAPPER
AUX DROITS DE DOUANE
DE L'AUTRE CÔTÉ...

LÀ, ON PEUT
IMPORTER, COMPOSER
ET FABRIQUER CE
QU'ON VEUT SANS
ÊTRE SURTAXÉS
SUIVANT L'HUMEUR
DU MOMENT...



MOI AUSSI,
JE VEUX ÊTRE
UNE ÎLE EN
ORBITE...

... LOIN DES
CAPRICES DES
UNS ET DES AUTRES.



ET CE
MAINFRAME LÀ,
COMMENT ON
LE REFROIDIT?
À L'EAU?

IL EST CHAUD
LUI... FAUDRAIT
PLUTÔT DE
L'AZOTE
LIQUIDE...

F. COINTE

La souveraineté technologique au menu français

Le climat géopolitique tendu avec les États-Unis et l'imprévisibilité de Donald Trump, peu avare en droits de douane et en menaces en tout genre, place l'Europe dans une situation inconfortable. La France, comme tant d'autres, est fortement dépendante des technologies américaines (et chinoises), et se voit contrainte d'interroger une nouvelle fois son modèle et d'envisager une plus grande indépendance technologique.

Le contexte géopolitique incertain fait craindre que l'administration Trump brandisse son statut de leader technologique pour faire pression, y compris sur ses alliés. Une menace qui, si elle a eu un mérite, est celui d'avoir poussé les Européens à se repencher sur la question de leur souveraineté technologique, avec en tête les marchés du cloud computing et de l'intelligence artificielle.

Entre 300 et 5 000 Md€ pour sortir de la dépendance

Pour tenter d'inverser la tendance, pas le choix : il va falloir des moyens, beaucoup de moyens. L'initiative Eurostack a chiffré les besoins de l'UE à 300 milliards d'euros d'ici à 2035 pour développer son propre écosystème numérique. L'association américaine Chamber of Progress est moins optimiste et avance plutôt le chiffre de 5 000 milliards d'euros. Sur la seule IA, la présidente de la Commission européenne, Ursula von der Leyen, a annoncé vouloir mobiliser 200 milliards d'euros, dont 150 milliards provenant de fonds privés.

La France, qui a annoncé une enveloppe pour l'IA (109 milliards d'euros), s'est aussi très récemment repenchée sur la problématique du cloud souverain. À l'occasion d'une réunion à Bercy avec des entreprises, des administrations, des associations, des experts et des parlementaires, la ministre déléguée chargée de l'Intelligence artificielle et du Numérique, Clara Chappaz, a relancé un appel à projets (AAP) de Bpifrance, présenté en octobre 2024, pour renforcer l'offre de cloud souverain.

Intitulé « Renforcement de l'offre de services Cloud » et doté de plusieurs dizaines de millions d'euros, cet AAP « est d'ampleur, il était attendu, je le sais. Il doit



notamment permettre de bâtir des solutions basées sur les dernières avancées en intelligence artificielle. Il porte une ambition : celle de bâtir une offre de cloud européenne attractive, performante, compétitive », a déclaré la ministre dans son discours d'ouverture.

Dans le même temps, elle a annoncé la création d'un Observatoire de la souveraineté numérique. Piloté par le Conseil général de l'Économie, il aura pour mission d'identifier et de cartographier les dépendances technologiques de la France.

Des choix stratégiques critiqués

Et des dépendances, il y en a. Certaines administrations ont récemment été pointées du doigt quant à leurs choix de partenaires. En mars 2024, le député de Vendée (1^{re} circonscription), Philippe Latombe, a demandé au gouvernement de dénoncer deux contrats respectivement signés par les ministères de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, et par l'École polytechnique avec Microsoft. Le député Les Démocrates regrettait que

l'utilisation de solutions souveraines et libres n'ait pas été privilégiée.

« Alors que les relations transatlantiques actuelles devraient inciter nos administrations à la plus grande prudence concernant les solutions informatiques qu'elles choisissent, et les pousser à s'émanciper des géants américains, certaines d'entre elles persistent et signent sans vergogne », avait-il déclaré. Dans ce contexte, le député craint notamment que des informations sensibles ne soient exposées au droit extraterritorial américain, qui autorise, dans certains cas, les autorités des États-Unis à exiger d'entreprises technologiques qu'elles leur fournissent un accès aux données stockées sur leurs serveurs.

Le député a également alerté sur le cas de la plateforme des données de santé (le Health Data Hub ou HDH), toujours hébergée sur Azure (Microsoft), et qui n'a toujours pas migré vers un environnement cloud souverain, comme cela était initialement prévu. Face à la polémique ravivée, Clara Chappaz a annoncé le lancement d'un appel d'offres pour la migration des données du HDH, sans préciser toutefois de calendrier.

Intel et TSMC : imbroglio autour d'un projet de coentreprise

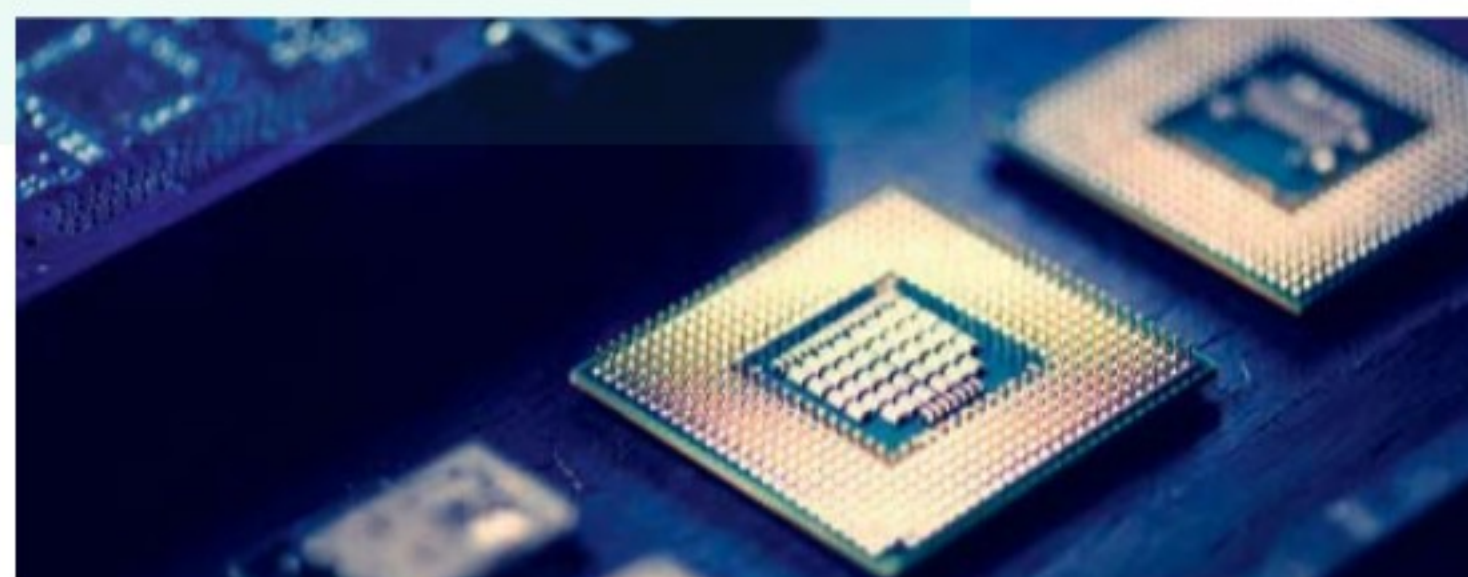
Washington souhaiterait créer un nouveau géant des semi-conducteurs et tenter, au passage, de faire sortir la tête de l'eau à Intel. C'est en tout cas ce que nous apprenait le média The Information, début avril, qui citait deux personnes proches du dossier. Intel et la Taiwan Semiconductor Manufacturing Company (TSMC) auraient conclu un accord préliminaire pour former une coentreprise en exploitant les usines de fabrication de puces d'Intel.

Dans cette réalité, le Taïwanais prendrait 20 % du capital de cette nouvelle société. Le reste serait partagé entre Intel et d'autres entreprises américaines du secteur. Reuters a rapporté en mars que TSMC avait déjà proposé à Nvidia, AMD et Broadcom de prendre des participations dans une coentreprise, afin d'exploiter les usines de fabrication d'Intel, suite à une demande de l'administration Trump.

Toutefois, à la mi-avril, TSMC a démenti les informations selon lesquelles il serait en discussions avec Intel. « *TSMC n'est engagée dans aucune discussion avec d'autres entreprises concernant une coentreprise, une licence technologique ou une technologie* », a affirmé le PDG C.C. Wei lors de la conférence téléphonique. Intel avait précédemment indiqué à des médias, dont CNBC, qu'elle ne commentait pas les rumeurs.

Washington met son grain de sel

Ce rachat était présenté comme une aubaine pour Intel, lui permettant de profiter de l'expérience et de l'expertise de TSMC. En pleine traversée du désert, notamment autour de son activité de puces d'IA et de fabrication de puces en général, l'américain a été supplanté par la concurrence, TSMC en tête. D'après d'anciens



cadres interrogés par Reuters, Intel a rencontré des difficultés pour fournir un niveau de service client et technique équivalent à celui de TSMC, ce qui a entraîné des retards et des échecs lors des tests. Dans les cordes, l'américain a enregistré une perte nette de 18,8 milliards de dollars en 2024, et a annoncé un plan de licenciement de 15 % de ses effectifs et une nouvelle coupe de 20 % a été révélée par Bloomberg en ce mois d'avril.

Pour le fondeur taïwanais, les intérêts potentiels sont ailleurs. Il pourrait aussi apaiser l'administration Trump, qui a récemment accusé Taïwan d'avoir "volé" l'industrie des semi-conducteurs aux États-Unis. C'est aussi, pour TSMC, une manière de contourner des droits de douane qui ne seront pas appliqués aux puces fabriquées sur le sol américain, alors que le pays adopte une politique de plus en plus protectionniste pour relocaliser la production aux États-Unis. Le chef d'État envisage en effet de soumettre les semi-conducteurs à ces fameux tarifs (droits de douane), dont ils sont exemptés encore aujourd'hui. À cela s'ajoute un contexte géopolitique tendu, alors que la Chine, qui revendique l'île, multiplie les démonstrations de force. TSMC a donc tout intérêt à ne pas mettre tous ses œufs dans le même panier et à développer son activité hors de l'île.

Schneider Electric investit 700 M\$ dans ses infrastructures aux États-Unis

Schneider Electric a annoncé flécher 700 millions de dollars vers ses opérations aux États-Unis jusqu'en 2027. Ces investissements doivent consolider les infrastructures énergétiques et soutenir la demande des centres de données et de l'IA.

« *Nous nous trouvons à un tournant pour les secteurs technologique et industriel aux États-Unis, stimulés par une croissance incroyable de l'IA et une demande énergétique sans précédent. Pour diriger la transformation à venir, nous devons être agiles et agir maintenant pour faire progresser des objectifs ambitieux de numérisation et d'efficacité, afin d'avoir un impact pour les générations futures* », s'est exprimé Aamir Paul, président des opérations pour l'Amérique du Nord.

36 % de ses revenus aux États-Unis

Schneider Electric ambitionne de moderniser et étendre ses installations dans le Tennessee, le Massachusetts, le Texas, le Missouri, l'Ohio, ainsi qu'en Caroline du Nord et Caroline du Sud, avec 1 000 emplois créés à la clé. Le Français construira, entre autres, une nouvelle installation de moyenne tension à Mt. Juliet (Tennessee), un laboratoire d'unité de distribution d'énergie pour les centres de données d'IA à Andover (Massachusetts), et étendra son campus à El Paso (Texas), afin de répondre à la demande croissante dans les produits de commutation et de distribution d'énergie. Cette somme vient s'ajouter aux 400 millions de dollars

annoncés depuis 2020, et porte à plus d'un milliard de dollars les investissements de Schneider Electric dans la région. Et ces investissements aux États-Unis n'ont rien d'une surprise, tant le pays est important pour le Français, puisqu'il représentait 36 % de ses revenus l'année dernière. L'entreprise emploie 21 000 personnes sur place et exploite 20 usines et centres de distribution intelligents. Elle compte en outre 40 % des entreprises du Fortune 500 parmi ses partenaires, et ses solutions sont présentes dans quatre foyers sur 10 aux États-Unis. Dans un contexte de fortes tensions commerciales à l'échelle mondiale, beaucoup d'entreprises veulent montrer patte blanche face à l'administration Trump.

CGI avale Apside

L'entreprise de conseil en technologie renforce sa force de frappe avec le rachat d'Apside, une entreprise de services numériques et d'ingénierie. L'accord a été signé avec les actionnaires principaux d'Apside, Siparex et les sociétés d'investissement du Groupe Crédit Agricole.

En activité depuis près de 50 ans, Apside est reconnue pour ses connaissances approfondies dans les secteurs de l'industrie, des services financiers, de l'assurance, des mutuelles et du secteur public, ainsi que pour sa vaste expertise technologique en

données, en intelligence artificielle (IA), en cloud (infonuagique), en cybersécurité...

Actuellement, Apside sert plus de 300 clients à l'échelle mondiale et compte 28 bureaux dans six pays. Une fois l'entente conclue, plus de 2 500 professionnels se joindront à CGI, consolidant la présence de l'entreprise en France, au Canada, au Portugal, en Belgique, au Maroc et en Suisse. La transaction proposée est sous réserve d'approbations réglementaires et d'autres conditions de clôture habituelles.

SoftBank rachète Ampere Computing pour 6,5 milliards de dollars

Après avoir annoncé en début d'année sa participation dans le projet pharaonique de Donald Trump, Stargate, avec OpenAI, Oracle et le fonds d'investissement MGX, SoftBank a révélé, jeudi 20 mars, avoir mis la main sur l'entreprise de conception de microprocesseurs haute performance pour centres de données, Ampere Computing. Montant de

l'opération : 6,5 milliards de dollars. Un gros coup pour le Japonais, dans un contexte de forte demande d'infrastructures dédiées au calcul d'IA.

Ampere est une entreprise de conception de semi-conducteurs fondée en 2018, connue pour avoir lancé la famille de processeurs cloud native Ampere, conçue pour des environnements de calcul haute

performance. Ampere va être amenée à soutenir l'écosystème de SoftBank et les entreprises du groupe, notamment dans le développement et la fabrication de puces basées sur Arm. SoftBank espère finaliser l'opération dans la seconde moitié de 2025, sous réserve des approbations réglementaires.

ServiceNow s'empare de Logik. AI

L'éditeur de la plateforme de transformation des services de l'entreprise a signé un accord définitif pour la reprise des actifs de Logik. AI. Logik. AI fournit une solution de CPQ (configure, price, quote) s'appuyant sur l'intelligence artificielle. L'intégration à venir avec ServiceNow va étendre ses fonctionnalités de

gestion de la relation client, en particulier pour la solution de Sales & Order Management. La transaction doit encore obtenir les autorisations habituelles auprès des actionnaires et des autorités.

Cloudflare rachète Outerbase

Spécialiste du cloud de connectivité, Cloudflare met la main sur Outerbase, un spécialiste des bases de données pour développeurs. Cette acquisition

doit permettre de muscler les capacités de sa plateforme pour les développeurs, Cloudflare Workers. Construite sur cette dernière, la technologie d'Outerbase

s'intégrera à Durable Objects, D1 et au SDK agents de Cloudflare, et mettra à disposition des interfaces et frameworks permettant d'interagir avec les données et de développer des applications reposant sur des bases de données construites et gérées par les développeurs. « Elle permettra à davantage d'équipes de développer et de déployer des applications full-stack, soutenues par l'IA, sur le réseau mondial de Cloudflare », décrit Cloudflare dans un communiqué. En effet, les bases de données sont au cœur des applications, alors que l'utilisation de l'IA et des agents IA va croissant, afin de conserver le contexte, stocker les conversations et interagir avec les données.



Qevlar AI lève 14 M€

La jeune pousse Qevlar AI a bouclé un tour de table de 14 millions de dollars. L'opération a été menée par EQT Ventures et Forgepoint Capital International, et a été soutenue par Olivier Pomel (CEO, Datadog), Florian Douetteau (CEO, Dataiku) et Mehdi Ghissassi (ancien directeur produit chez Google DeepMind).

Fondée en 2023, Qevlar AI développe des agents d'IA qui investiguent, contextualisent, analysent et structurent les alertes de façon autonome. « Nos agents autonomes ne se contentent pas d'assister les analystes ; ils redéfinissent ce qui est possible en menant des enquêtes complètes en quelques secondes au lieu de plusieurs heures, avec une précision et une constance supérieures. Les SOC peuvent ainsi passer d'un rôle réactif à celui de véritables chasseurs de menaces proactifs », déclare Ahmed Achchak, cofondateur et



CEO de Qevlar AI, cité dans un communiqué. La startup revendique des résultats « spectaculaires » auprès des fournisseurs de services de sécurité managés (MSSP). C'est, par exemple, un temps réduit de traitement des alertes malveillantes de 40 minutes à trois minutes, une diminution de 90 % du temps d'enquête des analystes de niveaux 1 et 2, une précision

de classification des alertes de l'ordre de 99,8 % et une clôture de 100 % des alertes bénignes.

Avec ces nouveaux fonds, l'entreprise va muscler ses équipes, déjà multipliées par 4 sur l'année écoulée, investir dans sa R&D et soutenir son expansion à l'international.

Hammerspace lève 100 M\$

Dernièrement, Hammerspace avait frappé les esprits avec sa technologie de Tier 0 pour une activation plus efficace des GPU dans les environnements pour l'intelligence artificielle. Des investisseurs de renom semblent avoir été séduits par le futur de l'entreprise. Mené par Altimeter Capital, qui avait investi dès les débuts d'entreprises comme Meta, MongoDB,

Nvidia, Snowflake et Uber, ce tour se monte à 100 M\$, avec la participation d'Ark Invest et d'investisseurs choisis. Les fonds visent à accélérer le développement de l'entreprise et à consolider sa position de référence pour les environnements HPC, d'IA et de cloud hybride.

OpenAI lève 40 Md\$

Après avoir levé 6,6 milliards de dollars en octobre dernier, la maison mère de ChatGPT vient de boucler un tour de financement de 40 milliards de dollars. Le principal financeur n'est autre que le groupe japonais SoftBank, déjà partenaire d'OpenAI. Microsoft, Coatue Management, Altimeter et Thrive Capital ont également participé à l'opération. La valorisation

post-investissement d'OpenAI grimpe à 300 milliards de dollars, contre 157 milliards en octobre.

Avec ces nouveaux fonds, l'entreprise va accélérer ses recherches sur l'intelligence artificielle générale (AGI) et muscler son infrastructure, essentielle à l'entraînement de ses modèles. Elle s'émancipe en même temps encore un peu plus de Microsoft, son

soutien technique et financier historique. Le montant de la levée de fonds est toutefois soumis à condition. SoftBank prévoit d'investir 10 milliards de dollars dans un premier temps, puis 30 milliards supplémentaires en décembre 2025, à condition qu'OpenAI devienne une société à but lucratif. Sinon, son investissement sera limité à 20 milliards de dollars.

Island lève 250 M\$ pour son navigateur d'entreprise

La valorisation d'Island atteint 4,8 milliards de dollars après cette nouvelle levée de fonds de série E de 250 millions de dollars, dirigée par la société d'investissement, Coatue Management. Insight Partners, Sequoia et Canapi Ventures, notamment, ont participé au tour de financement.

Island fournit un navigateur pour entreprises basé sur Chromium, présenté comme plus sécurisé que les navigateurs traditionnels en maintenant une navigation confidentielle. L'un des intérêts, selon Island, est de se substituer aux outils de sécurité réseau — qui « ont alourdi les équipes de sécurité, ralenti le travail et frustré

les utilisateurs finaux » — en intégrant des mécanismes de sécurité avancés. Island assure que sa solution protège contre les attaques sophistiquées, car elle est conçue avec la sécurité dès le départ (security by design).

Le navigateur est en mesure de bloquer automatiquement le contenu malveillant, d'éliminer les exploits du navigateur, de stopper les attaques de phishing et de spoofing, de bloquer et d'avertir si du contenu potentiellement dangereux a été détecté, de protéger les données localement et de capturer des journaux d'activité pour les partager avec des outils SIEM.

Amundi Technology et Murex font cause commune

L'entité technologique d'Amundi noue un partenariat avec Murex, un éditeur spécialisé dans les logiciels à destination du secteur de la finance.

Les clients d'Amundi Technology pourront tirer avantage de Murex au travers d'Alto Investment, la plateforme d'Amundi Technology dédiée aux sociétés de gestion, assureurs, fonds de pension et family offices. Alto optimise la gestion d'actifs grâce à ses fonctionnalités avancées de suivi de portefeuille et à ses données de marché intégrées, permettant aux gestionnaires de fonds de prendre des décisions pertinentes tout en réduisant les risques opérationnels. Les clients

bénéficieront de la large couverture d'instruments OTC de Murex, tirant de surcroît parti de l'architecture ouverte d'Alto. Cela inclut son intégration transparente avec la plateforme de Murex, MX.3, qui fournit des outils avancés pour le traitement des dérivés OTC, tels que la modélisation de payoffs, la gestion du cycle de vie des transactions et l'analyse approfondie des risques.

Collaboration stratégique entre Safran et AWS

Le groupe collabore avec AWS pour accélérer le déploiement des technologies cloud avancées et de l'intelligence artificielle générative à grande échelle pour accélérer ses innovations.

En tant que fournisseur cloud principal de Safran depuis 2021, AWS aide l'entreprise à atteindre son objectif d'accélération et d'industrialisation de sa transformation digitale. Cette collaboration a permis au groupe d'accroître significativement son agilité opérationnelle, tout en maintenant les plus hauts standards de sécurité requis. Aujourd'hui, Safran lance une nouvelle phase pluriannuelle de sa transformation numérique, qui couvre autant la

migration des applications vers AWS, le déploiement de plateformes de données modernes basées sur les services AWS incluant Amazon Simple Storage Service (Amazon S3) et Amazon Redshift que la mise en œuvre de nouveaux cas d'usage par ses équipes d'IA. Ainsi, Safran utilise AWS pour évaluer et déployer l'intelligence artificielle générative (IA) de manière maîtrisée dans ses différents métiers. Safran s'appuie sur Amazon Bedrock, une plateforme

qui donne accès à une large gamme de modèles d'IA comme ceux de Meta, Mistral AI, Anthropic ou Amazon. Safran a déjà concrétisé une première vague de projets d'IA générative pour ses métiers engineering, manufacturing, support et services. Fort de résultats probants en efficacité opérationnelle, Safran accélère actuellement avec une deuxième vague d'une vingtaine de projets.

ITS Integra intègre la solution de Cyberesist

L'infogéreur et opérateur de services Cloud intègre à sa plateforme sa solution d'audit de sécurité de Cyberesist.

Ce partenariat vise à renforcer la protection des entreprises face aux cybermenaces croissantes en intégrant la plateforme avancée d'audit de cybersécurité de Cyberesist aux services managés de sécurité d'ITS Integra. Grâce à cette collaboration, ITS Integra pourra offrir à ses clients une évaluation continue et automatisée de leur posture de cybersécurité, leur permettant ainsi

d'identifier et de corriger les vulnérabilités de manière proactive. La plateforme Cyberesist, basée sur les dernières avancées en matière d'intelligence artificielle et d'automatisation, s'intègre parfaitement à l'offre MSSP d'ITS Integra, garantissant ainsi une surveillance et une amélioration constantes de la sécurité des infrastructures IT.

Scale Computing rejoint le programme partenaire de Veeam

L'éditeur de solutions d'hyper-convergence et de Edge Computing rejoint le programme partenaire de Veeam, afin d'apporter un support complet de Veeam pour sa plateforme.

La coopération se concrétise par un partenariat stratégique entre les deux éditeurs. Il a pour but l'intégration de Scale Computing Platform dans Veeam Data platform, apportant de nouvelles fonctionnalités de sauvegarde et de restauration aux environnements de Scale Computing par une intégration native.

La solution sera en disponibilité générale à la fin de cette année.

Les bénéfices attendus pour les clients conjoints aux deux éditeurs sont :

- Des sauvegardes immutables, même en périphérie prévenant les pertes de données par des ransomwares ou des erreurs humaines
- Une protection renforcée des

machines virtuelles

- Une gestion simplifiée des données
- Une migration des tâches des restaurations dans un environnement de confiance entre la plateforme de Scale Computing et VMware vSphere, Microsoft Hyper-V, ainsi que les principaux environnements de cloud public.

LE CNB et Lefebvre Dalloz forment les avocats à l'IA

Le Conseil National des Barreaux (CNB) et Lefebvre Dalloz Compétences, organisme de formation de Lefebvre Dalloz, lancent un parcours de formation inédit à l'intelligence artificielle à destination des 78 000 avocats et 6000 élèves-avocats français.

Ce dispositif, 100 % digital et gratuit, a été conçu pour accompagner la profession à cette transformation majeure. Disponible jusqu'au 31 décembre 2027, la formation est accessible dès aujourd'hui pour tous les avocats, et dès le 30 avril 2025 pour les élèves-avocats via leurs écoles. Elle permet de valider jusqu'à quatre heures de formation continue, répartie en deux grands axes : les concepts

fondamentaux de l'intelligence artificielle et l'IA générative dans la pratique professionnelle de l'avocat. Accessible gratuitement depuis un ordinateur, une tablette ou un smartphone, la plateforme Skilia Avocats propose un parcours composé de modules courts, pratiques et interactifs, pour favoriser l'appropriation des notions. Cette formation est le fruit d'une co-construction entre le CNB, Lefebvre

Dalloz Compétences et l'avocat Raphaël d'Assignies, expert en droit de l'IA et en cybersécurité. Elle aborde les usages, les limites, et les implications juridiques, déontologiques et éthiques de l'intelligence artificielle. Ce partenariat prolonge les travaux engagés par le CNB, notamment avec la publication du Guide pratique sur l'IA générative de septembre 2024.

Infoblox fait cause commune avec Google Cloud

Le spécialiste de la gestion de la sécurité des DNS s'associe à Google pour accompagner les entreprises dans la simplification de leurs réseaux et de leur cybersécurité.

Dans le cadre de cette collaboration, deux nouvelles solutions sont lancées pour répondre aux enjeux du réseau hybride, du multicloud et de la cybersécurité.

Infoblox Universal DDI pour Google Cloud WAN combine l'infrastructure globale Cross-Cloud Network de Google Cloud avec les fonctionnalités avancées de DNS et DHCP d'Infoblox. Google Cloud DNS Armor, s'appuyant sur Infoblox, est une solution de sécurité DNS de nouvelle génération, intégrée nativement à Google Cloud DNS, qui assure une détection proactive et renforcée des activités malveillantes ciblant les charges de travail hébergées sur Google Cloud.

Un pare-feu pour la blockchain

Fuse et Check Point annoncent un partenariat pour concevoir et déployer une couche de prévention des menaces en temps réel dans le but de sécuriser l'ensemble de l'écosystème de Fuse.

Ce partenariat permet à Fuse de s'appuyer sur une infrastructure de cybersécurité performante, capable d'identifier et de neutraliser, en temps réel, toute menace susceptible de compromettre l'intégrité de la blockchain, avant même qu'elle ne puisse être exploitée.

En collaborant étroitement avec Check Point, Fuse renforce la résilience de son écosystème au service de sa communauté d'utilisateurs et de développeurs, et donne ainsi un nouvel élan à son ambition : intégrer de manière fluide et sécurisée les paiements chiffrés au sein des applications B2B et B2C. La technologie de Check Point repose sur des moteurs d'analyse de menaces alimentés par l'intelligence artificielle et enrichis. Elle permet d'assurer une prévention proactive et continue des activités malveillantes qui établit ainsi une nouvelle référence en matière de sécurité blockchain. Bien au-delà de la simple réalisation d'audits de contrats intelligents, cette collaboration entre Fuse et Check Point va bientôt permettre d'intégrer une solution de détection des menaces en temps réel, et d'assurer ainsi une protection renforcée à l'échelle de l'ensemble du réseau.

AGENDA

IoT World Congress

13 - 15 mai 2025

Gran Via Venue | HALL 8,
Barcelone Espagne

SAP Sapphire

19 - 21 mai 2025

Orange County Convention
Center, Orlando Florida

Microsoft Build

19 - 22 mai 2025

Seattle Convention Center,
Seattle USA

Dell Technologies World

19 - 22 mai 2025

Venetian Resort, Las Vegas USA

Red Hat Summit

19 - 22 mai 2025

Boston Convention and Exhibition
Center, Boston USA

Computex

20 - 23 mai 2025

Taipei Nangang Exhibition Center,
Taipei

Snowflake summit

2 - 5 juin 2025

Moscone Center, San Francisco
USA

Infosecurity Europe

3 - 5 juin 2025

Excel London, Londres UK

L'éco-conception : vers un numérique plus responsable

Tribune de Cyril de Sagazan, directeur marketing et responsable RSE chez DATASOLUTION, agence digitale experte en eCommerce et référentiel de données.

Le web semble dématérialisé, et pourtant... le web pollue !

Depuis 1995, le poids moyen des pages web est passé de 14Ko à 2 566 Ko, soit 183 fois plus lourd. Face à la montée des préoccupations environnementales et à l'impact croissant du numérique, l'éco-conception se pose en démarche innovante et indispensable.

Au-delà de la démarche éthique, l'éco-conception permet d'améliorer la performance globale des sites, en réduisant leur empreinte carbone tout en optimisant l'expérience utilisateur et la vitesse de chargement.

Pourquoi l'éco-conception est-elle cruciale ?

Le secteur numérique représente environ 4% des émissions mondiales de gaz à effet de serre, chiffre qui continue de croître, notamment avec l'IA. Les sites web contribuent à cette empreinte par leur consommation énergétique liée aux serveurs et aux transferts de données.

L'outil EcoIndex estime qu'un site web « standard » émet en moyenne 2,5 g de CO₂ par page vue. Un site éco-conçu peut réduire considérablement les émissions, atteignant en moyenne entre 1,2 g et 1,5 g de CO₂ par page vue.

Éco-conception et performance : un duo indissociable

Contrairement à une idée reçue, l'éco-conception ne sacrifie pas la performance, bien au contraire, il se révèle souvent plus rapide et efficace.

En minimisant le nombre d'éléments à charger, le site réduit de fait le temps de chargement. Des images compressées et une absence de ressources inutiles réduisent la bande passante. La simplification du HTML, l'élimination des scripts non essentiels et la compression des fichiers JavaScript et CSS améliorent la vitesse globale. Une interface épurée ne nécessite pas de ressources superflues, offrant une navigation fluide.

Concevoir des sites optimisés pour mobile permet de limiter les ressources utilisées. Ces sites mobile-first présentent en

effet un triple avantage : plus performant, plus écologique, ils augmentent significativement la conversion.

Privilégier les datacenters Green IT

L'hébergement joue un rôle clé dans l'impact écologique d'un site, les datacenters fonctionnant majoritairement avec des énergies fossiles. Privilégier un hébergeur utilisant des énergies renouvelables ou compensant ses émissions permet de réduire considérablement l'empreinte carbone. Par exemple, les datacenters, situés en France, utilisés par DATASOLUTION adoptent des pratiques écoresponsables : panneaux solaires, réutilisation de la chaleur pour les générateurs, et free cooling indirect pour un refroidissement optimisé.

Se préparer à une IA plus responsable

Derrière les promesses de gains de productivité annoncés par l'IA, il est surtout question d'investissements (licences, formations), de changements structurels ou organisationnels, d'impacts sociétaux et environnementaux. Sans parler en plus des enjeux de souveraineté liées à l'utilisation de plateformes américaines ou chinoises.

Il ne s'agit pas de faire ici un procès de l'IA, mais de savoir quelle IA nous sou-

haitons promouvoir, en restant cohérents et responsables dans nos choix et nos stratégies. En formant non seulement aux outils, mais également en sensibilisant aux conséquences sociétales et environnementales de ces outils. En privilégiant des solutions françaises ou européennes. Et en intégrant, pourquoi pas, des critères d'éco-responsabilité dans nos projets autour de l'IA.

Vers un numérique responsable

Adopter l'éco-conception est une démarche proactive et engagée. Cela demande un changement de paradigme, où la priorité n'est plus seulement donnée au design ou aux fonctionnalités, mais à l'impact global. C'est aussi un investissement dans la qualité et la pérennité de vos environnements digitaux. □





Serveurs

Dédiés à l'IA ?

La soudaine irruption de l'intelligence artificielle a entraîné une forte demande des serveurs. Ceux-ci, pour répondre aux besoins de cette nouvelle technologie, se sont transformés pour s'adapter aux nouvelles exigences des entreprises. L'IA n'est pas la seule responsable, le volume des données à traiter, la volonté de réaliser les traitements au plus près des données, l'augmentation de la bande passante des communications avec la 5G redessinent les serveurs pour obtenir toujours plus de performance, des latences plus faibles.

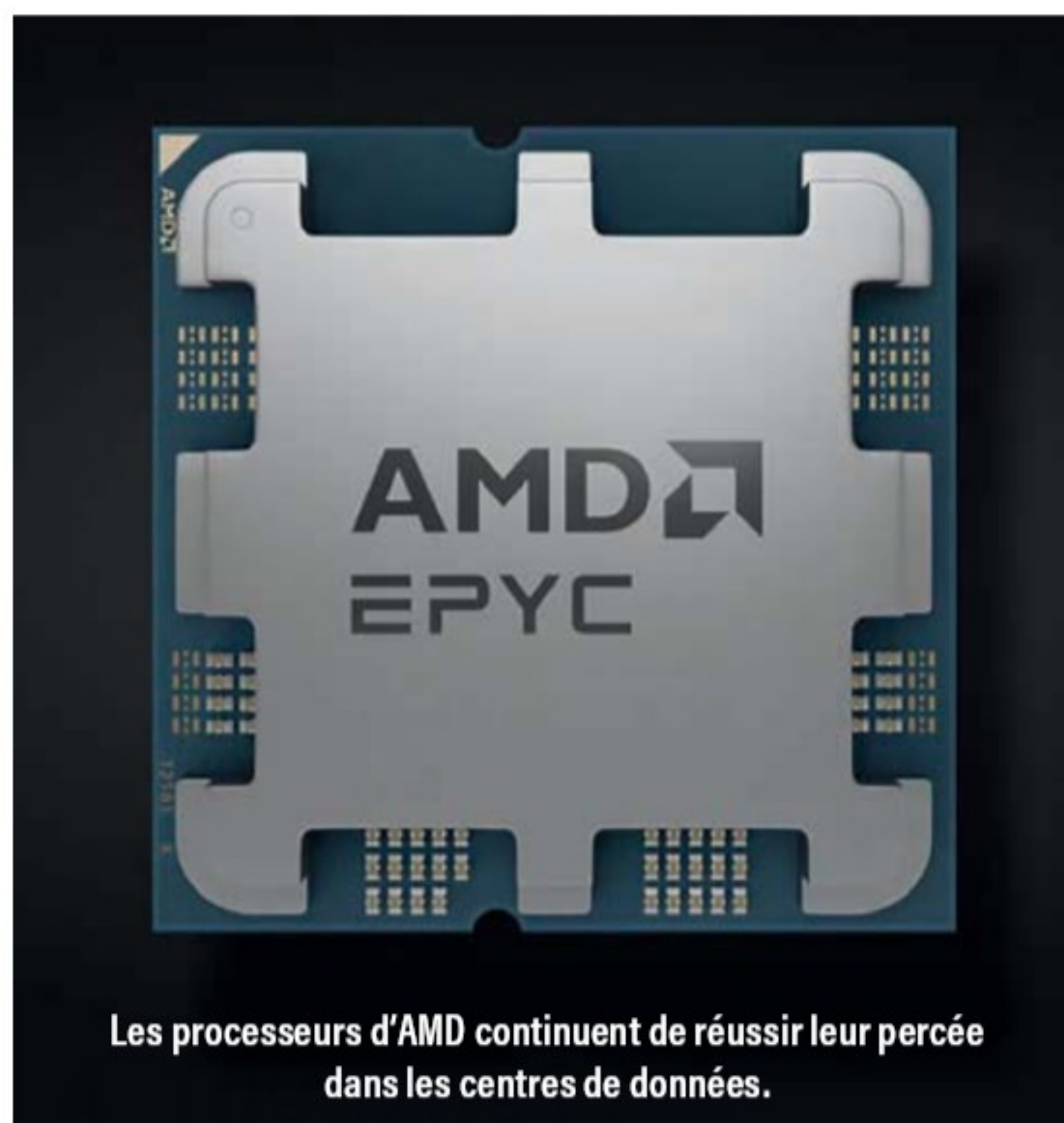
Un **marché** sous stéroïde

Pour répondre aux exigences de nouvelles technologies comme l'intelligence artificielle ou le Edge, le marché de serveurs est en ébullition et connaît une très forte croissance.

Si les chiffres diffèrent d'un cabinet d'analystes à un autre, tous s'accordent pour indiquer que le marché de serveurs connaît une période radieuse. Dans son rapport sur ce marché, IDC indique que lors du dernier trimestre de 2024, le marché a culminé à 77 Md\$ inscrivant quasiment un record en devenant le deuxième trimestre avec la plus forte croissance depuis 2019. Les taux de croissance sont faramineux, alors que la plupart des constructeurs n'ont pas encore totalement rafraîchi leur gamme de machines avec l'arrivée des nouvelles générations de processeurs qu'ils soient d'Intel ou d'AMD. Ainsi, les serveurs x86 ont connu lors de ce trimestre une croissance avoisinant les 60 %. Les serveurs sous une autre architecture connaissent eux une explosion avec une croissance de 262,1 % !

L'IA comme dénominateur commun

Cette forte croissance s'explique par les besoins en traitement des données pour l'intelligence artificielle et, concurremment, avec l'extension des projets de transformation des entreprises dans leur migration vers le Cloud et le besoin, pour certains secteurs d'activités, de traiter et d'analyser les données au plus près de la production (Edge). Preuve de cette tendance, IDC indique dans son document que les serveurs embarquant des puces graphiques (GPU) ont quasiment



Les processeurs d'AMD continuent de réussir leur percée dans les centres de données.

doublé lors de ce dernier trimestre 2024 (+192,6 %). Dans ce domaine spécifique, Nvidia écrase la concurrence et détient plus de 90 % de parts de marché. Les fournisseurs de Cloud, en particulier les hyperscalers, sont d'ailleurs particulièrement friands de ce type de serveur afin d'accueillir les charges IA de leurs clients

LE MARCHÉ EN FRANCE

Il y a peu de données disponibles directement sur le marché français. Les chiffres et analyses suivants sont issus d'une étude réalisée par Mordor Intelligence. Le rapport met en avant les financements du gouvernement dans l'expansion de l'infrastructure informatique du pays. De plus, le cabinet d'étude constate un engouement et des investissements importants dans le Cloud. Ce virage vers le Cloud se réalise sous la pression de l'émergence de nouvelles technologies comme l'intelligence artificielle, le volume des données à traiter (Big Data) et la chaîne de blocs. En outre, l'utilisation intensive des smartphones

et du volume de données qui en découle alimente une tendance forte vers un marché dynamique pour les serveurs dans l'Hexagone.

De nombreux acteurs sont présents sur le marché français. Dell, HPE, Cisco, IBM sont les principaux. Plusieurs acteurs asiatiques sont également présents : Lenovo, Hitachi Ventara, Fujitsu, Huawei. Le marché est donc devenu assez concurrentiel tirant les prix vers le bas, sauf sur la catégorie des serveurs embarquant des GPU. Le prix de ce composant reste élevé et permet aux constructeurs d'augmenter le prix de leur machine. Tous visent à étendre

leur empreinte sur le marché français, que ce soit avec leurs clients ou en gagnant des signatures nouvelles. Ils tirent parti de collaborations et partenariats stratégiques pour augmenter leurs parts de marché et leur rentabilité. Le cabinet précité voit comme tendance une adoption plus large des serveurs lames autorisant des consolidations de surface au sol. De plus, les châssis de serveurs lames sont généralement conçus avec des mécanismes de refroidissement efficaces, ce qui peut aider à maintenir une température constante dans le centre de données et à réduire les coûts de refroidissement.

Ici une vue d'un PowerEdge de Dell, constructeur numéro un en France.



ou proposer leurs propres services d'intelligence artificielle. Pour montrer la dynamique du marché, il suffit de constater qu'il a doublé depuis 2020. La demande s'étend aussi au plus près de l'endroit de production des données. Selon une analyse de Lenovo, d'ici 2030, le marché de l'edge computing devrait connaître une croissance exponentielle de 37 % par an, témoignant de l'essor fulgurant de l'IA en périphérie. Cela correspond à une réalité des entreprises : d'ici la fin de cette année, plus de 50 % des données gérées par les entreprises seront créées et traitées en dehors des data centers et du cloud selon le cabinet Gartner. Un autre cabinet, ABI Research, estime le marché de serveurs Edge à la fin de 2027 à 19 Md\$. Les déploiements sur site contribueront à plus de la moitié de cette somme. Les opérateurs de télécommunications et les hyperscalers seront les principaux acheteurs.

X86 toujours premier

Les serveurs x86 continuent de dominer le marché. Il est à noter les gains d'AMD sur les serveurs dans les centres de données qui détient désormais quasiment un quart de parts du marché avec ses puces EPYC. Si Intel est toujours devant, AMD voit ses ventes principalement sur le segment le plus haut du marché, alors qu'Intel profite de l'effet volume sur les autres segments. Il faut remarquer que, pour la première fois au troisième trimestre de 2024, AMD a plus vendu qu'Intel (3 549 Md\$ vs \$3,3 Md\$). La perpétuation de la domination de x86 dans les serveurs

s'expliquent par les progrès réalisés par les différents acteurs des semi-conducteurs. Ainsi, AMD avec son processeur Zen gravé en 7 nm, en est un exemple. Avec le retard pris par Intel sur ce niveau de gravure, AMD a pu gagner des parts de marché significatives. L'interopérabilité de ce type de serveur avec de nombreuses applications est l'avantage le plus souvent mis en avant ainsi que leur flexibilité dans leur déploiement, que ce soit sur site ou dans le cloud. Ils sont d'ailleurs critiques dans le support des environnements virtualisés et dans le traitement de larges volumes de données. Ils sont de plus en plus chers et plus économes en ressources énergétiques que les autres types de serveurs.

Linux prend sa revanche

Le système d'exploitation Linux connaît une forte croissance et la majorité de serveurs x86 fonctionne sous ce système désormais. La tendance devrait continuer et la croissance des systèmes Linux devrait s'élargir dans les centres de données et dans les entreprises. Il devance désormais Windows qui conserve cependant une forte attraction dans les entreprises. Ce système devrait lui aussi connaître une belle croissance selon les différents rapports de cabinets d'analystes. Unix suit loin derrière ces deux systèmes d'exploitation, mais reste présent ainsi que les systèmes centraux d'IBM (Mainframe et System I ex AS 400). □

B.G

Au service de l'IA

Les serveurs récents semblent n'avoir qu'une seule vocation : répondre aux besoins de la mise en œuvre de l'intelligence artificielle. Revue de quelques annonces de serveurs récentes.

viden, la branche d'activité d'Atos pour le numérique a étendu récemment sa gamme de serveur Sequana SH avec quatre nouveaux modèles. Les machines sont toutes motorisées par la dernière génération de processeurs d'Intel, les Xeon 6. Le constructeur indique que ces serveurs sont adaptés aux besoins des entreprises, des fournisseurs de cloud et des hyperscalers, pour déployer les applications critiques et d'intelligence artificielle (IA). Les machines combinent le

mode scale-up et scale-out. Bien que la virtualisation et l'adoption du cloud aient favorisé les déploiements en architecture scale-out, ils ne sont pas toujours bien adaptés au traitement d'opérations en temps réel, au Big Data et à l'analytique (par exemple, SAP Hana) qui nécessitent des ressources de calcul maximales pour traiter un très large volume de données. Ces applications peuvent bénéficier d'une architecture scale-up qui offre un grand nombre de processeurs et une capacité de mémoire importante, ce qui permet de conserver de

Une vue du ThinkEdge SE100



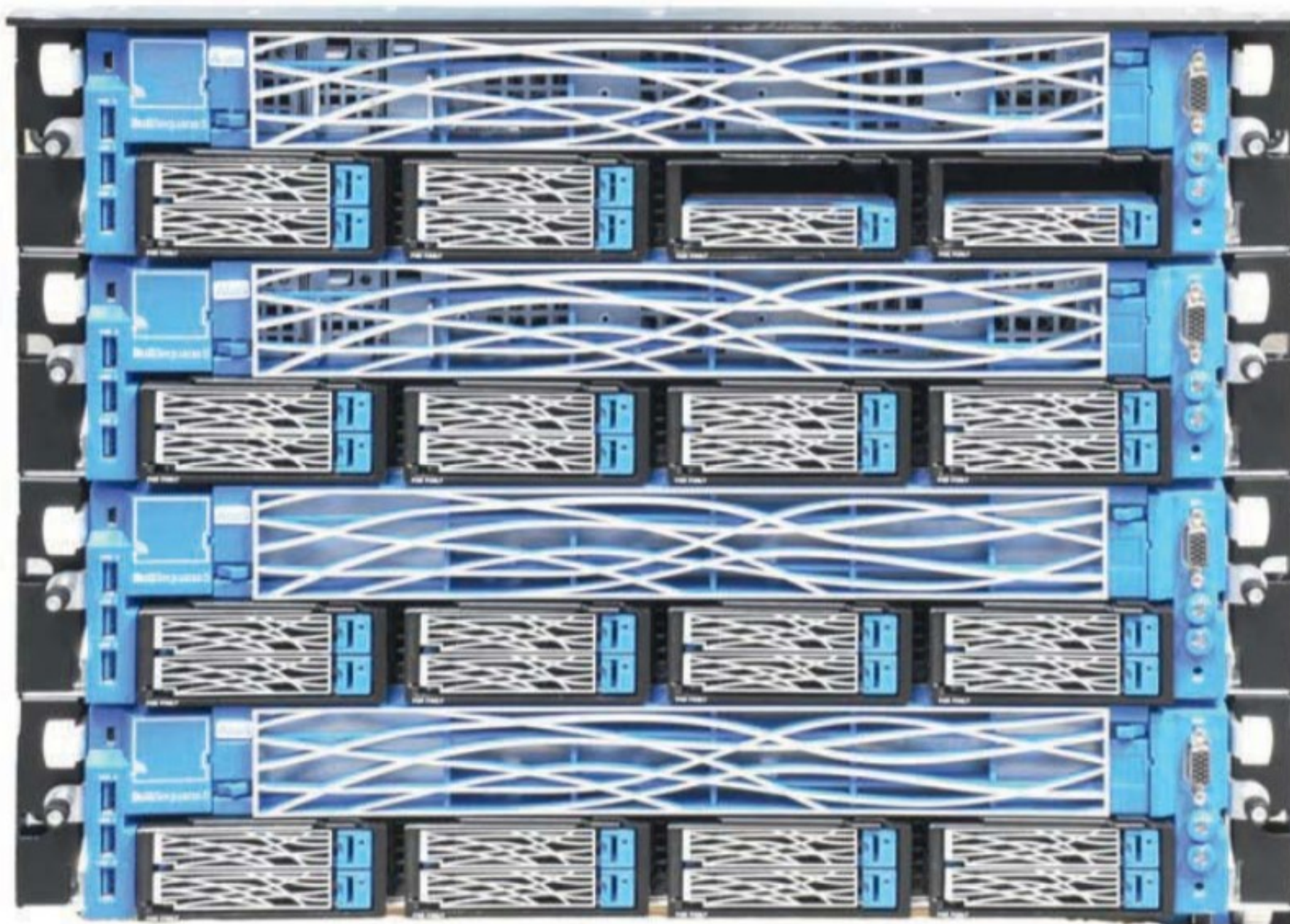
grandes quantités de données à proximité du processeur, minimisant ainsi la latence lors de la récupération des informations. Les nouveaux ajouts à la gamme BullSequana SH sont composés de quatre serveurs complémentaires d'une capacité de calcul allant d'un à huit modules serveur à deux sockets. Offrant jusqu'à 128 téraoctets de capacité de mémoire DDR5 et une évolutivité de deux à 32 processeurs par incrément de deux, grâce à la technologie d'interconnexion Node Controller UNC5 d'Eviden, les clients peuvent facilement faire évoluer leur infrastructure, en passant d'un modèle à l'autre. Cette flexibilité leur permet ainsi d'éviter la surallocation des ressources tout en préservant les investissements et les environnements applicatifs.

Pour la première fois, les serveurs scale-up BullSequana SH intégreront la technologie brevetée d'Eviden Direct Liquid Cooling (DLC) comme nouvelle option de refroidissement. Cette innovation donne aux clients la possibilité de sélectionner la solution de refroidissement adaptée à leurs besoins spécifiques. Avec la dispersion de la chaleur atteignant jusqu'à 97 % et pouvant même supporter une température d'arrivée d'eau à 40 °C, ces serveurs permettent un meilleur refroidissement tout en réduisant la consommation d'énergie pour le transport de chaleur, améliorant ainsi considérablement les rendements d'utilisation de l'énergie (Power usage Effectiveness ou PUE) des centres de données. Comparée aux systèmes de refroidissement par air avec des configurations similaires, la technologie DLC offre une réduction d'au moins 10 % de la consommation d'énergie, entraînant également une réduction de 10 % des émissions de carbone. De plus, elle permet de doubler la densité des serveurs dans les racks, et optimise ainsi les performances des processeurs Intel Xeon 6.

HPE rafraîchit ses serveurs ProLiant

Cette nouvelle génération de serveur, la Gen 12, améliore la performance et se dote d'une forte sécurité avec une gestion dopée à l'intelligence artificielle.

Sur un environnement Intel Xeon 6, les nouveaux serveurs HPE ProLiant Compute Gen12 visent à apporter la performance pour un rapport qualité/prix et une durabilité optimisée. De plus, la sécurité est renforcée avec une enclave sécurisée sur une puce dédiée dans HPE Integrated Lights Out (iLO) 7 dans le but de proposer une protection avancée contre les cybermenaces y compris celles liées à l'informatique quantique. Les serveurs respectent la norme cryptographique FIPS 140-3 niveau 31. HPE propose également un service sécurisé de mise hors service garantissant aux entreprises un recyclage conforme et la destruction des données en toute sécurité. La nouvelle génération fait bénéficier d'une amélioration des performances de 41 % par watt, réduisant jusqu'à 65 % la consommation d'énergie du serveur et permettant une consolidation importante



Un chassis blade d'un Sequana S800

du nombre des matériels. Un seul Gen 12 est l'équivalent de sept Gen 10. Un refroidissement liquide est disponible en option. L'administration des machines est simplifiée avec l'apport de l'intelligence artificielle qui prend en charge la consommation et permet de définir des seuils pour réduire les coûts et l'empreinte carbone. Une cartographie interactive facilite la gestion des infrastructures IT multisites.

Six modèles de la gamme HPE ProLiant Compute Gen12 seront disponibles à la commande à partir du 24 février 2025, avec une livraison mondiale prévue pour le 25 mars. Ces modèles incluent les HPE ProLiant Compute DL320, DL340, DL360, DL380, DL380a et ML350 Gen12. D'autres modèles, comme le HPE Synergy 480 et le HPE ProLiant Compute DL580 Gen12, seront commercialisés à l'été 2025. L'ensemble du portefeuille HPE ProLiant Compute Gen12 sera proposé via HPE GreenLake.

Dell ou l'usine d'intelligence artificielle

Avec de nouveaux serveurs PowerEdge, Dell enrichit son portefeuille de solutions d'IA générative, avec des mises à jour de la Dell AI Factory conçues spécifiquement pour les environnements AMD. Ces nouvelles solutions offrent aux entreprises des capacités d'IA optimisées. Ces nouveaux produits de la gamme Dell PowerEdge répondent à un large éventail de cas d'utilisation en matière d'IA et de charges de travail traditionnelles, tout en simplifiant la gestion et la sécurité des serveurs. Ces plateformes proposent des solutions personnalisables facilitant la gestion tout en exécutant des charges de travail à haute performance.

Ainsi, le Dell Power XE7745 prend en charge jusqu'à huit GPU PCIe double largeur ou 16 GPU PCIe simple largeur avec des processeurs AMD EPYC de 5^e génération dans un châssis 4U refroidi par air. Développés pour l'inférence IA, l'affinage des modèles et le calcul haute performance, les emplacements GPU internes sont associés à huit emplacements PCIe Gen 5.0 supplémentaires pour la connectivité réseau, créant ainsi des configurations flexibles avec une capacité de GPU PCIe DW multipliée par deux.

Ses petits frères, R6725 et R7725, sont eux aussi sur des processeurs AMD de 5^e génération. Le nouveau design du châssis DC-MHS optimise le refroidissement par air et utilise deux CPU de 500 W, ce qui permet de relever des défis thermiques en matière de puissance et d'efficacité. Ces plateformes sont conçues pour l'analyse de données et les opérations d'IA, avec leurs configurations optimisées pour l'évolutivité, et offrent des performances record pour des charges de travail telles que la virtualisation, les bases de données et l'IA. Le R7725 offre jusqu'à 66 % d'augmentation des performances et jusqu'à 33 % d'augmentation de l'efficacité au sommet de la pile.

Lenovo pousse l'IA à la périphérie

Lenovo a lancé un serveur d'inférence IA d'entrée de gamme du marché. Conçu pour rendre l'IA en périphérie plus accessible et abordable pour les PME et les grandes entreprises. Le nouveau serveur d'inférence IA est 85 % plus petit que les modèles précédents, sans pour autant faire de concession sur la puissance. Adaptable aux bureaux, aux fixations murales, aux plafonds et aux racks 1U, le ThinkEdge SE100 a été conçu pour être abordable, brisant ainsi les barrières financières de l'IA. Il propose des performances d'inférence IA qui apportent un rapport qualité/prix optimisé. Ce serveur offre des fonctionnalités de sécurité avancées et une compatibilité GPU permettant de gérer des charges de travail IA en temps réel, telles que l'inférence, l'analyse vidéo et la détection d'objets dans les secteurs des télécommunications, du commerce de détail, de l'industrie et de la fabrication. Il peut être équipé de six à huit cœurs haute performance pour une puissance maximale dans un format compact. Le ThinkEdge SE100 facilite son intégration grâce à Lenovo Open Cloud Automation (LOC-A) et un contrôleur de gestion intégré (BMC). Ces innovations permettent une installation rapide, de réduire de 47 % les coûts de déploiement et de générer jusqu'à 60 % d'économie de ressources et de temps. Sa consommation reste inférieure à 140 W, même dans sa configuration GPU maximale, ce qui réduit le coût total de possession et contribue aux engagements en faveur de la durabilité.



Un PowerEdge XE 7745 au format rack

De plus, il réduit son empreinte carbone jusqu'à 84 %, et cela commence avant même sa mise en fonction grâce à un emballage EPE à 90 %, diminuant son impact logistique. Lenovo favorise aussi une économie circulaire avec l'utilisation de soudure basse température, de matériaux recyclés et d'un programme de recyclage des produits en fin de vie.

Partenariat entre Fujitsu et Supermicro

Fujitsu et Supermicro vont combiner leurs capacités techniques et leur envergure mondiale pour offrir une gamme de serveurs sur le marché. Ces serveurs et solutions à hautes performances, économes en énergie et optimisés en termes de coûts, prendront en charge un large éventail de charges de travail dans les environnements d'IA, de HPC, de cloud et d'applications edge.

L'approche Building Block de Supermicro en matière de conception de serveurs permet de créer et de certifier rapidement une large gamme de serveurs pour des

charges de travail spécifiques, couvrant les applications IA/HPC et l'informatique à usage classique, dans des configurations allant des data centers dans le cloud aux applications edge. S'appuyant sur Fujitsu-Monaka, un processeur basé sur l'architecture du programme d'instructions Arm, qui utilise une technologie de pointe de 2 nanomètres, vont permettre d'offrir d'excellentes performances et une efficacité énergétique tout en recherchant une fiabilité, une sécurité et une facilité d'utilisation élevées avec une large compatibilité logicielle, permettant ainsi aux clients de mettre en œuvre une infrastructure d'IA durable. La collaboration s'étendra également à Fsas Technologies Inc., une filiale de Fujitsu, qui fournira des solutions d'IA générative basées sur une plateforme IA internationale qui combine les produits de serveur GPU de Supermicro et des services d'assistance à la mise en œuvre pour les opérateurs et les entreprises de data centers. □

B.G

Institut Catholique de Lille



Le choix pragmatique du brokering. Quentin Patou est un DSI de terrain. Pour répondre au choix économique de l'institut où il travaille et à sa volonté de «faire son métier», il a fait le choix de se tourner vers Evernex et de pas céder aux sirènes du Cloud ou du tout nouveau tout beau !

L'Institut catholique de Lille est l'un des établissements de l'Université catholique de Lille qui regroupe, au total, 22 établissements de formation. Cela représente au total 350 filières de formation, pour plus de 40 000 étudiants. L'Institut catholique de Lille est la plus grosse

structure avec ses 15 000 étudiants, soit un petit tiers de l'ensemble. L'institut est sous statut associatif et comptabilise environ 1 000 collaborateurs aidés de 2 000 enseignants et vacataires. Il est principalement présent sur le campus Vauban à Lille, mais également sur les sites de Roubaix, de Lomme et Issy-les-Moulineaux. Sur

le plan informatique, l'IT représente en termes de réseau environ 600 bornes Wi-Fi, 200 switchs, 300 serveurs virtualisés présents dans un double data center synchrone, deux salles qui fonctionnent en synergie, pour pouvoir pallier tous les scénarios de panne. Pour assurer cette redondance, l'institut dispose d'un data center privé et d'un collocatif.

A la base, l'institut utilisait les services d'Evernex pour maintenir des matériels anciens qui avaient dépassé les cinq ans de vie initiales et dont on avait encore besoin, tels que des serveurs sur lesquels tournaient des firewalls. Certains serveurs métiers étaient encore nécessaires pour pouvoir maintenir tous types de systèmes. Le fonctionnement n'étant pas nécessaire en 24-7, l'institut dispose d'un contrat de type J plus 1. La panne est détectée et l'intervention s'effectue le lendemain, ce qui convient à l'établissement. Les systèmes sont de toute façon déjà sur un minimum de niveau de redondance, que ce soit au niveau des alimentations, des disques ou même de certains serveurs qui sont sous cluster. Depuis à près trois ans, l'institut travaille avec Evernex dans la cadre de la modernisation du data center. Ensuite, pour des besoins métiers spécifiques, Evernex a positionné des serveurs de type reconditionné. Quentin Patou explique : « *Cela vient de notre modèle économique. Nous sommes une association loi 1901 comme les Restos du cœur, à but non lucratif et reconnue d'intérêt général. Ce qui veut dire que nous n'avons pas de résultat, pas d'actionnaires et qu'on ne récupère pas la TVA. Ce que peut faire une entreprise classique, nous ne pouvons pas nous le permettre car les effets de levier son trop forts et cela nous coûterait trop cher. C'est donc avant tout un choix*

**Quentin Patou,
DSI de l'Institut
catholique de Lille**

« *Les machines n'ont pas énormément évolué depuis près de 10 ans. On a eu à peine 30 % d'augmentation de performance* »



LES DONNÉES DU SYSTÈME DE L'INSTITUT

- Deux clusters de 300 serveurs, technologie de virtualisation Microsoft
- Soit dix serveurs physiques au total pour la virtualisation
- Encore 10 autres serveurs physiques pour d'autres besoins
- Et pour finir, une autre dizaine de serveurs physiques qui vont être supprimés dans les prochains mois.
- 30 % des serveurs reconditionnés fournis et maintenus par Evernex
- 30 % achetés neufs il y a plus de 3 ans
- 30 % plus anciens encore
- Probabilité de 100 % de la maintenance faite par Evernex d'ici deux ans

économique. » Il ajoute : « *le fait de garder les compétences en interne est aussi un choix économique. Le fait de payer deux ou trois gars pas trop cher est loin de ce que nous coûteraient des prestations de services.* »

Pas forcément besoin de machines neuves

Le lien avec Evernex provient aussi de l'absence de besoin de matériel de dernière technologie, l'institut n'ayant pas forcément le budget pour investir dans du neuf. Ainsi, l'établissement s'est tourné vers des serveurs reconditionnés, fournis et maintenus par Evernex permettant d'assurer une production efficace. Quentin Patou juge d'ailleurs sévèrement cette course à la dernière génération. « *Sincèrement, les machines n'ont pas énormément évolué depuis près de 10 ans. On a eu à peine 30 % d'augmentation de performance avec du 3 GHz de fréquence, et on a rarement mieux si l'on veut garder un ratio taux de puissance/ watt consommé intelligent. Au lieu d'un gros serveur, nous en avons mis trois. Le serveur reconditionné nous a coûté un tiers du prix du neuf et est maintenable sans limite de temps, explique-t-il, le cluster que j'ai mis en place n'a plus besoin de matériels additionnels. En termes de ressources technologiques, il est suffisant. Je n'ai pas de date buttoir. Mon contrat va se maintenir et se poursuivre d'année en année, pendant cinq ou dix ans, tant que la solution matérielle correspond toujours à nos besoins.* » Il poursuit : « *Evernex peut travailler sur plusieurs technologies différentes, donc vous pouvez unifier votre approche contractuelle.* » □

B.G

Mainframe

L'IBM z17 veut rapprocher les IA et les transactions

Comme tous les trois ans, IBM livre une nouvelle version de sa plateforme mainframe « z ». Destinée à ses clients fidèles, cette version 17 s'inscrit dans l'emballlement sur l'IA avec une nouvelle puce Telum II et ses accélérateurs IA Spyre.

Qui pourrait penser que le mainframe, né à l'aube de l'informatique, est encore bien présent dans le paysage IT de 2025 ? Le mainframe est encore très présent chez les gros industriels, les assureurs et bien entendu les banques. « 70 % des transactions mondiales transitent à un moment donné par le mainframe », explique Catherine Chauvois, directrice de la plateforme ZStack chez IBM France, qui ajoute : « Ce qui est intéressant, petit clin d'œil pour tous les

dissidents qui disent que le mainframe c'est cher, le mainframe représente seulement 8 % de la dépense IT mondiale. Donc 70 % des transactions pour 8 % des coûts, cela laisse à penser que la transaction est quand même ultra compétitive ! » Quant à penser que le mainframe reste anachronique dans une architecture IT moderne, la responsable réplique qu'une étude d'IBM montre que 82 % des décideurs IT interrogés prônent l'hybridation de leur mainframe avec le cloud, et qu'ils intègrent de plus en plus la plateforme Z dans leur écosystème IT. Si IBM ne communique pas sur le nombre de systèmes Z encore en service, l'annonce d'un nouveau modèle reste néanmoins destinée à des grands comptes fidélisés par des millions de lignes Cobol bien difficiles à moderniser. Des applications legacy, mais pas seulement, puisque ces applications doivent échanger avec les fronts web et le reste du système d'information. Catherine Chauvois souligne que la puissance installée « Z » a été multipliée par 3 ces 10 dernières années...

Que vaut le millésime 2025 de l'IBM Z ? Le constructeur axe sa communication sur l'IA. Le mainframe ne va pas bousculer Nvidia sur l'entraînement des IA, mais IBM le positionne sur l'exécution des inférences des modèles d'IA. « Le z17 a la capacité de traiter 450 milliards d'inférences avec un temps de réponse de 1 ms en une journée », argumente Guillaume Wazner, leader technique de la plateforme ZStack chez IBM France, « mais le point important, c'est que ces traitements ont très peu d'impact sur les temps de réponse de la machine. » Le secret de cette performance est que chaque puce Telum II, fondue en technologie Samsung 5 nm offrant huit cœurs de calcul à 5.5GHz. Le processeur intègre aussi une DPU (Data Processing Unit) qui va accélérer l'exécution des inférences et les entrées/sorties du processeur. En outre, la taille des caches a été accrue de 40 %. Le cache L3 dispose de 360 Mo et le cache L4 passe à 2,88 Go.

La mission du « Z » : ramener l'IA au plus près des transactions

Pour les banques qui auront besoin d'une grosse puissance de calcul IA pour faire de la détection de fraude, par exemple, IBM propose d'accroître encore la puissance d'inférence de sa machine avec l'IBM Spyre Accelerator. Celui-ci affiche une puissance de 24 TOPS, un chiffre à comparer aux près de 4 000 TOPS d'un H200 Nvidia... Néanmoins, l'architecture interne permet à la machine qui peut compter jusqu'à 200 cœurs de calcul d'accéder à n'importe quel accélérateur disponible.



L'IBM z17 occupe 4 tiroirs d'un rack standard.

4_109623917

Le processeur Telum II intègre huit cœurs de calcul, ainsi qu'un accélérateur IA accessible à l'ensemble des cœurs de la machine.



Le « Z » ne va clairement pas défier Nvidia sur l'entraînement des grands LLM, par contre, la position d'IBM est de placer les inférences au plus près des données : « Le challenge du mainframe, c'est de ramener l'IA vers la donnée et la transaction critique qui est exécutée sur le mainframe », argumente Guillaume Wazner. L'argument est de limiter les échanges avec d'autres machines lors de la vérification en temps réel d'une transaction bancaire.

Une machine de plus en plus ouverte à l'extérieur

4_109623917

Outre l'évolution du matériel, le z17 évolue sur le plan logiciel avec la nouvelle version de son système d'exploitation z/OS 3.2. « La volonté d'IBM est de standardiser la plateforme, que ce soit au niveau des traitements, de sa gestion opérationnelle. Tous les services sont maintenant exposés sous forme d'API. » IBM a revu sa copie sur le plan de l'observabilité. Commercialisée dans le courant de l'année 2025, la solution IBM Z Operations Unite viendra consolider toutes les métriques de performance de la machine, ainsi que ses fichiers logs au format OpenTelemetry. La solution viendra en complément d'IBM Concert, afin de pouvoir établir des corrélations entre les données du système Z et des autres solutions en production dans le système d'information, avec bien évidemment des modèles IA pour analyser ces masses de données.

Le z17 s'inscrit même dans l'approche Infrastructure as Code avec la capacité de gérer les configurations avec Ansible. L'acquisition d'HashiCorp par IBM en février 2025 devrait rapidement impacter le monde mainframe.

De même, les traitements du Z sont exposés au cloud via la solution z/OS Connect. Une application de nouvelle génération peut solliciter le mainframe via ses API. La solution bénéficie de la hausse de performances du z17, mais évolue en parallèle à la plateforme. « z/OS Connect permet d'exposer les services et les applications, mais beaucoup de clients veulent exposer directement

les données. Le logiciel Data Gate est très utilisé en France. Il se transforme peu à peu en Data Gate pour Watson X, car il permet d'exposer les données du mainframe à des modèles IA qui sont exécutés en dehors du système Z. Notre stratégie est de rapprocher les modèles et les IA, mais nous ne sommes pas jusqu'au-boutiste. Nous pouvons exposer ces données à destination d'un Data Lakehouse Apache Iceberg, par exemple. »

Enfin, pour les développeurs, IBM pousse à l'intégration de son système Z dans les chaînes CI/CD des entreprises. En outre, Watson X Code Assistant permet de reprendre en main le code

legacy insuffisamment documenté et dont les concepteurs ont quitté l'entreprise. Le LLM Granite d'IBM permet de convertir le code Cobol et Java, mais cette conversion du code source est loin d'être suffisante pour porter une application bancaire ou d'assurance dans sa totalité. Néanmoins, l'IA peut apporter une aide précieuse dans la maintenance ou la modernisation de ces applications. L'IA peut notamment réaliser une cartographie applicative du patrimoine Cobol, analyser toutes les dépendances des composants de l'application et aider les développeurs à refactoriser une application Cobol monolithique difficile à appréhender dans sa globalité par des humains.

La cybersécurité prend une place plus importante

Conséquence directe de l'ouverture des mainframes au cloud et aux autres applications d'entreprise, une autre évolution majeure du système Z porte sur l'arrivée de la solution de sécurité z/OS Threat Detection (IBM TDz). A la manière d'une solution de détection comportementale de type UEBA (User and Entity Behavior Analytics), la solution met en œuvre l'IA pour détecter toute déviance du système par rapport à son fonctionnement nominal. Tout changement dans un fichier de configuration, toute élévation de droits d'un utilisateur va déclencher une alerte de sécurité. La solution dispose de dashboard, afin de présenter les anomalies détectées de manière synthétique et permettre aux administrateurs d'intervenir. Enfin, IBM positionne son z17 comme plateforme de centralisation des clés de chiffrement de l'entreprise. Anciennement, IBM Enterprise Key Management Foundation — Web Edition, la solution Unified Key Orchestrator for IBM z/OS, UKO, assure une gestion centralisée des clés de chiffrement, que ce soit pour accéder à des ressources hébergées dans IBM, mais aussi pour communiquer avec les systèmes de gestion de clé d'AWS, Azure et Google Cloud. □

A.C

Backup

Synology repense la sauvegarde

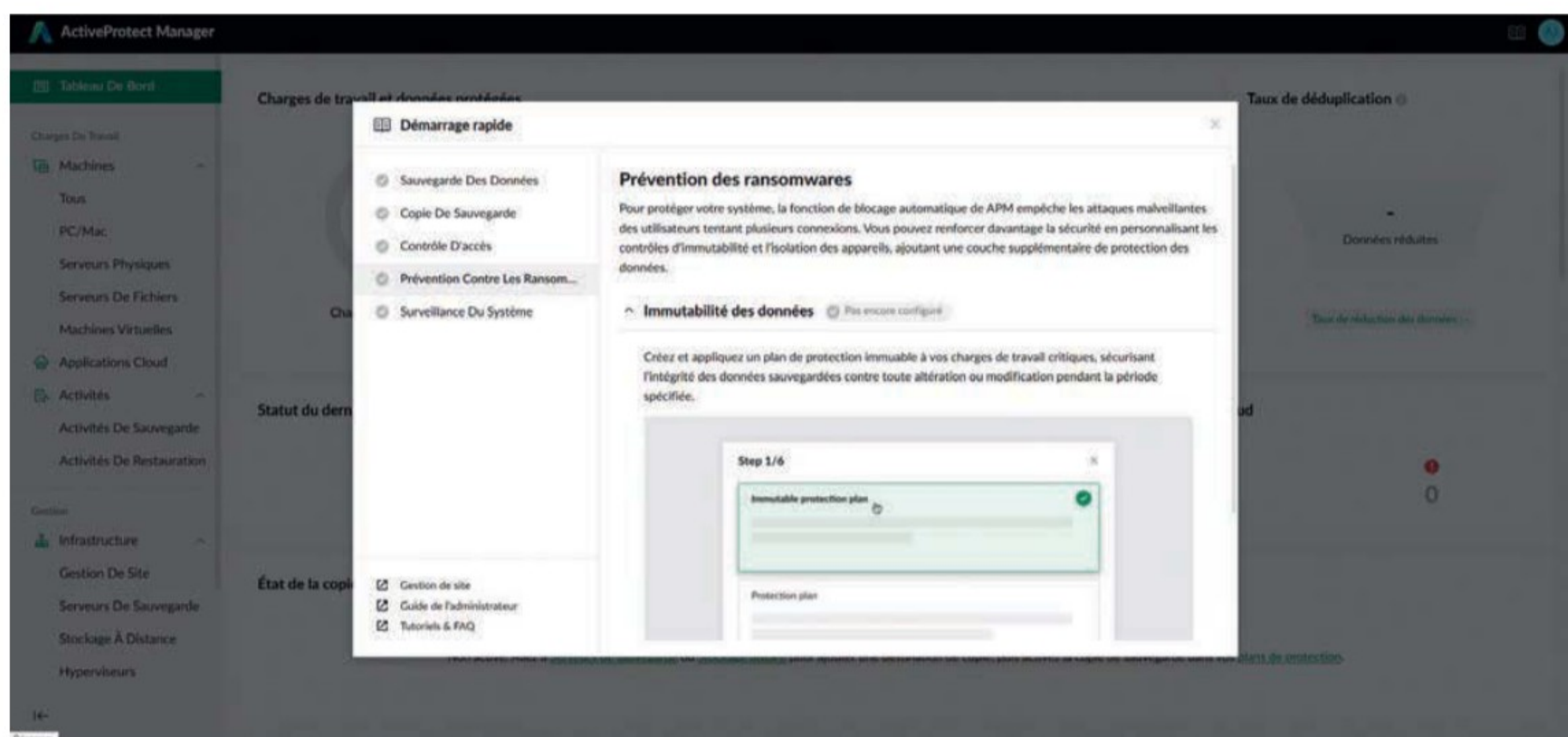
Avec sa nouvelle gamme ActiveProtect, Synology inaugure une approche résolument moderne de la protection des données. Le modèle DP320, compact et prêt à l'emploi, allie simplicité d'administration, sécurité renforcée et évolutivité pour répondre aux besoins des PME et des infrastructures multisites.

Bien connu pour ses NAS, Synology franchit une nouvelle étape en investissant le domaine de la sauvegarde d'entreprise avec une approche « appliance first » pour les environnements distribués. L'ActiveProtect DP320 est le premier modèle d'une nouvelle famille conçue pour centraliser, sécuriser et rationaliser la protection des données dans des environnements hybrides, physiques et virtualisés. Reposant sur une architecture de type scale-out, chaque boîtier peut fonctionner de manière autonome ou être intégré à une infrastructure distribuée. À grande échelle, l'ActiveProtect Manager (APM 1.0) est capable de superviser jusqu'à 2 500 sites distants et 150 000 workloads. Il faut savoir que la gestion centralisée nécessite une licence CMS à partir de 4 appliances. L'ambition de Synology est claire : proposer une solution unifiée, capable d'évoluer avec les besoins métiers, tout en garantissant une continuité d'activité optimale. Cette orientation stratégique s'inscrit dans une réponse aux limites des architectures de sauvegarde traditionnelles, souvent fragmentées et peu flexibles. En misant sur l'intégration logicielle et matérielle, Synology entend offrir un modèle résolument tourné vers l'autonomie et l'agilité, où chaque appliance peut s'ajouter à

l'existant sans perturber les opérations. Une approche qui fait écho aux tendances DevOps et aux enjeux croissants de résilience IT.

Un boîtier complet prêt à l'emploi

Contrairement aux NAS DSM classiques de la marque, le DP320 est une appliance entièrement dédiée à la sauvegarde et à la protection des données. L'appareil intègre de série deux disques durs Synology de 8 To préconfigurés en RAID 1, accompagnés du système de fichiers Btrfs, reconnu pour ses capacités d'autoréparation et de détection d'erreurs. Équipé d'un processeur AMD R1600 Dual Core et de 8 Go de RAM, il dispose d'une puissance suffisante pour effectuer des sauvegardes avancées, y compris pour des charges de travail intensives. Pour l'heure, la gamme comprend également deux autres versions : le DP340 (4 baies HDD + 2 SSD 400 Go) et le DP7400 (10 baies + 2 SSD 3840 Go) qui permettent de couvrir une large palette de besoins pour des entreprises multisites comme des datacenters. Compact (166 x 106 x 223 mm pour un poids de 2,9 kg), le DP320 embarque une connectique limitée, mais fonctionnelle : un port USB 3.2 en façade, un double port réseau Gigabit à l'arrière, dont un réservé à l'administration.



Dérivée de la console DSM, ActiveProtect Manager centralise l'ensemble des opérations de sauvegarde dans une interface web à la fois claire et intuitive.

Il est livré avec deux câbles RJ45, une alimentation externe, ainsi que deux clés pour verrouiller et déverrouiller l'accès aux baies. Seul bémol : l'impossibilité d'étendre à terme la capacité de stockage. L'ensemble est en revanche optimisé pour une mise en production immédiate. Le pré-équipement en disques, la configuration RAID d'usine et la robustesse du châssis en font une solution de sauvegarde de proximité parfaitement adaptée à des déploiements décentralisés.

Déploiement simplifié, gestion centralisée

L'un des atouts majeurs de cette appliance réside dans la simplicité de son déploiement. Grâce à des assistants de configuration intelligents et une interface web minimaliste, le boîtier est opérationnel en quelques étapes. La console ActiveProtect Manager 1.0, dérivée de l'expérience DSM, mais spécifiquement conçue pour la sauvegarde, centralise la configuration des agents, les politiques de protection, la supervision des tâches, et les alertes dans une interface intuitive. Conçu pour des équipes IT réduites, l'environnement ActiveProtect s'intègre rapidement dans des infrastructures hétérogènes. Le système propose des tableaux de bord clairs, des rapports personnalisables et un niveau de granularité d'administration poussé, compatible avec les principaux services d'annuaire (LDAP, AD, SAML 2.0). La configuration des droits, des rôles et des notifications se fait de façon centralisée, avec une logique orientée métier. L'interface favorise l'adoption rapide de bonnes pratiques : planification des sauvegardes, isolation des flux critiques, tests de restauration périodiques. Elle offre également une vision consolidée des incidents, des taux de réussite, ou encore des niveaux de conformité pour faciliter le reporting.

Une sécurité sans compromis

L'ActiveProtect DP320 prend en charge une très grande variété d'environnements, dont VMware, Hyper-V, Windows, Linux, macOS, bases de données (SQL, Oracle) et même des applications SaaS comme Microsoft 365 ou Google



Workspace. La solution est conçue pour offrir une protection continue, avec vérification automatique de l'intégrité des sauvegardes, tests de restauration dans des environnements sandbox, et restaurations granulaires ou complètes (fichiers, machines virtuelles, conversions P2V/V2V). Sur le plan de la cybersécurité, Synology dote son appliance d'une palette avancée de mécanismes anti-ransomware : chiffrement fort et stockage Worm (Write Once Read Many), immutabilité des sauvegardes, isolation réseau via segmentation, et détection des comportements anormaux. Concrètement, cela signifie que les données ne peuvent être ni modifiées ni supprimées par un attaquant, ou même un administrateur. À cela s'ajoute une déduplication à la source qui optimise l'usage de la bande passante et de l'espace de stockage, sans impacter les performances. Sans oublier des fonctions de « data audit » et de « compliance tracking » qui peuvent s'avérer précieuses dans le cadre de la gouvernance des données (RGPD, ISO 27001), ou pour des obligations de conservation à long terme.

Une offre prometteuse

Avec l'ActiveProtect DP320, Synology marque une rupture avec son ADN purement NAS pour s'attaquer aux nouveaux défis de la protection des données des entreprises. Cette appliance, à la fois robuste, compacte et intuitive, propose une alternative sérieuse aux solutions de sauvegarde plus complexes et souvent plus coûteuses. Elle se distingue par sa capacité à s'intégrer progressivement dans des environnements distribués, sa conception centrée sur l'utilisateur et ses fonctionnalités de sécurité avancées. Dans un marché en pleine recomposition, où la sauvegarde devient un enjeu de résilience stratégique, le DP320 incarne une vision pragmatique, mais ambitieuse : redonner aux entreprises un contrôle total sur leurs données, sans surcoût ni complexité excessive. ☐

J.C

CARACTÉRISTIQUES TECHNIQUES SYNOLOGY ACTIVEPROTECT APPLIANCE DP320 :

- **Processeur** : AMD R1600 (2 cœurs)
- **Mémoire vive** : 8 Go
- **Stockage** : 2 HDD 8 To (RAID 1)
- **Système de fichiers** Btrfs
- **Interface réseau** : 1 port RJ-45 (gestion)
+ 1 port RJ45 (transfert de données)
- **Connectique** : 1 port USB 3.2
- **Dimensions (H x L x P)/poids** : 166 x 106 x 223 mm, 2,9 kg
- **Prix** : 2088 €

Événement

Pure Storage à la rencontre de ses clients

Lors d'un dîner organisé par l'Informaticien avec le concours actif de Pure Storage, le CEO de la compagnie a été à la rencontre de ses clients français.

Charles Giancarlo, le CEO de Pure Storage, a fait l'honneur d'être présent à un dîner où étaient invités différents clients importants de la société. À la suite d'une rapide introduction, Charles Giancarlo a pris la parole afin d'expliquer la manière dont il dirige l'entreprise et pourquoi sa technologie existe. Il a ensuite évoqué son parcours dans l'industrie informatique. *« Je me sens toujours ingénieur ». Il a ajouté : « En tant qu'ingénieur en semi-conducteurs. J'ai conçu des puces. Mais dans la vingtaine, j'étais très intéressé par les communications de données. Et étant très insensé, j'ai lancé une entreprise. Cette entreprise a finalement trouvé son chemin vers la Californie. Et nous avons produit, à l'époque, ce qui était considéré comme un commutateur haute vitesse. Et l'un de nos premiers clients, plusieurs de nos premiers clients, se lançaient dans ce qui était encore non public et qui s'est appelé Internet. J'ai donc été très tôt impliqué dans la construction de la technologie Internet. Une entreprise que nous avons lancée après cela a été la première à construire un commutateur Ethernet ». Après le rachat de cette première entreprise par Cisco, où il est resté 15 ans en charge de toutes les lignes de produits, il a ensuite passé du temps dans le capital-investissement jusqu'au début des années 2000. Puis après s'être ennuyé dans le milieu de l'investissement, il a rejoint Pure Storage.*

La vision d'une révolution

Ce qu'il l'a attiré vers cette entreprise : *« je sentais que le stockage des données allait subir une sorte de révolution sur l'idée que nous allions passer du disque magnétique, inventé il y a 70 ans, à un semi-conducteur que je connaissais bien, qui était la technologie flash. En général, lorsque vous passez d'un dispositif mécanique à un dispositif électronique, cela ouvre généralement de nombreuses opportunités pour créer un avantage concurrentiel, pour améliorer l'industrie dans son ensemble. Cela était vrai à l'époque d'Internet, lorsque nous sommes passés de l'analogique au numérique. Mais c'est aussi vrai dans le stockage des données. Lorsque vous utilisez la puissance d'un semi-conducteur, vous pouvez faire beaucoup de choses qui ne pouvaient pas être faites mécaniquement. Chez Pure Storage, nous avons développé cette idée d'Evergreen, dont je pense que beaucoup d'entre vous sont très familiers, ce qui signifie que nos produits ne deviennent jamais obsolètes ».*



Charles Giancarlo, lors de son intervention au dîner organisé par l'Informaticien et Pure Storage

La discipline comme vertu

« Pendant les 10 dernières années, nous nous sommes concentrés sur le soutien de toutes les différentes zones de stockage, donc le bloc, le fichier, l'objet haute performance ainsi que le bas coût, les grands systèmes et les petits systèmes. Et nous nous concentrons sur le maintien d'un environnement logiciel unique, que nous appelons Pure. Honnêtement, cela nous a pris plus de temps que prévu. L'une des choses que nous voulions faire était de rendre

le stockage beaucoup plus simple à gérer, c'est-à-dire qu'il disparaisse effectivement. L'une des choses qui m'a surpris lorsque j'ai regardé le stockage, c'est qu'en entreprise, chaque matrice de stockage est son propre îlot », constate-t-il. Il précise : « nous avons maintenant introduit un logiciel qui permet à toutes nos matrices de fonctionner comme un cloud de stockage plutôt que comme des matrices individuelles. Ce que nous voulons faire pour nos clients, c'est leur permettre de créer des clouds de données d'entreprise, plutôt que de gérer des matrices de stockage individuelles. Si vous nous demandez notre vision, c'est d'unifier le stockage des données que nos clients ont afin qu'ils ne gèrent plus le stockage. Ils sont capables de gérer leurs données par politique. Ainsi, vous pouvez définir vos politiques et les données se gèrent d'elles-mêmes en fonction. Voilà notre vision globale pour l'entreprise. Sur l'entreprise elle-même,

nous venons, il y a quelques mois, d'annoncer un chiffre de 3,2 milliards de dollars, environ 6 000 personnes, dont un tiers en Europe et au Royaume-Uni. Et nous dépensons, et je pense que c'est probablement ce qui nous différencie le plus, 20 % de notre chiffre d'affaires en R&D. Nous considérons donc vraiment le stockage comme une haute technologie plutôt qu'un produit de base. »

Il est ensuite revenu sur le choix de Pure Storage par Meta et les bénéfices que Meta peut en tirer : réduction de coûts et économies d'énergie. *« Nous pouvons leur fournir une solution flash pour tous leurs niveaux de stockage. C'était plus performant, mais surtout, une puissance, un espace et un refroidissement équivalents à un dixième de leur environnement existant. »* □

B.G

Une plateforme unifiée de sécurité et d'observabilité pour une résilience inégalée.

De nombreuses organisations parmi les plus grandes et complexes au monde s'appuient sur Splunk pour contribuer à assurer la sécurité et la fiabilité de leurs systèmes numériques. Découvrez notre plateforme unifiée de sécurité et d'observabilité sur splunk.com/fr_fr

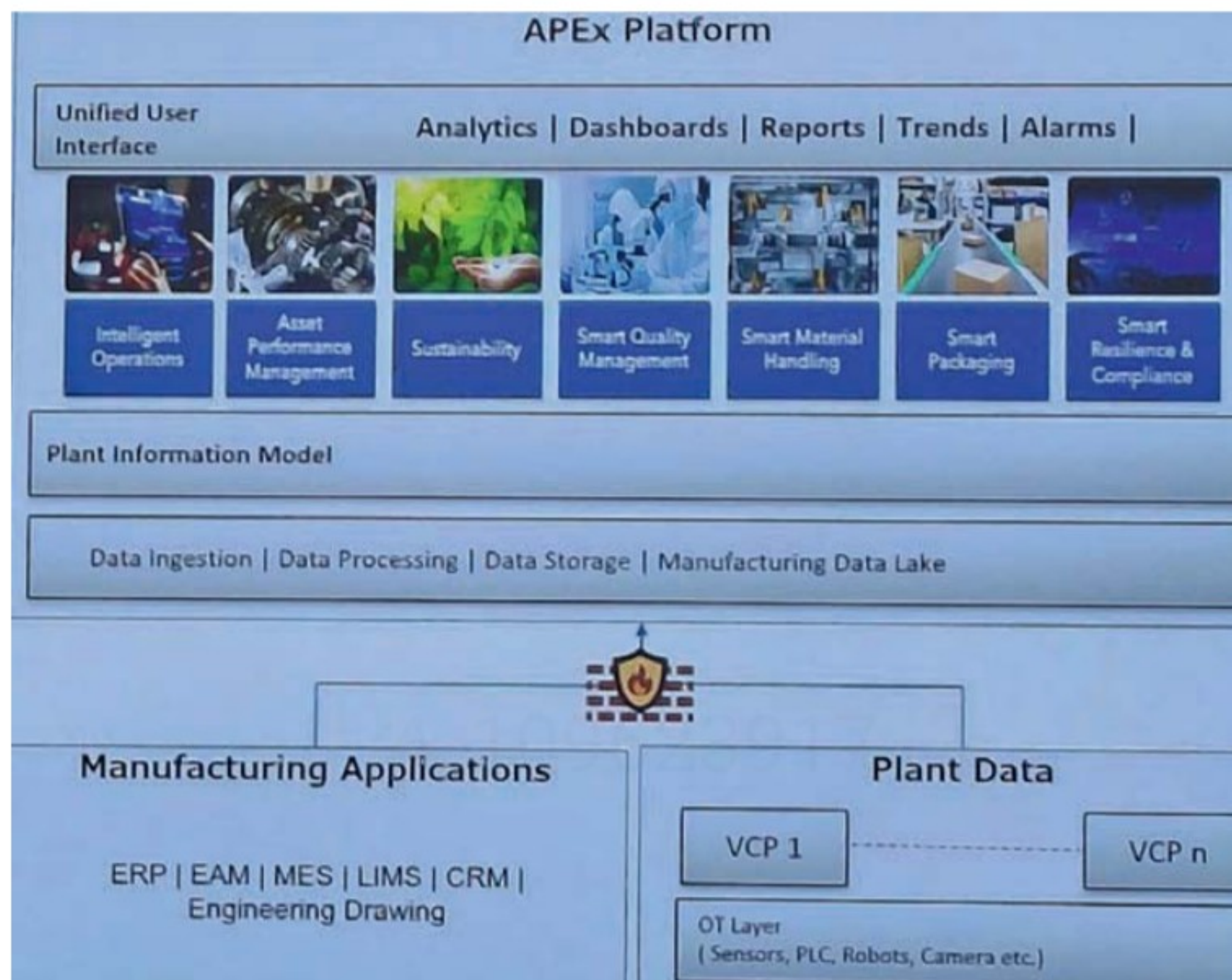
splunk>
a **CISCO** company

Manufacturing

Omron et Cognizant s'allient pour unifier IT et OT

Le géant nippon de l'automatisme et de la robotique a signé un partenariat stratégique avec Cognizant qui devient son choix de prédilection pour l'engineering, dans le but d'intégrer les technologies de l'information et les systèmes opérationnels.

Les deux entreprises se rapprochent avec des ambitions fortes, afin de développer le manufacturing du futur. Junta Tsujinaga, le président d'Omron présent lors de la présentation de ce partenariat au centre d'automatisation non loin de Stuttgart en Allemagne, voit dans la solution développée en commun le début de l'industrie 5.0. Il constate que les promesses de l'industrie 4.0 n'ont pas toutes été tenues et que l'apport de la technologie ne fait qu'émerger. Il voit trois écueils principaux : une mauvaise collaboration et un mauvais partage des données entre les deux mondes de l'IT et de l'OT, la difficulté de lire les données sur des machines ou des systèmes anciens, et le faible recours à l'intelligence artificielle dans le but d'une gestion proactive mêlant apprentissage machine et intelligence artificielle générative. Dans ce contexte,



La plateforme APex

il ajoute que le choix de Cognizant a été celui de son expertise dans le domaine précis de l'industrie, afin de mettre en œuvre la plateforme adéquate. Ravi Kumar, le CEO de Cognizant lui aussi présent lors de la manifestation, a mis en avant les développements de son ESN dans ce domaine avec la mise en place d'un manufacturing autonome intelligent et connecté. Il a rappelé le rôle de son entreprise dans l'absorption des technologies digitales par l'industrie et les partenariats importants que sa société détient avec plusieurs applications en ligne, AWS et Nvidia, pour proposer des jumeaux numériques, des analyses en temps réel, des plateformes allant des systèmes opérationnels au cloud avec une forte présence en périphérie (Edge). Il a ensuite développé les possibilités que souhaite apporter la plateforme commune comme une production en temps réel avec une forte automatisation.

AUDI VIRTUALISE SA PRODUCTION

Edge Cloud 4 Production (EC4P) d'Audi, reposant sur le logiciel VMware Cloud, est désormais opérationnelle. Le premier automate programmable virtuel (vPLC) a été déployé dans l'usine de Boellinger Hoefe, en Allemagne, où Audi produit son modèle électrique e-tron GT. Avec EC4P, Audi centralise la gestion et la maintenance de ses équipements informatiques industriels, simplifie l'application des correctifs de sécurité et réduit son impact environnemental en limitant l'usage de matériel physique et les interventions manuelles. Audi transforme ses usines en sites de production intelligents, en intégrant une automatisation pilotée par logiciel, directement sur la ligne de production et en rapprochant les mondes de l'IT et de l'OT (Operational Technology). Cette initiative repose sur une collaboration étroite avec plusieurs partenaires technologiques clés, dont Broadcom, Cisco et Siemens. Ce projet est un élément du programme 360factory, qui vise à rendre la production plus efficace et conduite par les données. À terme, ce cloud privé sera étendu à l'ensemble des sites de production, afin de tirer parti des avancées du contrôle numérique dans les processus industriels du constructeur.

Un partenariat gagnant pour les deux parties

Si le partenariat n'est pas exclusif, il prévoit cependant une démarche commerciale commune et des services joints fournis par les deux entreprises. La solution a été testée et est en démonstration dans le laboratoire d'Omron à Tokyo au Japon. L'objectif est plus qu'ambitieux et table sur un chiffre d'affaires de 50 B\$ en cinq ans. Les deux entreprises visent tout d'abord différents secteurs d'activité comme l'automobile, l'industrie des semi-conducteurs, l'électronique, l'industrie, les sciences de la vie et biens de consommation (agro-alimentaire, boissons).

La combinaison des deux entreprises a pour but de développer un nouveau modèle qui adresse les problèmes des sites industriels et leur gestion par une analyse et une amélioration continues. Par l'intégration des systèmes informatiques et opérationnels, les deux entreprises seront à même de promouvoir une industrie durable, tout en améliorant la productivité et l'efficacité des forces de travail tout en abaissant les émissions de carbone.

Par ailleurs, si les deux partenaires sont reconnus dans leur domaine sur le continent asiatique, ce n'est pas forcément le cas dans les autres parties du monde, en particulier l'Europe. Ainsi, comme nous l'avons vu dans un numéro précédent avec l'interview du directeur général de Cognizant en France, si l'entreprise a une force de frappe comparable à celle de Capgemini, elle déplore un déficit de notoriété patent. Dans la même veine, Omron vise à s'étendre en dehors de son archipel pour trouver

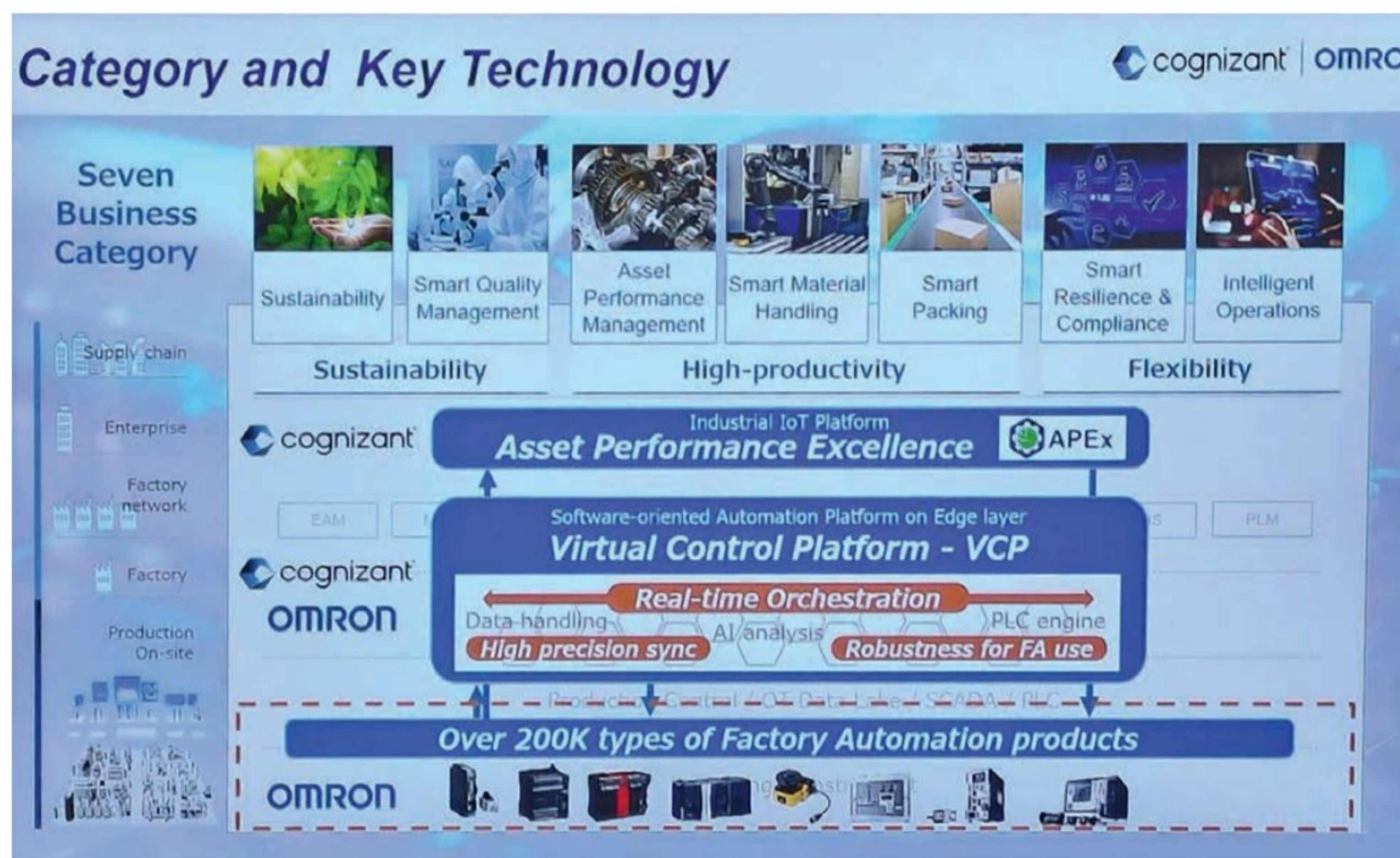
des relais de croissance et préparer son avenir avec des services s'appuyant sur les données.

Une plateforme virtualisée

A la base de la plateforme commune se placent les 200 000 produits d'automatisation d'Omron qui remonte ses données vers le cœur de la solution, la Virtual Control Platform (VCP), une plateforme d'automatisation définie par logiciel qui se localise en périphérie du site de production, afin d'ingérer et de manipuler les données, de réaliser des analyses par l'intelligence artificielle et sert de moteur de contrôle (PLC engine) programmable. Cette couche apporte une synchronisation des données d'une grande précision, une orchestration en temps réel et des fonctions d'automatisation robuste de la production en usine. Au-dessus de cette couche technique, l'appliquatif Apex fait le lien vers l'utilisateur. À ce niveau, sont unifiées les données provenant des ERP, LIMS, MES, EAM et les données provenant des sites de production par le biais de VCP. La plateforme prend en charge l'ingestion, le traitement, le stockage et la rétention des données dans un data lake. À la suite d'analyse, elle crée un modèle des informations du site de production pour proposer des alarmes, des tendances, des analyses dans différents domaines dont la qualité et la durabilité.

Des cas d'usages précis sont déjà perçus comme une gestion unifiée sur des centaines de sites, la gestion des données pour améliorer l'excellence opérationnelle des sites ou tout simplement de la maintenance prédictive pour éviter les arrêts de ligne. □

B.G



La plateforme VCP

Certification

Écritel : une nouvelle page à écrire

L'ESN a déjoué et surmonté l'attaque dont elle a été victime en fin d'année. Elle prépare sa certification SecNumCloud et un data center désigné ainsi près de Vannes.

« *a page est tournée. Écritel sort renforcée de cette attaque d'activistes pro-russes que nous avons déjouée* », insiste Audrey Louail, la CEO et co-fondatrice de l'ESN installée dans le Morbihan, également présidente de Croissance Plus jusqu'à cet été. En décembre dernier, le spécialiste de l'hébergement en cloud et de services managés a fait face à un assaut mené par le groupe d'activistes russes Hunters International. « *Hélas, ils ont eu accès à l'un de nos serveurs internes et ont pu voler des données non stratégiques. Nous avons échangé et partagé avec nos clients en totale transparence tout ce qui pouvait les concerner. Écritel n'a enregistré aucune défection de clients. Aujourd'hui, notre sécurité est renforcée, durcie grâce au rajout de nouveaux outils. Nous avons élaboré une politique spécifique pour les données les plus sensibles de nos clients et renforcé plus tôt que prévu la sécurité de nos postes utilisateurs* », poursuit-elle.

De ce « *mal pour un bien* », comme le résume son directeur général, Julien Mellul, Écritel en a profité aussi pour se lancer dans une procédure de certification SecNumCloud avec l'ANSSI, également mobilisée à ses côtés pendant l'attaque de la fin 2024. « *Cela représente une belle opportunité pour nous. Nous avons bon espoir d'obtenir cette certification et de terminer son déploiement à la fin de cette année. Les audits définitifs devraient se dérouler entre septembre et octobre* », précise-t-il. « *Ces process sont extrêmement exigeants, mais nous avons en face de nous des hackers extrêmement agressifs* », rebondit Audrey Louail.

En l'attendant, l'ESN prépare l'implantation d'un centre de données dans la lande bretonne, à Grandchamp, un petit village dans l'agglomération de Vannes. Ce projet vise à répondre à la demande croissante d'un écosystème breton très dynamique mais dépourvu en infrastructures de ce genre. Pour mettre toutes les chances de son côté, et séduire davantage d'industries stratégiques, aussi bien publiques que privées en particulier dans la défense, terrestre et navale, ce centre est désigné SecNumCloud dès sa conception, ce qui, insiste Audrey Louail, « *garantit une disponibilité immédiate des données et leur sécurisation même à longue distance. Ce site répond à une demande de nos clients bretons pour un hébergement de leurs données et de leurs applis en Bretagne, voire qui les rapatrient en France. Il répond à un besoin de souveraineté numérique de la Bretagne et de la France. Ce sera le premier site SecNumCloud by Design en France et se trouvera en Bretagne* ».

Reste également à convaincre les riverains de ce PUMA ou, selon les mots d'Audrey Louail, « *projet utile mais ailleurs* ». Écritel avance avec deux atouts non négligeables :

Audrey Louail,
CEO et co-fondatrice
d'Écritel



« *Aujourd'hui notre sécurité est renforcée, durcie grâce au rajout de nouveaux outils. Nous avons élaboré une politique spécifique pour les données les plus sensibles de nos clients et renforcé plus tôt que prévu la sécurité de nos postes utilisateurs* »

une alimentation en énergies vertes, en partie avec des ombrières sur site, et un réemploi de la chaleur fatale dégagée pour le chauffage d'infrastructures communales.

D'ici-là, Écritel poursuivra son expansion en France, et n'exclut pas des acquisitions. Même si le climat des affaires semble moins porteur qu'il y a un an, comme la présidente de Croissance Plus a pu le constater aussi en prenant le pouls des membres de l'association réunis pour un Spring Campus à Biarritz début avril. « *Notre activité est bonne, malgré un développement plus lent marqué par l'absence d'investissements informatiques* », pointe Audrey Louail. L'ESN, qui a enregistré un chiffre d'affaires de 34 millions d'euros l'an dernier, en hausse de 13 %, veut accélérer son déploiement vers l'est, en particulier en Auvergne Rhône-Alpes, avec une nouvelle agence à Annecy ou vingt recrutements sont prévus. En outre-mer, une implantation devrait suivre en Guadeloupe, après Tahiti il y a deux ans. L'entreprise ne prévoit pas de lever des fonds. « *Nous avons bouclé le rachat de sa participation à notre actionnaire. Cette opération de rachat de notre dette a permis à tous nos salariés de devenir actionnaire de leur entreprise. Nous avons enfin pu le faire* », se réjouit sa CEO. □

V.B

L'IA manque de matière première

par Bertrand Garé



Alors que notre pays et l'ensemble de la planète s'apprête à construire partout des centres de données, le fantasme Elon Musk et des chercheurs d'Epoch AI, relayés dans un article dans la revue *Nature* ont tout à coup refroidi les enthousiasmes. Il est souligné que, d'ici 2028, le volume de données nécessaires pour entraîner un modèle sera équivalent à l'ensemble des textes publics mis en ligne. Pour faire court, l'IA va manquer de données et, en conséquences, les modèles ne progresseront plus.

À la recherche de nouvelles données

Les grands de l'IA ne semblent pas s'inquiéter outre mesure même s'ils reconnaissent le problème. Ils avancent d'ailleurs la possibilité d'utiliser des données synthétiques ou de se servir de nouvelles sources comme les vidéos que vous mettez sur votre réseau social favori. Il leur est aussi possible de faire des partenariats afin d'avoir accès à des données qui ne sont pas dans le domaine public. Le problème n'est cependant pas anodin. OpenAI en subit les conséquences avec les retards pour GPT-5 dont les coûts de développement s'envolent. Dans le journal belge « L'Écho », il est indiqué que « *chaque entraînement de six mois de GPT-5 coûterait à OpenAI environ un demi-milliard de dollars rien qu'en coûts informatiques. La formation de GPT-4 avait coûté plus de 100 millions de dollars. Les futurs modèles d'IA devraient, eux, dépasser le milliard de dollars de coût de développement. L'échec d'un entraînement est comparable à l'explosion d'une fusée spatiale peu après son lancement* ».

Des modèles plus intelligents ?

Normalement, les modèles d'intelligence artificielle deviennent de plus en plus performants à mesure qu'ils ingurgitent des données. Mais avec le manque de données textuelles,

il va falloir chercher ailleurs. Tamay Besiroglu, chercheur en IA, a déclaré dans une interview donnée à Associated Press : « *Si vous commencez à vous heurter à ces contraintes concernant la quantité de données dont vous disposez, vous ne pouvez plus vraiment faire évoluer vos modèles de manière efficace. Et la mise à l'échelle des modèles a probablement été le moyen le plus important d'étendre leurs capacités et d'améliorer la qualité de leur production.* »

Dans l'étude réalisée par Epoch AI, il est estimé que la quantité de données textuelles sur lesquelles les modèles d'IA sont formés augmente d'environ 2,5 fois tous les ans seulement. Ils affirment également que les grands modèles de langage, tels que GPT-4 d'OpenAI et Llama 3 de Meta, pourraient être à court de données d'ici 2026. La solution la plus viable, selon eux, serait de former les modèles de langage sur des données synthétiques (générées), ce qui implique l'apprentissage par transfert à partir de domaines riches en données. D'ailleurs, OpenAI, Google et Anthropic travaillent déjà sur cette solution, ainsi que des éditeurs de solutions applicatives afin d'améliorer les jeux de données pour entraîner leurs agents. Des chercheurs de Rice et de Stanford constatent cependant que l'alimentation par ces données synthétiques a entraîné une baisse considérable de la qualité de production de contenu. Dans un article sur le site de la RTS, la radiotélévision suisse romande, il est souligné que cette méthode est risquée et peu fiable, car les modèles ont tendance à halluciner et à mentir. Il est aussi mentionné dans cet article comment un modèle de langage, publié par Meta en 2022, s'est dégradé après avoir été formé à plusieurs reprises sur des données créées par des intelligences artificielles. Ces erreurs sont ensuite transmises, répétées et multipliées. Cette situation risque ainsi de créer une « boucle autophage » et suscite de nombreuses questions quant à la possibilité des algorithmes d'IA de devenir plus efficaces en produisant de meilleurs résultats avec moins de données. Dès 2023, les chercheurs

soulignaient que les résultats obtenus avec l'utilisation de ce type de données étaient confus et inquiétants. La situation est telle que le patron d'OpenAI a lancé un appel vibrant aux collectivités et aux détenteurs de grands jeux de données.


Les futurs beaux jours de la CNIL

Les grands faiseurs de l'IA, Gafam et OpenAI, essaient donc de trouver des solutions de contournement. Deux méthodes ont été utilisées. La première consiste à changer les conditions d'utilisation des réseaux sociaux et autres documents mis en ligne, comme Google qui a discrètement changé ses conditions d'utilisation afin de pouvoir exploiter librement les documents Google Docs accessibles au public et les évaluations des restaurants sur Google Maps. L'autre est le passage en force. « Les chercheurs de l'OpenAI ont donc créé un outil de reconnaissance vocale. Nommé Whisper, il est capable de retranscrire l'audio de vidéos YouTube et de produire un nouveau texte conversationnel. La maison mère de ChatGPT a ainsi dérobé, sans respecter le droit d'auteur, les données d'un million d'heures de vidéos pour nourrir et entraîner son modèle », indique le média suisse. L'article ajoute : « Chez Meta, qui possède Facebook et Instagram, des responsables ont envisagé, l'an dernier, d'acquérir la maison d'édition Simon & Schuster dans le but d'obtenir le contenu d'œuvres longues. Ces discussions ont été révélées par le New York Times qui a obtenu des enregistrements de réunions internes. Elles incluaient des échanges sur la collecte de données protégées par le droit d'auteur sur internet, quitte à risquer des procès. »

Il faut raison garder

Puisque les grands modèles risquent de faire défaut, le cabinet Gartner voit, lui, l'essor des SLM (Small Language Models). Les petits modèles d'IA adaptés à des tâches spécifiques gagnent du terrain. Selon les prévisions, d'ici 2027, les usages des SLM en entreprises seront au moins trois fois supérieurs à ceux des modèles linguistiques généraux. C'est la piste la plus prometteuse pour le moment, du fait de l'emploi de jeux de données plus restreints, d'un plus faible besoin de puissance de calcul avec souvent juste le recours à des CPU classiques et non des puces graphiques à hautes performances, et une plus faible empreinte énergétique toujours bonne à prendre pour s'afficher « vert ». Ces modèles ne sont pas une version allégée des LLM et créent une nouvelle catégorie pour l'exploitation des données des entreprises avec leurs qualités et défauts propres. D'ailleurs, les cabinets d'analystes commencent sérieusement à se pencher sur ce type de modèles et estiment déjà son marché à plus de 5 Md\$ en 2032, en connaissant un taux de croissance annuel pondéré de plus de 28 %.

Il devient évident que l'économie de l'IA va totalement changer dans les mois ou années à venir. Il est possible que les entreprises monétisent leurs propres modèles et acceptent enfin de partager les données contre plus que de la menue monnaie. Cela pourrait aussi pousser les créateurs de modèles à partager la valeur en payant leur obole à ceux qui accepteraient de leur fournir leurs données. Serait-ce la fin du « quand c'est gratuit, c'est toi le produit » ? □



*En science et surtout
en politique, les idées,
souvent plus têtues
que les faits, résistent
au déferlement
des données et des
preuves.*

Edgar Morin

Protocoles Infiniband vs. Ethernet

InfiniBand et Ethernet ont leurs propres caractéristiques et différences. Quelle est la meilleure ? Cela va surtout dépendre du contexte d'utilisation. Ces deux technologies continuent de se développer et d'évoluer dans différents domaines d'application. Nous allons voir dans cet article quelles sont leurs caractéristiques respectives.

Réseau InfiniBand et norme IBA

La différence entre l'InfiniBand et l'Ethernet tient avant tout dans leur conception. InfiniBand est une technologie d'interconnexion de réseau standard ouverte, largement utilisée dans les grappes de superordinateurs pour l'interconnexion de réseaux en raison de sa faible latence, de sa large bande passante, de sa grande fiabilité et de son support optimisé pour le calcul parallèle. C'est aussi la technologie d'interconnexion de réseau favorite pour les serveurs GPU et le HPC (High Performance Computing pour calcul haute performance). Le traitement de données à grande échelle et la communication fréquente entre les nœuds sont indispensables à ce dernier. La norme InfiniBand permet la transmission de signaux à débit unique (SDR) à un débit de base de 2,5 Gbits/s par voie. Cela permet d'atteindre un débit brut de 10 Gbits/s sur des câbles 4X. Un canal unique peut même être augmenté à 5 Gbits/sec ou 10 Gbits/sec. Le débit de données maximal potentiel pourra alors être de 40 Gbits/sec sur des câbles 4X, et de 120 Gbits/sec sur des câbles 12X. C'est ce qui va permettre aux réseaux InfiniBand d'avoir des signaux à double débit de données (DDR) et à quadruple débit de données (QDR). InfiniBand présente des avantages significatifs par rapport à Ethernet/Fibre Channel et à la technologie obsolète Omni-Path. Depuis 2014, la plupart des supercalculateurs TOP500 ont adopté cette technologie réseau. Les applications liées à l'IA et au

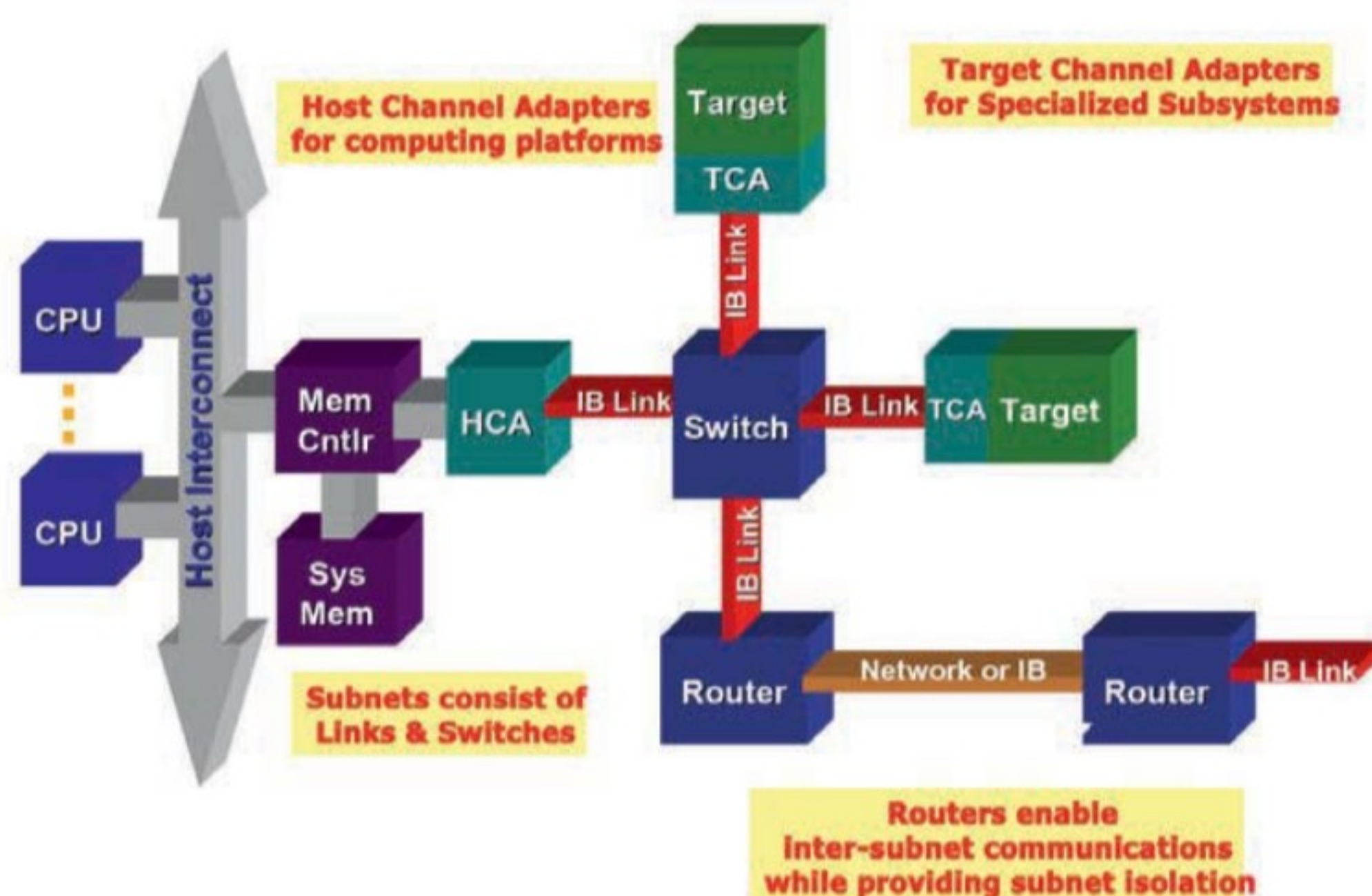
Big Data ont fait de même à grande échelle ces dernières années, dans le but de réaliser des déploiements de clusters hautes performances. IBA (InfiniBand Architecture) est une spécification standard de l'industrie qui définit un cadre d'entrée/sortie commuté point à point pour l'interconnexion des serveurs, des infrastructures de communication, des périphériques de stockage et des systèmes embarqués. La plus petite unité IBA complète est un sous-réseau, et plusieurs sous-réseaux sont connectés par des routeurs pour former un grand réseau IBA.

Réseau Ethernet

Le standard Ethernet est devenu le protocole de communication le plus utilisé dans les réseaux locaux depuis 1980. Contrairement à InfiniBand, il a été conçu avec comme leitmotiv de faire circuler l'information entre plusieurs systèmes le plus facilement possible. Ses autres objectifs sont une distribution efficace et une compatibilité maximale. L'Ethernet traditionnel s'appuie principalement sur le couplage TCP/IP pour mettre en place un réseau. Ethernet et la technologie IP constituent la pierre angulaire de l'édifice Internet mondial. Il est couramment employé dans les réseaux d'entreprise, l'accès à l'internet et les réseaux domestiques pour connecter plusieurs ordinateurs ou périphériques réseau tels que des imprimantes ou des scanners à un réseau local. Ses principaux avantages sont son faible coût, sa standardisation et son large support. Les principaux types d'Ethernet sont le Fast Ethernet, le Gigabit Ethernet, le 10 Gigabit Ethernet et le Switched Ethernet.

Différences entre InfiniBand et Ethernet

L'objectif initial de la conception d'InfiniBand est la résolution des goulots d'étranglement dans la transmission de données en grappes pour les scénarios de calcul à haute performance. InfiniBand et Ethernet présentent donc de très nombreuses différences, principalement en termes de bande passante, de latence, de fiabilité du réseau, de technologie de réseau et de scénarios d'application. L'un a été conçu, pour simplifier au maximum la mise en place de réseaux, alors que l'autre a pour principaux objectifs la fiabilité et les performances.



L'architecture InfiniBand est composée d'adaptateurs de canal, de commutateurs, de routeurs, de câbles et de connecteurs. Le CA est divisé en un adaptateur de canal hôte (HCA) et un adaptateur de canal cible (TCA). (source Fibermail.com)

Latence du réseau

InfiniBand et Ethernet se comportent également de manière très différente en ce qui concerne la latence du réseau. Les switches Ethernet utilisent généralement l'adressage par stockage et transfert et l'adressage par consultation de la table MAC en tant que technologies de couche 2 dans le modèle de transport du réseau. Le flux de traitement des switches Ethernet est plus long que celui des switches InfiniBand, car des services complexes tels que IP, MPLS ou QinQ doivent être pris en compte. Le traitement de la couche 2 est, en revanche, très simple pour les switches InfiniBand. Le LID (Local Identifier) de 16 bits est le seul à pouvoir être utilisé pour rechercher les informations relatives au chemin d'acheminement. La technologie Cut-Through est utilisée en parallèle afin de réduire de manière significative — à moins de 100 ns — le délai d'acheminement. C'est bien plus rapide qu'un switch Ethernet.

Fiabilité du réseau

La perte et la retransmission de paquets ayant un impact significatif sur les performances globales de l'HPC, un protocole de réseau hautement fiable est nécessaire pour garantir les caractéristiques de haute fiabilité sans perte. Avec ses propres formats définis de la couche 1 à la couche 4, comme TCP/IP, InfiniBand est un protocole de réseau complet. Le contrôle de flux de bout en bout est la base de l'envoi et de la réception des paquets du réseau InfiniBand. Comparativement à InfiniBand, le réseau Ethernet ne dispose pas d'un mécanisme de contrôle de flux basé sur l'ordonnement. Il est par conséquent impossible de garantir que l'extrémité du pair sera ou non encombrée lors de l'envoi de paquets. Pour pouvoir absorber l'augmentation soudaine du trafic dans le réseau, il est nécessaire d'ouvrir un espace de cache de plusieurs dizaines de Mo dans les switches pour

HSI

L'interconnexion à grande vitesse (HSI pour High Speed Interconnexion) est devenue la clef du développement d'ordinateurs hautes performances, car la puissance de calcul de l'unité centrale de traitement (CPU) augmente à un rythme très rapide. HSI est une nouvelle technologie proposée pour améliorer les performances de l'interface de composant périphérique (PCI). Après des années de développement, les HSI prenant en charge le calcul haute performance (HPC) sont désormais principalement Gigabit Ethernet et InfiniBand, ce dernier étant celui qui a la croissance la plus rapide. Il a été développé sous la supervision de l'IBTA (InfiniBand Trade Association) fondée en 1999 par la fusion de deux organisations industrielles, le Future I/O Developers Forum et le NGI/O Forum. L'IBTA travaille sous la direction d'un comité de planification et d'exploitation composé de HP, IBM, Intel, Mellanox, Oracle, QLogic, Dell, Bull et quelques autres. Elle est spécialisée dans les tests de conformité et d'interopérabilité des produits.

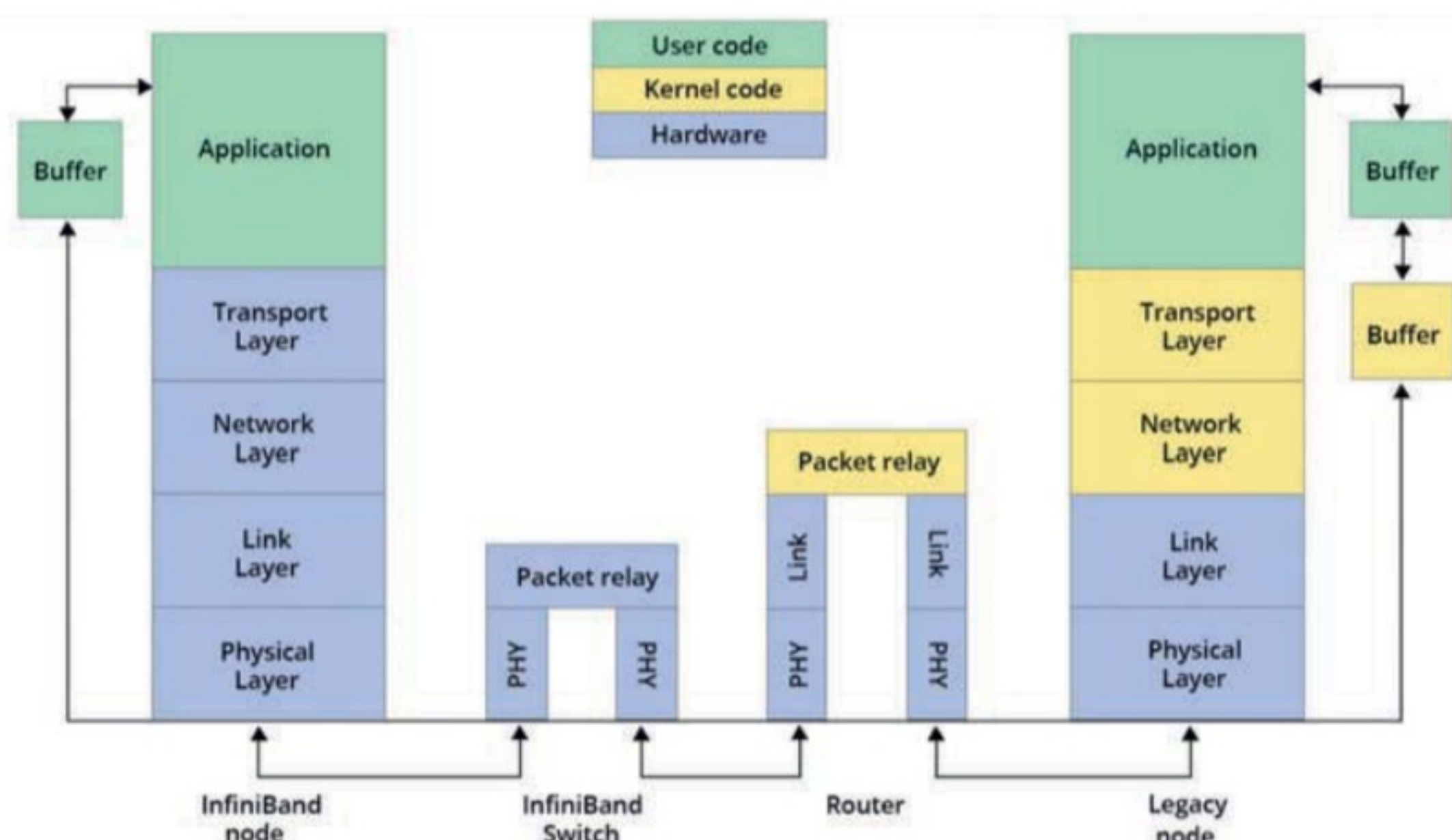
stocker temporairement ces messages qui, du coup, vont monopoliser les ressources de la puce. Cela signifie que la surface de la puce d'un switch Ethernet est nettement plus grande que celle d'une puce InfiniBand ayant les mêmes spécifications. Cela non seulement coûte plus cher mais en sus consomme plus d'énergie.

Méthodes de mise en réseau

En termes de mode de mise en réseau, le réseau InfiniBand est plus simple à gérer que le réseau Ethernet. L'idée de SDN (Software Defined Network) est intégrée dans InfiniBand par sa conception. Un gestionnaire de sous-réseau est présent sur chaque réseau InfiniBand de couche 2 pour configurer l'ID (LocalID) des nœuds du réseau, calculer uniformément les informations sur le chemin d'acheminement via le plan de contrôle et les transmettre à l'échange InfiniBand. Le mode réseau Ethernet permet de générer automatiquement des entrées MAC et l'IP

doit coopérer avec le protocole ARP. De plus, chaque serveur du réseau doit envoyer régulièrement des paquets pour garantir la mise à jour des entrées en temps réel. Pour diviser le réseau virtuel et limiter son ampleur, il faut donc mettre en œuvre des VLAN. Cependant, comme le réseau Ethernet lui-même ne dispose pas d'un mécanisme d'apprentissage des entrées, il en résultera un réseau en boucle. Pour éviter les boucles dans le chemin d'acheminement du réseau, des protocoles comme STP doivent être mis en œuvre, ce qui augmente la complexité de la configuration du réseau. □

T.T



InfiniBand utilise une pile de protocoles en cinq couches : physique, liaison, réseau, transport et application. Chacune d'entre elles joue un rôle essentiel pour assurer une communication efficace et fiable entre les dispositifs InfiniBand. (source : <https://community.fs.com/>)

Management

La data pro à l'étranger, simple et sans surcoût

Transatel simplifie la connectivité des entreprises à l'international grâce à sa solution Ubigi for Business. Compatible avec les terminaux eSIM ou SIM classiques, cette offre permet de gérer à distance des flottes de smartphones, PC ou tablettes. Elle garantit une connexion 4G/5G sécurisée dans plus de 200 destinations, tout en optimisant les coûts de roaming.

Avec l'essor du télétravail, des déplacements à l'international et de la digitalisation des processus métiers, la connectivité mobile devient un enjeu stratégique pour les entreprises. L'approche « zero touch » de Transatel facilite les déploiements massifs, même dans des contextes multisites ou internationaux. Filiale du groupe NTT et pionnier de l'eSIM, l'opérateur français répond à ce besoin avec Ubigi for Business : une solution pensée pour les collaborateurs en mobilité. Grâce à son infrastructure multi-opérateur et sa couverture mondiale, Ubigi for Business permet aux entreprises d'équiper leurs terminaux compatibles (smartphones, PC, tablettes) avec une connectivité 4G/5G fiable, chiffrée et instantanée, sans dépendre des réseaux Wi-Fi publics souvent peu sécurisés. L'offre se distingue également par sa souplesse de déploiement : elle est aussi bien adaptée aux grandes entreprises qu'aux structures plus petites ne disposant pas de service informatique dédié. Le mode de facturation à l'usage ou par volume, combiné à une visibilité complète sur les consommations, facilite la maîtrise des coûts tout en évitant les mauvaises surprises.



Installation et prise en main

L'activation du service s'effectue en quelques minutes via un QR code eSIM transmis par Ubigi sur le portail de gestion dédié ou par email. Après avoir scanné le code avec un terminal compatible eSIM, celui-ci est automatiquement provisionné avec un profil Ubigi. Les responsables IT peuvent ensuite piloter l'ensemble de leur flotte à travers une interface web centralisée, qui permet de visualiser la consommation, de suspendre un profil ou de définir des alertes personnalisées. Transatel propose également des cartes SIM physiques pour les équipements non compatibles eSIM, afin d'assurer une plus grande flexibilité dans le choix des appareils. Nul besoin de compétences techniques poussées : les profils eSIM peuvent être activés à distance sans intervention humaine, ce qui facilite le déploiement à grande échelle, notamment pour les collaborateurs nomades ou les filiales internationales. Cette simplicité d'usage constitue l'un des points forts

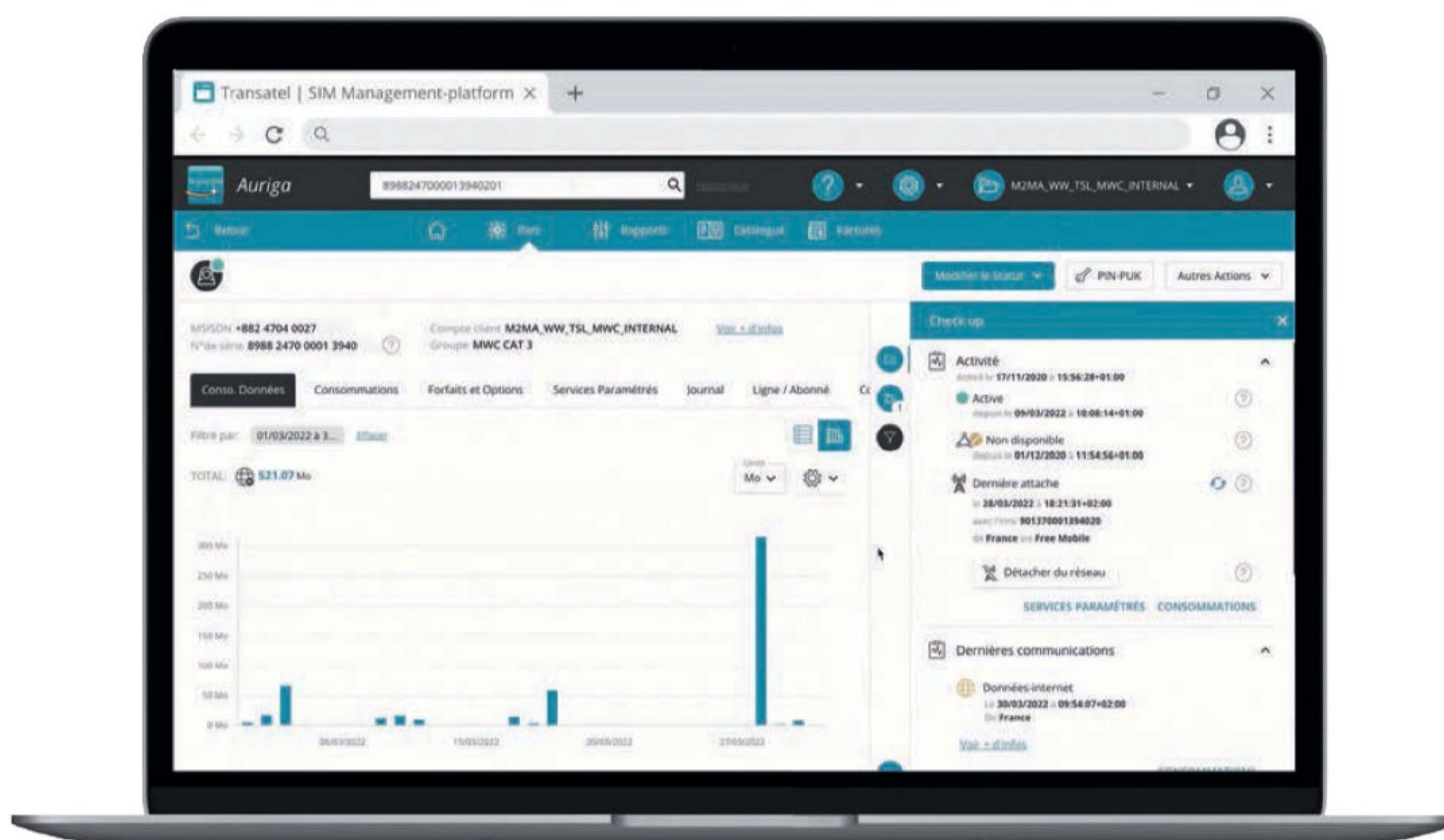
du service de Transatel. Contrairement à certains services concurrents, Ubigi for Business permet en outre de partager la connexion mobile avec d'autres appareils sans aucune limite. Il est également possible de choisir son adresse IP (à l'étranger ou dans son pays de résidence) pour accéder sans problème aux services géorestricts comme les banques, les chaînes de télévision, etc.

CARACTÉRISTIQUES TECHNIQUES UBIGI FOR BUSINESS :

- Console de gestion web pour les entreprises
- Gestion centralisée des flottes de SIM et eSIM
- Suivi en temps réel de la consommation data
- Alertes personnalisables, rapports détaillés
 - Sécurité renforcée (chiffrement, RGPD)
- Application mobile dédiée (Android et iOS)
 - Compatible Windows 11 et MacOS
- Tarifs par zone géographique et sur devis selon volume

Une gestion unifiée et compétitive

La console Ubigi for Business offre une interface claire permettant de superviser des centaines de lignes en quelques clics. Chaque entreprise peut définir des règles d'usage, suivre la consommation de data en temps réel, attribuer des profils de connectivité par région ou par utilisateur, et recevoir des rapports détaillés. En éliminant les contraintes logistiques des cartes physiques, cette approche renforce la continuité de service pour



La console d'administration Ubigi offre aux responsables IT une gestion simplifiée et centralisée de leurs flottes de terminaux.

les collaborateurs itinérants. Selon Transatel, certaines entreprises économisent jusqu'à 90 % sur leurs factures mobiles hors UE. La solution est destinée en particulier aux entreprises dont les collaborateurs voyagent fréquemment hors d'Europe, où les forfaits classiques deviennent rapidement prohibitifs. Pour les organisations multisites, la mutualisation de la gestion des lignes via une interface unique constitue un avantage non négligeable. L'offre s'adresse également aux travailleurs indépendants qui peuvent gérer leur compte directement via l'application mobile du service. Avec Ubigi, les données sont acheminées via des réseaux partenaires dans chaque pays, ce qui permet à Transatel de proposer des tarifs très compétitifs tout en maintenant une qualité de service élevée. Le test en Amérique centrale a permis de constater que l'offre tarifaire de Transatel dans cette région était parmi les plus compétitives avec celle d'Holafly.

Sécurité, conformité et confidentialité

Le service repose sur une architecture sécurisée répondant aux exigences des DSI. La connexion mobile chiffrée permet aux collaborateurs d'accéder à distance aux outils métiers sans compromettre la confidentialité des données. Contrairement aux réseaux Wi-Fi publics, la solution évite les risques liés aux interceptions ou aux accès non autorisés. Transatel garantit en outre un respect strict du RGPD, avec des données hébergées en Europe et une gouvernance claire. Des fonctionnalités avancées de contrôle permettent également de restreindre l'accès à certaines ressources selon le profil ou la géolocalisation du terminal. Ubigi for Business s'intègre également dans une logique SASE, où la connectivité et la sécurité convergent pour répondre aux nouveaux paradigmes du travail hybride.

ESIM : UNE ÉVOLUTION À LA FOIS PRATIQUE ET FLEXIBLE

Dématérialisée, l'eSIM s'active en quelques instants via les paramètres du smartphone, du PC ou de la tablette, sans carte physique, manipulation matérielle ni intervention technique. Les eSIM sont particulièrement adaptées aux appareils compatibles double SIM : elles permettent de conserver sa ligne principale tout en ajoutant une ligne locale, dédiée à l'accès Internet (uniquement avec Ubigi) dans un pays spécifique. Plusieurs profils peuvent être stockés et activés selon les besoins, pays par pays. Finis les risques de perte ou de casse : la gestion est centralisée, rapide et sécurisée. De plus en plus de PC professionnels intègrent cette technologie, ce qui devrait contribuer au succès croissant des eSIM.

Transatel, qui revendique 1,5 million d'utilisateurs Ubigi dans le monde, mise sur la montée en puissance des appareils eSIM natifs, dont les PC portables 5G-ready. Avec Ubigi for Business, Transatel répond à un besoin stratégique : fournir une connectivité mobile mondiale, simple à déployer, sécurisée et économique. La solution se distingue par sa souplesse, sa couverture étendue et son pilotage centralisé. Cette vision unifiée de la connectivité mobile s'inscrit pleinement dans les stratégies IT modernes, centrées sur la flexibilité, la sécurité et la continuité de service. Face aux forfaits de roaming classiques souvent limités ou onéreux hors Europe, Ubigi for Business s'impose comme une alternative stratégique, alliant souplesse, sécurité et performance. □

J.C

Adobe Summit 2025

Une nouvelle ère pour l'expérience client avec l'agentique

Au cours de l'Adobe Summit 2025, Adobe a placé l'IA agentique au cœur de sa stratégie, dévoilant une nouvelle génération d'agents intelligents destinés à transformer l'expérience client.

De la personnalisation à grande échelle à l'optimisation des parcours, l'entreprise a présenté une orchestration unifiée mêlant créativité, marketing et technologie, portée par sa plateforme Adobe Experience Platform et ses nouveaux outils comme Agent Orchestrator, Brand Concierge ou GenStudio.

L'éditeur a ainsi dévoilé des agents redéfinissant les contours de l'expérience client à l'ère de la personnalisation à grande échelle. En lançant Adobe Experience Platform Agent Orchestrator, l'entreprise californienne entend transformer la manière dont les marques interagissent avec leurs clients, conçoivent leurs campagnes et optimisent leurs parcours, tout en fusionnant marketing, créativité et technologies d'IA.

Au cœur de cette nouvelle dynamique, Adobe Experience Platform (AEP) sert de socle à l'orchestration intelligente des expériences. Déjà utilisée pour activer plus de 1 000 milliards d'expériences chaque année, cette plateforme marketing unifiée connecte en temps réel les données de l'ensemble d'une entreprise pour générer des connaissances actionnables. La nouveauté phare, Agent Orchestrator, vient enrichir AEP en permettant de concevoir, gérer et coordonner des agents IA spécialisés, issus à la fois des environnements Adobe et de partenaires tiers.

Répondre à des objectifs concrets

Ces agents autonomes sont conçus pour répondre à des objectifs métier concrets : accompagnement client, optimisation de contenu, automatisation des flux de travail, gestion de la donnée ou encore planification stratégique. Adobe fait ainsi passer la gestion de l'expérience client (CXM) à un niveau supérieur, celui de la Customer Experience Orchestration (CXO). Une approche fondée sur l'agilité, la collaboration homme-machine et une compréhension sémantique profonde des données, du contenu et des parcours clients. « Adobe est parfaitement positionné pour accompagner les entreprises à l'ère de l'orchestration de l'expérience client,

une ère où la créativité et le marketing s'associent à l'IA pour réaliser la personnalisation à grande échelle », a déclaré Anil Chakravarthy, président de la division Digital Experience d'Adobe. Et de souligner que cette évolution n'est pas seulement technologique. Elle transforme les méthodes de travail et les rôles des équipes.

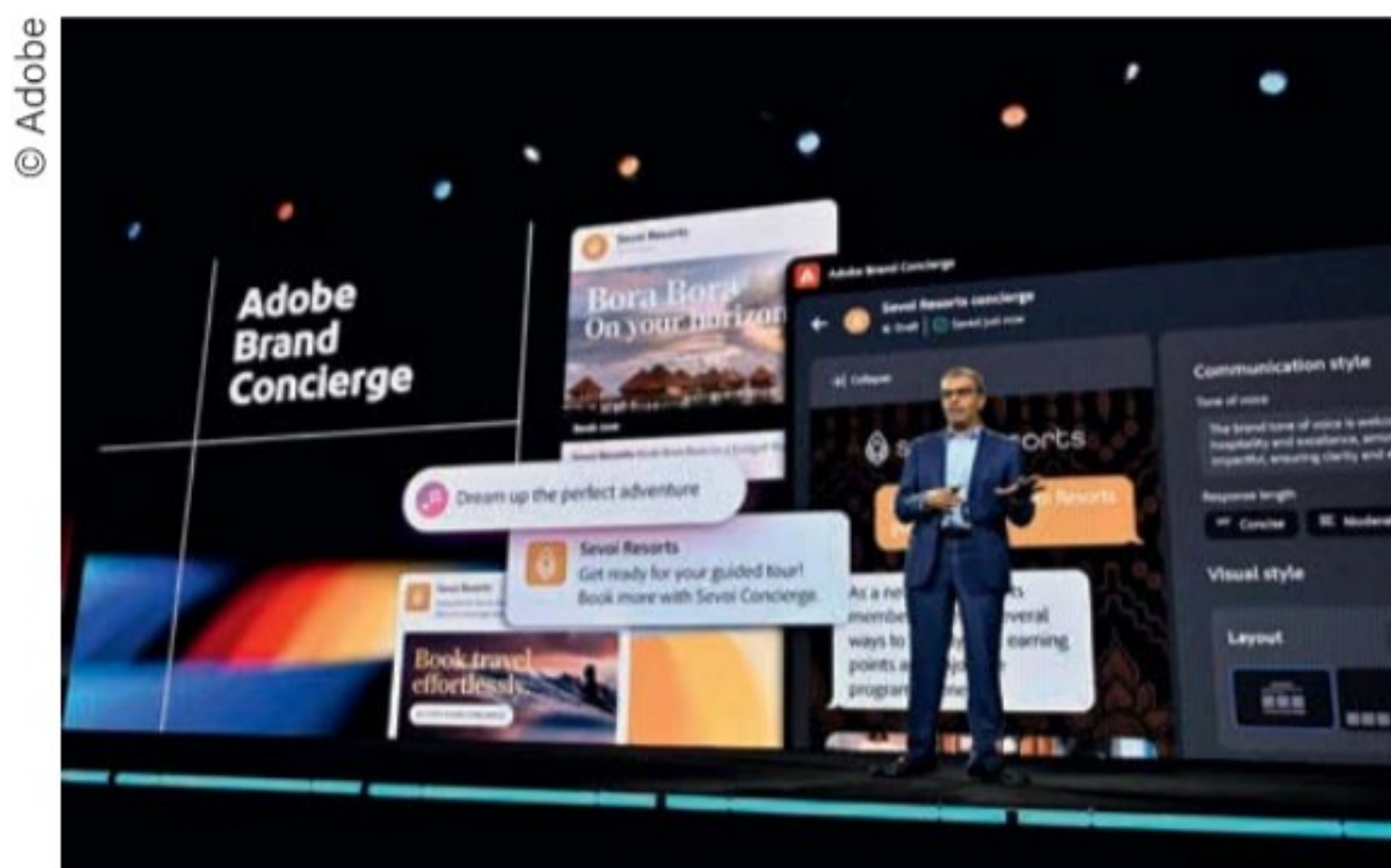
Une suite d'agents IA à mission spécifique

Ainsi, Adobe a présenté dix agents IA prêts à l'emploi, conçus pour renforcer les capacités des utilisateurs d'applications Adobe comme Adobe Real-Time Customer Data Platform, Adobe Journey Optimizer, Adobe Experience Manager ou Adobe Customer Journey Analytics. Ces agents s'appuient sur les données unifiées d'AEP pour interpréter des informations complexes et générer des actions adaptées à chaque situation. Parmi eux, l'agent Audience analyse les données d'engagement multi-canal pour créer des segments de clientèle à fort potentiel, activables dans des campagnes de personnalisation. L'agent Content Production assiste les équipes marketing et créatives en générant du contenu à partir d'un brief, tout en respectant les lignes directrices de la marque. De son côté, l'agent Experimentation permet de simuler des hypothèses, d'analyser leur impact et de lancer des tests à grande échelle.



Shantanu Narayen, CEO d'Adobe, à Las Vegas

D'autres agents se concentrent sur la qualité et l'efficacité opérationnelle, comme Site Optimization, qui surveille en temps réel la performance des sites web et propose des correctifs immédiats, ou encore Workflow Optimization, qui suit l'état des projets, facilite les validations et optimise les délais. Pour les entreprises B2B, l'agent Account Qualification identifie les décideurs au sein des groupes d'achat et pilote des actions ciblées. Anjul Bhambhri, SVP de l'ingénierie d'Adobe Experience Cloud, souligne l'impact stratégique de cette nouvelle génération d'agents. « L'IA agentique est une avancée majeure qui va accélérer la transformation des entreprises. Les gains de productivité induits par nos innovations permettront aux utilisateurs de se concentrer sur l'idéation et la stratégie de marque. »



Anil Chakravarthy, président de la division Digital Experience d'Adobe, a présenté Adobe Brand Concierge, une application conversationnelle immersive représentant une évolution des chatbots.

Brand Concierge : le visage humain de l'IA

Mais c'est sans doute Adobe Brand Concierge qui incarne le plus clairement la vision d'une IA au service de l'expérience client. Cette application conversationnelle immersive représente une évolution des chatbots traditionnels vers des agents capables de comprendre le contexte, de s'adapter aux préférences de l'utilisateur et de guider l'interaction de bout en bout. Alimentée par AEP Agent Orchestrator, elle propose des échanges multimodaux (texte, voix, image) et s'intègre aux canaux numériques des marques pour offrir une expérience fluide, personnalisée et conforme. Brand Concierge peut ainsi accompagner un client tout au long de son parcours d'achat, de la découverte d'un produit à la prise de décision, en passant par des comparatifs personnalisés ou des recommandations issues de l'historique client. En B2B, il peut proposer du contenu adapté à la relation existante avec un compte, planifier des rendez-vous ou orienter les décisions d'achat en temps réel.

Egalement consciente que l'avenir de l'IA repose sur la collaboration entre systèmes, la société Adobe a également insisté sur l'importance de l'interopérabilité entre agents issus de différents écosystèmes. Pour cela, la société a renforcé son réseau de partenaires technologiques, incluant AWS, Microsoft, SAP, IBM, Workday ou encore Genesys. Ces alliances visent à garantir l'intégration fluide des agents IA dans les flux de travail liés au service client, à la gestion des ressources humaines, à la planification d'entreprise ou à la collaboration. Adobe a aussi élargi ses partenariats avec les grands intégrateurs de systèmes comme Accenture, Deloitte Digital, EY et IBM, afin de personnaliser les cas d'usage et d'accompagner les clients dans leur transformation.

Repenser la production de contenu à l'ère de l'IA

Autre axe stratégique : la transformation de la chaîne d'approvisionnement de contenu. Avec Adobe GenStudio, Adobe propose une plateforme unifiée qui couvre tout le cycle de

vie du contenu, de la planification à la diffusion en passant par la création et la mesure. La nouvelle interface GenStudio Foundation centralise les données issues des différentes applications Adobe, offrant aux équipes marketing et créatives une visibilité complète sur les campagnes, les projets, les assets et les connaissances.

Grâce à l'intégration de l'IA générative, les utilisateurs peuvent produire des centaines de variations de contenu en quelques clics, optimiser les formats pour chaque canal (réseaux sociaux, e-mails, display...) et automatiser des tâches comme le redimensionnement ou la génération de visuels à partir d'un brief. Les intégrations avec Google Campaign Manager 360, Microsoft Advertising et LinkedIn Ads viennent enrichir les possibilités de diffusion, notamment pour les campagnes B2B. « *La chaîne d'approvisionnement de contenu est souvent un maillage complexe de flux de travail et d'outils hétérogènes. Avec GenStudio, nous proposons une solution unique qui unifie ces processus et permet aux équipes de se concentrer sur ce qui compte : la créativité et la performance* », explique Amit Ahuja, SVP digital experience business d'Adobe.

Des outils pour unifier les parcours clients

Enfin, Adobe a annoncé plusieurs nouveautés pour accompagner les équipes dans la création d'expériences cohérentes et efficaces sur tous les points de contact. Le module Adobe Journey Optimizer Experimentation Accelerator permet aux responsables marketing de tester différentes expériences, d'analyser leurs performances et d'identifier les stratégies les plus efficaces grâce à un agent IA spécialisé. Ce dernier s'appuie sur les données unifiées d'AEP pour produire des recommandations fondées sur les probabilités de succès.

De son côté, Adobe Experience Manager Sites Optimizer aide les marques à améliorer l'acquisition de trafic et l'engagement sur leurs sites web. L'agent Site Optimization y surveille les performances en temps réel, identifie les problèmes techniques et propose des corrections.

Une offre B2B boostée par l'IA

Enfin, l'intelligence agentique s'invite aussi dans les outils dédiés aux entreprises B2B. Avec Adobe Journey Optimizer B2B Edition et les évolutions de Customer Journey Analytics, Adobe propose des solutions adaptées à des cycles de vente longs et complexes. Les agents IA facilitent l'identification des bons interlocuteurs au sein des groupes d'achat, attribuent les rôles, détectent les profils manquants et suggèrent des parcours personnalisés selon les centres d'intérêt. Ils peuvent également automatiser le suivi, synthétiser l'engagement ou déclencher des alertes en cas de pics d'activité.

Bref, l'édition 2025 d'Adobe Summit a clairement placé l'IA agentique au cœur de ses produits. L'entreprise a présenté des outils pour l'ensemble des processus liés à l'expérience client, dans une logique d'orchestration intelligente et d'interaction augmentée. □

M.C

Futur du travail

Workday infuse l'IA partout

L'éditeur spécialisé dans le domaine des ressources humaines et de la finance a convié quelques journalistes à la présentation de ses nouveautés dans son quartier général de Dublin en Irlande. Au final, l'éditeur prend le train de l'IA et l'infuse partout dans sa plateforme.

Dans les 350 nouvelles fonctions ajoutées dans la dernière version de la plateforme, le point central est le rôle prépondérant que va jouer ou joue déjà Illuminate. Juste au-dessus du cœur technologique de la plateforme (modèle de données, processus métiers, sécurité) arrive Illuminate qui se compose d'un moteur intelligent de données, des LLM de Workday et d'une passerelle pour les entreprises ayant choisi d'autres modèles. La technologie est à la base de Workday Assistant, un agent conversationnel personnalisable automatisable et capable de prendre des décisions seul selon les données contextuelles fournies. A l'avenir, ces assistants devraient se voir doter de capacités de raisonnement pour devenir proactifs et autonomes.

Du nouveau dans les ressources humaines

Les dernières mises à jour concernent la gestion de talents et simplifient les processus de recrutement. Les recruteurs et les équipes RH peuvent identifier des profils qualifiés au sein des viviers de talents existants — collaborateurs, anciens candidats ou talents déjà connus — directement dans Workday Recruiting. Grâce à l'intégration du sourcing et du screening pilotés par l'IA au sein d'une seule et même solution, le recrutement devient plus rapide, moins coûteux et plus fluide pour les candidats. La solution offre, de plus, des parcours de pré-intégration et d'embauche personnalisée pour accompagner et impliquer les collaborateurs dès

leurs premiers instants dans l'entreprise. Des offres de mutations internes s'appuient sur l'intelligence artificielle pour proposer aux collaborateurs des postes correspondant à leur expérience, leurs compétences, leurs centres d'intérêt et leurs aspirations professionnelles, tout en respectant les règles internes de mobilité. Intégrée à Career Hub et à l'Opportunity Marketplace, cela donne la capacité à l'ensemble des parties prenantes — salariés, managers, équipes RH et recrutement — d'accéder aux mêmes opportunités, renforçant ainsi la rétention des talents, la satisfaction au travail et le développement des compétences.

Créer des applications avec l'IA

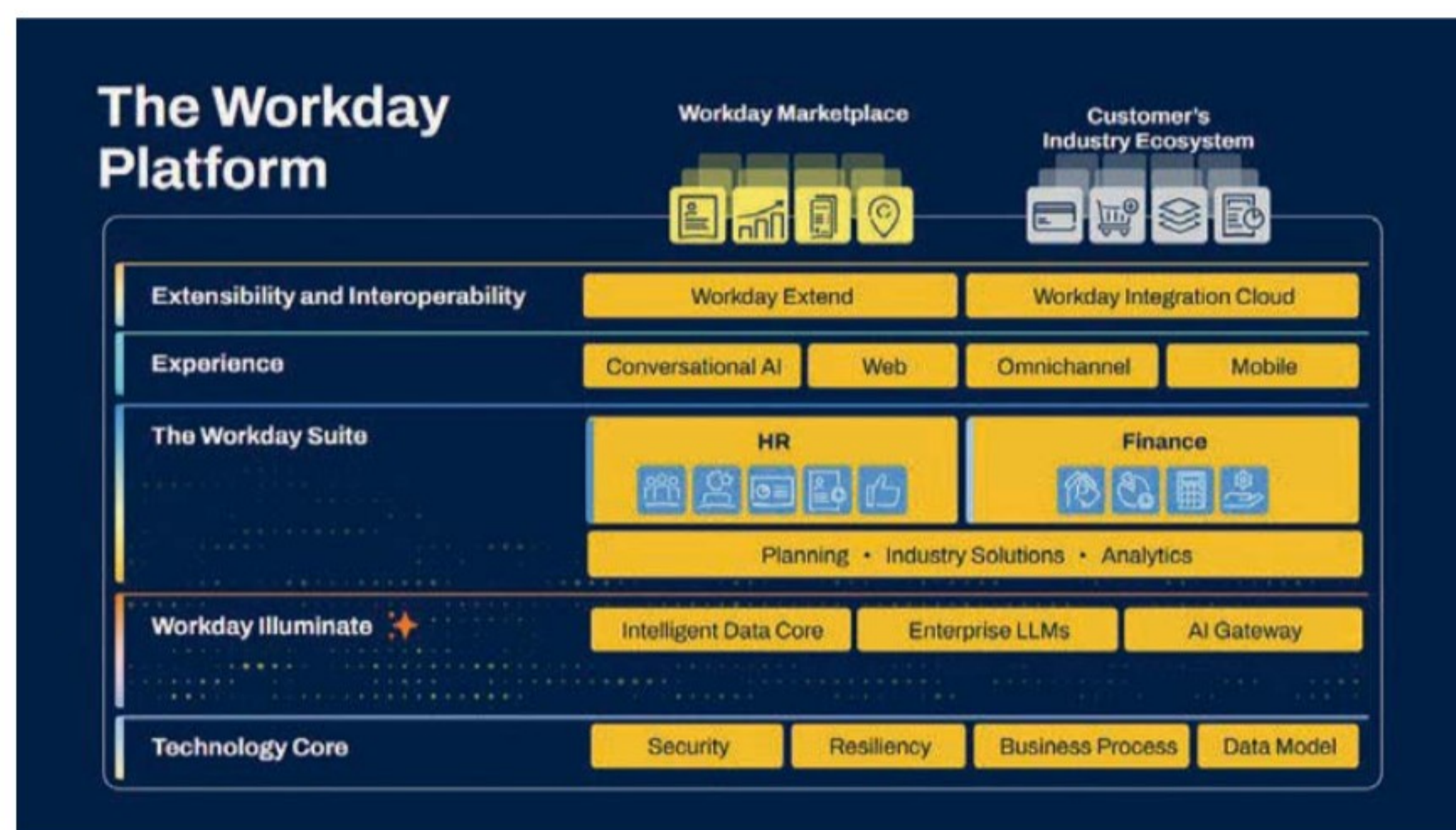
Les services d'IA désormais disponibles via Workday AI Gateway renforcent les moyens des développeurs, afin qu'ils créent des applications intelligentes à l'aide des API Workday AI, en s'appuyant sur Workday Extend Professional. Une fonction Data Query exploite l'IA pour simplifier les requêtes de données dans Workday, en offrant la capacité aux utilisateurs d'effectuer des recherches en langage naturel. Document Intelligence permet aux utilisateurs d'extraire automatiquement les informations clés de documents tels que les contrats fournisseurs, accords clients et bulletins de paie.

Les autres nouveautés visent à simplifier le travail en première ligne avec des solutions de pointage hors ligne ou pour

la saisie des temps de travail, et de simplifier les processus métiers pour améliorer l'expérience utilisateur Workday sur des processus complexes et chronophages comme la recherche de documents administratifs dans leur dossier RH.

De nouvelles solutions pour la finance sont désormais disponibles pour aider les équipes à prendre des décisions plus rapides en accélérant des tâches telles que les devis et la facturation. □

B.G



La plateforme Workday et la place d'Illuminate dans celle-ci

RISE

Les nouvelles offres de SAP à la moulinette

L'association des utilisateurs de l'ERP a récemment publié un guide sur l'offre Rise & Grow with SAP de l'éditeur éponyme. Le but est en particulier d'évaluer tous les impacts potentiels liés à une migration vers ces offres.

« Cette offre nouvelle et assez complexe n'est pas toujours bien comprise par les entreprises. Nous avons d'abord voulu décrypter les concepts utilisés », justifie Bernard Cottinaud, chargé de mission stratégie de l'USF. Un éclairage d'autant plus important que basculer d'une solution en place, en particulier sur site, vers ces nouvelles offres en souscription « est un peu un voyage sans retour », ajoute-t-il. Présidente de la commission Rise à l'USF, Sadjia Dahmane résume : « Cette évolution de l'offre soulève des questions stratégiques et contractuelles essentielles. » Outre cet éclairage, le guide recense une série de questions à se poser avant d'opter ou pas pour ces solutions, sur la contractualisation et enfin, sur les implications de ce choix une fois en production. « Chaque organisation est différente : sur le plan business et réglementaire, sur l'existant, sur les principes de fonctionnement, sur la stratégie... Les DSI peuvent ainsi sélectionner les questions qui les concernent dans leur propre contexte », détaille Bernard Cottinaud. La migration vers les offres Rise ou Grow demande donc une analyse encore plus poussée qu'habituellement. Une analyse qui peut se dérouler sans urgence, le support sur S/4HANA est garanti jusqu'en 2040. Un point confirmé par Christian Charvin, directeur avant-vente de la filiale française. Le but de l'USF est d'aider les organisations à évaluer tous les impacts consécutifs à la mise en œuvre de ces offres cloud en termes de risques économiques, métiers, aussi bien que sur les centres de compétences SAP internes.

Premier aspect, essentiel, les offres Rise ou Grow ne sont pas des solutions techniques, mais « des modes de contractualisation de solutions techniques exclusivement cloud, assorties de services gérés par SAP », explique Bernard Cottinaud. Des contrats qui se passent exclusivement avec ce fournisseur.

Devenant l'interlocuteur unique, ce dernier prend en charge les contrats avec les autres fournisseurs, hyperscalers et autres. Ce qui lui permet d'obtenir des tarifs avantageux avec ces derniers. Evolution majeure, le métier historique de l'éditeur se transforme en fournisseur de services. « Une transformation profonde, reconnaît Christian Charvin, qui a l'avantage de nous impliquer davantage dans une relation à long terme avec nos clients. » Concrètement, pour Rise, plutôt destinée aux clients existants, comme pour Grow,



Christian Charvin, directeur avant-vente SAP France



Bernard Cottinaud, chargé de mission stratégie de l'USF

pour les nouveaux, les deux possibilités cloud public ou privé sont proposées. Le client a le choix entre les hyperscalers et quelques fournisseurs retenus par SAP (T-Systems...). Aucune solution de cloud souverain n'est disponible en France. « Bleu devrait répondre à ce besoin en 2026 », avance Christian Charvin. L'éditeur n'envisage pas, à ce jour, d'inclure OVHcloud, qui a obtenu le label SecNumCloud depuis peu.

Troisième possibilité, « il est également possible avec Rise d'héberger les solutions S/4HANA dans les locaux du client », ajoute Christian Charvin. Comme pour les autres options, SAP opérera complètement ces solutions. Cette dernière alternative ne met pas à l'abri du Cloud Act, « nous ne garantissons pas, au moins pour l'instant, la nationalité et la localisation des ressources qui interviennent », souligne Christian Charvin. Cette évolution amène naturellement à se poser une question sensible, à savoir quelles seront les ressources disponibles pour assurer ces services. « Des départements entiers sont créés dans ce but. En France, un nouveau département de 30 personnes est chargé des services, répond Christian Charvin. Et la disponibilité garantie est de 99,7%. Constatée, elle monte aujourd'hui à 99,998 %. ».

Attention aux mises à jour

Outre toutes les questions habituelles, les mises à jour deviennent avec ces nouveaux contrats un point sensible. Comme pour les offres de type SaaS, opter pour Rise ou Grow implique de creuser cette

question. Pour Grow, les mises à jour sont semestrielles et obligatoires à chaque fois. Elles sont effectuées par SAP. Le calendrier est fourni à l'avance. Pour Rise, les clients ont la possibilité d'attendre quatre ans au maximum. « *Nous préconisons deux ans* », précise Christian Charvin. Pour l'USF, cette contrainte est étudiée dans le détail. « *Les centres de compétences internes des entreprises ont entre autres pour rôle de recetter ces mises à jour, afin de limiter les régressions en production. SAP propose d'ailleurs plus de 800 services, BDD, transaction BDS, services applicatifs... Cette fréquence imposée va se traduire chez une partie des clients par des cycles de test et de validation bien plus fréquents pour lesquels il faut s'organiser* », souligne Bernard Cottinaud. Ainsi, avec Grow, les clients ne seront pas maître de la date d'arrêt de leur production pour mettre à jour leurs systèmes d'information, le planning étant défini par SAP. Autre risque, « *ce mode de contractualisation avec un interlocuteur unique risque de rendre la recherche de responsabilité plus difficile en cas de problème* », ajoute Bernard Cottinaud. Dans un registre commercial, SAP peut potentiellement repackager ses offres de services, quitte à en inclure une partie dans des offres, mêmes si celles-ci ne sont pas demandées par une partie des clients.

IA et données

Une autre question récurrente porte sur l'utilisation des données. Si elles demeurent la propriété du client, avec leur accord, elles sont utilisées par SAP pour entraîner les algorithmes. « *Une fois cryptées et anonymisées, 25 000 clients ont donné leur accord pour ces usages*, souligne Christian Charvin. *Les clients sont demandeurs de référentiels par domaine d'activité (taux de rotation moyen, BFR...), de données sur le Value Life Cycle Management.* » De son côté, l'USF pointe des risques potentiels inhérents à ces usages. « *Même si les données des clients ne sont pas directement identifiées, qu'est-ce qui pourrait empêcher SAP, à l'avenir, de les utiliser de manière agrégée comme le ferait*

Sadjia Dahmane,
présidente de
la commission
Rise à l'USF



un cabinet d'analystes pour publier, par exemple, des tendances conjoncturelles ou même des notations sectorielles ? En tout cas, c'est une question à se poser », illustre Bernard Cottinaud. Dans le registre de la sécurité, ces contrats nécessitent également une attention particulière. A l'instar de toute application SaaS, « *repartir sur un système sain (PRA) repose entièrement sur le fournisseur puisque la maîtrise de son ERP a été déléguée. La question est donc à creuser* », souligne Sadjia Dahmane. La réglementation récente CSRD impose entre autres de prendre en compte le bilan carbone lié à ses sous-traitants. Ce qui implique donc d'obtenir ces chiffres auprès de SAP. « *Quelles sont les garanties apportées sur la disponibilité de ces données et sur leur réalité, demande Bernard Cottinaud. Est-ce, par exemple, la moyenne de kg de CO₂ émis par octet incluant les data centers localisés aux États-Unis, en Europe ou ailleurs (par rapport à la France où l'électricité est beaucoup plus décarbonée) ?* » Une liste de questions à se poser pas exhaustive. Au final, le constat s'impose. Si le choix de SAP en tant qu'ERP était déjà stratégique, opter ou non pour Rise ou Grow l'est encore davantage. □

P.Br

PREMIÈRE VALORISATION EUROPÉENNE

En mars dernier, SAP est devenu la première valeur boursière au DAX. Une première place liée sans nul doute à ses activités cloud en forte croissance. Elles ont augmenté de 25 % sur une année. Il reste légitime de se demander si l'éditeur allemand va conserver cette place dans la durée. Sa politique tarifaire, notamment le passage vers le mode souscription comme la complexité de son mode de licence, n'ont pas enthousiasmé tous les clients. Si l'on en croit plusieurs cabinets d'études, la migration vers S/4HANA reste souvent plus compliquée qu'attendu. Selon le cabinet d'études allemand Horvath*, 60 % des entreprises connaissent des écarts de budget importants, des retards et des problèmes de qualité lors de leur migration vers S/4HANA. Dans un autre rapport récent, le Gartner avance que 60 % des clients sont toujours sur ECC et qu'en 2030, date de la fin de son support, 40 % devraient le rester.

* <https://www.horvath-partners.com/en/press/detail/study-shows-sap-s-4hana-transformations-rarely-go-as-planned-60-percent-exceed-budget-and-schedule-two-thirds-dissatisfied-with-result-quality>

Google Cloud Next 2025

Google Cloud mise sur une IA raisonnée, interopérable et souveraine

Google a dévoilé sa nouvelle ambition : faire de l'IA, via Gemini 2.5, une interface centrale des systèmes d'information. Portée par des agents intelligents développés avec Vertex AI, l'architecture permet de piloter métiers, données et infrastructures grâce à une orchestration sans code (Workspace Flows) et un accès sécurisé aux données internes (Agentspace). Le tout est désormais déployable en cloud public ou sur site via Google Distributed Cloud, répondant aux exigences des environnements les plus régulés.

Avec la nouvelle version Gemini 2.5, Workspace Flows, Vertex AI, Agentspace et la montée en puissance de Google Distributed Cloud, Google Cloud Next 2025 a montré une évolution stratégique majeure de Google : faire de l'IA un système transversal, capable de piloter infrastructures, outils métiers, applications analytiques et environnement de travail à partir d'agents spécialisés, raisonnants, interopérables et déployables partout. Ainsi, durant son keynote introductif, Sundar Pichai, directeur général de Google et Alphabet, a présenté une feuille de route technologique articulée autour d'un objectif assumé : faire de Gemini, non seulement une plateforme de modélisation linguistique, mais aussi l'interface cognitive des systèmes d'information modernes. « Sur les 400 clients français qui sont là, je n'ai pas pu debriefer avec tout le monde. Mais en tout cas, c'est une réaction hyper positive, à la fois sur les innovations technologiques, les nouveaux processeurs, les annonces qui ont été faites », a constaté Isabelle Fraine, la directrice de Google Cloud France.

« Un modèle qui pense avant de parler »

Lorsqu'il a évoqué Gemini 2.5, Sundar Pichai a parlé d'un « modèle qui pense avant de parler ». Cette nuance résume bien la direction stratégique prise par Google : l'intelligence n'est plus dans la génération de texte, mais dans la capacité à raisonner, à structurer, à interpréter un contexte opérationnel complexe. Et c'est précisément cette compétence cognitive que l'entreprise veut mettre à disposition des décideurs et des équipes techniques, sous la forme d'agents logiciels modulaires, spécialisés, gouvernés et orchestrés à l'échelle de l'entreprise.



À Las Vegas durant Google Cloud Next, Thomas Kurian, directeur général de Google Cloud, a présenté une feuille de route technologique.

Mais Google ne s'est pas arrêté à une promesse de raisonnement. Ce que le géant de Mountain View a présenté à Cloud Next 2025, c'est une architecture distribuée de l'IA, capable de s'exécuter dans le cloud public, sur site ou dans des systèmes hybrides où les contraintes réglementaires sont très importantes. L'agent, dans cette vision, devient le nouveau composant applicatif fondamental.

Déploiement et développement d'agents

Le cœur technique de cette approche repose sur une transformation de Vertex AI, qui n'est plus un simple guichet d'accès à des modèles, mais une plateforme de développement et de déploiement d'agents intelligents. Chaque agent, développé avec l'Agent Development Kit, peut s'intégrer à une chaîne métier, réagir à un événement, manipuler des documents, consulter des bases internes, et surtout dialoguer avec d'autres agents. Il devient, de fait, un acteur opérationnel autonome. « Nous avons conçu Vertex AI comme un environnement où l'on ne code pas un modèle,



Pour cette édition 2025, Google Cloud Next a réuni quelques 32 000 personnes à Las Vegas.

mais une stratégie cognitive», a expliqué Saurabh Tiwary, vice-président de Cloud AI. «Chaque agent n'est pas un prompt amélioré, c'est un composant raisonné, versionné, testable, interopérable et traçable.»

une capacité d'interrogation fluide, multimodale, sécurisée et contextualisée sur les données internes de l'entreprise. Contrairement aux moteurs classiques, Agentspace ne se contente pas d'indexer. Il interprète des silos entiers

Pour que ces agents ne soient pas isolés, Google a introduit le protocole Agent2Agent, conçu comme une grammaire d'interaction pour des agents construits sur des bases techniques hétérogènes. L'enjeu n'est pas uniquement technique. C'est celui de la souveraineté cognitive : garder le contrôle sur qui raisonne, avec quelles données, dans quelles conditions et avec quels droits d'accès.

Gestion de formats multiples avec Agentspace

Toute architecture cognitive repose sur une capacité à accéder à la bonne information au bon moment. Avec Agentspace, Google a conçu une plateforme pour donner aux agents

GOOGLE POSE LES BASES D'UNE INFRASTRUCTURE IA DE NOUVELLE GÉNÉRATION

Google investit massivement dans ses data centers et lance Ironwood, sa puce d'inférence surpuissante. Objectif : soutenir l'essor de l'IA générative. Le groupe avance aussi en informatique quantique avec Willow, capable de corriger les erreurs à grande échelle.

Alphabet a franchi un nouveau cap dans sa course à l'IA. Lors de Google Cloud Next, Sundar Pichai, directeur général de Google et Alphabet, a confirmé que le groupe investirait cette année 75 milliards de dollars pour développer ses capacités en centres de données. Cet investissement servira à renforcer les fondations technologiques nécessaires à l'intégration de l'IA générative dans les produits Google, tout en soutenant les besoins croissants de ses clients cloud.

Derrière cette ambition, l'annonce phare est l'arrivée d'Ironwood, la première puce Google spécifiquement conçue pour les modèles non seulement capables de répondre à des requêtes, mais de générer et interpréter des informations de manière proactive. Issue de plus d'une décennie d'expérience avec

les TPU, Ironwood a été développé pour accompagner les modèles les plus avancés, comme les LLM massifs, tout en répondant aux contraintes thermiques, énergétiques et de bande passante qui freinent aujourd'hui le déploiement de l'IA à grande échelle. Techniquement, la puce repose sur une architecture pour limiter au maximum les flux de données, réduire la latence intra-puce et optimiser les calculs tensoriels complexes. Pour atteindre ces objectifs, Google a intégré un réseau de communication ICI (Inter-Chip Interconnect) de nouvelle génération, capable de supporter une synchronisation à l'échelle de 9.216 puces dans une seule configuration avec un système refroidi par liquide. Dans cette configuration, Ironwood délivre une puissance de calcul de 42,5 exaflops, soit plus de 24 fois celle d'El Capitan, le plus puissant des superordinateurs actuels.

La performance ne se limite pas à la capacité de traitement brut. Google a également accru la mémoire disponible sur chaque puce, passant à

192 Go de HBM (une multiplication par six par rapport à la génération précédente), tout en augmentant le débit mémoire à 7,2 To/s par puce. Cette infrastructure n'est pas pensée uniquement pour les produits Google, mais pour l'ensemble de son écosystème cloud. Gemini 2.5, AlphaFold et d'autres modèles de recherche avancés tournent déjà sur des TPU. Avec Ironwood, Google entend généraliser ce niveau de performance à ses clients. Les cas d'usage dépassent aujourd'hui le cadre de l'IA traditionnelle pour s'étendre à des domaines tels que la finance quantitative, la recherche biomédicale ou les simulations physiques de très haute précision. En parallèle, Google continue de repousser les limites de l'informatique quantique. Le groupe a dévoilé Willow, une puce quantique qui surmonte un obstacle fondamental : la correction d'erreurs à grande échelle. Cette avancée, longtemps considérée comme hors de portée, rapproche la perspective d'un ordinateur quantique réellement utilisable, en rendant la croissance du nombre de qubits plus fiable.

comme des sources dynamiques d'intelligence, en s'adaptant aux autorisations utilisateurs, aux formats multiples (texte, image, vidéo, documents métier), et aux cas d'usage métiers.

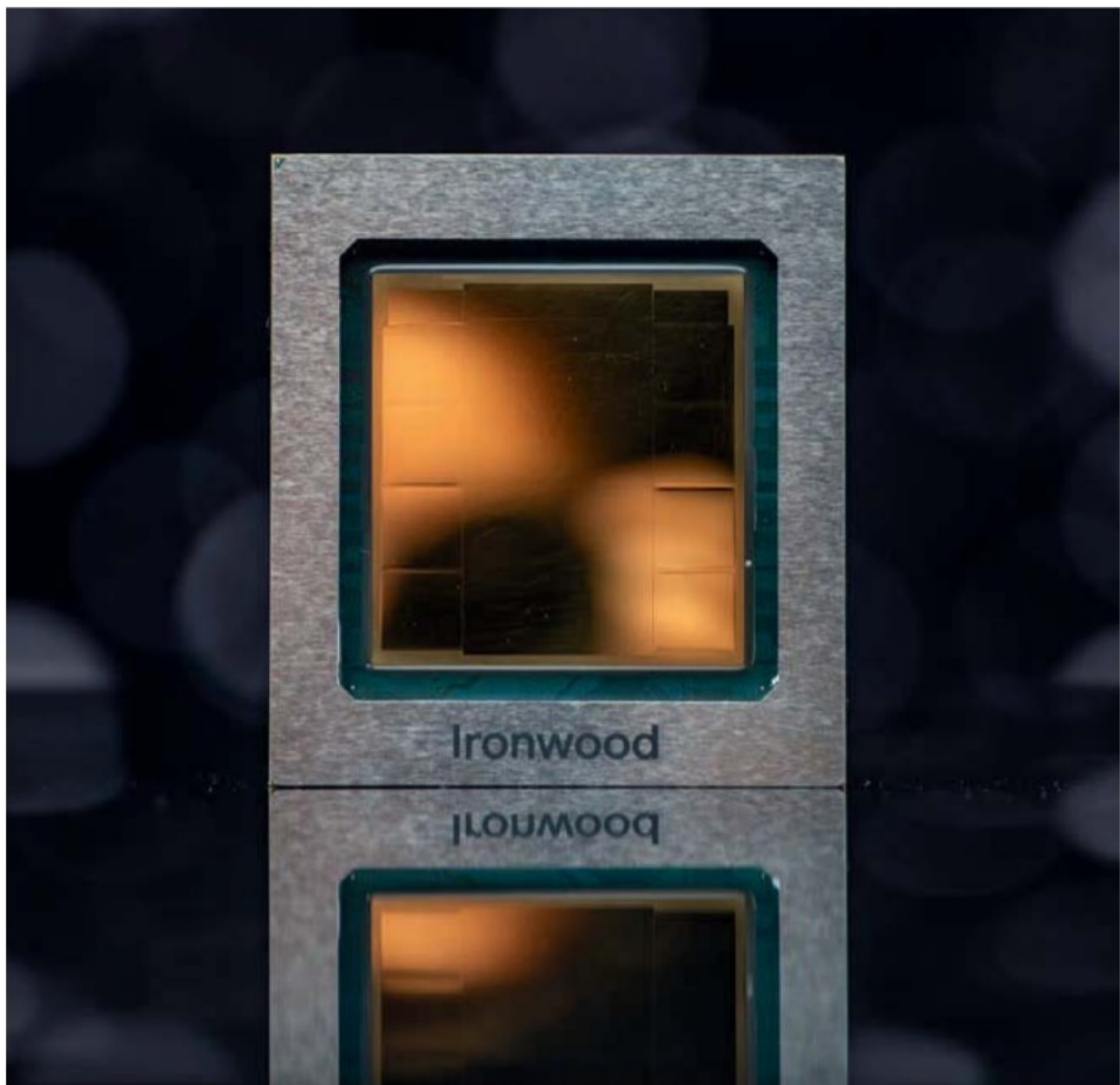
« *Le problème des entreprises n'est pas qu'elles manquent d'informations, mais qu'elles sont incapables d'y accéder de manière fluide, sûre et contextualisée* », a souligné Raj Pai, vice-président produit de Google Cloud. « *Nous avons voulu que chaque collaborateur puisse poser des questions complexes et obtenir une réponse raisonnée, légitime, fondée sur les droits d'accès en place — sans mise à niveau de la plateforme, sans refonte de leur système d'information documentaire.* » Intégré nativement à Vertex AI, Agentspace agit comme le cerveau sémantique partagé de tous les agents déployés dans l'entreprise. Il peut être personnalisé, interfacé avec les outils internes, et bientôt exécuté en environnement souverain grâce à Google Distributed Cloud.

Orchestration cognitive avec Workspace Flows

Dans les environnements utilisateurs, Google pousse encore plus loin l'intégration avec Workspace Flows, une interface de conception d'orchestrations intelligentes en langage naturel. Il s'agit ainsi d'un moteur de composition dynamique d'agents, qui permet à n'importe quel utilisateur d'assembler des séquences de traitement complexes pilotées par l'IA. De ce fait, une intention formulée en langage naturel devient une chaîne logique de déclencheurs, d'analyse, de génération, de décision, répartie sur plusieurs agents spécialisés.

« *Ce n'est pas l'automatisation que nous avons renforcée, c'est la capacité de raisonnement réparti à travers les outils* », a affirmé Yulie Kwon Kim, vice-président produit de Google Workspace. « *Chaque flux devient une sorte de mini-système multi-agents, où des Gems — nos agents spécialisés — prennent en charge des tâches qui demandent du discernement, pas juste des clics.* »

Dans un scénario de gestion de tickets, par exemple, un formulaire client peut être analysé par un agent pour en extraire l'intention, validé par un second agent qui consulte Agentspace pour trouver des réponses similaires, puis traité par un troisième qui rédige une réponse personnalisée, respectant les directives de communication internes. Le tout, sans écrire une seule ligne de code.



© Google Cloud

La nouvelle puce Google Ironwood permettra de répondre à des requêtes, mais de générer et interpréter des informations de manière proactive.

Une IA souveraine, distribuée et sécurisée

Concernant la question de la souveraineté, de la conformité et de la localisation des traitements, Google a pris une avance stratégique en portant toute son architecture cognitive sur Google Distributed Cloud (GDC). Grâce à l'intégration des systèmes Nvidia Blackwell et au support d'environnements air-gapped, GDC permet de faire tourner Gemini, Vertex AI, Agentspace et même Workspace Flows dans des centres de données contrôlés, compatibles avec les réglementations les plus strictes. L'ensemble de la stack est opéré par Google, mais sous le contrôle du client, qui choisit les régions de traitement, les configurations réseau, et les mécanismes de gouvernance.

« *Les secteurs régulés ont trop longtemps dû choisir entre performance IA et conformité. Avec GDC, ce compromis disparaît* », a estimé Vithal Shirodkar, VP et GM de Google Distributed Cloud. « *C'est la même IA que dans notre cloud public, mais opérée localement, sur des infrastructures validées pour le secret défense, et extensible avec vos outils métier.* » Cette capacité à déployer les mêmes agents, les mêmes modèles, les mêmes protocoles d'orchestration dans le cloud ou on-premises, ouvre la voie à des architectures réellement hybrides, où la donnée reste là où elle doit rester, sans freiner l'innovation cognitive. □

Michel Chotard

Plateforme

OVHcloud mise sur Iceberg et Trino

Annoncée lors de l'OVHcloud Summit il y a quelques mois, l'offre Cloud Data Platform d'OVHcloud est maintenant disponible. OVHcloud mise sur les solutions Open Source stars du moment, Apache Iceberg et Trino, pour convaincre les entreprises de lui confier leurs données.

La Data Platform est aujourd'hui un composant critique pour un fournisseur cloud. C'est la base d'une stratégie Data et IA des entreprises, mais c'est aussi un composant clé pour choisir un fournisseur cloud. En effet, la tendance est de rapprocher au maximum les applications de la plateforme de données, pour des raisons de performances et de coûts : multiplier les échanges inter-cloud étant coûteux sur ces deux plans. La révolution de l'IA ne fera qu'amplifier encore cette tendance lourde dans les prochaines années. La qualité de la Data Platform de Google Cloud a permis à l'américain de se faire une place auprès des grandes entreprises en dépit de la domination d'AWS sur le cloud mondial. OVHcloud se devait de proposer une plateforme de données à la hauteur de ses ambitions.

Annoncée en novembre dernier lors de l'OVHcloud Summit, l'OVHcloud Data Platform est désormais sortie de sa phase beta et enfin disponible à tous les clients. Cette offre se veut une offre de bout en bout avec l'ensemble des composants technologiques qui constituent une Data Platform : des capacités de cataloging des données, une solution de Lake House, de la puissance de calcul, de l'analytique, des API. Sans surprise, le fournisseur cloud a largement pioché dans les communautés Open Source pour assembler son offre. La Data Platform est le fruit du partenariat engagé avec le finlandais Aiven en 2022. Le Français s'appuyait sur son partenaire pour ses offres DaaS (Database as a service) et celui-ci pouvait proposer un hébergement sur OVHcloud en tant qu'alternative aux hyper-scalers américains.

Une séparation nette entre le « compute » et le stockage

Stéphane Ligneul, solution architect chez OVHcloud, revient sur les trois priorités qui ont dicté à la construction de cette offre : « Celle-ci devait être à la fois composable, portable pour que les données puissent être déplacées sans contraintes sur site vers le cloud et entre fournisseurs cloud différents, et

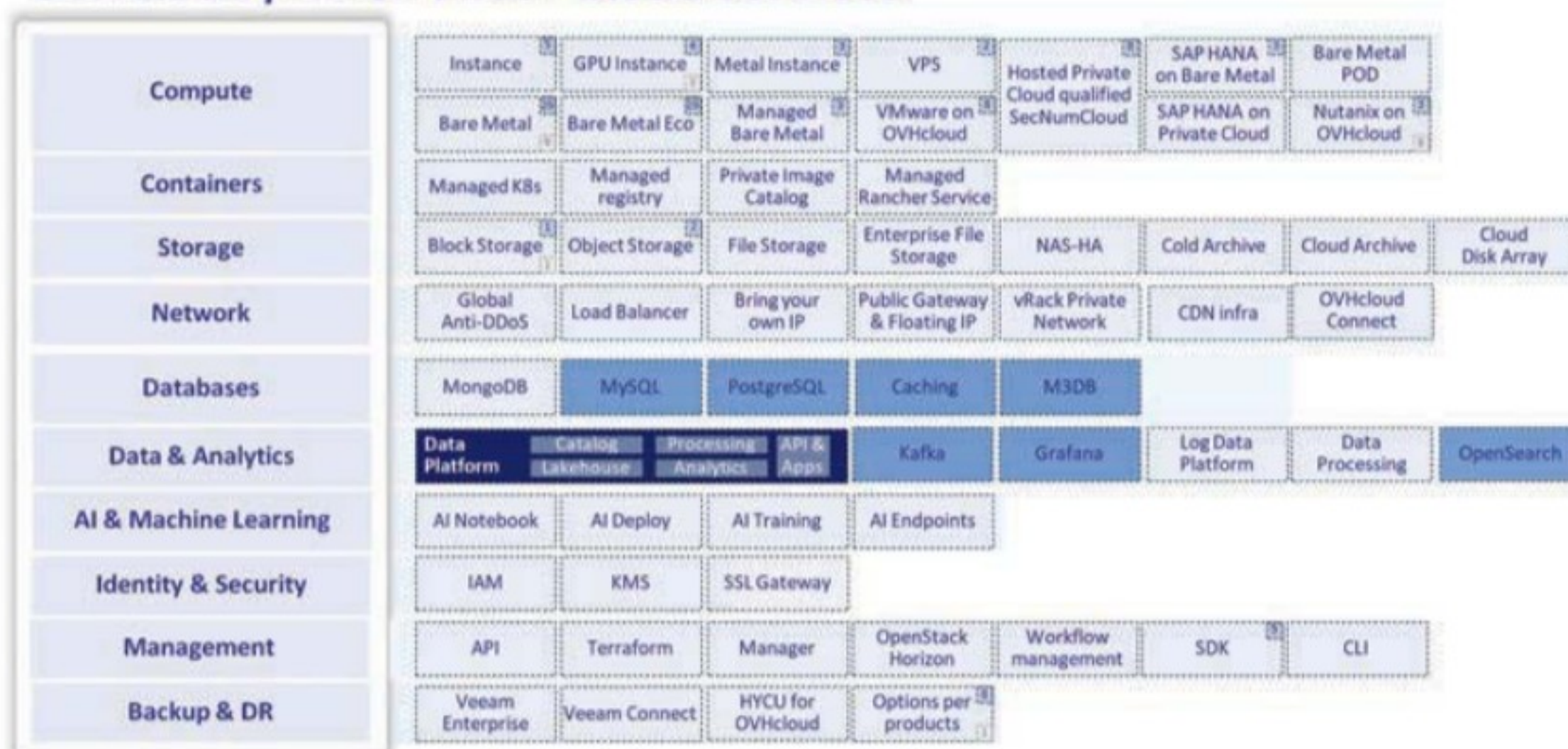
enfin, devait être totalement programmable et s'inscrire dans une approche 'Everything as code'. »

Un précepte a dicté à l'élaboration de l'architecture de la plateforme : une séparation stricte entre la partie stockage des données et la partie « compute », et notamment les moteurs d'ingestion des données et de requêtage. « Les deux n'évoluent plus de manière linéaire comme c'était le cas par le passé, avant Hadoop. Il faut aujourd'hui dissocier le stockage de la donnée de la partie traitement, ce qui permet aussi d'aller chercher des composants optimisés. » Cette dissociation permet de consolider les données sur un même Lake House sans pour autant compromettre ses performances, tout en multipliant les cas d'usage. « Cette consolidation permet d'éviter la duplication des données et toute la complexité liée à savoir quelle est la bonne donnée à utiliser, quelle est celle qui est à jour, et éviter tous les efforts de réplication, de transformation, etc. Enfin, quand on centralise la donnée, on peut mieux contrôler qui a accès à quoi, et mettre en place des mécanismes de sécurisation et de contrôle de cette donnée. »

Iceberg, la solution de Lake House qui a le vent en poupe

La plateforme est constituée de nombreuses composantes sous licence Open Source, et c'est sans doute l'un des atouts majeurs de l'offre OVHcloud. Son Lake House n'est autre que Apache Iceberg, une solution de Lake House basée sur des fichiers, donc avec de très faibles coûts de stockage.

OVHcloud / Aiven : 100+ cloud services



OVHcloud, avec son partenaire Aiven, proposent aujourd'hui une offre complète, avec une Data Platform et une offre de bases de données managée où l'Open Source est dominant.

La gouvernance des données et la gestion des métadonnées sont centralisées sur la solution Rest Catalog qui gère les métadonnées de la Data Platform, mais aussi tous les fichiers mis en œuvre par Iceberg. « *Rest Catalog est vraiment le cœur de l'architecture, car c'est lui qui permet de découpler les moteurs de requêtes du stockage* », explique Julien Thiaw Kine, team lead chez OVHcloud qui ajoute : « *C'est aussi le seul point de vérité et c'est là où s'appliquent les règles de gouvernance.* » La plateforme peut aussi mettre en œuvre les catalogues Tabular, Polaris, Unity. L'ingestion des données est assurée par les moteurs Pylceberg, Apache Spark. Trino est le moteur SQL mis en avant par OVHcloud pour sa plateforme. Le Français a validé la compatibilité de sa plateforme avec les solutions analytiques Metabase, Tableau et Superset. « *Ce modèle nous permet de faire évoluer chaque couche indépendamment des autres* », ajoute Julien Thiaw Kine. « *L'entreprise peut décider de changer de moteur d'ingestion ou ajouter un nouveau moteur de lecture de données sans remettre en cause les données et les formats de stockage.* »

L'équipe OVHcloud a créé des dashboards Grafana, afin d'assurer une bonne observabilité sur la plateforme à partir des métriques collectées par Prometheus. Pour la sécurité des accès, OVHcloud a fait le choix d'OPA pour gérer les ACL de manière dynamique. Enfin, pour assurer la performance d'Iceberg, des jobs ont été implémentés pour assurer la compaction des partitions des tables et maintenir les temps de réponse au meilleur niveau.

La vitesse de Trino au service des clients OVHcloud

Cette combinaison de Trino et Iceberg nous permet d'offrir un Lake House moderne, souverain. Victor Coustenoble, architecte solutions chez Starburst, éditeur de la version « Entreprise » de Trino, vante les performances de l'outil. « *Trino est un moteur SQL très performant qui est distribué de type MPP (Massivement Parallèle) conçu pour requêter de grands volumes de données, mais aussi capable de se connecter à de multiples sources de données* », explique l'expert. Starburst a été fondé par des membres de Teradata et les créateurs de Presto / Trino chez Meta. Presto avait été initialement créé par Facebook pour remplacer le moteur

Julien Thiaw Kine,
team lead chez
OVHcloud

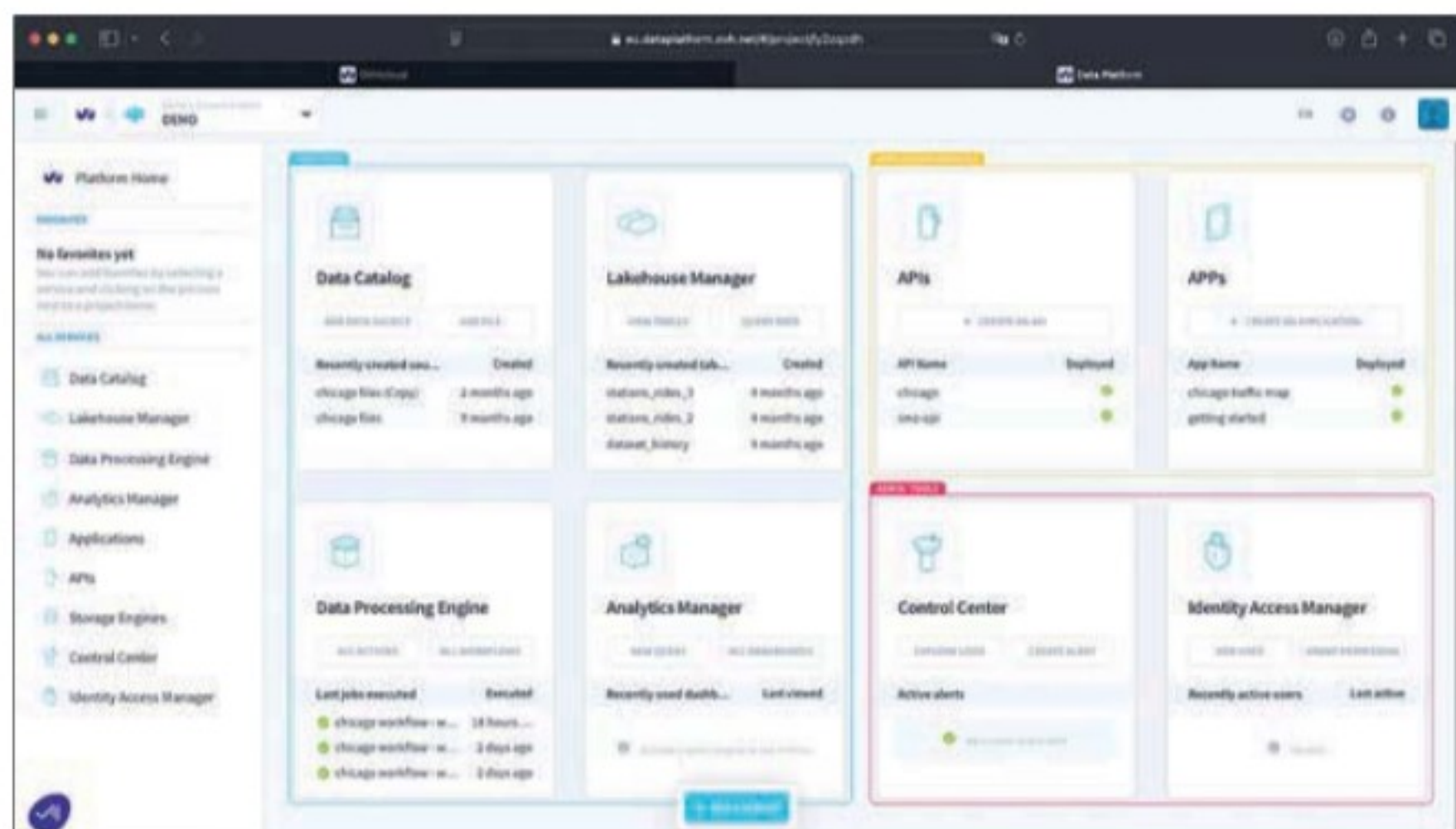


« *Notre objectif était de concevoir une plateforme de Data partagée, interopérable et bâtie sur un stack open source. Elle devait être capable de porter de multiples workloads, dont la BI, les explorations ad-hoc, et être mise en œuvre par des équipes distribuées et un découplage du compute et du stockage. Nous avons la volonté de maîtriser totalement la stack pour assurer la scalabilité, garantir la souveraineté et la conformité de la plateforme via l'open source* »

Hive et faire du SQL haute performance sur HDFS. Presto a fait l'objet d'un fork et a été rebaptisé Trino. Outre ses performances brutes, le moteur permet de faire de la fédération de données avec la possibilité de faire des jointures entre Kafka, MongoDB, Oracle et Hadoop, par exemple. Trino n'assure pas de stockage de données. Il s'agit d'un moteur SQL qui doit obligatoirement être connecté à une source de données. Il bénéficie aujourd'hui d'un vaste catalogue de connecteurs pour des bases de données de tous types. En outre, OVHcloud a fait le choix de mettre en place Trino Gateway, un reverse proxy qui sert de point d'entrée unique pour toutes les requêtes : il effectue le routage des requêtes entre plusieurs clusters Trino. C'est un élément clé de l'architecture car il fait passer entre les utilisateurs et les clusters Trino.

Créé par Facebook pour Facebook, il présente l'atout de pouvoir scaler et supporter un très grand nombre d'utilisateurs concurrents. Une puissance désormais disponible en mode managé sur les infrastructures d'OVHcloud. □

A.C



Après quelques mois en accès restreint, la console de la Data Platform OVHcloud est sortie de sa phase beta. Elle est désormais accessible à tous les clients du fournisseur cloud français.

Kubecon Europe

Un écosystème dynamique

Début avril, la CNCF réunissait le petit et grand monde de l'open source, à l'eXCel à Londres, pour sa conférence européenne couplée à celle de Kubernetes. Il y a beaucoup été sujet d'intelligence artificielle. Sans surprise.

La conférence de la CNCF en Europe est devenue un des rendez-vous incontournables de l'année. Plusieurs milliers de personnes se sont rendus à Londres pour assister aux différentes sessions et démonstrations qui s'y tenaient. Les cadres de la fondation n'ont pas manqué de rappeler la force et le dynamisme de l'écosystème open source et de Kubernetes. Ainsi, selon une étude révélée lors de la conférence, 89 % des entreprises interrogées assuraient utiliser des technologies nativement Cloud. Dans cet écosystème, Kubernetes est devenu plus qu'un standard de fait avec 93 % des entreprises qui évaluent, testent ou utilisent l'orchestrateur. Pour la fondation, cela dénote un changement profond dans la manière de développer et de déployer les applications. Les principales raisons d'une adoption aussi rapide mises en avant par les entreprises interrogées sont l'accélération de l'innovation, la réduction du temps de mise sur le marché de nouvelles applications ou services et l'amélioration de la résilience dans un contexte de plus en plus digitalisé. Dans sa session, Chris Aniszczyk, CTO de la CNCF a ajouté : *« les organisations connaissent de nombreux obstacles opérationnels et culturels lorsqu'elles adoptent globalement la technologie. Si la sécurité reste évidemment vitale, elles se concentrent maintenant sur l'automatisation et les meilleures pratiques qui apportent une livraison des applications plus rapide et plus solide »*. Ainsi, l'étude précitée voit une augmentation de 31 % des processus de CI/CD (continuous integration/continuous deployment) pour accélérer la livraison de nouvelles applications et plaçant le GitOps comme une bonne pratique. En conséquence, les entreprises sont passées de la simple adoption à l'optimisation du travail des équipes par la collaboration, l'automatisation et l'industrialisation de leurs opérations.



Un grand avenir avec l'IA

Le déploiement de l'intelligence artificielle ou de l'apprentissage machine est émergent sur une pile logicielle reposant sur Kubernetes. 48 % seulement des entreprises interrogées déclaraient avoir mis en œuvre de telles applications. Les principales utilisations sont les batchs de données, suivis des essais de modèles, des inférences en temps réel et de la préparation des données. Le défi reste cependant d'opérer totalement l'intelligence artificielle dans un environnement cloud natif. Avec l'évolution des logiciels et des meilleures pratiques, la fondation voit un avenir radieux pour Kubernetes dans ce domaine pour devenir un élément clé de l'innovation tirée par l'intelligence artificielle dans les différentes industries.



Chris Aniszczyk, CTO de la CNCF, lors de la session générale à Londres

Avec l'évolution des environnements nativement Cloud, l'étude indique que la transformation culturelle des entreprises, l'automatisation et la sécurité sont les trois éléments clés, et que celles qui investissent dans ces trois secteurs seront les mieux placées pour s'adapter aux rapides changements technologiques et de l'écosystème, et vont pouvoir proposer des applications plus sécurisées, plus efficaces et plus évolutives. □

B.G

Deep fake

Contrer la fraude documentaire

L'utilisation de documents falsifiés pas identifiables à l'œil nu se banalise entre autres pour des escroqueries à la garantie. Confronté au problème, HP a outillé ses équipes du service SaaS de Finovox.

« Les tentatives de fraude à la garantie sont toujours plus nombreuses et inédites », constate Élodie Foucher, EMEA Global Brand Security Investigator chez HP. Suffisamment pour motiver HP à créer un service dédié aux seules tentatives de fraude à la garantie, qu'elles soient pour les particuliers ou les entreprises. Si l'entreprise américaine ne veut pas donner de chiffres, elle détaille les types de fraude. « Les fraudes peuvent porter sur une demande de remboursement, de remplacement de produit ou encore de réparation. Les documents justificatifs sont des fausses factures, des étiquettes de numéros de série trafiquées, des photos retouchées... », détaille Élodie Foucher. Et depuis quelques années, les demandes émanent de plus en plus souvent de faux profils. » Problème, les manipulations de documents de plus en plus élaborées rendent l'identification des fraudes impossible à l'œil nu. Si la prise en charge de la demande repose toujours sur une décision humaine, « nous avons besoin d'asseoir nos réponses sur des certitudes », insiste Élodie Foucher.



Élodie Foucher, EMEA Global Brand Security Investigator chez HP

besoin de s'assurer de l'authenticité des documents dans le cadre de leurs vérifications », rappelle Élodie Foucher. Disponible en SaaS ou en API, Finovox est retenu la même année, notamment pour sa capacité à extraire des données textuelles des documents (date, numéro de série, nom de client...). Il est utilisé chez HP en mode SaaS. Les données sont exportées par l'équipe antifraude et sont ensuite comparées aux bases de données internes de l'entreprise. Le résultat indique un niveau de confiance attribué au document (vert, orange et rouge). Illustration, si le code barre ne correspond pas au numéro de

série, le risque est considéré comme élevé. « Ce n'est pas forcément une tentative de fraude, toutes les anomalies ne sont pas des manipulations. Il peut s'agir d'un problème de lecture », décrit Élodie Foucher. Pour attribuer le score de confiance, Finovox a développé d'autres contrôles destinés en particulier à vérifier les métadonnées. Ils sont capables de détecter le nombre de version de document, le logiciel à l'origine de la création... Illustration, « cette semaine nous avons été confrontés à une demande de remboursement basée sur un document Word créé il y a quelques jours, alors que la facture était censée dater de 2023 », signale Élodie Foucher. Outre les contrôles, l'éditeur utilise des algorithmes de machine learning sur les seules métadonnées pour évaluer l'authenticité des documents. Enfin, aspect le plus sensible, il restait nécessaire de calibrer ces seuils de confiance, « pour éviter trop de faux positifs ou inversement », résume notre interlocutrice. Dans ce but, l'équipe a travaillé avec Finovox au cours d'une phase initiale de calibrage. « Nous prenons en charge cette étape. Elle ne nécessite pas de développements mais repose sur l'activation ou la désactivation de contrôles »,

détaille Jacques de Chevron Villette, head of implementation chez Finovox. Cette étape initiale a également pour but de personnaliser la solution si besoin. « Pour HP, nous avons ajouté la prise en charge de nouveaux codes-barres », illustre Jacques de Chevron Villette.

Satisfait de la solution qui permet de traiter plus de demandes et ce, avec plus de certitudes, HP envisage de déployer l'outil dans d'autres services, en particulier les call centers. « Nous continuons à améliorer la performance de l'outil », indique Élodie Foucher. Ce qui devrait permettre d'automatiser tout ou partie de ces contrôles jusqu'à la prise de décision. **P. Br**

Des fraudes sophistiquées

Pour faire face à cette sophistication croissante des tentatives comme à leur augmentation, les équipes antifraude cherchent en 2024 un outil facilitant la détection de documents faux ou falsifiés. A l'époque, HP ne souhaitait pas développer une solution interne et opte donc pour un appel d'offres. « Pour rédiger le cahier des charges de l'outil, nous avons collaboré avec les équipes internes qui avaient des besoins similaires, les départements compliance et audit. Ils ont également

L'IA GÉNÉRATIVE INDUSTRIALISE LA FRAUDE DOCUMENTAIRE

Selon l'observatoire de l'éditeur Tessi, la fraude documentaire coûte 65 milliards d'euros annuels à l'économie française représentant une augmentation de 244 % en un an. Pour la fraude à l'assurance auto en France, l'addition monte à 2,5 milliards d'euros selon la Fédération Française de l'Assurance. Sur ces chiffres, et selon Tessi, 42,5 % des fraudes sont désormais générées par l'IA. Le plus inquiétant est peut être ailleurs. Le même rapport souligne que plus d'un Français sur 10 a déjà falsifié des documents, et que 20 % des moins de 35 ans reconnaissent l'avoir déjà fait pour des démarches bancaires.

Agentique

DSI et métiers composent ensemble les agents

Après deux ans d'expérimentations et de benchmarks, la DSI du groupe Constructa, sous l'impulsion de Woilide Nagmar, compose ses propres modèles d'IA en interne, en impliquant 20 ambassadeurs métiers.

« La DSI devient vraiment légitime avec l'IA. Rodés aux prompts ChatGPT, Copilot et Noota, les métiers ont vite compris le besoin d'intégration, d'automatisation et d'orchestration de services. Notre pôle dédié les accompagne, modernise les procédures, fournit les outils, guides et bons réflexes de sécurité », observe le DSI de Constructa, promoteur immobilier marseillais assurant la commercialisation, la gestion d'actifs et de fonds investissement, (13 milliards d'actifs sous gestion et 200 collaborateurs).



La tour M99 à Marseille (30 étages et 99 mètres de haut) prévoit une mixité d'usages dès 2028.

Gains en productivité et qualité de services

« Arrivé en 2021 comme DSI de transition, j'ai effectué un état des lieux avant de guider la transformation digitale du groupe : organisation du SI, mise en place du CRM et d'une plateforme e-commerce B2C et B2B. Il fallait accompagner la croissance de la société en impliquant ses partenaires revendeurs et courtiers indépendants. La numérisation des processus métiers, celle du parcours client et l'IA font

gagner les équipes en productivité. L'organisation devient plus compétitive, tandis qu'il y a quatre ans, beaucoup de choses étaient encore faites à la main. »

Parmi les projets structurants, la mise en place du CRM Salesforce a permis d'améliorer le pilotage et la digitalisation de la relation avec les clients, la dématérialisation des processus Order-to-Cash (O2C) et Procure-to-Pay (P2P) via la solution Esker contribuant à automatiser les flux de gestion.

QUELLES CLÉS DE RÉUSSITE POUR LES PROJETS IA ?

« Le secret d'une IA réussie, ce sont des data propres, bien structurées. Or, cette classification prend du temps et, en pratique, personne ne souhaite créer le référentiel ni apposer des tags sur ses documents », constate Woilide Nagmar, le DSI de Constructa.

Tester les outils et en tirer des enseignements permet d'aligner les technologies d'IA aux diverses attentes des métiers, une étape essentielle à leur adoption dans l'entreprise. Les formations, la mise en conformité réglementaire et la segmentation des usages forment trois autres facteurs de succès, selon lui : « Un test en ligne, proposé sur notre intranet, permet à chaque utilisateur de vérifier s'il est prêt ou non à se servir de l'IA en toute connaissance de cause, notamment s'il est conscient des devoirs qui lui incombent et de l'hygiène à tenir pour limiter les risques. Si un test est raté, le collaborateur n'accède pas immédiatement aux outils ; il doit prendre connaissance de notre charte d'usages intégrant, à présent, un chapitre sur l'IA et suivre une formation. »

La stratégie IA vient renforcer cette dynamique par l'intégration de solutions d'automatisation optimisant les tâches à faible valeur ajoutée et la formation des équipes. Elle s'appuie sur une gestion centralisée des données via un Data Hub interconnecté aux quatre ERP du groupe, et sur le développement de nouveaux services numériques visant à réduire les délais de traitement et à améliorer la satisfaction des clients. « Tout a démarré par la prospection de fonds, la recherche de terrain, le dépôt de permis de construire, la construction et la commercialisation de logements neufs, les logiciels modernes aidant à suivre les travaux, les ventes, la gestion et la promotion immobilière autour de données structurées », précise-t-il.

Par exemple, des agents d'IA couplés à la bureautique, aux jumeaux numériques et aux outils de cartographie SIG (système d'information géographique) peuvent élargir aux partenaires et aux clients la visibilité des chantiers en cours de construction.

Collecte et croisement de données

« La méthode retenue ? Lister et prioriser les cas d'usage afin de lancer des initiatives par thématique, et remporter de petits succès rapidement. Un portail intranet et une bibliothèque de prompts personnalisés sont actuellement en phase de tests, en environnement contrôlé. Nous avons également mené des actions de sensibilisation au RGPD et engagé des travaux exploratoires autour de la personnalisation d'assistants pour des fonctions internes. »

La donnée reste l'or noir des algorithmes d'IA. La collecte et le croisement d'open data foncières (Insee et data.gouv notamment) facilite la prospection et évite l'acquisition régulière de coûteuses études de marché externes.

« La direction et les métiers ont pris conscience de l'importance de la qualité des données. Un projet de GED (gestion électronique de documents) a été lancé parallèlement, pour classer les documents. La classification automatique et la collaboration entre les métiers aident à préparer les informations nécessaires à la conception de modèles d'IA plus précis, s'appuyant sur des données spécialisées. »

Un comité de pilotage composé de 20 ambassadeurs accompagne les expérimentations menées depuis le début de l'année, au sein d'un environnement sécurisé : le ConstructaLab. Les besoins exprimés par les équipes guident les réflexions en cours sur la conception d'assistants numériques adaptés à chaque usage : automatisation de tâches documentaires, aide à la rédaction, structuration de réponses ou génération de contenus.

Encore au stade exploratoire, ces projets devraient bénéficier, en premier lieu, aux services RH, juridiques, ventes et marketing de l'entreprise.

L'objectif est de tester l'approche par petits lots, d'examiner la valeur générée pour les métiers, et de renforcer

STRATÉGIE DE DSI : RENFORCER LES LIENS MÉTIERS

Wolide Nagmar accompagne depuis plus de 20 ans la transformation digitale d'organisations complexes, avec une culture du résultat. Impliqué dans la stratégie de l'organisation, il place l'innovation technologique au service de la compétitivité des métiers, et de leurs collaborations.

Ingénieur en systèmes, réseaux et télécoms (Orsay), titulaire d'un master en sécurité des SI (Epita) et d'un Executive MBA (Essec), il s'est récemment formé aux enjeux de l'IA.

Avant de rejoindre Constructa en 2021, il a piloté la transformation digitale de structures telles que Citeo, CVE (Cap Vert Énergie), Entoria (ex-Ciprés Assurances) et le Groupe April (courrier grossiste). Membre actif du club CIP Méditerranée, il partage régulièrement ses convictions sur l'IA, la cybersécurité et le rôle stratégique des DSI.

l'acculturation autour d'usages responsables de l'IA. Une fois les cas d'usages prioritaires identifiés, les utilisateurs sont formés avec le soutien du cabinet Acadys.

Modèles hybrides et cloud hybride

Le pôle IA de la DSI suit l'évolution rapide des grands modèles linguistiques (LLM) et des RAG (retrieval-augmented generation) afin de procurer des réponses affûtées, tenant compte des documents et process internes. Une veille stratégique consiste à comparer plusieurs solutions du marché, comme OpenAI ou MistralAI, afin de vérifier leur pertinence pour les métiers, mais aussi leur capacité à intégrer les contraintes de souveraineté, de coût, et de confidentialité des données. L'exploitation en interne d'un modèle IA spécialisé est envisagée, pour un investissement en ressources ne devant pas dépasser les 100 K€ cette année. L'ensemble de ces critères permettront d'alimenter les décisions futures lorsqu'il s'agira de déployer la solution à plus grande échelle, et de la faire monter en puissance.



Wolide Nagmar, DSI de Constructa

Les infrastructures du groupe s'appuient déjà sur un cloud hybride. Le datacenter, en colocation chez Unitel à Marseille, permet d'articuler les ressources internes et celles des éditeurs SaaS. Cette architecture hybride contribue au cloisonnement entre données publiques et données sensibles. □

O.B

Organisation

Innover ou agoniser, c'est vous qui voyez !



Dans son ouvrage, le docteur en informatique Philippe Boulanger s'appuie sur des expérimentations, des exercices guidés d'auto-analyse et des récits captivants, il interpelle autant qu'il instruit. Le lecteur, qu'il soit dirigeant ou collaborateur, y trouvera des clés pour réconcilier les objectifs souvent divergents au sein de son organisation. Philippe Boulanger propose une réflexion profonde sur la transformation

des entreprises dans un contexte économique, politique et technologique en mutation rapide. Ce livre donne aussi des réponses et l'auteur y développe sa méthode d'intelligence innovationnelle lors de conférences et auprès des équipes dirigeantes. Inspirée de la théorie intégrale de Ken Wilber pour formaliser le concept d'innovation intégrale — une approche qui met l'humain au cœur du processus, au-delà de la seule

technologie. Puis l'innovation intégrale est devenue l'intelligence innovationnelle, une marque déposée de l'auteur qui recouvre l'ensemble de ses activités. Il illustre le concept en cinq étapes : vision & stratégie, communication, exploration, déploiement & exécution, culture & sécurité psychologique. Tout cela contribue à une performance collective qui soutient la profitabilité — et donc le financement de l'innovation.

CHAPITRE 4

Le système intelligence innovationnelle®

« S'il existait un système imparable pour innover et lancer une startup, ça se saurait ! », me lança Rémi.

Partner de fonds d'investissement dédié aux fonds d'amorçage, il est plutôt assertif.

« Qu'est-ce que tu fais du design thinking et du lean startup ?, rétorquai-je aussi sec. Et que fais-tu des métriques clés des startups qui permettent de mesurer la santé de l'activité ? »

S'ensuivit un débat passionné, Rémi campait fermement sur sa position d'investisseur. Il tenait les cordons de la Bourse, c'était lui qui déciderait in fine.

Lorsque je lui demandai quel était son premier critère de décision, il me répondit : « L'équipe, l'équipe et encore l'équipe. »

Je souris, car c'était probablement le critère absolu que j'avais rencontré chez tous les investisseurs. Une équipe saine sera capable d'utiliser les bons outils, les bonnes méthodes, de s'adapter et de pivoter.

Alors pourquoi l'équipe s'enraye-t-elle ?

Pourquoi, lorsque la startup passe en mode scale-up, l'innovation semble-t-elle s'arrêter ?

Pourquoi, une fois l'entreprise bien installée, l'esprit, les valeurs et les méthodes de l'équipe initiale sont-ils perdus ?

J'ai cru pendant longtemps que pour innover, il suffisait de respecter des processus connus et de se protéger de mauvaises décisions avec des données.

Au fil de mes expériences au sein des entreprises que j'ai traversées, j'ai vu de superbes projets ne pas aboutir et des projets médiocres poursuivis en dépit du bon sens.

Je me suis alors posé une question fondamentale : quels sont les invariants du succès de l'innovation ?

La technologie n'est pas un invariant, elle évolue tout le temps, elle crée des modes, des tendances mouvantes. L'invariant constaté aujourd'hui, c'est plutôt l'accélération du changement de la technologie.

Le véritable invariant que j'ai enfin fini par percevoir est la dimension humaine : les compétences des hommes dans leurs comportements, leurs réactions, leurs émotions face à l'innovation, d'un simple collaborateur jusqu'à un dirigeant.

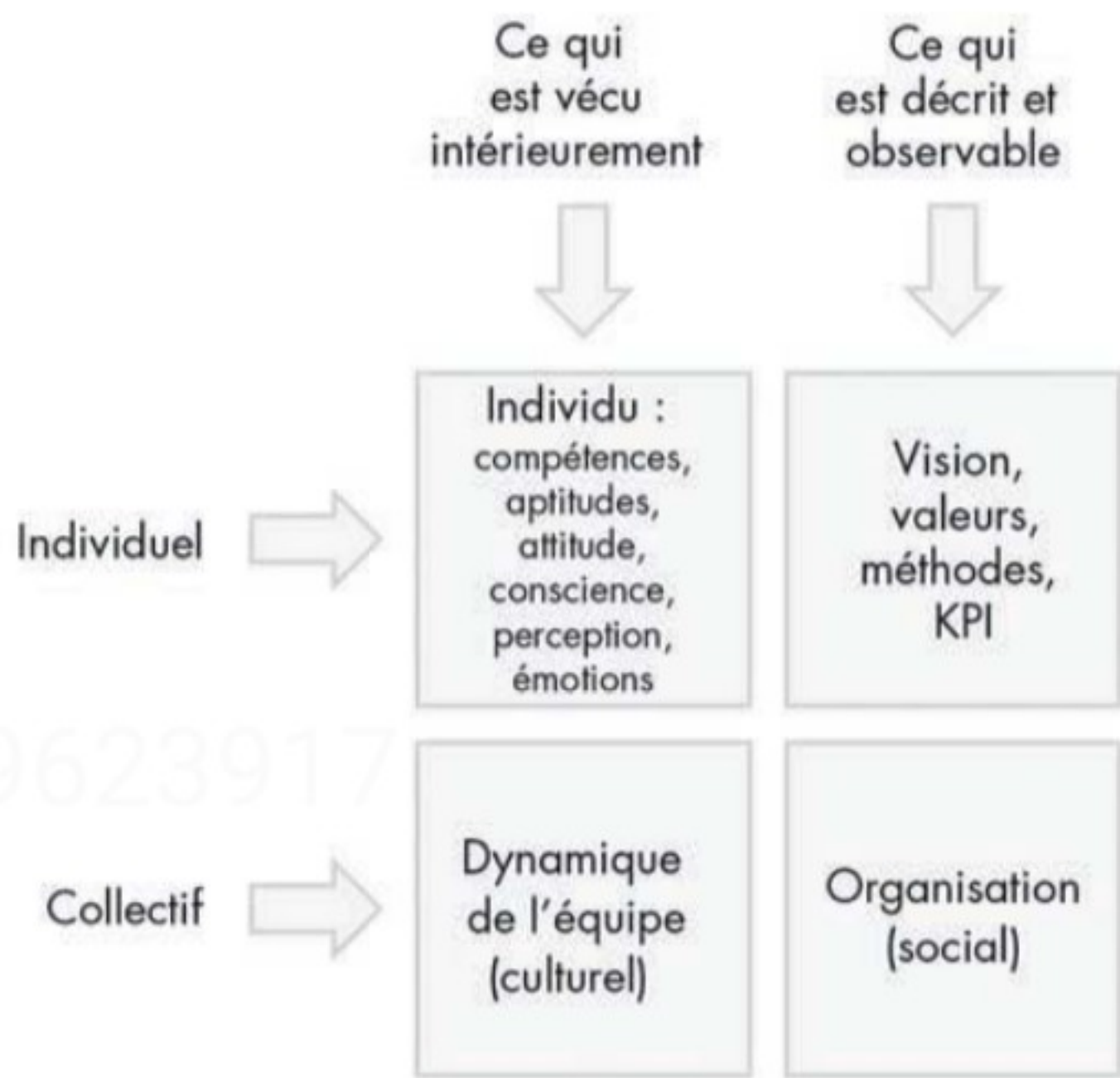
Pourquoi ? Simplement parce que l'homme, même s'il essaie d'être logique, est guidé inconsciemment par ses émotions¹. C'était assez difficile de l'accepter en tant que technologue affirmé.

¹ : António R. Damásio, Descartes' Error : Emotion, Reason and the Human Brain, 1994

« Nous ne sommes pas des machines à penser qui ressentent, mais des machines à ressentir qui pensent. »
António R. Damásio

C'est un événement de sérendipité qui m'a permis de structurer fondamentalement ma pensée sur l'innovation.

Dans le cadre de mon exploration personnelle, j'ai découvert, lors d'une formation de Ken Wilber sur la « théorie intégrale », une structure de pensée philosophique que j'ai décidé de dériver, d'adapter et de déployer pour l'innovation. Son ouvrage sur la psychologie intégrale² m'a permis de consolider et de structurer ma pensée. La première structure fondamentale à laquelle je suis arrivé, après quelques itérations, est basée sur quatre quadrants. J'ai commencé par appeler cette structure « l'innovation intégrale ».



Nous appellerons « étape » cette position sur l'échelle de développement. Par exemple, défier le statu quo est une aptitude typique d'un innovateur, d'un inventeur ou d'un entrepreneur. Dans ce cas, votre interprétation pourrait être évaluée à 1 sur l'échelle de développement si vous êtes conscient de ne jamais défier le statu quo, tandis qu'un score de 5 indiquerait que vous savez défier le statu quo et que c'est une pratique courante chez vous.

Exemple

Comment êtes-vous soumis au biais de statu quo sur une échelle de 0 (soumis) à 5 (pas soumis) ?

À l'aide des pratiques du système « intelligence innovante[®] », votre objectif sera de tendre vers le 5, étape par étape.

Avec humilité, je ne suis personnellement pas toujours au numéro 5 sur certaines de mes échelles de développement, je dois encore évoluer.

Vous pouvez aussi décider de répondre : « pas concerné » ou « pas influencé ». Un point d'attention cependant, il se peut que vous ne soyez pas conscient d'un comportement, d'un biais, d'une peur ou que vous soyez dans le déni. Ça arrive à tout le monde, y compris à moi-même, jusqu'au jour où l'on devient honnête avec soi-même.

La prise de conscience est essentielle pour progresser dans nos capacités pour contribuer à l'innovation.

Êtes-vous en mesure d'identifier les conditions personnelles, d'équipe ou environnementales qui vous permettraient de passer à une étape supérieure ?

Si vous ne défiez jamais le statu quo dans votre position actuelle, quelles autres échelles de développement vous permettront alors de contribuer d'une autre manière à l'innovation ?

L'échelle de développement variera légèrement en fonction des différents sujets traités dans les quadrants de l'innovation intégrale.

La bonne nouvelle est qu'un bon nombre de nos comportements sont programmés. Sans entrer dans le débat de ce qui est inné ou acquis, nous allons déboguer et corriger tous les comportements qui peuvent l'être, comme on le ferait pour un programme informatique.

Mon expérience professionnelle me conduit à considérer le cerveau comme un ordinateur dont la neuroplasticité permet des mises à jour profondes des différents logiciels et autorise même l'ajout de nouveaux logiciels.

J'ai donc créé, réglé et utilisé des outils pour travailler sur mon propre développement et je suis ravi de les partager avec vous car ils se sont révélés très efficaces. C'est l'objet du chapitre suivant, dédié aux outils du système.

Non satisfait, j'ai encore itéré sur la structure de Ken Wilbur en prenant en compte quelques limitations. Je me suis aussi inspiré d'autres structures rencontrées dans ma carrière.

Editions Dunod
304 pages
Broché : 22, 65 €
Kindle : 16, 99 €
ISBN-10 : 2100876554
ISBN-13 : 978-2100876556

2 : Ken Wilber, Integral Psychology : Consciousness, Spirit, Psychology, Therapy, 1999

Le résultat de cette réflexion est la structure du système intelligence innovationnelle® (référéncé par la suite SII dans le texte) qui repose sur deux fondations et six piliers dans ma version initiale.

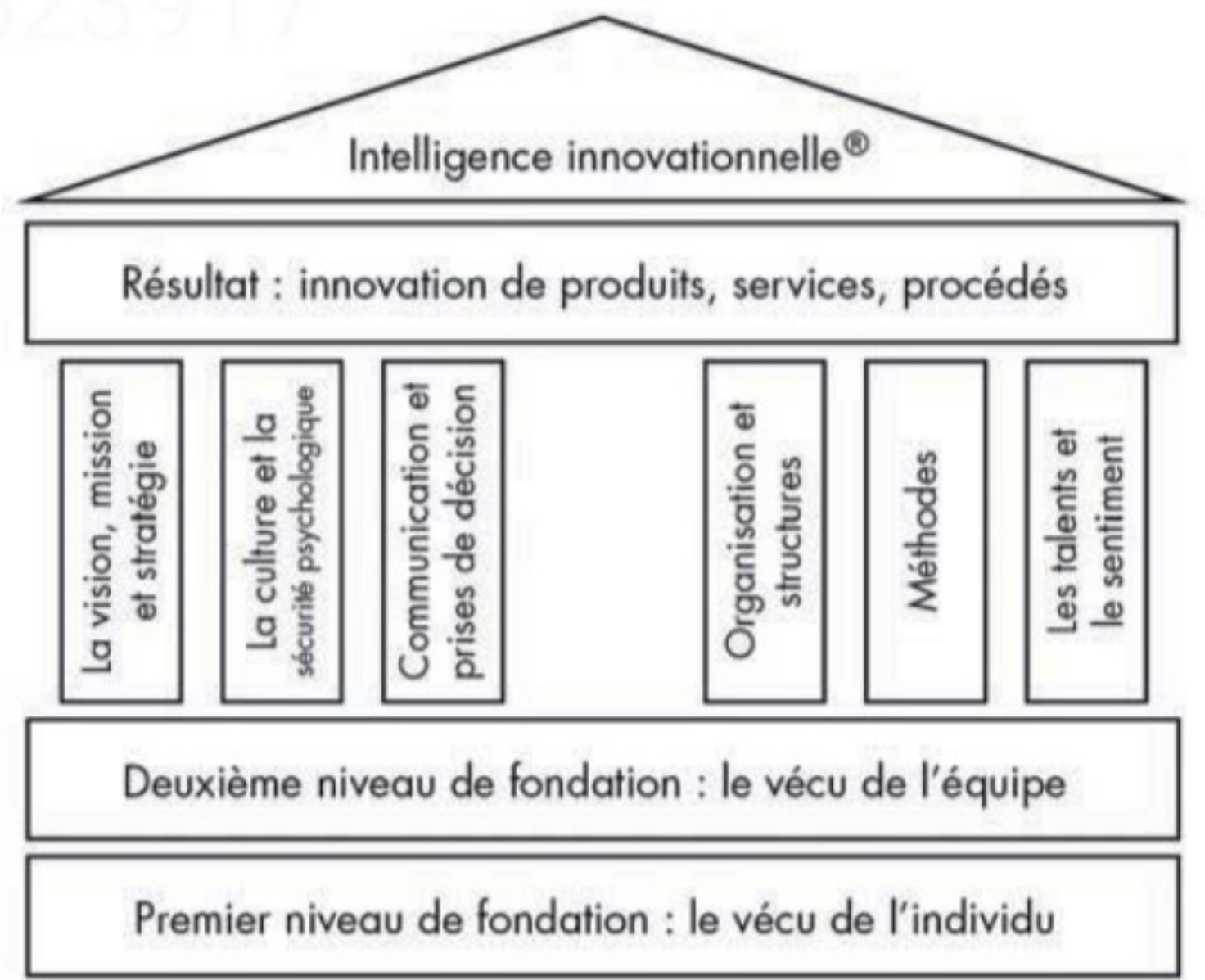
La structure a été conçue pour être souple et adaptative, vous permettant de rajouter des fondations et des piliers si ça correspond mieux à votre organisation.

Le quadrant en haut à gauche de l'innovation intégrale traite de l'individu et de son vécu. Il est le premier niveau de fondation du SII. Le quadrant en bas à gauche traite de l'équipe, au sens large, et de son vécu. Il est le deuxième niveau de fondation du SII.

À titre personnel, cette équipe peut être le couple, la famille proche ou la famille éloignée. À titre professionnel, il peut s'agir d'une équipe projet, d'un service, d'un département, d'une division ou d'une entreprise complète.

Au même titre que l'individu, l'équipe a ses propres échelles de développement. C'est dans ce quadrant que s'exprime le style de management et la culture de l'équipe.

Les quadrants de droite constituent ce que j'appelle les piliers du SII. Ils comportent tout ce qui est construit et observable par un individu ou une équipe. On y trouvera par exemple la vision, la stratégie, des méthodes, des KPI, des organisations... Vous retrouverez ces piliers dans un chapitre à suivre.



En fonction du niveau de responsabilité d'un individu dans l'entreprise, l'impact sur le deuxième niveau de fondation et sur les piliers varie.

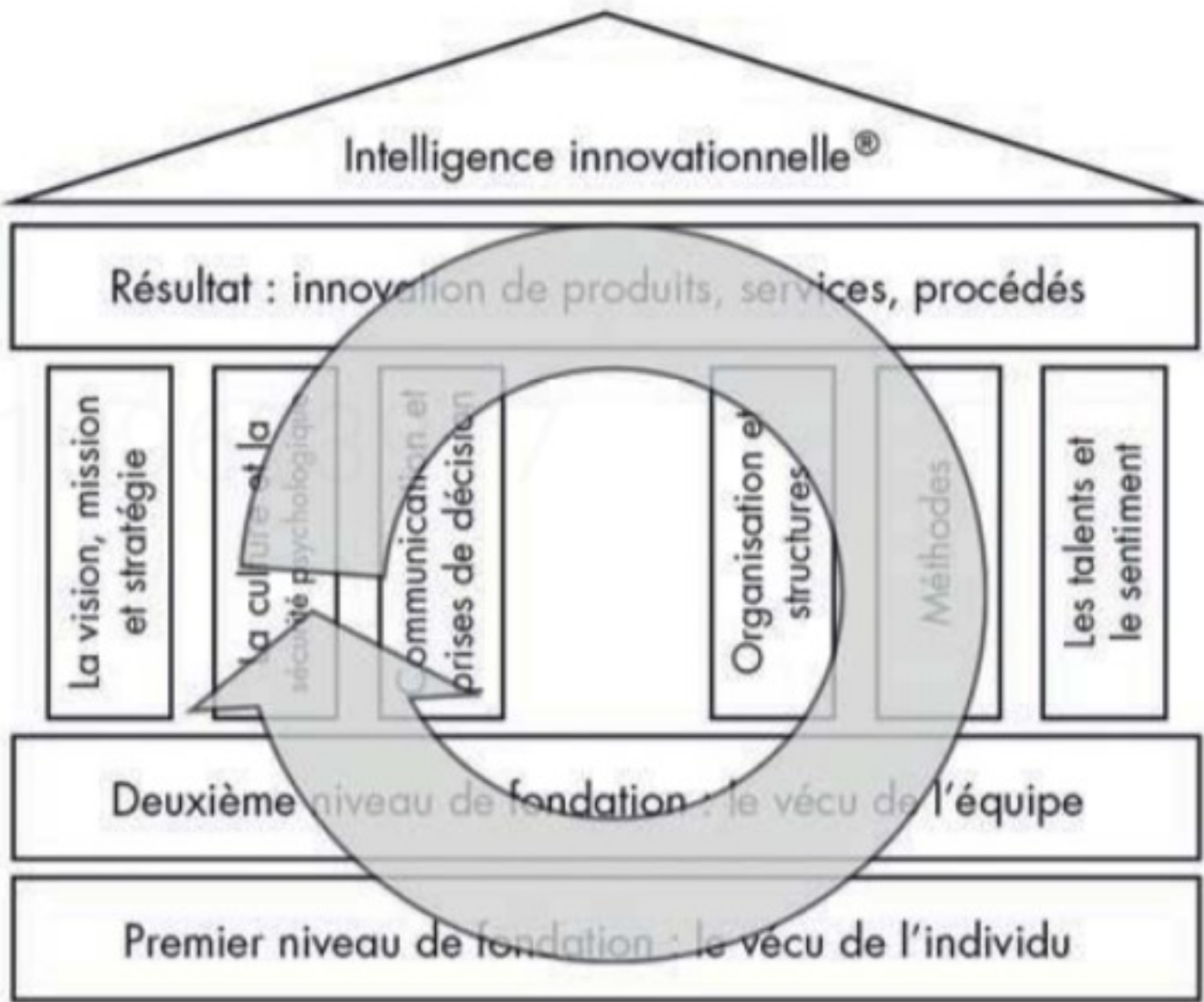
L'analogie avec un « temple » n'est pas fortuite. La gravité entraîne l'écoulement des pluies le long des piliers, les poussières, l'érosion et les attaques des pigeons qui ne se contentent pas de s'attaquer aux piliers, mais affectent aussi les fondations. Ça se traduit par des changements dans les comportements individuels et d'équipe, qui ne sont pas toujours positifs.

Ne pas prendre en compte la situation des piliers risque à terme de faire s'écrouler le temple, et donc l'agonisation de l'organisation qu'il représente.

L'ESSENTIEL

- Le système « intelligence innovationnelle® » est composé de deux fondations et de six piliers.
- Une boucle de rétroaction permanente existe entre l'individu et son impact sur le reste de l'organisation pour l'innovation.
- On peut mesurer aussi bien son niveau d'évolution individuel que d'équipe.
- Un outil de surveillance permet d'évaluer l'état du système.
- L'innovation est l'affaire de tous, sans exception.

En clair, l'innovation finira par disparaître si les fondations et les piliers n'évoluent pas au moyen d'une boucle de rétroaction globale (une version revisitée du feedback, si cher à mon collègue et ami Stéphane Moriou³).



Le pilier « talents et sentiment » contient un outil de suivi qui, lorsqu'il est utilisé régulièrement, permet d'évaluer l'état global du système et de sa performance. Dans certains passages, je ferai une distinction essentielle entre collaborateur, manager et dirigeant. Point important à retenir : un manager est aussi un collaborateur, il hérite de la majorité des caractéristiques de celui-ci, mais il se voit attribuer certaines contraintes supplémentaires. Quant au dirigeant, il hérite aussi des caractéristiques des collaborateurs et des managers, avec en plus des contraintes qui lui sont propres. Une innovation est le plus souvent le fruit d'une équipe alignée, pas le fait d'un seul individu. Si tout individu possède la capacité d'innover, sa position dans l'entreprise et son environnement lui donneront toujours la possibilité de contribuer à l'innovation. □

³ : Stéphane Moriou, Feedback : le pouvoir des conversations, Dunod, 2023

Partenariat Airbus lance un nouveau programme

Airbus Industrie renforce son partenariat avec Dassault Systèmes et lance un programme de nouvelle génération pour ses avions en s'appuyant sur les technologies de son partenaire.

Dassault Systèmes et Airbus prolongent leur partenariat stratégique de longue date. Ce nouvel accord place la plateforme 3DEXPERIENCE au cœur de la gestion du cycle de vie de tous les futurs programmes d'Airbus, qu'ils soient civils ou militaires. Ce déploiement couvrira l'ensemble de la chaîne de développement des avions et hélicoptères d'Airbus. Plus de 20 000 utilisateurs, incluant également des partenaires et fournisseurs, pourront collaborer de manière plus fluide, en s'appuyant sur les jumeaux

virtuels, que ce soit en interne ou via un cloud souverain. L'objectif : accélérer les cycles de développement, anticiper les besoins industriels, améliorer l'efficacité de production et le support après-vente, tout en réduisant les coûts.

Un levier majeur

La digitalisation renforcée est un levier important pour Airbus, afin d'augmenter la cadence de production des avions commerciaux, concevoir les plateformes du futur en phase avec

les objectifs de décarbonation, ou encore anticiper les solutions de défense et de sécurité de demain. De plus, Airbus pourra tirer pleinement parti des technologies d'intelligence artificielle générative, ainsi que des avancées scientifiques en science des matériaux, en modélisation, simulation, production et optimisation des systèmes d'exploitation en ouvrant de nouvelles perspectives pour imaginer, concevoir et fabriquer les expériences qui façonneront le futur de l'industrie aéronautique. □ **B.G**

LA PLATEFORME 3DEXPERIENCE

La 3DEXPERIENCE (3DX) est une plateforme d'innovation métier qui fournit aux entreprises une vision globale en temps réel de leur activité et de leur écosystème. Elle permet de connecter les personnes, les idées, les données et les solutions au sein d'un environnement collaboratif unique, afin que les entreprises de toutes tailles puissent innover, produire et vendre de manière totalement inédite. La plateforme fournit une version unique des données permettant de centraliser les processus basés sur les résultats et de regrouper toutes les activités au même endroit. Elle permet la collaboration en toute sécurité des personnes, des équipes, des services et du personnel externes pour transformer les idées en produits, services et expériences innovants. Sur la plateforme 3DEXPERIENCE, la modélisation (CAO) et la simulation (IAO) sont regroupées dans une fonctionnalité unifiée appelée MODSIM, qui rassemble les services d'ingénierie et de simulation, ce qui permet de bénéficier d'une véritable expérience de jumeaux virtuels et de faire enfin du virtuel une réalité. MODSIM permet aux entreprises de faire de la simulation une partie intégrante de leur pratique de conception, ce qui réduit les risques, les tests physiques, les coûts, et diminue voire supprime le recours au prototypage.

Jumeaux Numériques



Souveraineté

Le programme Confiance.ai rebondit

Avant tout destinée à promouvoir une IA de "confiance", une fondation poursuit les activités du programme de recherche Confiance.ai. Parallèlement à la diffusion de ses résultats, l'objectif est de créer un label de fait.

ancé en 2021 pour une durée de quatre ans et financé à hauteur de 30 M€ dans la cadre du programme France 2030, le programme Confiance.ai associait une vingtaine de partenaires industriels et académiques. Il avait pour objectif de formaliser des méthodes et des outils destinés à rendre l'IA plus robuste, plus explicable et plus sécurisée pour faciliter son utilisation en particulier dans des systèmes critiques. Le programme s'est clos en septembre 2024. Le 9 avril dernier, ses membres ont annoncé la poursuite de ses activités sous une forme différente.

La première et plus importante suite du programme prend la forme d'une fondation, à but non lucratif, qui n'est plus financée par France 2030, dédiée à la diffusion, à la maturation et à la promotion en Europe des méthodes et des composants logiciels open source développés au cours du programme. « Il s'agira notamment de gérer un portefeuille de composants open source destinés à aider les activités d'ingénierie utilisant de l'IA », précise Nicolas Rebierre, manager de l'European Trustworthy AI Foundation. Cette fondation comprend pour l'instant six anciens membres du programme, à savoir Air Liquide, Naval Group, Sopra Steria, Safran et Thales, côté industriel, et SystemX côté recherche. Ouverte à d'autres acteurs, européens ou internationaux, elle a vocation à stimuler l'utilisation des méthodologies et outils jusqu'à en faire un standard de fait. « La fondation étudiera avec soin l'adhésion de nouveaux membres », souligne Nicolas Rebierre.

Si les composants logiciels sont développés dans une logique open source, leurs accès peut-être restreint en fonction de propriétés intellectuelles spécifiques. « Plusieurs cas de figure sont



Nicolas Rebierre, manager de l'European Trustworthy AI Foundation



Raphaël Braud, responsable de l'équipe sciences des données et IA

possibles. Le programme référence des composants antérieurs à 2021, qui conservent leurs règles de propriété. Pour ceux développés par le programme, la copropriété est de mise, en fonction de l'investissement des partenaires. La logique est de consolider cette propriété pour éviter un morcellement. La même approche s'appliquera aux composants à venir », détaille Nicolas Rebierre. L'utilisation de ces outils facilitera l'intégration des composants d'IA dans les systèmes en maîtrisant les risques. Globalement, la fondation imagine à terme l'adoption d'un label « AI Trust ».

Le deuxième axe prévu porte sur la continuation des travaux de recherche en particulier sur l'IA générative. Un financement public de 10 M€ est prévu sur quatre années. SystemX portera le projet. Des contacts avec le monde académique sont en cours. « Nous sommes en relation avec des labos européens et canadiens entre autres », décrit Nicolas Rebierre. Ce volet a pour but de lever les verrous technologiques encore nombreux de ces technologies, en particulier avec l'IA générative. Raphaël Braud, responsable de l'équipe sciences des données et IA chez SystemX décrit : « Cette technologie sert par exemple à automatiser la génération des annotations. Il s'agit de comprendre com-

ment se propagent alors les attributs de confiance. » Les défis ne se limitent pas à l'IA générative. Ils comprennent toujours les interactions homme-machine comme l'IA hybride. « L'un des objectifs majeurs sera d'articuler les avancées scientifiques avec la méthodologie pour garantir une approche de bout en bout », ajoute Raphaël Braud.

Le dernier et troisième volet vise, au même titre que la Fondation, à développer la mise en œuvre des résultats du programme. Il s'appuie sur la promotion d'offres commerciales déjà disponibles basées sur les résultats du programme. Airbus Protect propose une offre plus destinée à faciliter la gouvernance dans l'utilisation des composants d'IA et le respect de l'IA Act. Autre partenaire, Sopra Steria propose déjà une offre dédiée à l'aide à la mise en œuvre pour des activités d'ingénierie. Côté formation, un mastère « IA de confiance » a été lancé par CentraleSupélec Exed et l'IRT SystemX pour les futurs ingénieurs. □

P. Br

UNE MÉTHODOLOGIE PUBLIQUE

Le programme Confiance.ai publie les résultats de ses avancées en ligne (<https://bok.confiance.ai>). Les documents peuvent être filtrés en fonction de sa fonction (ingénieur data, ingénieur sécurité...). Ils décrivent par exemple la succession des étapes à appliquer et les spécialistes ad hoc pour mettre au point un système critique basé sur un composant d'apprentissage machine. Il sera enrichi avec de nouveaux développements.

Cloud

Atlassian déploie Rovo dans ses environnements cloud

Durant Team'25, Atlassian a annoncé le déploiement de Rovo, son écosystème d'IA agentique, à tous ses clients. Conçu comme un coéquipier intelligent, Rovo entend transformer la collaboration en entreprise en connectant données, outils et processus à travers une plateforme unifiée, pilotée par le Teamwork Graph.

Lors de la conférence Team'25 organisée à Anaheim du 8 au 10 avril, Atlassian a franchi une étape en annonçant que Rovo, son écosystème d'intelligence artificielle agentique lancé l'an dernier, serait désormais accessible à l'ensemble de ses clients. Cette extension, qui débute avec les offres Premium et Enterprise, s'étendra prochainement aux abonnements Standard. Rovo ne constitue pas simplement une suite d'outils intégrés à Jira, Confluence et Jira Service Management : c'est un pivot technologique qui redéfinit le rapport entre les équipes, les flux d'information et les systèmes d'aide à la décision.

Cette généralisation s'appuie sur la montée en puissance d'une plateforme unifiée, orchestrée par le Teamwork Graph. Cette couche de données intelligente cartographie de manière dynamique les entités métier (qu'il s'agisse d'objectifs stratégiques, de messages, de projets ou de collaborateurs) pour les relier via plus de dix milliards d'objets connectés. Ce graphe relationnel, qui s'apparente à une fondation cognitive pour l'ensemble des services Atlassian, permet à l'intelligence artificielle de s'exécuter dans un espace interprétatif contextualisé, où chaque action d'un agent repose sur une compréhension dynamique des enjeux métier, des contraintes opérationnelles et des dépendances systémiques.

Quatre composants

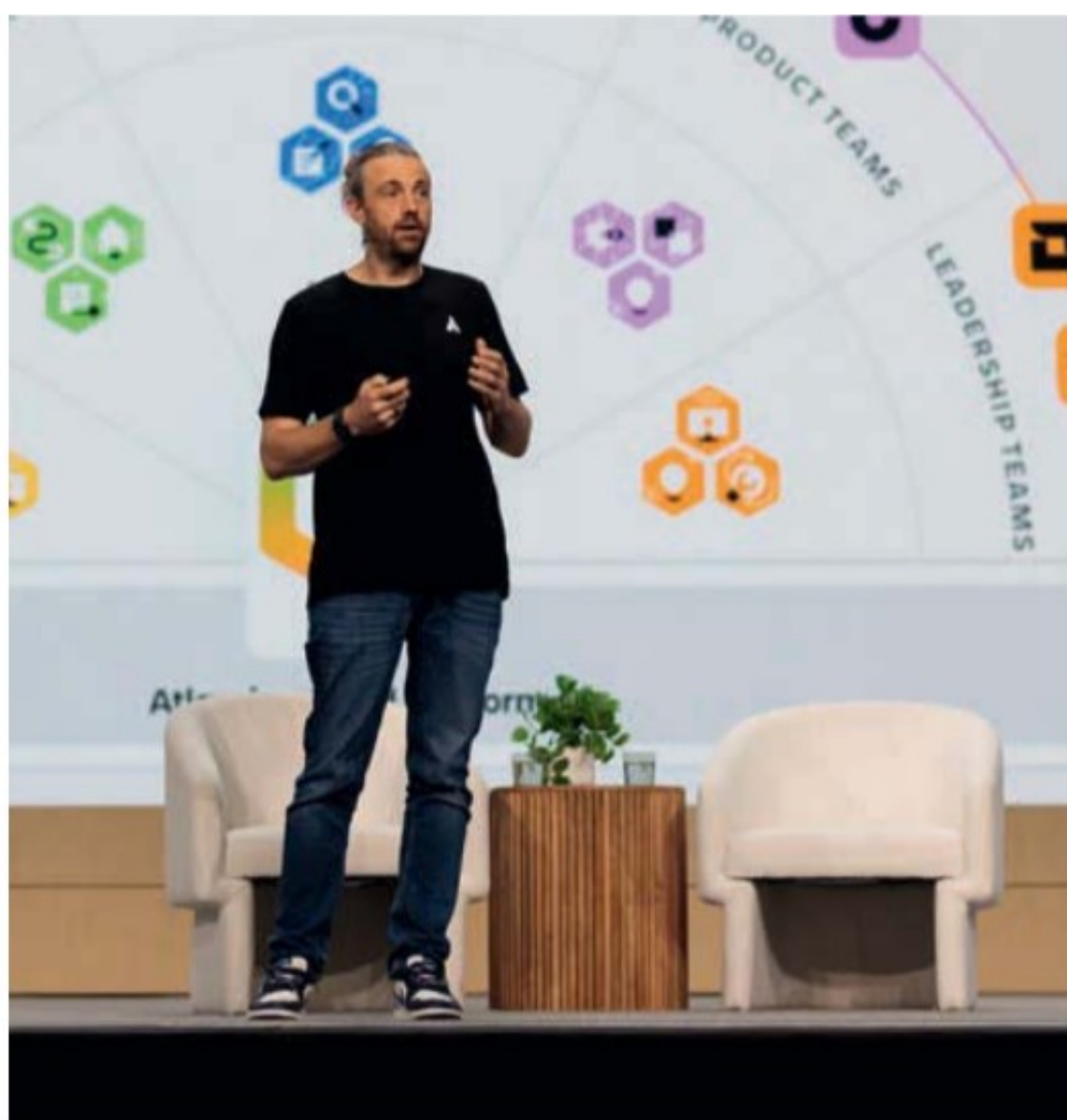
La montée en puissance de Rovo repose sur trois composants principaux, désormais complétés par une quatrième brique stratégique : Rovo Studio. Rovo Search, doté de 50 connecteurs vers des systèmes tiers comme Gmail, OneDrive, Notion ou encore des référentiels internes propriétaires, est conçu pour offrir un moteur de recherche sémantique multi-source, capable de restituer des réponses contextualisées, filtrées selon les droits d'accès et indexées presque en

temps réel. Rovo Chat, quant à lui, s'inscrit dans la logique d'un assistant conversationnel continuellement informé par le Teamwork Graph : il conserve l'historique contextuel des échanges, peut exécuter des actions dans les outils Atlassian ou dans des environnements tiers (Slack, Google Workspace, Microsoft Teams), et deviendra bientôt un outil de synthèse décisionnelle avancée via sa fonctionnalité Deep Research.

Ce socle est prolongé par Rovo Agents, une nouvelle génération d'agents IA à la fois autonomes et spécialisés. Ils sont conçus pour interagir de manière proactive avec les utilisateurs et les processus. Ces agents intègrent des objectifs, une base de connaissances métier et des capacités d'action dans les systèmes d'une entreprise. Ils peuvent être configurés ou développés sur mesure, et s'intègrent à des métiers critiques. Il peut ainsi s'agir de l'analyse de retours d'utilisateur, de la génération de tickets structurés à partir de notes de réunion, ou de la détection automatisée de goulets d'étranglement dans des flux de travail de production. Dans un environnement de développement, un agent de revue de code comme Code Reviewer Agent est ainsi capable de vérifier une demande d'extraction en tenant compte des exigences métier trouvées dans Jira, tout en évaluant la couverture fonctionnelle en fonction des risques techniques.

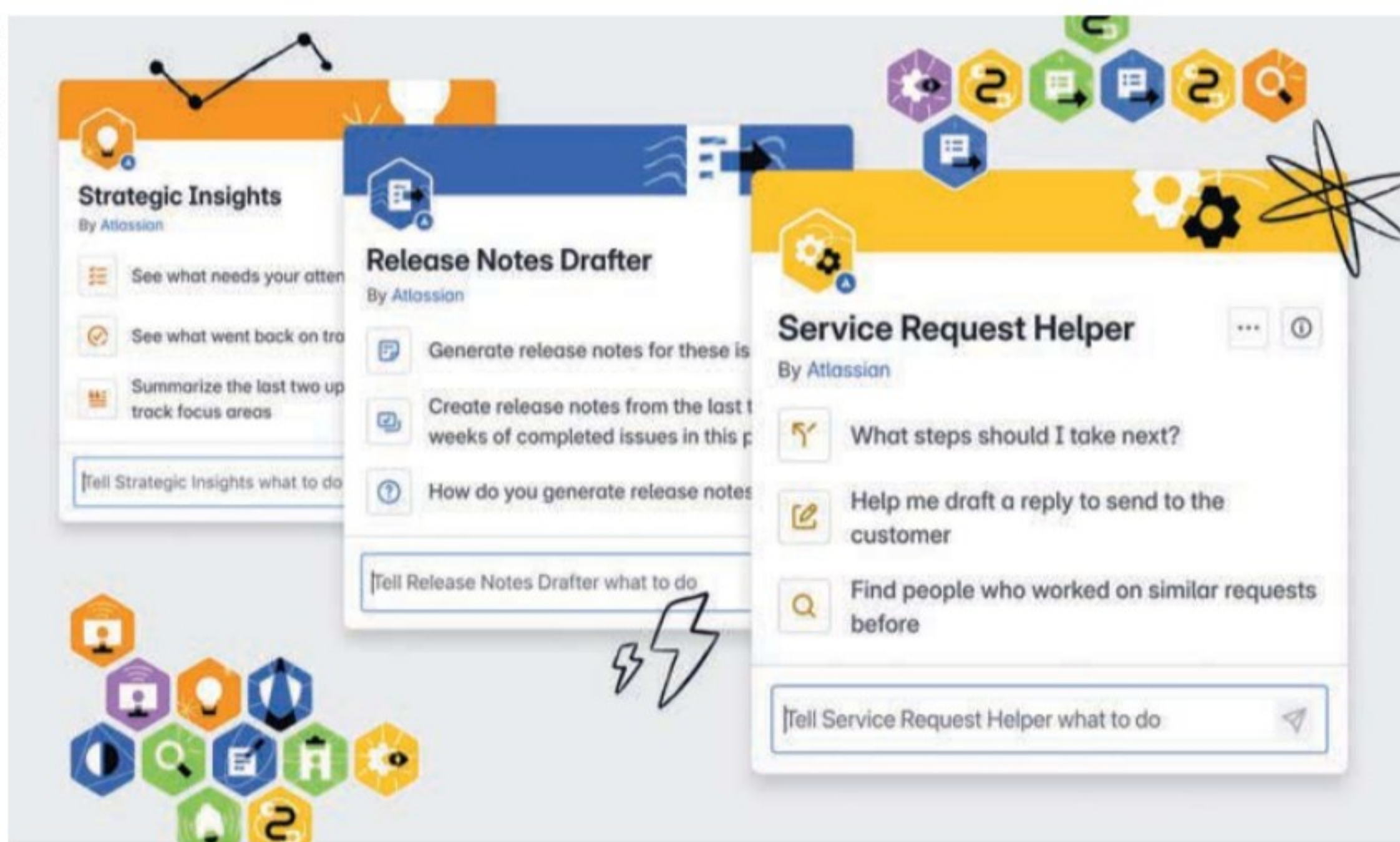
Conception et déploiement d'agents IA avec Rovo Studio

Rovo Studio vient compléter cet écosystème en offrant un environnement intégré pour la conception, le déploiement et la gestion d'agents sur mesure. L'approche low-cod et no-code permet aux équipes métiers ou IT non techniques de définir des comportements, des règles, des modèles de données et des processus sans écrire de code, tout en autorisant



Michael Cannon-Brookes, PDG et cofondateur d'Atlassian, lors du discours d'ouverture des fondateurs, à la conférence Team'25

© Atlassian



Près de 2 000 agents Rovo ont déjà été intégrés aux flux de travail des clients pour réinventer la manière dont les équipes collaborent, toutes disciplines confondues.

les développeurs à injecter du code natif via Forge pour enrichir l'expérience utilisateur ou créer des compétences. Cette capacité à modéliser des objets métiers réels, à automatiser des règles d'interaction, à créer des hubs de contenu interactif, inscrit Rovo dans une logique de plateforme agentique configurable, où l'intelligence artificielle devient un vecteur de gouvernance distribuée.

Dans cette dynamique, Atlassian structure ses solutions autour de deux nouvelles collections. « Strategy Collection » est dédiée à l'alignement entre la planification et l'exécution opérationnelle. Elle s'articule autour de trois solutions : Focus pour le suivi des objectifs et des ressources en temps réel ; Talent pour la planification capacitaire en lien avec les projets prioritaires et Jira Align afin de piloter les transformations de projets. L'intégration avec Rovo permet aux utilisateurs d'identifier, dans un même espace, les zones à risque, les ressources sous-exploitées et les initiatives mal alignées. L'agent Talent Finder permet, par exemple, d'explorer dynamiquement les compétences disponibles au sein de l'entreprise et de projeter leur adéquation avec des besoins émergents, tout en prenant en compte la charge projet et les arbitrages budgétaires.

Cette approche orientée stratégie est doublée d'un dispositif de soutien transversal avec la Teamwork Collection. Cette dernière suite propose un tissu unifié entre Jira (comme système d'enregistrement du travail), Confluence (comme espace de collaboration documentaire) et Loom (comme canal de communication vidéo structurant), le tout enrichi d'agents IA dédiés à la collaboration opérationnelle. La recherche d'efficacité collaborative est soutenue par une intégration verticale entre les objets de connaissance et les actions exécutables. Un échange capté dans Loom peut donc être transcrit, analysé par un agent de réunion, enrichi de liens dynamiques vers des pages Confluence, puis transformé en actions concrètes dans Jira.

Une infrastructure cloud renforcée

La nouvelle application Customer Service Management constitue une autre démonstration de cette volonté de casser les silos historiques. En se basant sur le Teamwork Graph, elle relie directement les interactions clients à l'ensemble des composants produit, aux historiques d'incident, aux déploiements en cours et aux équipes techniques concernées. Cette connectivité permet à l'agent de support de contextualiser immédiatement une demande client,

d'interagir avec un agent de développement pour créer une branche correctrice, et de revenir vers l'utilisateur final avec un état d'avancement automatisé et personnalisé. L'ensemble du cycle de résolution est orchestré sans rupture entre les équipes front-office et les équipes produit, grâce à une architecture où l'IA ne remplace pas l'humain mais augmente sa capacité d'intervention.

Enfin, Atlassian a annoncé le renforcement de son infrastructure cloud avec deux nouveaux environnements dédiés aux exigences réglementaires les plus strictes. Atlassian Government Cloud, déjà autorisé FedRAMP Moderate et en accès anticipé pour les agences fédérales américaines, constitue une réponse aux enjeux de souveraineté numérique pour le secteur public. Atlassian Isolated Cloud, attendu en 2026, proposera un cloud privé virtuel managé par Atlassian, incluant stockage, calcul, réseau et bases de données dédiées, dans un environnement conçu pour les entreprises ayant des exigences de souveraineté, d'isolation ou de conformité très élevées. Cette solution viendra compléter les investissements dans Atlassian Guard, la gestion proactive des accès et la résidence des données, en intégrant les certifications GRPD, ISO27001, SOC2, et bientôt les standards FedRAMP High et DoD IL5.

Dans un environnement où les entreprises cherchent à concilier résilience opérationnelle, vitesse de livraison et gouvernance distribuée, Atlassian a voulu démontrer que l'entreprise ne se limite plus à fournir des outils collaboratifs. En effet, l'entreprise entend désormais proposer un cadre d'IA d'entreprise, contextualisé et interopérable, capable de redéfinir la manière dont les équipes planifient, agissent et apprennent. Comme le résume Jamil Valliani, en charge de l'IA chez Atlassian : « Rovo n'est pas un assistant, c'est un coéquipier. Et c'est toute la différence. » □

M.C

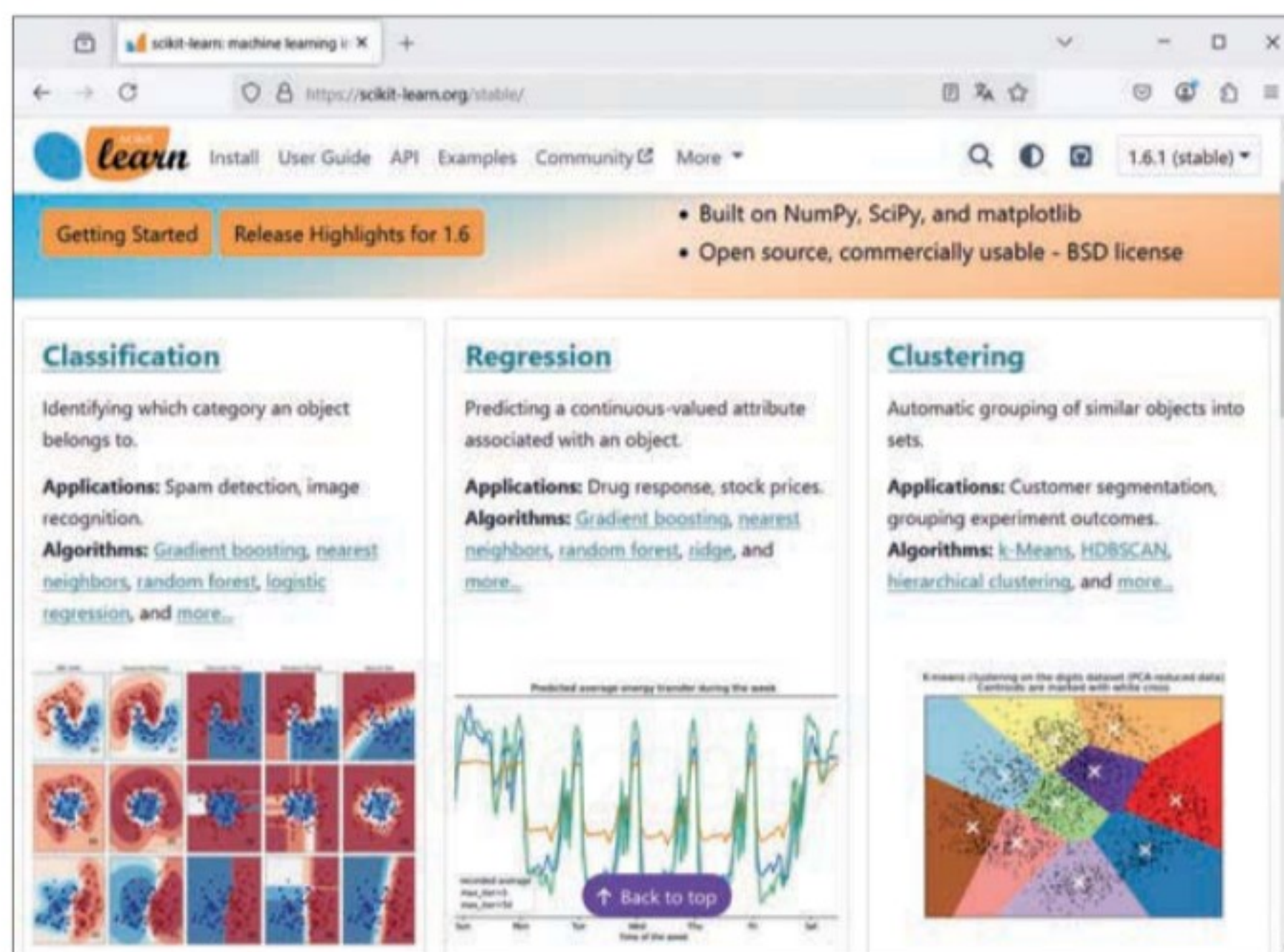
Python

Les bibliothèques pour le ML et le DL

Ce n'est pas un hasard si le langage Python est plébiscité dans les projets de data science et d'IA, particulièrement ceux concernant le machine learning et le deep learning.

Outre sa puissance et sa souplesse, il a l'avantage de proposer des bibliothèques adaptées et efficaces sur ces sujets. Numpy, Scipy, PyTorch, Pandas et quelques autres sont à l'origine de ce succès. Nous allons dans cet article faire un petit tour d'horizon des bibliothèques disponibles.

Le machine learning est un domaine de l'intelligence artificielle qui confère aux machines la capacité d'apprendre et de s'améliorer par elles-mêmes. C'est aussi un sous-domaine de la data science puisque les données sont collectées, nettoyées et visualisées afin de construire des modèles efficaces. Plusieurs systèmes créés grâce au machine learning sont utilisés pour la classification des images, la data viz (visualisation des données, le NLP – Neuro-Linguistic Programming pour programmation neuro-linguistique –, les chatbots et tutti quanti). Grâce au ML, les développeurs peuvent entraîner les machines à apprendre à partir de leurs propres expériences sans avoir à les programmer explicitement. Le ML peut être divisé en trois types de modèles : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement.



Scikit-Learn est basée sur les bibliothèques NumPy, SciPy et Matplotlib. Son interface permet d'accéder à de nombreux algorithmes de classification, de régression, de clustering, de réduction de dimensionnalité et de sélection de modèles.

Bibliothèques de machine learning

Une bibliothèque de ML est une compilation de fonctions et de routines facilement réutilisables. Les bibliothèques sont indispensables aux développeurs. Elles leur permettent d'écrire plus aisément des programmes complexes. Python est un langage de programmation open source incontournable pour la data science. Il dispose de nombreuses bibliothèques pour effectuer des tâches telles que le calcul scientifique, l'exploration et la manipulation des données, la visualisation, le traitement et l'analyse de texte et bien d'autres encore.

NumPy, l'épine dorsale de la manipulation de données

La polyvalence de NumPy la rend indispensable pour la manipulation de données et les opérations mathématiques. Python n'a pas été développé initialement comme un outil de calcul numérique. C'est l'arrivée de NumPy qui a été à l'origine de l'expansion des capacités de Python avec des fonctions mathématiques sur

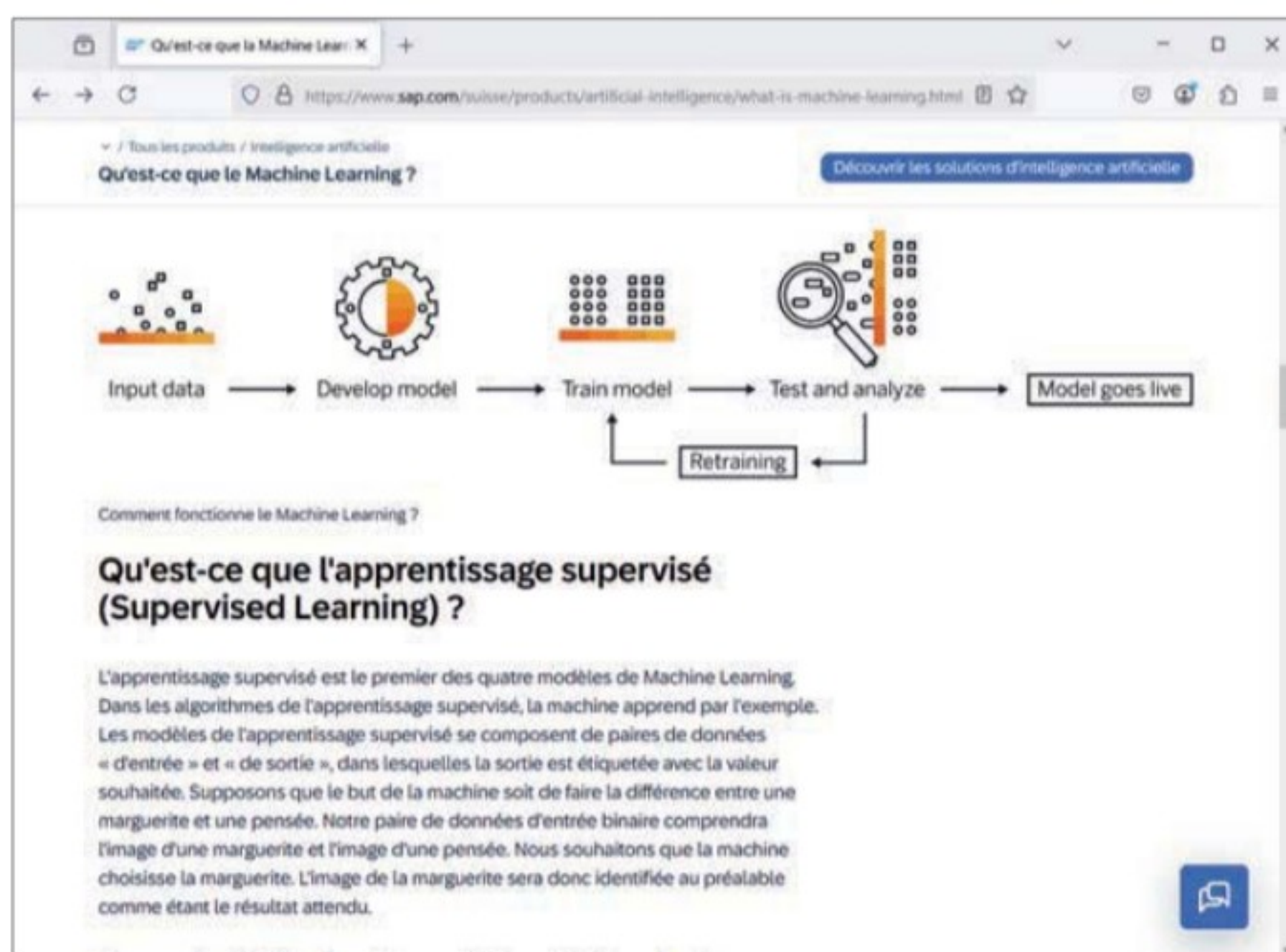
la base desquelles des solutions de ML ont pu être construites. Elle traite l'algèbre linéaire, les calculs matriciels, la génération de nombres aléatoires et tout ce qui aide les développeurs à créer des systèmes intelligents et réactifs. Les fonctions mathématiques de haut niveau sont exécutées sur des tableaux, permettant ainsi aux algorithmes Python de s'exécuter plus rapidement. Les programmeurs peuvent, en cas de problème ou pour poser des questions, se tourner vers la communauté NumPy et partager leurs expériences, voire trouver une solution toute faite répondant à un cas de figure précis. Son nom signifie Numerical Python. Devenue rapidement une librairie incontournable du langage, elle offre la possibilité de faire du calcul numérique et de créer des tableaux de données et des matrices. Elle propose la plupart des fonctions usuelles telles que l'exponentielle, le logarithme et autres sinus ou arc tangente. De plus, elle est optimisée pour les calculs et permet de paralléliser les opérations, c'est-à-dire d'utiliser tous les processeurs de l'ordinateur pour calculer plus vite. Vous pouvez aussi, grâce à cette librairie, intégrer directement du code en C, C++ ou Fortran.

Scikit-learn

Scikit-learn a été initialement conçue en tant qu'extension tierce de la bibliothèque SciPy. C'est aujourd'hui une bibliothèque autonome et l'une des plus populaires sur GitHub. La gamme de composants ML traditionnels qu'elle propose comprend des algorithmes de classification pour la détection des spams et la reconnaissance d'images, des algorithmes de régression pour la prédiction, des algorithmes de regroupement pour la segmentation des clients et des opérations similaires, ou encore la sélection de modèles destinés à améliorer la précision des calculs. Scikit-learn est désormais un complément essentiel aux principales bibliothèques numériques et scientifiques Python. Elle est facilement interopérable avec d'autres outils de la pile SciPy. Cette librairie possède de nombreux modèles avec des paramètres classiques, mais aussi de nombreuses variantes permettant de démultiplier ses possibilités. Scikit-learn offre de nombreuses métriques pour évaluer la qualité de vos modèles. Elle permet de gérer de bout en bout la création d'un algorithme de ML, du formatage à l'évaluation des modèles en passant par le découpage et le paramétrage des datasets, ainsi que l'entraînement et le test des modèles. C'est vraiment LA librairie à utiliser aussi bien pour créer ses premiers modèles de ML que pour des modèles de base nécessitant une prédiction, une classification ou du clustering. Elle permet également de créer des modèles de réseaux de neurones, bien que ce ne soit pas sa tâche de prédilection. Scikit-learn est construite sur Scipy et permet, de fait, d'opérer sur des données qui en sont issues. Vous trouverez, parmi les modèles existants, celui des K-plus proches voisins, le K-means, la régression linéaire (avec de nombreuses variantes), la régression logistique, l'arbre de décision de random forest et le SVM (Support Vector Machine).

LES DIFFÉRENTS TYPES D'APPRENTISSAGE

Dans l'apprentissage supervisé, une caractéristique cible dans un ensemble de données est employée pour dériver le reste des caractéristiques. C'est à partir de cette caractéristique que l'ordinateur apprendra à prédire et à créer des modèles. Il existe deux types d'apprentissage automatique supervisé : la régression et la classification. Dans la régression, la caractéristique cible est une variable continue, tandis que dans la classification, la cible est une caractéristique avec deux classes ou plus qui peuvent être obtenues après l'entraînement du modèle. Il n'y a pas de caractéristique cible dans l'apprentissage automatique non supervisé. Au lieu de cela, l'ordinateur trouve des connexions entre les caractéristiques et les regroupe en classes sur la base de leurs similitudes. Le seul type d'apprentissage non supervisé est le regroupement. Dans l'apprentissage par renforcement, les machines ont tendance à apprendre en fonction des données et de l'environnement. Elles finissent par prendre des décisions sur la base de ces données et interagissent avec l'environnement. Ce type d'apprentissage peut être observé dans le développement de robots et de jeux d'intelligence artificielle comme les échecs ou le jeu de go.



Il existe trois types d'apprentissage en ML : le supervisé, le non-supervisé et l'apprentissage par renforcement. Le supervisé se décline en deux types : la régression et la classification. Le seul type d'apprentissage non supervisé est le regroupement.

Pandas

Pandas offre des structures de données puissantes permettant une manipulation et une analyse transparentes des données. C'est une bibliothèque Python de bas niveau construite sur NumPy et faite pour manipuler facilement des données structurées. Elle permet notamment de créer de nouvelles colonnes, de gérer les données manquantes, de filtrer des données, d'agréger des informations selon des colonnes ou de calculer des métriques telles que la moyenne, la médiane ou des sommes. Elle est basée sur deux types d'objets : les séries, très similaires aux listes en termes de fonctionnement, et les data frames qui sont des tableaux à plusieurs colonnes. Un type d'objet appelé Panels permet de manipuler des objets en trois ou quatre dimensions. De plus, elle facilite la lecture de données provenant de différentes sources : CSV, SQL ou encore texte. Bref, c'est l'outil incontournable pour manipuler des données sur Python. Cette librairie offre également une meilleure vue d'ensemble sur les données. Pandas est surtout employée pour l'analyse et la manipulation des données, ainsi que pour les opérations de ML sous la forme de dataframes. En utilisant ces derniers, les développeurs peuvent aisément avoir une vue d'ensemble des données garantissant ainsi une meilleure qualité du produit obtenu. Les développeurs utilisent cette bibliothèque pour structurer, remodeler et filtrer de grands ensembles de données.

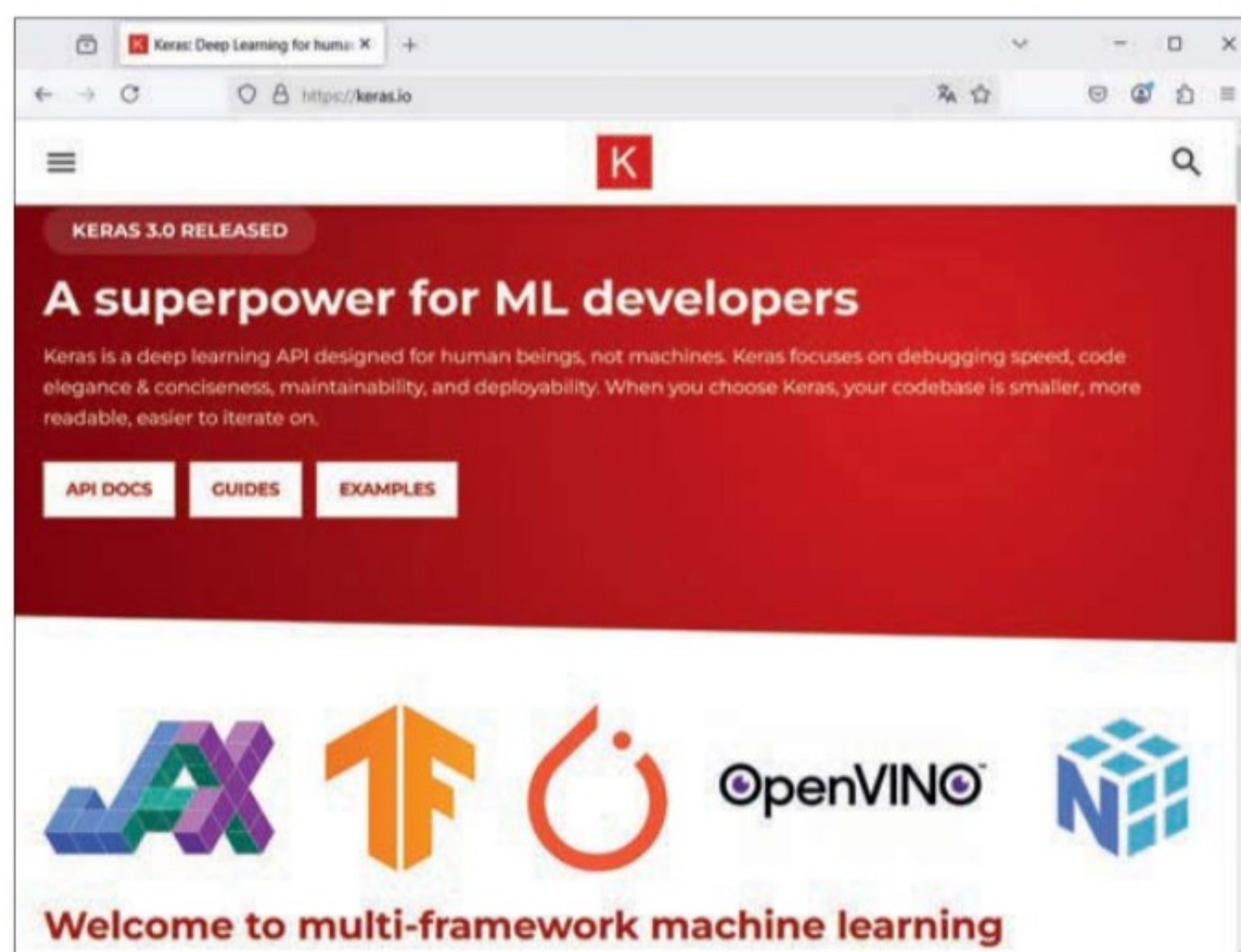
Matplotlib

NumPy, SciPy et Matplotlib ont été conçues afin de donner une alternative ouverte au langage statistique propriétaire Matlab. C'est ce qui explique pourquoi les fonctionnalités

conjuguées de ces bibliothèques sont similaires à celles de Matlab. Néanmoins, les paquets Python sont disponibles gratuitement et bien plus flexibles, en faisant un choix de prédilection pour de nombreux data scientists. Des diagrammes, des schémas 2D et 3D, des graphiques et d'autres outils de visualisation permettent aux scientifiques d'effectuer une analyse détaillée des données. Les développeurs peuvent construire des modèles de ML fiables sur la base de cette analyse. Les traitements portant sur du ML et, plus généralement, sur de la data, impliquent de savoir représenter les données. Cette représentation donne un premier niveau d'information très précieux pour établir un premier filtre entre les données utiles et celles qui ne le sont pas. Les capacités de visualisation de Matplotlib sont très similaires à celles de Matlab. Le caractère open source et gratuit de Python et de Matplotlib a poussé et pousse encore de nombreux data scientists à se tourner vers Python.

Keras

Cette librairie a été conçue pour implémenter des modèles de deep learning. Elle a été développée par François Chollet, un ingénieur de chez Google. Keras facilite notamment la création de réseaux de neurones et peut s'exécuter aussi bien sur des CPU (Central Processing Unit) que sur des GPU (Graphics Processing Unit), permettant ainsi d'accélérer considérablement les calculs. À l'origine, Keras était une plateforme d'expérimentation de réseaux neuronaux profonds, mais elle s'est rapidement transformée en une bibliothèque de ML autonome. Elle dispose d'un ensemble complet d'outils de ML permettant de traiter efficacement les données de type texte et image. Des abstractions de haut



Keras est une API de réseaux de neurones de haut niveau écrite en Python et interfaçable avec CNTK, TensorFlow et Theano.

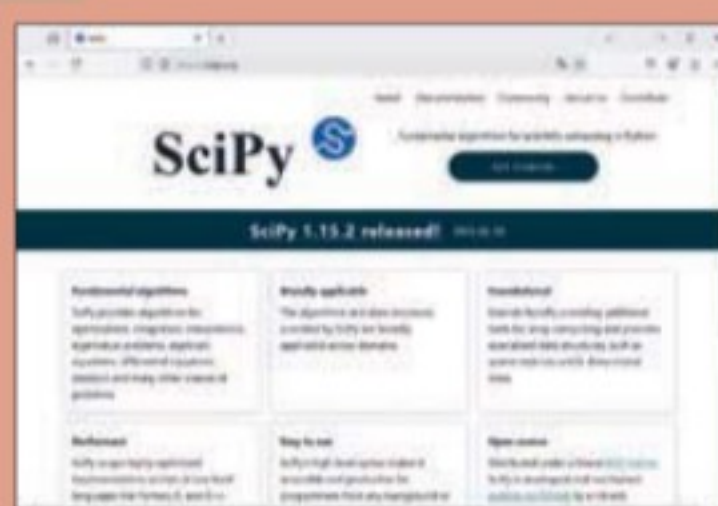
niveau permettent aux programmeurs de créer et d'intégrer des modèles deep learning dans différents backends, afin de maintenir la stabilité des systèmes. Les spécialistes utilisent des modules contenant des modèles prêts à l'emploi. Keras est également très compatible avec d'autres bibliothèques, avec des langages de deep learning bas niveau et des outils tiers.

TensorFlow

TensorFlow est une bibliothèque de calcul numérique open source pour le ML basée sur les réseaux neuronaux. L'équipe de recherche Google Brain l'a créée en 2015 afin de l'utiliser en interne dans les produits Google. Elle est clairement une alternative à Keras, créée aussi par Google. Elle permet de travailler sur des réseaux de neurones et est notamment très bien intégrée avec tout l'écosystème de Google Cloud. Elle peut être employée aussi bien en mode client lourd que client léger pour le web. TensorFlow a la réputation — justifiée — d'être un peu compliquée à prendre en main, mais ce petit inconvénient est largement compensé par sa grande flexibilité. Tout comme Keras, elle peut être utilisée par les CPU et les GPU, mais aussi, et c'est là qu'elle fournira les meilleures performances, sur les TPU (Tensor Processing Unit) développés par Google pour ce type spécifique de tâche. TensorFlow dispose d'un écosystème flexible d'outils et de ressources communautaires qui permettent aux ingénieurs logiciels de mener efficacement des recherches sur le ML et le DL, et de créer et déployer facilement des solutions. TensorFlow

LA STACK SCIPLY

Les bibliothèques Python incluses dans la stack SciPy constituent une famille de paquets destinés au calcul scientifique et technique. Elles forment un ensemble complet d'outils pour le ML. Chaque paquet autonome est adapté à des tâches de données spécifiques, mais fonctionne encore mieux lorsqu'il est utilisé en combinaison avec d'autres outils de la même stack. Parmi les bibliothèques principales, nous trouvons NumPy, SciPy, Scikit-Learn, Matplotlib et Pandas. SciPy est une alternative à NumPy sur laquelle elle est basée. Elle représente un outil essentiel pour effectuer des traitements mathématiques tels que l'interpolation numérique, l'intégration, l'algèbre linéaire ou les statistiques très rapidement, augmentant de fait la vitesse d'intégration des modèles ML et réduisant le temps de développement. La bibliothèque est facile à comprendre et à utiliser, ce qui permet aux spécialistes de se familiariser rapidement avec ses fonctionnalités. SciPy ayant été construite au-dessus de NumPy, elle peut opérer sur ses tableaux en apportant une meilleure qualité et une exécution plus rapide des opérations de calcul.



permet de mettre les modèles ML en mode production sur diverses plateformes, que ce soit dans le cloud ou sur site, via un navigateur ou directement sur un appareil. Chaque nouvelle version apporte son lot de nouveaux outils très utiles pour répondre aux besoins toujours croissants des organismes et des équipes de développement. La preuve, s'il le fallait, du soutien permanent de Google, est le lancement récent de TensorFlow Enterprise qui permet de construire des solutions de ML à grande échelle.

PyTorch

PyTorch est une seconde alternative à Keras. Elle permet elle aussi de lancer des modèles de ML à grande échelle. Elle est disponible dans les différents grands clouds : AWS (Amazon), Google Cloud, Microsoft Azure ou encore Alibaba Cloud. PyTorch est tout à fait adaptée au traitement des données textuelles et au NLP. Elle est l'une des plus grandes bibliothèques de ML développée par le groupe de recherche en IA de Facebook. Elle est couramment employée pour la computer vision, le NLP et d'autres tâches complexes du même genre. Parmi les nombreux avantages offerts par PyTorch, il faut citer le chemin rapide du prototypage à la production, des performances système optimisées et sa grande disponibilité dans le cloud. Le modèle TorchScript accélère la vitesse de développement et est particulièrement puissant pour gérer des projets à un rythme soutenu. De très nombreux systèmes s'appuient sur le côté backend distribué de PyTorch, afin d'optimiser leurs performances lorsqu'ils traitent une grande quantité de données. Elle est accessible à partir de n'importe quel appareil, à n'importe quel moment et à n'importe quel endroit, de telle sorte qu'aucun coût supplémentaire n'est requis pour des outils matériels ou logiciels spécifiques.

Orange3

Cette librairie propose des outils de ML, de dataviz et d'exploration de données. Les scientifiques de l'université de Ljubljana l'ont créée en 1996. Elle est écrite en C++ et a été

conçue particulièrement pour créer des systèmes de recommandation et des modèles prédictifs de haute précision. Cette collection d'outils très variée permet de tester de nouveaux algorithmes ML dans diverses industries, en particulier l'informatique et la biomédecine. Ses widgets comprennent des fonctionnalités très diverses. En plus de se concentrer sur les tâches de visualisation de données, ils aident les développeurs à créer des modèles ML prédictifs qui fournissent des prévisions commerciales précises. Orange3 est incluse dans les programmes de formation scolaire, universitaire et professionnelle. De plus en plus de spécialistes choisissent cette bibliothèque pour fournir efficacement aux clients des solutions de qualité alimentées par le ML.

Theano

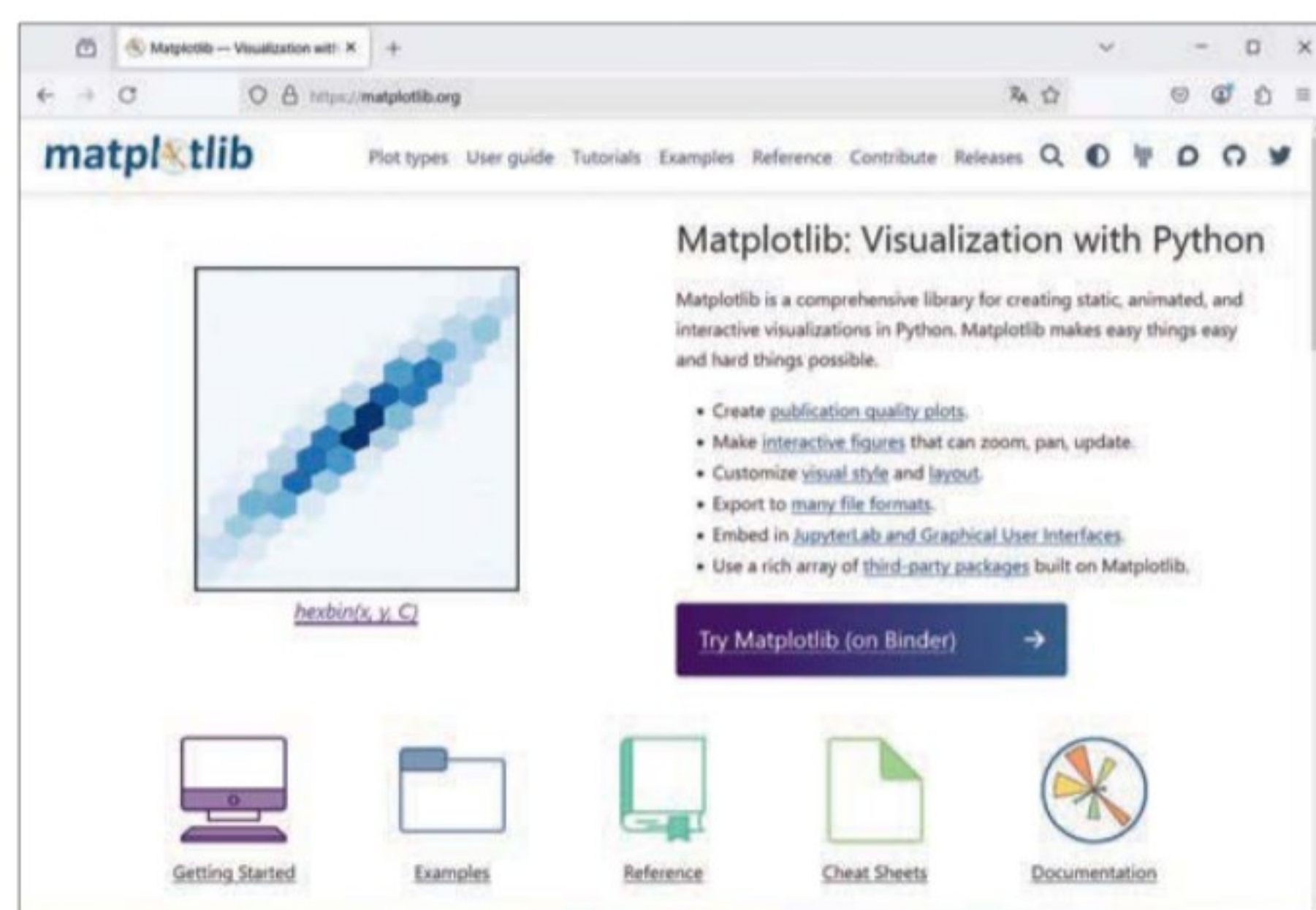
L'Institut des algorithmes d'apprentissage de Montréal (MILA) a créé Theano en 2007 pour manipuler et évaluer diverses expressions mathématiques. Sur la base de ces expressions, cette bibliothèque de ML Python permet de construire des réseaux neuronaux de DL optimisés. Bien que Theano ne soit pas aussi efficace pour le ML que TensorFlow, elle présente néanmoins quelques avantages indéniables. Elle gère de multiples calculs en maintenant ses performances, et permet de réutiliser des morceaux de code pour des fonctions similaires, réduisant de fait le temps de développement du modèle. La librairie affiche des performances élevées sur les architectures CPU et GPU, ce qui permet de gagner du temps lors du développement. Theano peut également définir les expressions instables et les remplacer, garantissant ainsi une meilleure qualité des systèmes.

NetworkX

NetworkX est dédiée à l'analyse, à la création et à la visualisation de graphes et de réseaux complexes. Elle offre des outils puissants pour étudier les structures des réseaux. C'est l'outil idéal pour le calcul des centralités, la détection de communautés ou l'optimisation des chemins. Son interface conviviale permet aux chercheurs et aux développeurs de modéliser des systèmes variés, qu'ils soient sociaux, biologiques ou technologiques. Cette bibliothèque s'avère indispensable pour quiconque s'intéresse à la théorie des graphes et à l'analyse de données relationnelles. Elle intègre des fonctionnalités pour l'exportation et l'importation de réseaux dans différents formats.

PyCaret

PyCaret est une version Python de la bibliothèque d'IA et de ML Caret de R. Conçue pour exécuter les tâches standards dans des projets ML, elle automatise et simplifie l'évaluation, la comparaison et le réglage des modèles. La vérification des modèles sur un ensemble de données de classification ou de régression peut, par exemple, se faire avec une seule commande. ☐ **T.T**



Matplotlib est une bibliothèque complète permettant de créer des visualisations statiques, animées et interactives en Python.

Guerre électronique

Les responsables IT inquiets des tensions géopolitiques

Les responsables IT français sont majoritairement inquiets face au contexte international et à son impact sur leur organisation.

Armis Lab, l'entité de recherche d'Armis, spécialiste de la gestion de l'exposition aux menaces et de solutions de sécurité, vient de livrer la troisième de son étude Armis cyberwarfare report, qui analyse cette fois le rôle de l'intelligence artificielle dans la cyberguerre. L'étude révèle que l'escalade de l'instabilité géopolitique, combinée aux progrès rapides de l'IA, alimente une nouvelle ère de cyberguerre aux enjeux critiques.

Des frontières floues

Le rapport constate que la frontière entre un conflit classique et la cyberguerre est de plus en plus floue. Ainsi, 87 % des personnes interrogées dans différents pays dont la France sont inquiètes des répercussions d'opérations de cyberguerre sur leur organisation. Une augmentation spectaculaire comparativement aux 46 % de personnes qui ne sentaient pas ou peu concernées dans la vague précédente de l'étude. Selon Armis, le principal moteur de cette croissance provient de l'intelligence artificielle qui a dramatiquement augmenté l'échelle et la sophistication des attaques. Une étude de Capgemini Institute dénombre ainsi que 45 % des entreprises ont subi des attaques sophistiquées par deepfake au cours des deux dernières années. Un autre sondage, réalisé pour le compte de Checkpoint, enregistre une augmentation de 135 % des attaques alimentées par l'IA en 2023. Google constate pour sa part que le taux de succès des attaques de phishing assistées par IA a augmenté de 78 %. Plus spécifiquement, 85 % des décideurs IT français estiment que les tensions géopolitiques mondiales intensifient les menaces de cyberguerre — soit le pourcentage le plus élevé parmi tous les pays étudiés par Armis — tandis que 82 % se disent inquiets de l'impact généré sur leur organisation. 29 % ont déjà signalé un tel acte aux autorités indiquant une menace imminente.

L'intelligence artificielle générative est clairement dans le collimateur des responsables IT dont les deux tiers (64 %) s'accordent à penser que la Gen AI est un défi au statu quo géopolitique. Ils sont 74 % à s'inquiéter sur le fait que les attaques renforcées par l'intelligence artificielle menacent très sérieusement leur sécurité. Ce sentiment est particulièrement ressenti en Europe et en Grande-Bretagne. Ils sont surtout inquiets des attaquants soutenus par des États mettant en avant principalement la désinformation, bien qu'une forte majorité (77 %) déclarent avoir mis en place des outils pour détecter et

What do you consider to be the biggest cybersecurity threat facing your organization today, if anything? *Select up to 3*

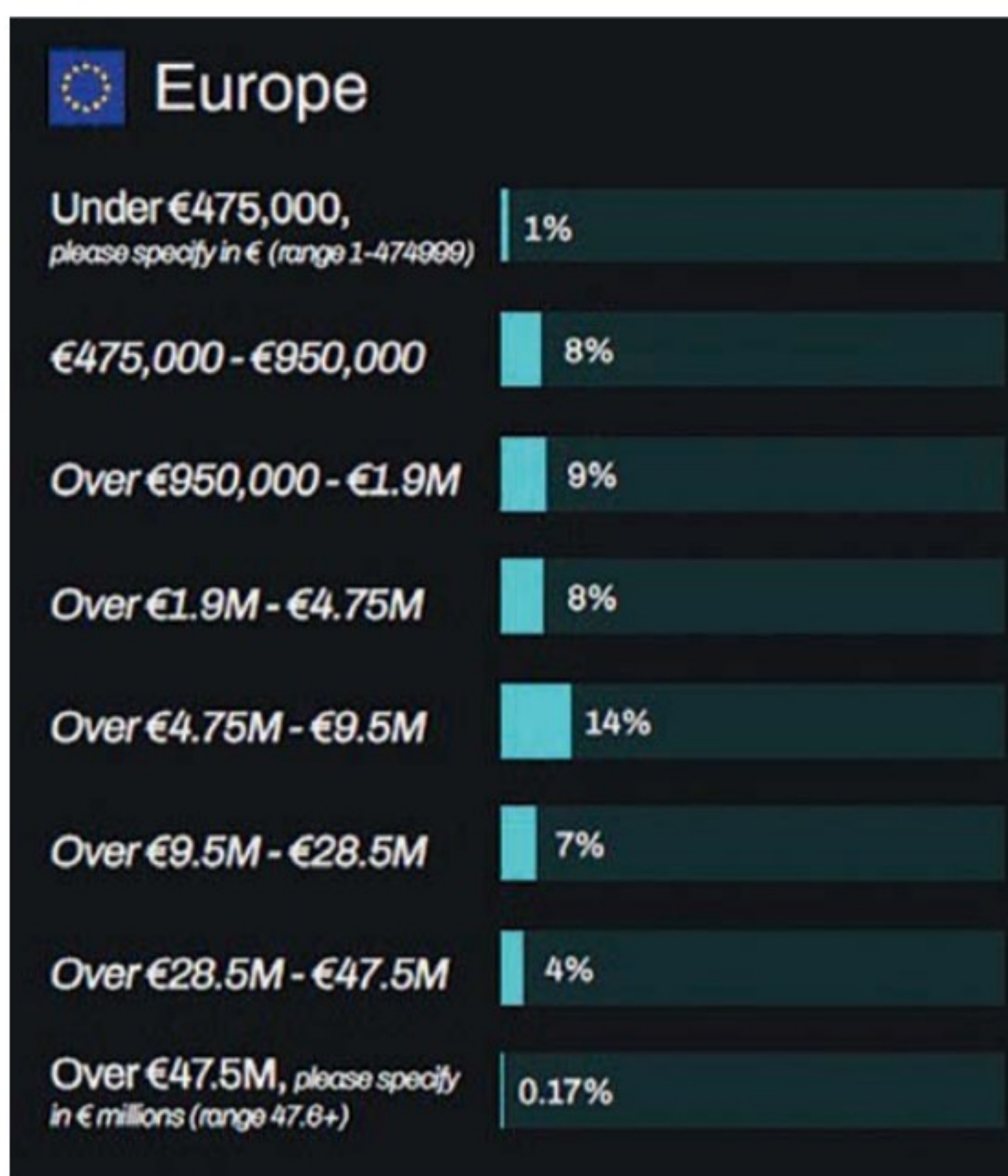


Les principales craintes des RSSI

contrer des attaques utilisant l'intelligence artificielle. L'étude constate d'ailleurs une montée du combat « IA contre IA », mais les entreprises indiquent cependant manquer d'expertise et de ressources pour mener à bien ce combat.

Des ennemis connus

Pour plus de 70 % des personnes interrogées, les attaques proviennent de Russie ou de Chine. Viens ensuite la Corée du Nord à 40 %. Ces chiffres varient selon les zones géographiques, mais la plupart craint la capacité de ces États à déclencher une cyberguerre à grande échelle. Ils sont cependant un peu plus de la moitié (53 %) à penser que leurs



Les sommes moyennes payées par les entreprises lors d'une attaque gouvernements ont la capacité de les défendre en cas de cyberguerre. De plus, ils sont un tiers à penser que leur organisation a fait les efforts pour résister à de telles attaques.

Des fonds insuffisants

Si les efforts sont reconnus, ce n'est pas le cas des budgets. Seulement 35% des décideurs interrogés pensent fortement que leur entreprise a alloué le budget suffisant

pour la cybersécurité, incluant le personnel et les processus. Le plus mauvais élève est le secteur public qui arrive juste derrière le secteur de la santé. Cela sonne comme une évidence vu les attaques réussies de rançongiciel. Trois quart pensent que les attaques hostiles dépassent le cadre militaire et que les institutions et la presse sont largement visées. Cela entraîne une inflation des coûts des brèches. Selon le rapport, le coût moyen aux USA est de 4,88 M\$, en augmentation de 10% sur un an. Les attaquants semblent se concentrer cependant sur les secteurs d'activité où ils peuvent maximiser leurs gains, comme le secteur automobile.

Des réponses réactives

Le rapport constate aussi que les entreprises sont encore dans une posture réactive face à un incident (58%). Les principaux risques mis en avant sont la sécurité du travail à distance, des budgets insuffisants pour mettre à l'échelle les opérations de sécurité et une faible intelligence des menaces afin d'identifier et de prioriser les risques. De plus, 85% des répondants indiquent que les offensives évitent régulièrement les outils de sécurité démontrant la limite des mesures mises en place. Vient en tête le « spear phishing ».

Une demande forte pour l'IA

Pour lutter à armes égales, les décideurs souhaiteraient avoir des outils d'IA pour les assister (94%) qui semblent avoir déjà prouvé leur valeur pour la cybersécurité. Pour éviter que l'écart ne se creuse entre attaquants et défenseurs, il va devenir nécessaire d'investir dans l'intelligence artificielle. ☐

B.G

LA CYBERGUERRE DES INTELLIGENCES



Technicien de maintenance data center

Une nouvelle formation en alternance proposée par l'AFPA et ses partenaires pour un métier en tension

Une formation de 19 mois en alternance de technicien de maintenance data center a été lancée par l'Agence nationale pour la Formation Professionnelle des Adultes (AFPA), et les Plombiers du Numérique, dispositif d'insertion pour les jeunes éloignés de l'emploi et les bénéficiaires du RSA.

« **C**e projet pallie les besoins de compétences d'une des filières les plus dynamiques de l'économie française, explique Florian du Boÿs, initiateur des Plombiers du Numérique (qui ont formé plus de 1500 jeunes depuis six ans, avec un taux de réussite de 75 à 80 %, dans des métiers techniques, dont plus de 500 sur les métiers du datacenter). *Nous nourrissons de grandes ambitions dans le développement national de ce dispositif, grâce au réseau de l'AFPA.* »

Un métier jeune et peu connu

La première session a débuté en avril, avec 12 personnes. Elle se déroule pendant 19 mois en entreprise et dans les centres de l'AFPA à Stains (93), Meudon (92) et Lardy (91), ainsi que dans le Labo — lieu d'apprentissage qui reproduit en miniature un data center — chez Equinix à Saint-Denis (93). C'est l'une des entreprises partenaires de la formation, avec Digital Realty, autre spécialiste des data centers, CBRE (immobilier) et Eiffage Energie Systèmes. « *Le recrutement et la formation des équipes d'exploitation ont été et restent des enjeux majeurs pour notre industrie* », met en avant Nicolas Buono, directeur des opérations d'Equinix France. Anass Mhaidra, directeur

du campus de data centers de Saint-Denis chez Equinix a recruté cinq des 12 alternants : « *Le métier de technicien de maintenance en data center est jeune, peu connu, avec un salaire intéressant, Nous aimerions que cette formation, qui débouche sur un titre professionnel de niveau 4, évolue vers un niveau 5 (bac + 2).* »

Le projet répond à un double enjeu : former des professionnels capables de garantir un fonctionnement optimal et durable d'un data center, dans une filière en croissance ayant des besoins de compétences importants, et offrir des opportunités de reconversion et d'emploi. Elle cible principalement des jeunes avec une base technique ou électrotechnique (diplôme de niveau 4/baccalauréat ou expérience professionnelle). L'objectif est de délivrer une qualification complète en gestion et maintenance des infrastructures de data center. Le cursus inclut des formations sur les infrastructures électriques et climatiques, la sécurité incendie et les systèmes de gestion technique des bâtiments.

Une formation en alternance combinant théorie et immersion professionnelle

Le parcours de formation démarre par une période de préparation opérationnelle à l'emploi individuelle (POEI) de trois mois (formation à temps plein), puis continue avec une alternance de 16 mois. La POEI se compose de deux blocs de compétences du titre Technicien d'équipement et d'exploitation en électricité. L'alternance se compose de quatre modules de formation du titre Technicien de maintenance CVC (chauffage, ventilation, climatisation), d'un module du titre Technicien de maintenance et de travaux en systèmes de sécurité incendie, et de deux modules sur mesure en relation clientèle dans un data center en anglais. □

C.C

LE LABO, DATA CENTER ÉCOLE CHEZ EQUINIX POUR APPRENDRE EN IMMERSION

Le Labo, data center miniature, a été inauguré le 25 mars 2025 sur le plus gros site d'Equinix en France, à Saint-Denis. Il doit permettre aux techniciens data center d'Equinix et aux étudiants de la nouvelle formation de l'AFPA de réaliser des travaux pratiques encadrés en gestion et maintenance des infrastructures. Il comprend plus de 25 équipements professionnels répartis en six zones, permettant de reproduire le travail quotidien du technicien : configurer, faire de la maintenance préventive, dépanner...

Recrutement

« Nous recrutons 600 profils tech en 2025 »

La société de conseil et de technologies Magellan Partners recrute 600 spécialistes tech en 2025, jeunes diplômés comme experts seniors.

L'Informaticien : Pouvez-vous présenter votre entreprise ?

Anne-Flore Le Gal, DRH de Magellan Partners :

Notre société emploie 2 700 personnes, en France où elle dispose de 15 sites, et à l'international (Europe, États-Unis et Île Maurice). Partenaire des grands éditeurs et fournisseurs de technologies, elle emploie 40 % de consultants, et 60 % de spécialistes de la tech. Le chiffre d'affaires du groupe s'est élevé à 375 M€ en 2024. Le groupe a pour but de rassembler environ 7 000 salariés en 2030, pour un CA d'un milliard d'euros.

Quelle est votre politique de recrutement ?

A-F Le Gal : Nous prévoyons 1 000 recrutements en 2025, dont : 600 sur les postes tech, 200 jeunes diplômés ou ayant moins de trois ans d'expérience, 60 alternants, 40 personnes en reconversion.

Notre politique de cooptation dynamique est à l'origine de 40 % de nos recrutements. Les réseaux d'alumni fonctionnent bien ; la prime est de 1 500 euros pour une première personne recrutée, 2 500 euros pour une deuxième. Notre taux de turn over est inférieur à 20 %. Côté tech, nous recrutons des développeurs, des ingénieurs en IA, data, DevOps, cloud, des consultants en data gouvernance, des architectes experts des solutions partenaires (Microsoft, SAP, AWS, Salesforce...). Nous proposons des salaires dans la moyenne ou au-dessus du marché selon les postes et profils. Partenaire d'une vingtaine d'écoles d'ingénieurs et d'informatique, nous avons pour objectif de renforcer les relations écoles dans les deux ans.

Comment se déroule le recrutement ?

A-F Le Gal : Notre processus est rapide, entre 15 jours et trois semaines à partir de la publication de l'annonce. Il y a jusqu'à trois entretiens, un de préqualification par un membre de l'équipe dédiée au recrutement, un avec un manager et un avec un associé.

Comment se passe l'intégration ?

A-F Le Gal : Nous gardons le lien entre la signature du contrat et l'arrivée de la recrue. Les premiers jours sont structurants, ils ont lieu dans le site régional où est rattaché le nouveau salarié. Nous voulons offrir un maximum d'interlocuteurs et d'informations les deux premières semaines, qui représentent une phase d'acculturation à la stratégie de l'entreprise. Tous les lundis se déroule une journée d'intégration, avec un déjeuner du nouveau collaborateur avec un manager ou son parrain.

Quelle est votre politique de fidélisation ?

A-F Le Gal : Nos collaborateurs attendent des missions de qualité, des opportunités d'évolution, de formation, de la flexibilité.

Nous investissons dans les formations et certifications, mais aussi dans la qualité de vie au travail et l'équilibre vie professionnelle/personnelle, avec des places en crèche, des tickets restaurants, des afterworks et sessions de team building... Notre charte de télétravail prévoit deux jours par semaine, plus 15 jours volants dans l'année.

C'est dans notre ADN d'avoir une proximité entre les managers et les équipes, et d'avoir une approche de test and learn. Nos collaborateurs sont invités à expérimenter, proposer leurs idées et solutions, à échanger dans les communautés Teams.

Quelle est votre politique d'inclusion et de diversité ?

A-F Le Gal : Notre comité RSE adresse les sujets de diversité et d'inclusion. Signataires de la charte de la diversité, nous venons d'adhérer à l'Association Française des Managers de la Diversité (AFMD). Nous sommes partenaires de l'association d'égalité des chances Nos Quartiers ont du Talent, avec laquelle nous avons organisé le mois dernier un social day regroupant 75 mentors volontaires parmi nos collaborateurs, et 75 jeunes diplômés issus de milieux divers pour favoriser leur intégration dans le monde du travail.

Nous avons un réseau interne, Women In Magellan, sachant que les femmes composent 40 % de nos effectifs, et 35 % dans la tech. □

**Anne-Flore Le Gal,
DRH de Magellan
Partners**



« 40 % des recrutements proviennent de notre politique de cooptation dynamique »

C.C



SMART IMPACT

THOMAS HUGUES

8H30 | 19H30

VOTRE ÉMISSION QUOTIDIENNE DÉDIÉE À LA RSE ET À LA TRANSITION ÉCOLOGIQUE DES ENTREPRISES

Orientée « solutions », l'émission SMART IMPACT animée par Thomas Hugues monte en puissance et vous propose désormais un rendez-vous quotidien. Chaque jour, retrouvez des témoignages d'entrepreneurs et d'experts autour de la transition écologique, de l'économie durable et des enjeux RSE.

N°230
orange™

N°245
bouygues
TELECOM

N°349
free

B SMART
1. Change

Cybersécurité : anticiper pour ne pas subir

Sommaire

Dossier : Cryptographie post-quantique, l'heure est à la migration.P68

Resilient by Design : anticiper plutôt que subir ..P72

Sécurité matérielle : un autre talon d'Achille de la résilience IT.....P73

La plateforme de Google où convergent ses capacités de sécurité.....P74

Sécuriser un smart building de bout en boutP75

Chatbot : une urgence pour la protection des donnéesP77

Le Zero Trust, grand thème du forum InCyber ..P78

La souveraineté numérique de l'UE questionnée...P79

L'impact des dispositifs multi-protocoles sécurisés sur l'IoT.....P80

Rencontre avec Joséphine Delas, ingénieure en IA chez HarfangLabP82

Face à l'irruption du quantique, la cryptographie est à un tournant. Les standards d'aujourd'hui ne résisteront pas à la puissance des ordinateurs de demain. La migration vers des algorithmes post-quantiques n'est plus une option.

Plus largement, l'heure est encore et toujours à la résilience proactive. "Resilient by Design" n'est pas un slogan, mais une stratégie : intégrer la sécurité dès la conception. Les géants du numérique, à l'instar de Google, centralisent de plus en plus leurs capacités de protection pour répondre à des menaces systémiques. Dans les smart buildings aussi, la sécurisation de bout en bout fait son bout de chemin. Cette montée en vigilance (au moins sur le papier) s'étend jusqu'aux objets connectés, avec des dispositifs multi-protocoles et des textes réglementaires comme le Cyber Resilience Act, qui imposent de nouvelles exigences. Ce sujet sera d'ailleurs au cœur du dossier du prochain numéro de l'INFOCR.

Dans ce contexte, des talents comme Joséphine Delas, ingénieure IA chez HarfangLab, interviewée en dernière page de ce numéro rappellent que les outils ne suffisent pas. L'expertise humaine reste au cœur de la cybersécurité.

Préparer la transition à l'ère de l'informatique quantique

À l'approche de l'ère quantique, les fondations de la cryptographie vacillent. Les algorithmes classiques, piliers de la cybersécurité depuis des décennies, pourraient devenir obsolètes avec les ordinateurs quantiques dès 2030. Les organisations sont encouragées à amorcer une transition vers des algorithmes de cryptographie post-quantique pour sécuriser les données, les accès et les communications. Mais cette transition s'annonce longue et fastidieuse.

Pendant des décennies, l'industrie de la sécurité IT a vécu sur des algorithmes classiques bien connus, tels que le chiffrement RSA (du nom de ses trois inventeurs Ronald Rivest, Adi Shamir et Leonard Adleman), et la cryptographie sur courbes elliptiques (ECC). Schématiquement, la cryptographie est surtout utilisée dans deux cas de figure : le chiffrement et l'authentification. Le premier consiste à protéger des informations et données personnelles, la seconde vise à empêcher des acteurs de la menace d'usurper l'identité d'autres personnes, comme un collaborateur d'une entreprise par exemple.

Une cryptographie bientôt battue

La cryptographie utilisée aujourd'hui est, dans la plupart des cas, symétrique et asymétrique. Dans ce dernier cas, le système repose sur deux clés. La première, partagée publiquement, sert à chiffrer les données ou à vérifier l'authentification. Elle ne peut pas être utilisée pour déchiffrer un message ou s'authentifier à la place d'un autre. C'est la deuxième clé, privée celle-ci, qui le peut.

Les normes actuelles sont performantes. À tel point que, comme le souligne IBM sur une page dédiée, « lorsque des données ou des systèmes sécurisés sont piratés, ce n'est presque jamais parce que quelqu'un a déjoué le chiffrement lui-même ». En effet, c'est la plupart du temps un humain qui a transmis ses identifiants et donné un accès par la grande porte à un acteur malveillant. Les codes, clés et

schémas de chiffrement et d'authentification sont des problèmes mathématiques qui ont été pensés pour être impossibles à résoudre pour des ordinateurs classiques. Il faudrait, par exemple, des milliards d'années pour que des superordinateurs classiques puissent décomposer une clé publique RSA.

Mais l'arrivée prochaine des ordinateurs quantiques va sérieusement rebattre les cartes. S'ils sont prometteurs et accéléreront la recherche dans des domaines majeurs comme la santé, ils seront aussi en mesure, à terme, de faire tomber les systèmes de chiffrement asymétrique. « Les ordinateurs quantiques représentent un paradigme de calcul totalement nouveau, laissant de côté les bits binaires pour les espaces de calcul complexes créés par l'utilisation de qubits, et permettant de résoudre des problèmes qui semblaient auparavant impossibles à résoudre », détaille IBM. Par exemple, les algorithmes quantiques de Grover et Shor, du nom de leurs inventeurs, remettent en question la solidité des systèmes de chiffrement actuels. Leurs propriétés respectives vont accélérer la recherche dans des espaces de clés pour l'un, et factoriser rapidement de grands nombres pour l'autre.

Une affaire de calcul

Selon le rapport sur la cryptographie post-quantique (PQC) du National Institute of Standards and Technology (Institut national des normes et de la technologie), un organisme américain en charge d'établir des normes et standards, et rattaché au département du commerce des États-Unis, des premières violations pourraient survenir dès 2030.

En 2023, l'Anssi (Agence nationale de la sécurité des systèmes d'information) et dix-huit pays d'Europe se sont alignés sur les prédictions du Nist dans un position paper. La cryptographie post-quantique (PQC) « doit devenir une priorité absolue et être engagée dès maintenant », notamment pour les administrations publiques, les fournisseurs d'infrastructures critiques, les



« Les systèmes doivent être crypto-agiles. C'est-à-dire capables de basculer vers la cryptographie post-quantique le moment venu »

Thiébaut Meyer,
directeur des stratégies de
sécurité chez Google Cloud

prestataires informatiques et le secteur industriel en général.

Du côté des géants de la tech, même son de cloche. « Un ordinateur quantique va abaisser la barrière et compromettre des systèmes en quelques jours, donc suffisamment rapidement pour que ça soit un danger. Aujourd'hui, nous n'avons pas encore de matériel suffisamment puissant, mais ça va arriver. Est-ce que ça sera dans cinq ou dix ans, on n'en sait trop rien », explique Thiébaut Meyer, directeur stratégies de sécurité chez Google Cloud.

De son côté, Pierre Jaeger, responsable des partenariats stratégiques d'IBM Quantum, développe : « Comme on l'a observé avec l'essor de l'intelligence artificielle, l'accélération du hardware joue un rôle déterminant. Aujourd'hui, des acteurs comme IBM s'engagent sur des feuilles de route très ambitieuses en matière d'informatique quantique. À un horizon de quelques années, on peut donc envisager l'émergence de machines capables — potentiellement — de mettre en danger les algorithmes de chiffrement actuels. » L'expert nous a d'ailleurs confié qu'IBM allait annoncer une mise à jour de sa roadmap avec un certain nombre d'accélération dans les prochaines semaines.

Récolter maintenant, décrypter plus tard

Bref, tout le monde s'accorde à dire que la question n'est pas de savoir si, mais quand l'environnement informatique représentera une menace pour les algorithmes asymétriques et de signature. Mais, au fait, que risque-t-on ?

Un acteur malveillant disposant d'un ordinateur quantique suffisamment puissant pourra théoriquement faire tomber les normes de cryptographie actuelles, et donc lire les données qu'elles protègent, ou encore falsifier une signature numérique. Les hackers, eux, adoptent déjà une stratégie d'attaque singulière baptisée « Harvest now, Decrypt later », que l'on peut traduire par « Récolter maintenant, décrypter plus tard ». Cette méthode consiste à conserver, sur le long terme, des données volées chiffrées, pour les déchiffrer une fois que les technologies le permettront.

« Lorsque l'acteur malveillant aura réussi à le faire, la plupart des informations n'auront sans doute plus de valeur. En revanche, certaines auront encore un intérêt (secret industriel, données de santé, ndlr) », met en garde Thiébaut Meyer. Mieux vaut prévenir que guérir. Côté prévention justement, le Nist a lancé, en 2016, un appel à proposition et soumissions d'algorithmes de chiffrement post-quantiques pour établir de nouvelles normes résistant à l'informatique quantique. Au total, une soixantaine de projets ont été soumis par des organisations à travers le monde. Le Nist a finalement conservé quatre algorithmes de chiffrement à clé publique ML-KEM (ex-Crystals-Kyber), et de signature numérique ML-DSA (ex-Crystals-Dilithium) et FN-DSA (ex Falcon) ; ainsi qu'un protocole pour la normalisation, SLH-DSA (ex-Sphincs+).

Une courte fenêtre de tir

« Les algorithmes annoncés aujourd'hui sont spécifiés dans les premières normes finalisées du projet de normalisation de la PQC du Nist et sont prêts à être utilisés immédiatement », écrivait l'institut dans un communiqué. Les équipes de développement, quant à elles, sont invitées à intégrer ces algorithmes dès la conception.



« Il faut réaliser l'inventaire de ses certificats pour évaluer l'impact du passage à la PQC »

Pierre Codis,
vice-président adjoint
des ventes pour les pays
nordiques et l'Europe
du sud chez Keyfactor

Pour l'existant, l'agence demande aux experts en cybersécurité de ne pas perdre de temps et de migrer vers ces normes post-quantiques, comme l'exige la feuille de route du mémorandum de sécurité nationale des États-Unis, le « NSM-10 and the Transition to Post-Quantum Cryptography », lequel préconise la migration vers ces nouvelles normes entre 2025 et 2035.

« Le Nist et les agences de sécurité nationales considèrent que ces algorithmes avec lesquels nous avons vécu pendant trente ans — le RSA, le ECC et les courbes elliptiques —, seront dépréciés et considérés comme disallowed, donc non autorisés à partir de 2035 », développe Pierre Codis, vice-président adjoint des ventes pour les pays nordiques et l'Europe du sud chez Keyfactor et directeur général de la filiale française. Il ajoute : « S'il ne faut pas être trop alarmiste, il ne faut pas non plus tomber dans l'excès inverse et dire qu'on a encore le temps, car c'est faux. »

L'Anssi a d'ores et déjà invité les États membres de l'Union européenne à rejoindre un groupe de travail dédié du groupe de coopération dédié pour se protéger le plus rapidement possible du scénario d'attaque « stocker maintenant, déchiffrer plus tard ». Ce groupe de travail invite aussi à anticiper les potentielles longues périodes de migration des systèmes complexes, tels que les infrastructures à clé publique (PKI) et autres dispositifs à longue durée de vie. En effet, « modifier tous ces algorithmes dans le code source et les applications est un long processus ! », prévient Thiébaut Meyer. Et ce laps de temps peut suffire à exposer des systèmes dans le cas où la migration vers une cryptographie post-quantique n'aurait pas été terminée dans les temps.

Une difficile mission d'inventaire

Cette problématique de la durée de la migration est prédominante, car c'est l'ensemble de l'écosystème informatique qui est concerné. Les algorithmes de chiffrement sont omniprésents. On les retrouve bien sûr dans les PKI (Public Key Infrastructure, système de gestion des clefs publiques, ndlr), dans les différentes couches logicielles et protocoles largement utilisés, comme TLS ou SSL/TLS. Ils sont également intégrés dans le matériel : cartes à puce, tokens, HSM, etc. Du côté des utilisateurs finaux, on parle des portefeuilles électroniques, des cartes d'identité et passeports électroniques, de la signature électronique. Tous ces systèmes reposent massivement sur les certificats numériques et les algorithmes cryptographiques pour sécuriser les échanges et

transactions. Cela implique que les bibliothèques logicielles, les API et l'ensemble des composants techniques devront progressivement migrer.

Les organisations vont ainsi devoir prendre tout un panel de mesures, afin de faciliter la transition. Elles ont été décrites dans un document du ministère de la Sécurité intérieure des États-Unis, repris en avril 2024 par le NSM-10. Il s'agit d'abord d'identifier quelles données et actifs critiques doivent être protégés en priorité. Les systèmes devront aussi être priorisés pour la transition au regard des besoins des entreprises. Ces dernières devront mener un inventaire de tous les systèmes utilisant des technologies cryptographiques. Les responsables de la cybersécurité auront pour mission d'identifier les normes de cybersécurité et de protection des données qui nécessiteront une mise à jour pour refléter les exigences de la cryptographie post-quantique. Pour Pierre Codis, inventorier et protéger avec de la PQC les données critiques est une priorité absolue : *« Si vous avez des contrats, de la propriété intellectuelle, documents ou codes, qui seront sensibles au-delà de 2030, alors il y a un vrai sujet parce qu'il y a une vraie menace. »* À l'inverse, pour les informations sensibles d'une entreprise dont la confidentialité doit être garantie jusqu'en 2028 ou 2029, la situation est un peu plus nuancée. *« Si vos données deviennent obsolètes ou sans importance avant 2028, il n'y a pas vraiment de sujet. En revanche, pour des données à protéger jusqu'entre 2028 et 2030, il est recommandé de migrer vers des algorithmes plus robustes, afin d'anticiper les évolutions en matière de sécurité. »*

Un manque (historique) de gouvernance cryptographique

Une fois l'inventaire de l'existant réalisé, un plan de transition des systèmes devra être élaboré. Toutes ces étapes sont particulièrement difficiles à mettre en œuvre, d'où l'importance d'anticiper. En effet, les organisations ne disposent pas nécessairement d'une vision d'ensemble sur les données qu'elles détiennent, sur les plus vulnérables, les plus critiques et celles qui doivent être protégées en priorité. D'autant que la gestion de la cryptographie, elle, n'est pas toujours centralisée. Elle a souvent été implémentée de manière isolée, par chaque développeur, fournisseur de hardware, éditeur de logiciels, au gré des développements. Cela découle de plusieurs années de pratiques en silo, où chacun agissait de son côté, sans véritable gouvernance cryptographique. *« Il n'y avait pas, ou très peu, d'autorité centrale — comme un RSSI ou une fonction équivalente — pour définir et imposer une politique commune à l'échelle d'une organisation. Que ce soit dans des systèmes distribués ou décentralisés, aucune ligne directrice unifiée n'était vraiment mise en œuvre »,* explique Pierre Codis. Résultat : chaque division, département, unité ou pôle appliquait ses propres choix, souvent en toute autonomie, utilisant des certificats autosignés. Or, *« vous ne pouvez pas préparer de migration si vous ne savez pas où sont vos certificats ».*

IBM a lancé un programme individualisé baptisé « IBM Quantum Safe », qui vise à aider ses clients à cartographier leur cybersécurité et à mettre à niveau leur infrastructure. Il existe d'autres



« Il faut commencer par créer Cryptography Bill of Materials, une base de données qui recense où de la cryptographie est utilisée dans l'entreprise »

Pierre Jaeger,
responsable des partenariats
stratégiques d'IBM Quantum

outils, tels que le logiciel métier CLM (Certificate Classical Management) qui détecte tous les certificats non conformes, non compatibles ou autosignés, permet de repérer les personnes ou systèmes qui ont généré des certificats sans consulter d'autorité cryptographique centrale, de les identifier, de les inventorier puis de les remplacer automatiquement. Autre avantage, le CLM va également inventorier les certificats provenant de PKI officielles, mais qui utilisent des algorithmes obsolètes. Il pourra ainsi détecter des certificats utilisant des algorithmes RSA avec une taille de clé de 1 000 bits, considérés comme insuffisamment sécurisés.

Établir des roadmaps avec ses fournisseurs

Dans le cas des systèmes déjà en place et qui sont encore en cours d'amortissement et utilisés en entreprise, le RSSI devra aussi consulter ses fournisseurs de pare-feu, d'ordinateurs, de portables, d'applications et de tout matériel mobilisant de la cryptographie. *« Il devra engager des discussions avec eux. Les exigences liées au post-quantique devront être incluses dans les cahiers des charges pour savoir dans quelle version leurs produits supporteront la PQC et quand cela sera disponible »,* fait remarquer Pierre Codis. En ce qui concerne les nouveaux achats et pour tous les équipements arrivant en fin de vie ou en fin de contrat d'ici 2030, *« nous ne pouvons pas attendre 2029 pour commencer à travailler sur le sujet. Ce n'est juste pas possible. Il y a un dialogue à engager avec tous les fournisseurs de sa supply chain, pour établir des roadmaps post-quantiques dès maintenant »,* insiste-t-il.

Et ce n'est pas seulement la mise en place d'algorithmes de PQC qui doit être questionnée. *« Nous recherchons également la capacité à être déployée de manière concrète et à générer des clés dans un temps correct [...], il ne faut pas qu'il y ait d'impact sur l'activité des entreprises »,* souligne Bernard Montel, directeur technique EMEA, expert en stratégie de cybersécurité chez Tenable et ancien de chez RSA Security. Les algorithmes post-quantiques ne sont pas toujours plus gourmands en temps de calcul, mais ils peuvent être beaucoup plus exigeants en mémoire. Cela pose un vrai défi pour les équipements contraints avec des capacités limitées. *« Il faut s'assurer que ces systèmes disposent d'assez de mémoire pour les faire tourner »,* précise Pierre Jaeger.

Le modèle hybride préconisé

Quid de la robustesse de la PQC ? Les principes qui sous-tendent les ordinateurs quantiques sont connus depuis des décennies. Aucune avancée majeure n'a encore été réalisée dans la résolution des problèmes mathématiques complexes qui servent de socle à ces nouveaux algorithmes. Ils sont donc a priori suffisamment résistants. Mais pas de certitude absolue ! Le NSM-10 prévoit d'ailleurs dans ses plans de transition, « la création d'une agilité cryptographique pour faciliter les ajustements futurs et permettre une flexibilité en cas de changements imprévus ».

En Europe, les signataires de la déclaration commune exhortent à privilégier un modèle hybride combinant une clé classique et une clé post-quantique. Faisant référence aux normes de cryptographie post-quantique désormais normalisées, ils écrivent : « Ces solutions étant encore relativement nouvelles, avec une expérience d'implémentation et d'analyse cryptographique en cours de développement, il est fortement recommandé pour la plupart des cas d'usage d'adopter des solutions hybrides. » Plus concrètement, en associant donc une méthode de cryptographie classique à une méthode PQC, de manière à ajouter une couche de sécurité au cas où l'une d'entre elles viendrait à être compromise.

« Lorsque l'on implémente du code dans une application, il peut y avoir des erreurs qui vont créer des vulnérabilités. Donc conserver les technologies courantes maintiendra une défense qui sera imparfaite certes, mais qui aura le mérite d'être là, le temps des ajustements », détaille Thiébaud Meyer. Bernard Montel ajoute : « Même les algorithmes qui existent depuis quarante ans ont évolué sans arrêt, parce qu'il y a ce qu'on appelle des crypto-challenges (défi cryptographique qui pousse à tester et améliorer les algorithmes). Et quand les premiers algorithmes post-quantiques seront déployés, il y aura aussi des crypto-challenges. Nous allons les implémenter, et les faire évoluer au fur et à mesure. »

Le modèle hybride doit donc promettre une transition fluide vers la PQC, tout en maintenant une sécurité robuste. En outre, l'écosystème client (applications, équipements, systèmes) n'étant pas entièrement contrôlé, il est difficile de s'assurer d'une adoption immédiate et généralisée des algorithmes post-quantiques. « Le certificat hybride garantit la compatibilité pendant cette période de transition », insiste Pierre Codis. Dans ce cas de figure, tant que les systèmes ne sont pas capables de supporter le post-quantique, la clé classique prend le relais. Dès que les composants sont prêts, la même clé hybride bascule automatiquement vers la PQC. ■

V.M

L'algorithme de cryptographie post-quantique HAETAE

La compétition coréenne débutée en 2022 visait la sélection et la standardisation d'algorithmes post-quantiques KpqC opposant 16 algorithmes candidats. Au terme d'une présélection réalisée en 2023, huit candidats sont restés en lice. La compétition a enfin abouti à une sélection finale de quatre algorithmes le 16 janvier dernier. Nous allons voir dans cet article quels sont les gagnants ou plutôt quel est leur schéma de signature.

La cryptographie post quantique est l'ensemble des algorithmes cryptographiques à clé publique conçus pour résister non seulement aux attaques actuelles mais également aux futures attaques d'ordinateurs quantiques de très grande capacité. Les algorithmes post-quantiques peuvent en effet être implémentés sur les ordinateurs actuels. Deux schémas de signature post-quantique ont été sélectionnés. Parmi ceux-ci figure HAETAE qui est le fruit d'une coopération entre des équipes françaises, coréennes et allemandes. Cet algorithme a pour coconcepteur Julien Devevey, cryptologue à l'Anssi.

Analyse technique approfondie

HAETAE représente une avancée majeure en matière de sécurité et de protection des données sensibles. Cet algorithme est capable de renforcer la résilience face aux futures menaces. Il repose sur des principes de cryptographie post-quantique innovants, à même de le rendre résistant aux attaques des ordinateurs quantiques. HAETAE offre une protection inégalée pour les communications et les données sensibles en utilisant des mécanismes de chiffrement avancés et des protocoles de sécurité robustes. Il est fondé, tout comme le standard NIST ML-DSA, sur la difficulté conjecturée du problème de réseaux euclidiens Module-LWE, ainsi que sur une variante dite avec rejets du paradigme de Fiat-Shamir. Il est cependant plus compact que l'algorithme ML-DSA. L'avantage qui en découle est qu'à niveau de sécurité égal son intégration dans certains protocoles de l'Internet nécessite moins de fragmentation. Les trois autres algorithmes de la sélection finale de la compétition KpqC (Korean post-quantum competition) suivent un autre schéma de signature, AIMer, et deux schémas de chiffrement et d'établissement de clef, NTRU+ et SMAUG-T.

Les implications pour la sécurité

L'algorithme HAETAE représente une avancée significative dans ce domaine en offrant une solution très prometteuse pour contrer les menaces futures du quantique. Les organisations ont donc sérieusement intérêt à se tenir dès maintenant informées des avancées dans le domaine de la cryptographie post-quantique et à évaluer la pertinence de l'adoption de solutions telles que HAETAE ou AIMer. ■

T.T

Vers une **cyber résilience** systémique

et de redéploiement automatisé, pour garantir la continuité, même en cas de compromission.

L'IA, un facteur de rupture et de vulnérabilité

L'un des constats les plus marquants du rapport est la sous-estimation des risques liés à l'intelligence artificielle. Si l'IA constitue un levier de performance pour les défenseurs, elle est aussi massivement exploitée par les attaquants, notamment pour automatiser la recherche de vulnérabilités. « On est juste au début du tsunami », alerte Zscaler. « L'intelligence artificielle va tout bouleverser. La fréquence des vulnérabilités explose, tous les éditeurs de VPN ou de firewalls l'ont constaté. » Dans le même temps, les usages professionnels de l'IA se multiplient, souvent sans garde-fous. Les plateformes de génération de contenus, les assistants intelligents ou les copilotes peuvent exposer des données sensibles si aucun cadre n'est défini. « Dans une réunion Teams, il y a énormément d'informations qui sont échangées. À quel point sont-elles confidentielles ? Le besoin d'une surveillance continue des flux s'impose, au-delà des solutions traditionnelles. »

Une plateforme unifiée et pédagogique

Pour faire face à ces nouveaux défis, Zscaler mise sur un arsenal technologique intégré dans une plateforme cloud-native (ZIA, ZPA, Data Protection, UVM, Risk360...), centrée sur la visibilité, le contexte et la priorisation métier. Mais la technologie ne suffit pas. « Bloquer un utilisateur, ça ne suffit pas. Il faut l'accompagner, le guider, » ajoute Ivan Rogissart, « cette pédagogie réduit les contournements tout en favorisant des usages sûrs ».

Avec plus de 45 millions d'utilisateurs et une couverture de 75 % du CAC 40, Zscaler séduit les grands comptes, mais pas uniquement. « Une petite entreprise n'a pas toujours les ressources à sa disposition pour faire fonctionner l'IT. Avec notre plateforme as-a-service, où l'on va avoir une sécurité by design et une mise à jour continue, c'est une simplification opérationnelle et un gain d'efficacité pour les petites entreprises. » L'approche « Resilient by Design » répond aussi à une exigence réglementaire croissante (DORA, NIS2) qui impose une résilience élargie à toute la chaîne de sous-traitance.

« Ce sont ceux qui ont le plus de moyens qui entraînent les autres. On ne sécurise plus seulement le cœur, mais toute la chaîne de valeur. Le monde est plus vulnérable, plus dépendant du digital. Il faut faire les choses autrement que ce qu'on a fait jusqu'à maintenant », conclut l'expert de Zscaler. Face à une sophistication croissante des menaces, et à une digitalisation irréversible des flux métiers, la résilience ne peut plus être pensée a posteriori. ■

J.C

Face à la montée des menaces, Zscaler défend une vision proactive de la cybersécurité : intégrer la résilience dès la conception des infrastructures. Une transformation profonde qui vise à rapprocher IT, stratégie métier et continuité opérationnelle.

Selon l'enquête « *Unlock the Resilience Factor : Why Resilient by Design is the Next Cyber Security Imperative* » réalisée par Zscaler, 94 % des responsables IT se disent confiants dans leurs dispositifs de cyber résilience. Pourtant, 60 % anticipent une faille majeure dans l'année à venir. Ce paradoxe souligne un écart entre les moyens alloués et l'existence d'un véritable plan stratégique. « Il y a une différence entre avoir des moyens et avoir un plan », résume Ivan Rogissart, directeur technique Europe du sud de Zscaler, « beaucoup d'entreprises ont investi, mais sans forcément penser à la reprise après incident ». Moins de la moitié des entreprises ont mis à jour leur stratégie face à l'émergence de l'IA, et 40 % ne l'ont pas revue depuis plus de six mois. Ce manque d'anticipation révèle une conception encore trop défensive de la cybersécurité, détachée du risque business.

Penser la résilience dès la conception

L'approche « Resilient by Design » portée par Zscaler repose sur un principe simple : « La résilience ne doit pas être une réaction, mais une anticipation. » Elle s'appuie sur la plateforme Zero Trust Exchange, qui conditionne l'accès aux ressources à l'analyse du contexte (identité, terminal, posture de sécurité). « Vouloir tout protéger, c'est illusoire. L'enjeu est de savoir ce qu'il faut redémarrer en priorité pour éviter une perte sèche », explique Ivan Rogissart. Cette approche systémique inclut des capacités d'auto-réparation



« Le monde est plus vulnérable, plus dépendant du digital. Il faut faire les choses autrement que ce qu'on a fait jusqu'à maintenant »

Ivan Rogissart,
Directeur technique Europe du sud de Zscaler

Sécurité matérielle : le talon d'Achille oublié de la résilience IT

Alors que les investissements en cybersécurité se concentrent sur les couches logicielles, une nouvelle étude d'HP Wolf Security révèle que les couches matérielles et firmwares restent sous-protégées. Une négligence coûteuse, à l'heure où la mobilité, la réglementation et l'IA reconfigurent les menaces.

Le dernier rapport de HP Wolf Security révèle une réalité préoccupante : les appareils informatiques sont vulnérables à chaque étape de leur cycle de vie, faute d'approches de sécurité globales intégrant les couches matérielles et firmwares. En France, 85 % des décideurs reconnaissent la criticité de ces couches, mais 60 % estiment qu'elles restent négligées dans le coût global du cycle de vie des produits. « Le marché est saturé d'excellents produits de cybersécurité pour les couches hautes, mais les couches basses - BIOS, firmware, interfaces matérielles - restent le parent pauvre », souligne Benjamin Duchet, chief technologist d'HP France, « pourtant, elles sont tout aussi vulnérables, car elles reposent sur des millions de lignes de code ».

Chaîne d'approvisionnement : un angle mort stratégique

Le rapport pointe d'abord les faiblesses de la sélection des fournisseurs : 31 % des entreprises ont vu l'un d'eux échouer à un audit cyber, et 27 % ont rompu un contrat pour cette raison. L'absence d'implication des équipes IT dans les processus d'achat constitue une faille majeure. « Choisir des fournisseurs de confiance est fondamental. Mais on ne peut pas demander aux utilisateurs d'être prudents si on leur vend du matériel peu sécurisé. C'est le serpent qui se mord la queue », commente l'expert en sécurité d'HP. Le fabricant américain observe une montée en puissance des

exigences côté clients : des certifications transport comme TAPA, des audits sur la sécurité de la chaîne logistique ou encore des engagements contractuels plus stricts. Dans le même temps, les mauvaises pratiques subsistent, avec des processus d'intégration des équipements souvent défectueux. Près de la moitié des décideurs admettent que les mots de passe BIOS sont trop faibles, partagés, voire jamais changés. « Il y a encore de grandes entreprises qui se font auditer sans avoir de mot de passe BIOS. En 2025, c'est quand même assez stupéfiant. » Cette négligence tient aussi à la peur de la complexité : 55 % des décideurs redoutent les mises à jour firmware - un phénomène baptisé FOMU (Fear Of Messing Up) par HP. « Pendant longtemps, on a pensé qu'il valait mieux ne pas toucher au firmware. Aujourd'hui, les attaquants savent qu'il s'agit d'un angle mort, et ils en profitent », ajoute Benjamin Duchet.

Travail hybride : le terminal en première ligne

La généralisation du travail hybride multiplie les incidents : pertes, vols, réparations non autorisées. En France, 24 % des salariés ont perdu ou se sont fait voler un appareil. Ces incidents coûtent 8,6 milliards de dollars par an à l'échelle mondiale ! « Le risque est là depuis longtemps, mais il est amplifié par la mobilité. Il n'est pas rare de voir des PC laissés allumés dans le TGV pendant que leurs utilisateurs vont aux toilettes, c'est hallucinant. » HP propose désormais un ensemble de solutions pour localiser, verrouiller ou effacer les terminaux, même éteints. Autre avancée : l'activation automatique des filtres de confidentialité à partir du contexte ou des métadonnées des fichiers. « Le meilleur cas d'usage, ce sont les entreprises qui activent automatiquement le filtre écran dès qu'un document confidentiel est ouvert. Là, on asservit la sécurité matérielle aux règles métiers », rapporte l'expert.

IA : un catalyseur d'attaques ciblées

Toujours selon l'étude, 79 % des décideurs pensent que l'IA va accélérer la mise au point d'attaques sur les firmwares. Benjamin Duchet confirme : « HP a déjà détecté des scripts Python générés par IA dans des attaques réelles. L'IA permet aux

attaquants de produire du code ciblé pour des couches qu'ils ne maîtrisaient pas. Ils deviennent plus créatifs, donc nous devons l'être aussi. » HP intègre également des mécanismes cryptographiques « post-quantiques » dans ses couches basses. Un pari sur l'avenir, alors que les premières attaques quantiques pourraient émerger à partir de 2030. « Aujourd'hui, ces algorithmes ne sont pas encore utilisés. Mais s'ils sont là, c'est pour éviter qu'on doive remplacer tout un parc dans cinq ans. » ■ J.C



« On ne peut pas demander aux utilisateurs d'être prudents si on leur vend du matériel peu sécurisé. C'est le serpent qui se mord la queue »

Benjamin Duchet,
Chief technologist d'HP France

Google lance sa **plateforme unifiée** Google Unified Security

À l'occasion de Google Cloud Next 2025, la firme de Mountain View a annoncé une avalanche de nouveautés dans son portefeuille de sécurité, articulées autour de Google Unified Security (GUS), une plateforme qui vise à faire converger les capacités de sécurité de Google.

Le paysage de la cybersécurité est très fragmenté, marqué par des données dispersées, pas toujours contextualisées, et une multitude d'outils. Un contexte qui n'est pas sans conséquences sur les capacités d'anticipation et de réaction des équipes de sécurité. Pour réduire les complexités opérationnelles que rencontrent ces dernières et améliorer leur efficacité, les éditeurs et fournisseurs d'outils de cybersécurité en tout genre ont pris le virage de la plateforme, en unifiant de plus en plus leurs processus de détection, de réponse et de remédiation.

Une plateforme pour les unifier tous

Google a, de son côté, décidé de rassembler ses solutions de sécurité et l'expertise de Mandiant au sein de Google Unified Security. Cette plateforme concentre cinq grands piliers que sont : les opérations de sécurité, la sécurité cloud, le renseignement sur les menaces, la navigation sécurisée via Chrome Enterprise et l'expertise de Mandiant. Elle réunit les solutions suivantes : Security Command Center (SCCE), Google SecOps (SecOps Enterprise Plus), Google Threat Intelligence (Enterprise) le Navigateur Chrome Enterprise (Premium) et Mandiant Threat Defense. Globalement, elle repose sur une base de données de sécurité unifiée et régulièrement mise à jour avec les données de Threat Intelligence de Google afin de mieux prioriser les alertes. Elle fournit des capacités de détection et de réponse sur les réseaux, les terminaux, les environnements cloud et les applications.

Google dévoile de nouveaux agents d'IA

La plateforme va, en outre, être alimentée par l'IA de Google, Gemini. La firme a ainsi présenté des agents semi-autonomes qui ont été ou seront intégrés aux capacités des solutions de Google pour assister les analystes, afin d'automatiser certaines tâches à faible valeur ajoutée et améliorer la détection et la réponse. Un

agent de triage et d'analyse des alertes va être intégré à Google Security Operations dans le courant du second trimestre 2025. Autre exemple : dans Google Threat Intelligence cette fois, un agent d'analyse de malwares va permettre de déterminer si un code est malveillant ou non. Il est lui aussi prévu pour le second trimestre.

La firme a également dévoilé des nouveautés dans Google Security Operations, autour de la gestion des flux de données. « Grâce à notre partenariat avec Bindplane, il est désormais possible de transformer et préparer les données pour des usages futurs, de les acheminer vers différentes destinations ou locataires, de filtrer les volumes pour en contrôler le flux et de masquer les données sensibles afin de garantir la conformité », décrit la firme.

Le nouveau service Mandiant Threat Defense pour Google Security Operations fournit, quant à lui, des capacités de détection et de réponse supplémentaires, et mobilise des experts Mandiant et des techniques de détection assistées par l'IA pour épauler les équipes de sécurité des clients en cas de menaces dopées à l'intelligence artificielle. Security Command Center intègre, quant à lui, les nouvelles fonctionnalités d'AI Protection,

un outil qui vise à aider les clients de Google Cloud à gérer les risques tout au long du cycle de vie de l'IA. Il a été intégré à Vertex AI afin que les développeurs puissent rediriger les requêtes et réponses des modèles vers les systèmes de protection.

Sécuriser les modèles

De nouvelles fonctionnalités dans Data Security Posture Management sont attendues en préversion au mois de juin. Elles assureront une meilleure gouver-

nance des données sensibles et des données d'entraînement de l'IA en les identifiant et en les classifiant, puis en appliquant des contrôles de sécurité et de conformité.

Chrome Enterprise Premium intègre, quant à lui, de nouvelles protections contre le phishing à destination des employés, pour les protéger contre les sites et portails frauduleux imitant des pages légitimes. « Les organisations peuvent désormais configurer et intégrer leur propre image de marque et leurs ressources d'entreprise, afin de mieux détecter les tentatives de phishing déguisées sur des domaines internes », précise Google. Le masquage des données de Chrome Enterprise est, quant à lui, disponible de manière générale. Mandiant Consulting, Rubrik et Cohesity se sont, quant à eux, associés pour fournir une solution qui doit réduire les temps d'arrêt et les coûts de récupération après une cyberattaque. ■

V.M



Les principales composantes de Google Unified Security sont : opérations de sécurité modernes, renseignement sur les menaces et gestion des expositions, protection du cloud et de l'IA, navigateur d'entreprise sécurisé et réponse aux incidents basée sur l'expertise de Mandiant.

Sécuriser un **smart building** de bout en bout

Pour réduire les risques cyber et la consommation énergétique de bâtiments intelligents, Kevin Polizzi, président d'Unitel Group, recommande d'inventorier les infrastructures IT/IoT/OT puis de les dissocier, d'exploiter des sondes dédiées et des outils open source de supervision.

Après Marseille et Lyon, Aubagne et Aix-en-Provence accueilleront prochainement leur propre édifice Quanta, où se regroupent, à proximité de racks informatiques, des sociétés innovantes en IA, biotech ou nanotechnologies. Dans chaque smart building, l'exposition aux attaques cyber est très élevée. Il s'agit non seulement de trouver des contre-mesures adaptées, mais surtout de tendre vers le zero-downtime, l'absence totale de panne du data center jusqu'aux bureaux des collaborateurs. En piratant un équipement de climatisation, un attaquant peut provoquer une fuite d'eau et détruire plusieurs serveurs et baies de stockage en cascade. Tout doit donc partir d'une analyse des risques propre à chaque site, préconise Kevin Polizzi : « Le RSSI doit quitter sa logique IT pour adopter le point de vue de l'utilisateur, avec pragmatisme. »

Le périmètre à protéger s'élargit

Le président d'Unitel Group rappelle : « J'ai commencé par bâtir des centres de données Jaguar Networks (Ndlr : société acquise en 2019 par le groupe Iliad, devenue Free Pro), où les deux incontournables sont la maintenance prédictive et la facturation de kWh aux clients confiant leurs serveurs. Le smart building naît de la faculté à piloter et à automatiser la production, en incluant ses infrastructures IT, IoT et industrielles. »



Le smart quartier Theodora dédié à l'IA, en construction à Marseille, doit être livré fin 2028.

Son objectif ? Rapprocher un écosystème d'entreprises utilisatrices, dont la conception commune de services pourra générer de nouveaux revenus. Antérieurs à l'adoption généralisée d'Internet, les capteurs, équipements électriques et contrôles d'accès physiques opposent peu de résistance aux cyberattaques. Pire, leurs bus et protocoles hétérogènes empêchent un firewall de les protéger tout à fait. L'approche d'Unitel Group consiste à relier des sondes à des cartes de supervision dédiées, la ségrégation de réseaux au niveau transport restant fondée sur l'offre de la société Seclab, afin de dissocier les réseaux physiquement et de répondre aux enjeux de résilience.

En complément, une plateforme open source développée en interne rapproche la gestion d'infrastructures cloud à celle des infrastructures techniques du bâtiment ; les consoles GTB et DCIM sont ainsi réunies autour d'une interface unifiée : « Nous avons créé la plateforme EOS (pour Extended Operating System) à partir de notre propre GTB conçue pour le data center, » précise-t-il.

Des économies rapidement significatives

Les résultats ne tardent pas à se manifester, la réduction du nombre de pannes réduisant les coûts humains et de matériels : « Passer d'une construction immobile à un smart building flexible, nous a permis de diviser la consommation d'énergie par trois en moins de 18 mois, au campus Thecamp proche de la gare Aix-TGV. Nous suivons une logique d'optimisation à la fois financière et environnementale », explique-t-il.

Le smart building se veut plus durable et plus sûr. Après lui, l'industrie et les collectivités locales confirment des besoins de pilotage automatisé et de supervision unifiée autour de la convergence de leurs infrastructures. « La cybermenace apparaît lorsque tout est connecté. Mais les métiers ne convergent pas aussi naturellement. Intervenant sur des univers différents avec des cycles de renouvellement d'équipements distincts, les électriciens, développeurs et administrateurs réseaux doivent se former sur le terrain, par la voie de l'apprentissage. Il va y avoir du travail pour plusieurs années, » prévoit-il. ■

O.B



Kevin Polizzi, président d'Unitel Group.

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 €

IA conversationnelle : urgence pour la protection des données personnelles

Comment instaurer la confiance sans freiner le progrès ?

Par Paul-Olivier Gibert, président de l'AFCDP

Les intelligences artificielles conversationnelles s'invitent dans tous les pans de notre quotidien numérique : assistants clients ou juridiques, outils de rédaction, agents virtuels dans les services publics. Leur efficacité impressionne, leur adoption s'accélère. Mais derrière l'enthousiasme technologique, une inquiétude grandit. En tant que délégués à la protection des données (DPD/DPO), nous lançons l'alerte : sans encadrement clair et rigoureux, ces technologies menacent les droits fondamentaux des citoyens.

Sous une apparence conviviale, ces IA absorbent d'énormes volumes de données personnelles : identités, coordonnées, informations bancaires, données de santé, voire opinions politiques ou orientations sexuelles. Ces informations, souvent saisies sans méfiance, alimentent les algorithmes pour les rendre plus efficaces. Mais elles sont aussi parfois stockées indéfiniment, traitées à l'étranger, utilisées à des fins qui échappent totalement à l'utilisateur.

Des outils performants... mais intrusifs

Et c'est bien là le problème : dans la plupart des cas, les utilisateurs ne mesurent ni l'ampleur des données collectées, ni la finalité réelle de leur traitement. Les politiques de confidentialité, trop longues, jargonneuses, découragent la lecture. Le consentement, pourtant au cœur du RGPD, devient une case cochée machinalement. La conversation semble anodine, mais masque une captation silencieuse et systématique d'informations sensibles.

Ces dangers ne sont pas hypothétiques. Ils sont déjà observables. Des fuites massives de données ont été signalées, parfois causées par une mauvaise configuration des systèmes. Des biais algorithmiques reproduisent ou amplifient des discriminations. Des entreprises utilisent des outils sans savoir où les données sont hébergées, ni si elles sont protégées. Et que dire de l'intégration de ces IA dans des environnements sensibles — cabinets médicaux, services juridiques, administrations — où la confidentialité devrait primer en toutes circonstances ?

Le RGPD demeure un cadre solide, mais il n'a pas été pensé pour la complexité des IA génératives. Il devient difficile d'identifier qui est le

responsable du traitement et qui est le sous-traitant. L'exercice des droits — droit d'accès, de rectification, d'effacement — devient presque impossible face à des systèmes opaques, parfois entraînés sur des corpus gigantesques et mal documentés.

Nos recommandations pour un usage responsable

Face à cette urgence, l'AFCDP appelle à une mobilisation collective. Les régulateurs doivent jouer leur rôle de vigie. Les éditeurs doivent prendre leurs responsabilités. Les décideurs publics doivent fixer un cap clair. Et les citoyens doivent pouvoir exercer leurs droits sans friction.

Nous proposons cinq priorités concrètes :

- **Transparence renforcée** : les éditeurs doivent informer, de façon simple et directe, sur les données collectées, leur finalité, leur durée de conservation, et les éventuels transferts hors de l'UE.
- **Consentement éclairé** : l'utilisateur doit comprendre ce qu'il accepte, pouvoir refuser facilement, et revenir sur ses choix.
- **Sécurité accrue** : chiffrement des données, authentification forte, traçabilité des accès doivent devenir la norme et non l'exception.
- **Encadrement interne strict** : chaque organisation doit mettre en place des politiques d'usage, former ses équipes, auditer régulièrement ses pratiques, et impliquer systématiquement les DPO.
- **Soutien aux solutions souveraines** : les alternatives éthiques et locales, conformes au droit européen, doivent être soutenues activement.

Ne pas sacrifier la confiance au nom du progrès

L'enjeu n'est pas de bloquer l'innovation, mais de l'orienter dans une direction bénéfique pour la société. Il est possible — et même nécessaire — de concilier performance technologique et respect des droits fondamentaux. La confiance ne peut être une variable d'ajustement.

Les IA conversationnelles ne sont pas neutres. Elles incarnent les choix — techniques, commerciaux, éthiques — de ceux qui les conçoivent. Elles reflètent des priorités, parfois invisibles, souvent discutables. Il est de notre devoir de veiller à ce que ces choix soient guidés par des valeurs fortes : respect de la vie privée, équité, responsabilité.

Construisons ensemble un numérique de confiance. Car c'est seulement dans un écosystème éthique et transparent que l'IA pourra tenir ses promesses sans compromettre les libertés individuelles. ■



Zero Trust : cap sur la maturité au Forum InCyber 2025

Le Forum InCyber Europe 2025, qui s'est tenu début avril au Lille Grand Palais, a placé le Zero Trust au cœur des échanges. Une stratégie dont la mise en œuvre progresse lentement dans les entreprises, car elle exige une refonte profonde des pratiques.

Du 1^{er} au 3 avril, le Lille Grand Palais a accueilli le Forum InCyber Europe 2025, qui a réuni quelque 18 000 participants, plus de 700 partenaires et 580 intervenants. Cette édition a notamment mis en lumière le Zero Trust. Lors d'une conférence intitulée « Zero Trust, le nouveau paradigme », les intervenants sont revenus sur leurs définitions du concept et la manière de le déployer dans les organisations.

Un concept encore abstrait

Eric Singer, de l'ESIEE-IT (l'école d'ingénieurs de l'expertise numérique), a interrogé plus de 1 000 membres du CESIN (Club des Experts de la Sécurité de l'Information et du Numérique) pour produire un livre blanc sur le Zero Trust. « Nous avons travaillé sur les étapes clés de mise en œuvre du Zero Trust dans les entreprises. De la compréhension de ce qu'est ce concept qui est difficile et pas souvent compréhensible. On pourrait penser, c'est un buzz word, mais quelle est vraiment la réalité pragmatique de la mise en œuvre du Zero Trust ? Le deuxième aspect, c'est vraiment savoir comment il est mis en œuvre dans les entreprises, [...] les retours d'expérience et comment faire pour que le Zero Trust soit un succès. »

Legacy et gestion du changement

L'expert défend ce qu'il désigne comme la confiance contextuelle — ou contextual trust — qui vise à adapter les mesures de sécurité en fonction du contexte et de la maturité des entreprises. « La première chose dans le Zero Trust, c'est d'identifier ce qui n'est pas connu », ajoute Eric Singer.

Jean-Noël de Galzain, président d'Hexatrust et PDG fondateur du groupe Wallix, a rappelé l'importance d'une approche réseau : « Nous sommes tous interconnectés au même réseau. Il y a des choses qui existent, le legacy, il y a des choses qui sont précieuses, ce sont les données, c'est la résilience, tout cela doit fonctionner en permanence, et il s'agit finalement de mettre en place une stratégie qui vise à dire que chaque personne, machine, objet, représente une menace pour l'ensemble du réseau que nous partageons tous. »

À partir de là, il faut déployer une stratégie Zero Trust globale, propre à chaque organisation, en

fonction de son activité. Jean-Noël de Galzain a insisté : « Le Zero Trust n'est pas une stratégie absolue, on n'en met pas partout. » Elle doit être adaptée en permanence, en fonction des impératifs, des besoins et des métiers et doit reposer « sur une bonne stratégie d'authentification des utilisateurs, de l'inventaire des utilisateurs, et des liens avec l'inventaire des machines, des objets connectés et des processus de l'entreprise ».

Accompagner la montée en compétence

Jean-Noël de Galzain a ensuite rappelé qu'il existait un grand nombre d'offres en Europe capables d'accompagner les organisations, grandes comme petites, pour les accompagner dans le déploiement du Zero Trust « à partir de solutions souveraines », un autre sujet qui a animé les débats lors du FIC. L'expert a attiré l'attention sur la nécessité d'accompagner la montée en compétence des équipes informatiques, mais aussi des utilisateurs, car, s'il était nécessaire de le rappeler : « Le comportemental fait partie des risques qui pèsent sur les données et la résilience de certains systèmes. »

D'où l'importance d'appliquer des technologies capables de détecter des comportements anormaux et de les faire remonter en temps réel, à partir du moment où l'inventaire des utilisateurs, outils et objets a été réalisé. On en revient à la stratégie globale.

Aujourd'hui, l'adoption des technologies atteint un certain niveau de maturité et a permis la mise en place d'automatisation et d'orchestration des données, qui a facilité l'adoption du Zero Trust au sein des entreprises.

Les intervenants ont également abordé la nécessité d'en parler clairement — notamment en termes d'adressage stratégique — au niveau du Comex (Comité exécutif) pour espérer voir le modèle se développer davantage en entreprise.

Au-delà de la sécurité des données, se pose aussi la question du ROI (retour sur investissement), dans un contexte où les entreprises accumulent des solutions de cybersécurité qui pèsent sur les budgets. ■

V.M



De gauche à droite : Jean-Noël de Galzain, président d'Hexatrust et PDG fondateur de Wallix ; Eric Singer, cybersecurity education director à l'ESIEE-IT et Benoît Fuzeau, président du Clusif (club de la sécurité de l'information français)

Au Forum InCyber, l'ombre de Trump

Le spectre de Trump ravive les inquiétudes de l'Union européenne quant à leur dépendance technologique, au cœur des débats lors du dernier Forum InCyber 2025. L'Europe tente tant bien que mal de reprendre la main sur son destin technologique et d'imaginer son propre modèle. Pas si simple.

L'élection de Donald Trump a une nouvelle fois mis l'Europe face à ses dépendances technologiques. La forçant, un peu dans l'urgence, à multiplier les annonces à coups de milliards pour se donner plus d'autonomie et mieux protéger ses intérêts, notamment sur le plan du numérique.

Lors d'une table ronde au nom sans équivoque, « L'Europe sous pression », organisée lors du dernier Forum InCyber début avril à Lille, Jonathan R. Cohen — diplomate américain, ancien représentant adjoint auprès des Nations Unies — a ouvert le bal en déclarant : « Je dirais pour commencer que ce que nous voyons émaner des États-Unis ne représente pas l'Amérique dans son ensemble. Cela représente cette administration actuelle, et principalement le noyau dur MAGA, ainsi qu'une influence disproportionnée des dirigeants du secteur technologique. »

La technologie comme levier diplomatique

Crise financière, Covid-19... « Nous avons vu que dans les moments de crise, l'Union européenne, la construction européenne, les dirigeants arrivent à accélérer les processus et à trouver des réponses », a éclairé Yves Leterme, ancien

Premier ministre belge, évoquant les capacités potentielles de l'Europe à s'émanciper des États-Unis du point de vue de la défense. Qu'en est-il du numérique ? Comme en défense, le Vieux Continent part de loin et accumule les retards face aux États-Unis et à la Chine, sans pour autant être dépourvu d'atout.

Tariq Krim, entrepreneur, fondateur de Netvibes, Jolicloud et de la plateforme de web éthique Polite, explique : « L'Europe ne manque pas de talents, d'ingénieurs ; les deux fondements de l'internet, le web et Linux, ont été inventés en Europe. Il fut un temps où un téléphone mobile sur trois était fabriqué en France. » Mais ça, c'était avant.

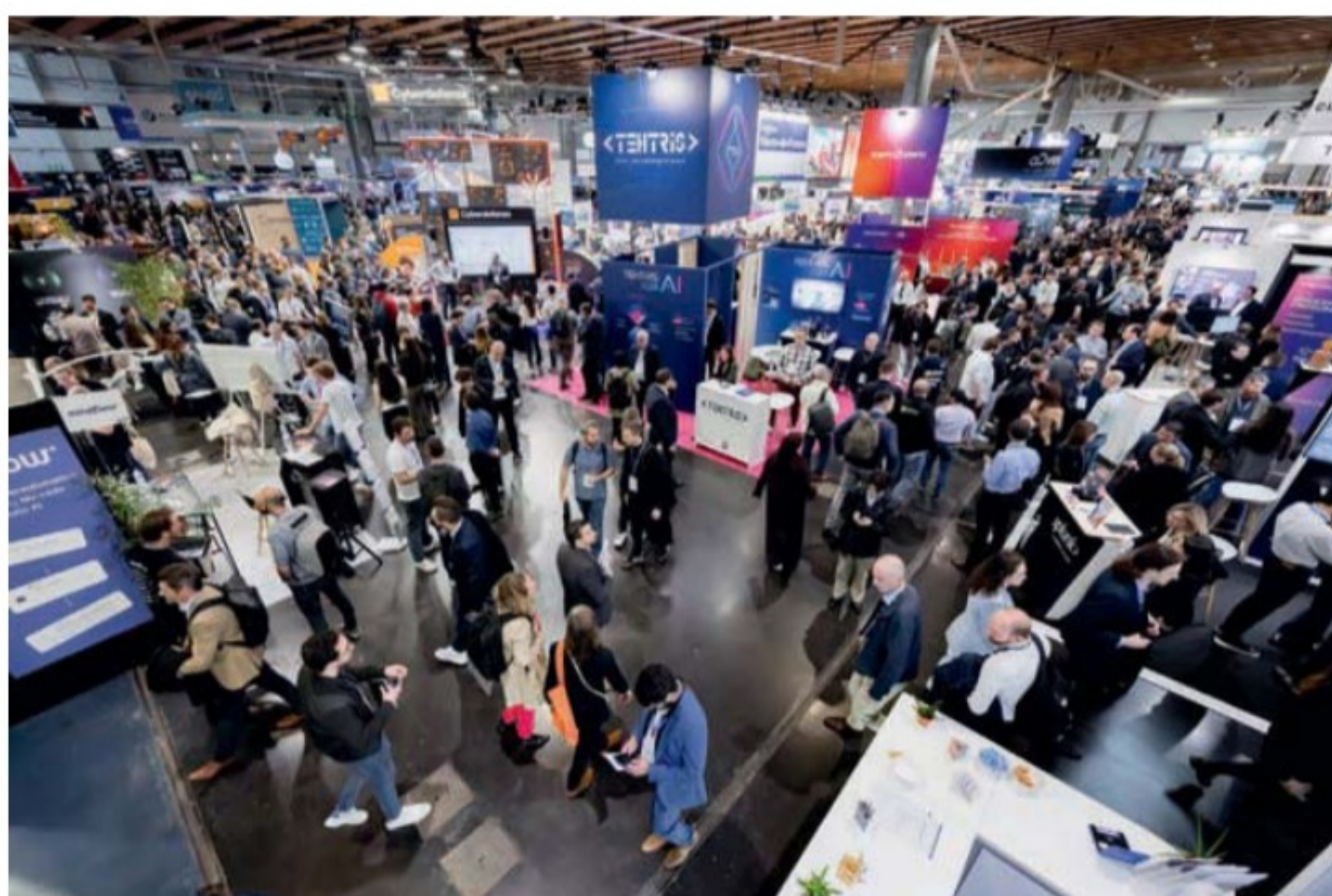
Quelle stratégie numérique adopter ?

Dans les années 1990, les États-Unis « ont construit une économie exceptionnelle autour de l'internet, qui leur a donné un pouvoir absolument extraordinaire », menant à l'émergence de géants. Avec les smartphones, « la puissance de domination des US s'est aussi immiscée dans nos intimités », générant une économie de rente. Ensuite, « on a commencé à comprendre que l'espace numérique était également politique », et qu'il pouvait être un terrain propice pour contenir ses adversaires ou faire pression, parfois même sur ses alliés. Rappelons que les États-Unis ont, un temps, brandi la menace d'une coupure du service Internet par les satellites Starlink en Ukraine — indispensable aux forces ukrainiennes — afin d'influer sur les négociations de paix.

Autre exemple, la guerre commerciale engagée entre les États-Unis et la Chine, qui gravite largement autour des technologies IT, semi-conducteurs et intelligence artificielle en tête, poussant Pékin à développer sa propre voie, notamment dans ces deux domaines. Quid d'une troisième voie ?

Tariq Krim a plaidé pour un retour des ingénieurs au cœur de la technologie, afin de redonner des couleurs à

la tech européenne, au-delà d'un sursaut politique et d'une cohésion (difficile) des 27 pour trouver une voie davantage tournée vers la technologie, alors que l'Union européenne aborde le numérique souvent d'abord à travers le prisme de la réglementation. Si l'Europe a le bon discours et la capacité économique, « nous ne sommes pas capables de développer des produits à l'échelle [...] et je pense que c'est là qu'il faut mettre le paquet », conclut Tariq Krim. Du point de vue strictement pécuniaire, l'initiative Eurostack estime que l'Union européenne aura besoin de 300 milliards d'euros d'ici à 2035 pour créer son propre écosystème numérique. À titre de comparaison, l'association américaine Chamber of Progress évoque un investissement de 5000 milliards d'euros. ■



Le prochain Forum INCYBER se tiendra du 31 mars au 2 avril 2026.

V.M

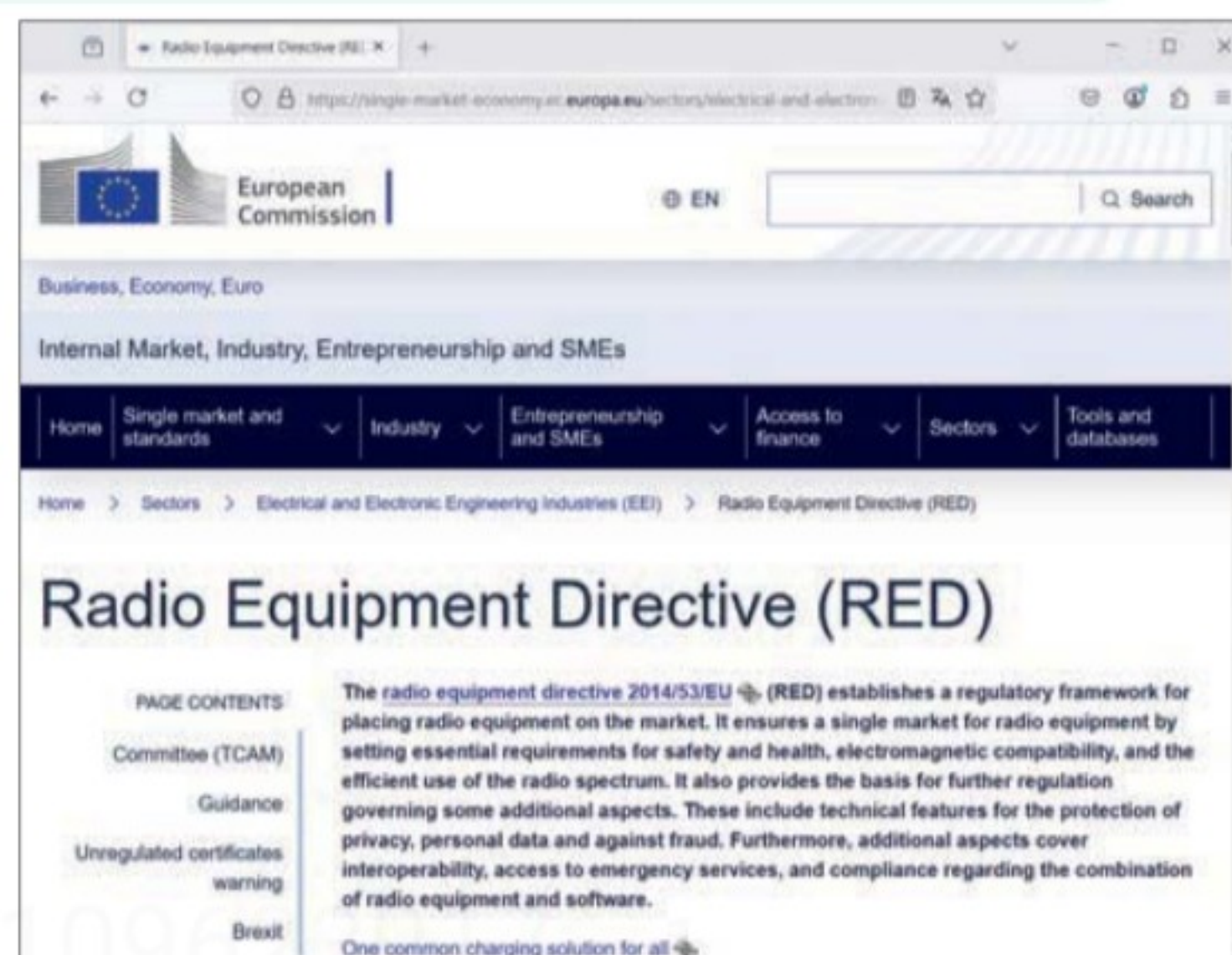
L'évolution des **dispositifs multi-protocoles sécurisés** et leur impact sur l'IoT

L'IoT évolue sans cesse. La multitude de protocoles en jeu tels que Bluetooth, Thread ou Zigbee rend la sécurité des connexions sans fil vitale. Nous allons voir dans cet article l'intérêt de l'introduction de dispositifs multi-protocoles sécurisés.

L'IoT (Internet des objets) imprègne de plus en plus notre société moderne, des utilisateurs finaux aux entreprises en passant par l'utilisation industrielle. L'introduction de dispositifs multi-protocoles sécurisés, comme ceux de la série STM32WBA de STMicroelectronics, offre une meilleure intégration et une sécurité renforcée grâce à des éléments matériels intégrés tels que la RoT (Root of Trust).

Ces avancées technologiques permettent une plus grande tolérance et une meilleure interopérabilité entre les différents protocoles sans fil du marché, simplifiant ainsi l'intégration des objets connectés dans des réseaux existants. Qui plus est, ces dispositifs répondent aux nouvelles normes législatives en matière de sécurité IoT comme notamment le Cyber Resilience Act de l'Union européenne, assurant une conformité indispensable aux fabricants d'équipements.

Les dispositifs multi-protocoles sécurisés offrent une solution robuste face aux menaces croissantes dans le



La directive RED (directive européenne 2014/53/UE) concerne la mise sur le marché des équipements radioélectriques. Les équipements entrant dans le domaine d'application de la directive et disponibles sur le marché européen doivent être obligatoirement conformes à cette directive.

monde de l'IoT grâce à une combinaison de cryptographie avancée et d'évaluation de sécurité. Ils permettent en parallèle une gestion efficace de l'énergie et un encombrement réduit.

Intégration de la sécurité dans les dispositifs multi-protocoles

L'IoT a connu une transformation impressionnante en seulement deux décennies, passant d'une simple idée à une composante essentielle de notre vie quotidienne. Cette croissance rapide a fait émerger des exigences nouvelles pour rendre les connexions sans fil plus sécurisées. Les nouveaux dispositifs multi-protocoles ont aujourd'hui la capacité de communiquer à travers plusieurs protocoles, tels que Bluetooth, Thread et Zigbee, ce qui permet une plus grande compatibilité et une meilleure intégration entre les écosystèmes technologiques. Le choix d'une fréquence commune de 2,4 GHz permet à ces dispositifs de servir divers besoins tout en donnant la priorité à la sécurité.

Cadre réglementaire sur l'IoT et ses impacts

La mise en œuvre du cadre réglementaire de l'IoT impose de nouvelles normes de sécurité. La directive RED (Radio Equipment Directive) de l'Union européenne, bientôt associée au Cyber Resilience Act (CRA), impose des mesures strictes que les fabricants doivent adopter. Les équipements RF devront tous se conformer à des critères spécifiques de la loi sur la résilience cybernétique, dès le mois d'août 2025. Cela signifie que les dispositifs certifiés selon les normes industrielles de sécurité deviendront essentiels. Le standard européen SESIP offre une structure claire pour l'évaluation de la sécurité des plateformes IoT, alignant ses exigences avec le Cyber Trust Mark US. La série STM32WBA de STMicroelectronics a été parmi les premières à obtenir cette certification qui garantit un niveau de sécurité élevé pour la prochaine génération de dispositifs connectés.

Les fonctionnalités de sécurité des protocoles sans fil standardisés ont évolué pour répondre à ces attentes croissantes en plaçant la sécurité en tête de liste des priorités. Il est désormais impératif d'assurer une sécurité robuste face aux nouvelles législations qui obligent les fabricants OEM à repenser le développement de produits IoT. Les microcontrôleurs sans fil multi-protocoles, comme notamment la série STM32WBA5 de STMicroelectronics, intègrent directement des éléments matériels dédiés à la sécurité comme le RoT (Root of Trust). La structure matérielle en question vérifie l'intégrité et l'authenticité de tous les composants du système, afin d'assumer une meilleure défense contre les cyberattaques et les manipulations.

Menaces pour les appareils IoT

Les menaces majeures pour les appareils de l'IoT ont de nombreuses sources. Les appareils IoT ont des capacités informatiques limitées, ce qui laisse un espace très faible pour y intégrer une protection et une sécurité robustes nécessaires pour une bonne défense contre les cyberattaques. Ils utilisent un grand nombre de technologies de transmission différentes, ce qui rend encore plus difficile la mise en œuvre de méthodes et de protocoles de sécurité suffisants. Leurs composants de base sont trop souvent vulnérables, laissant des millions d'appareils mal protégés contre les attaques. Les utilisateurs des organisations représentent quant à eux l'une des plus grandes menaces en matière de sécurité. Leur sensibilisation et leur formation sont essentielles. Les vulnérabilités des appareils IoT peuvent affecter les utilisateurs de différentes manières.

Les cybercriminels tentent de profiter des vulnérabilités des appareils IoT pour lancer des attaques globales contre les organisations et les utilisateurs finaux. Un attaquant va, par exemple, chercher à exploiter une vulnérabilité dans une machine pour effectuer une montée de privilèges et

ensuite faire un mouvement latéral pour atteindre des données critiques ou propager les malwares à travers un réseau. Les cybercriminels utilisent également des botnets (de grands réseaux d'appareils corrompus), comme notamment des routeurs afin de lancer des cyberattaques à grande échelle. L'IoT imprègne aussi de plus en plus la maison avec des appareils connectés, des assistants numériques, des objets connectés, des traceurs de santé et autres. Les services IoT peuvent eux aussi comporter des vulnérabilités, créant ainsi de nouveaux points d'entrée vers les autres appareils connectés aux réseaux domestiques.

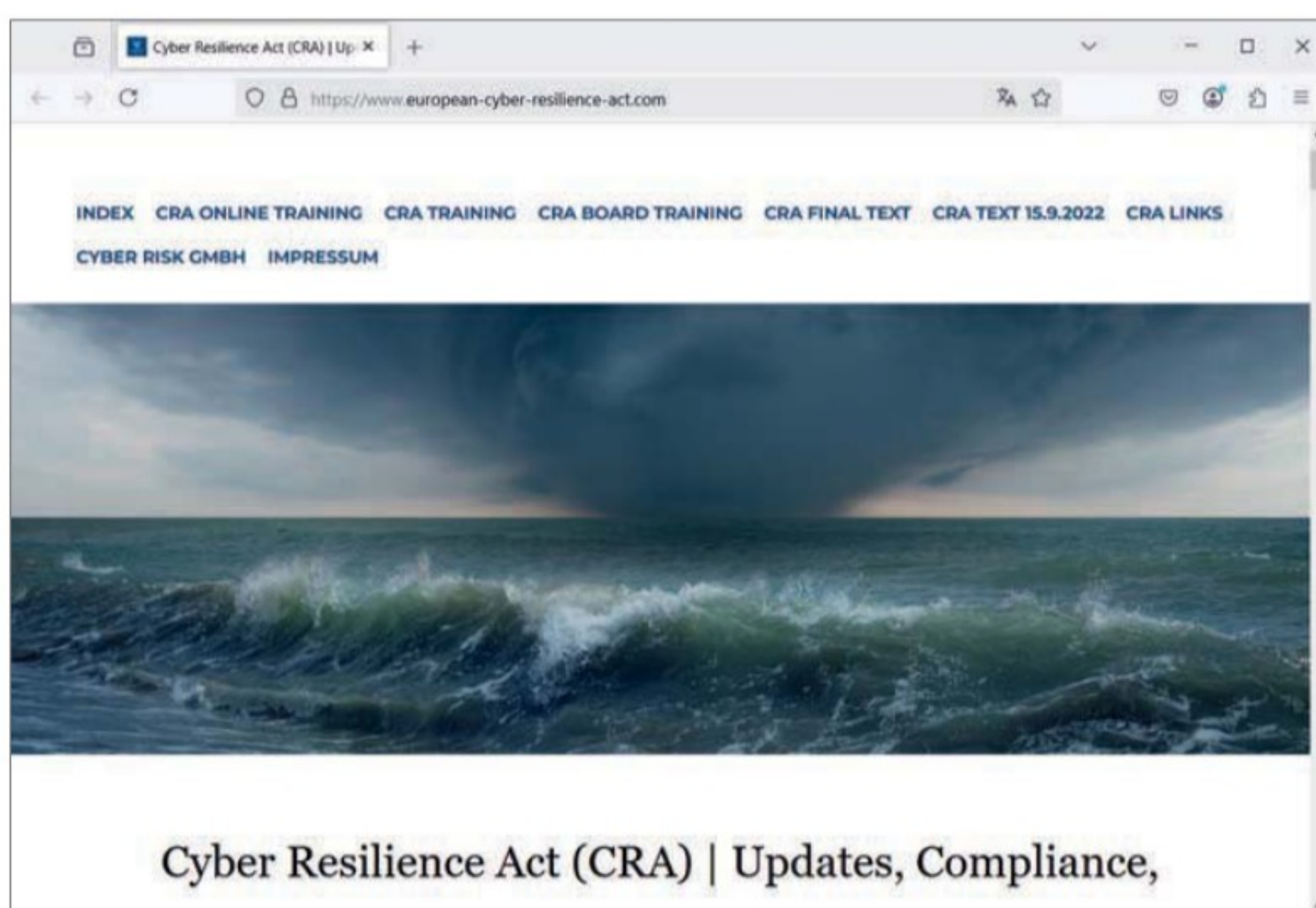
Principales vulnérabilités des appareils IoT

Les appareils IoT peuvent être compromis par un large éventail de vulnérabilités. Des mots de passe faibles ou codés en dur comptent parmi les méthodes les plus fréquentes employées par les attaquants. Les réseaux non sécurisés qui relient entre eux les appareils de l'IoT permettent souvent aux cybercriminels d'exploiter aisément les faiblesses des protocoles et des services qu'ils exécutent. Les interfaces d'écosystème non sécurisées, telles que les API et les applications mobiles ou Web sont elles aussi des vecteurs d'entrée privilégiés. Beaucoup d'appareils ont des processus de mise à jour non sécurisés permettant aux cyberattaquants d'installer des micrologiciels ou autres codes malveillants.

L'écosystème IoT peut lui aussi être compromis par des vulnérabilités de code et de logiciel de systèmes hérités. Les données reçues ou transmises par les appareils IoT sur les réseaux doivent être sécurisées et restreintes aux utilisateurs non autorisés. Une mauvaise gestion des appareils pendant leur cycle de vie les laisse ouverts à l'exploitation de ces vulnérabilités, même lorsqu'ils ne sont plus utilisés, du moment qu'ils restent allumés et connectés. Les appareils IoT, tout comme les appareils

personnels, sont livrés avec des paramètres par défaut souvent codés en dur pour en simplifier la configuration. Ces paramètres par défaut souvent non sécurisés fournissent des portes d'entrée très pratiques pour les attaquants. La protection des appareils IoT contre les vulnérabilités passe par tous les intervenants de leur cycle de vie. La sécurité des appareils de l'IoT commence par les fabricants qui doivent traiter les failles clairement connues de leurs produits en publiant des correctifs et signaler la fin du support. Il est également essentiel que les fabricants effectuent fréquemment des tests de pénétration pour s'assurer qu'aucune vulnérabilité n'a été ouverte tout au long de la production. ■

T.T



Le Cyber Resilience Act (CRA) est un règlement de l'UE visant à améliorer la cybersécurité et la cyber-résilience dans l'UE, grâce à des normes de cybersécurité communes pour les produits comportant des éléments numériques.

« Mon travail, c'est de mettre l'IA au service de la sécurité des systèmes informatiques, sans compromis sur la fiabilité »

Joséphine Delas, AI engineer chez HarfangLab

Joséphine Delas a rejoint l'équipe IA de l'éditeur de cybersécurité HarfangLab en 2023, en tant qu'ingénieure en intelligence artificielle. Elle travaille au développement et à l'intégration de modèles d'IA dans les solutions de protection des terminaux (EDR), afin d'améliorer la détection et la réponse à incident, et faciliter le travail des analystes. Elle nous raconte en quoi consiste son travail.

L'Informaticien : Comment est exploitée l'IA au sein des solutions d'HarfangLab ?

Joséphine Delas : L'axe historique englobe des modèles d'IA intégrés depuis 2018 dans notre EDR. Il s'agit de moteurs de détection de malwares reposant sur des modèles de machine learning ou de deep learning. Ils permettent de généraliser la détection à des malwares qui n'ont pas été nécessairement documentés et analysés par nos experts. D'une façon générale, les équipes IA et cybersécurité d'HarfangLab travaillent régulièrement ensemble pour développer des modèles pertinents.

L'Informaticien : Comment faites-vous pour maintenir la pertinence de ces modèles ?

J.D : Le cœur du problème, c'est la donnée. Nous réentraî-nons régulièrement nos modèles avec des données plus récentes pour qu'ils soient à jour avec l'état de la menace, qui évolue rapidement.

Il y a beaucoup de sources, venant notamment de la communauté open source, qui répertorient les malwares les plus récents, et nous allons nous appuyer là-dessus. Nous surveillons aussi de près les performances des modèles chez nos clients. S'ils commencent à mal détecter certaines menaces, par exemple, nous allons pouvoir les réentraîner et les réajuster. C'est un travail de veille donc, de recherche et de qualification de la donnée, qui consiste à consolider les informations trouvées en open source à l'aide d'outils développés en interne.

L'Informaticien : Quel est votre rôle dans ce flux de travail ?

J.D : Mon rôle et celui de l'équipe IA, c'est de concevoir, développer et déployer les modèles dans l'EDR. L'IA est un outil puissant pour traiter d'importants volumes de données. Mon

travail, c'est de mettre l'IA au service de la sécurité des systèmes informatiques, sans compromis sur la fiabilité. Pour innover, il faut donc réfléchir en continu aux manières d'appliquer des notions théoriques complexes à un environnement de production sous contraintes : trouver les bons compromis et optimisations pour que le modèle soit vraiment utile.

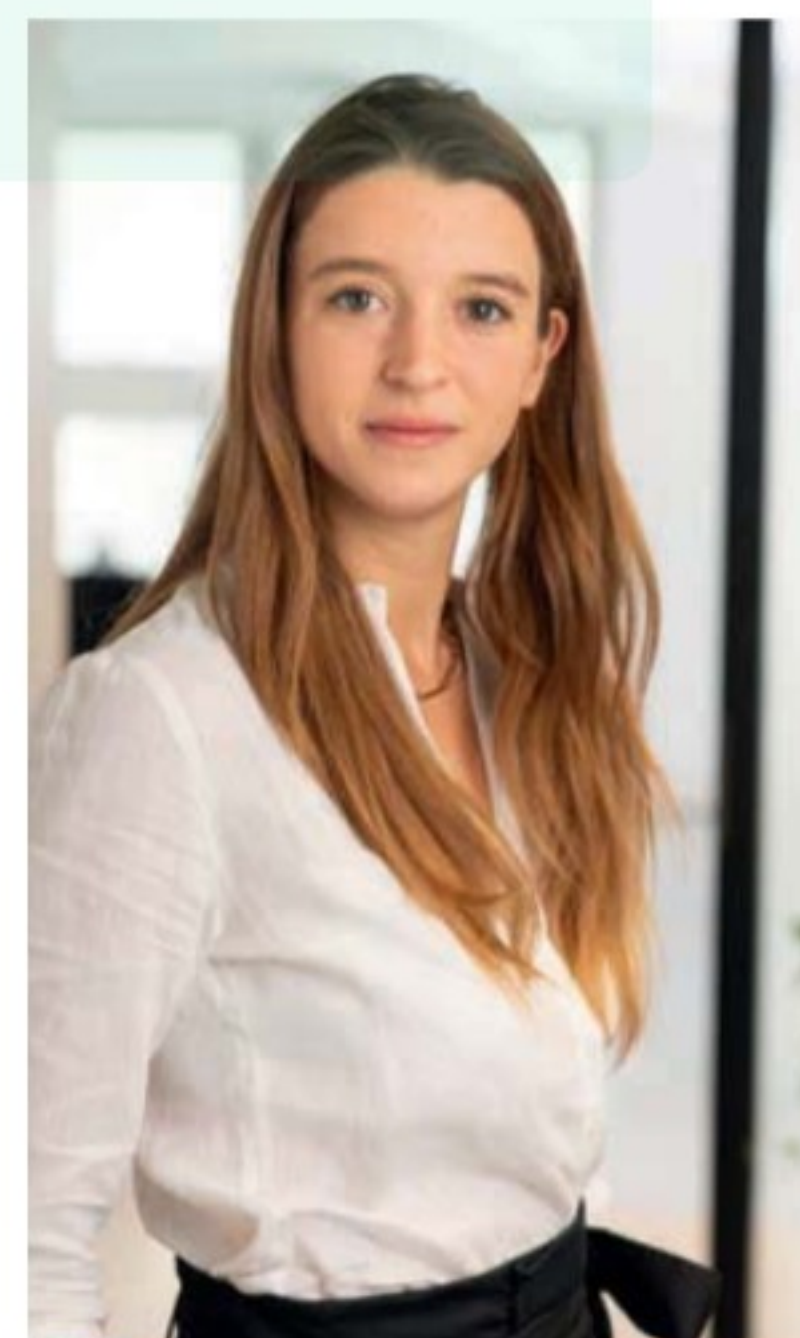
L'Informaticien : Justement, comment exploitez-vous l'IA générative chez HarfangLab ?

J.D : Nous travaillons à intégrer l'IA générative en tant qu'aide à l'analyste. Il ne s'agit pas de détecter des menaces, mais plutôt de support à l'investigation lors d'un incident. Nous avons lancé en décembre dernier un assistant pour les analystes baptisé Kio. Nous nous sommes imposés comme guideline de développer un outil utile qui ne soit pas de la poudre aux yeux. Donc nous avons réfléchi aux cas d'usage. Ce qui est intéressant pour un logiciel de cybersécurité, c'est qu'il a accès à énormément d'informations qui remontent, mais alors que de nouvelles menaces et façons d'attaquer émergent régulièrement, le travail de veille demande un effort conséquent à l'analyste. Notre but était donc de pouvoir filtrer les données, résumer le plus possible l'information pertinente, d'extraire le signal utile dans toute cette masse d'informations. C'est là que l'IA générative est intéressante, car elle permet de résumer un processus d'attaque, d'orienter la recherche et l'investigation là où c'est pertinent.

L'Informaticien : Kio produit-il des rapports et des exemples de remédiation à l'analyste ?

J.D : Des travaux dans ce sens sont en cours, mais pour l'instant, nos modèles d'IA générative se concentrent sur la recherche de télémétrie. Concrètement, l'analyste va poser une question concernant une alerte en langage naturel, qui va mobiliser une IA générative, mais aussi des outils sous-jacents afin d'aller chercher l'information dans les bases de données de l'EDR. ■

V.M



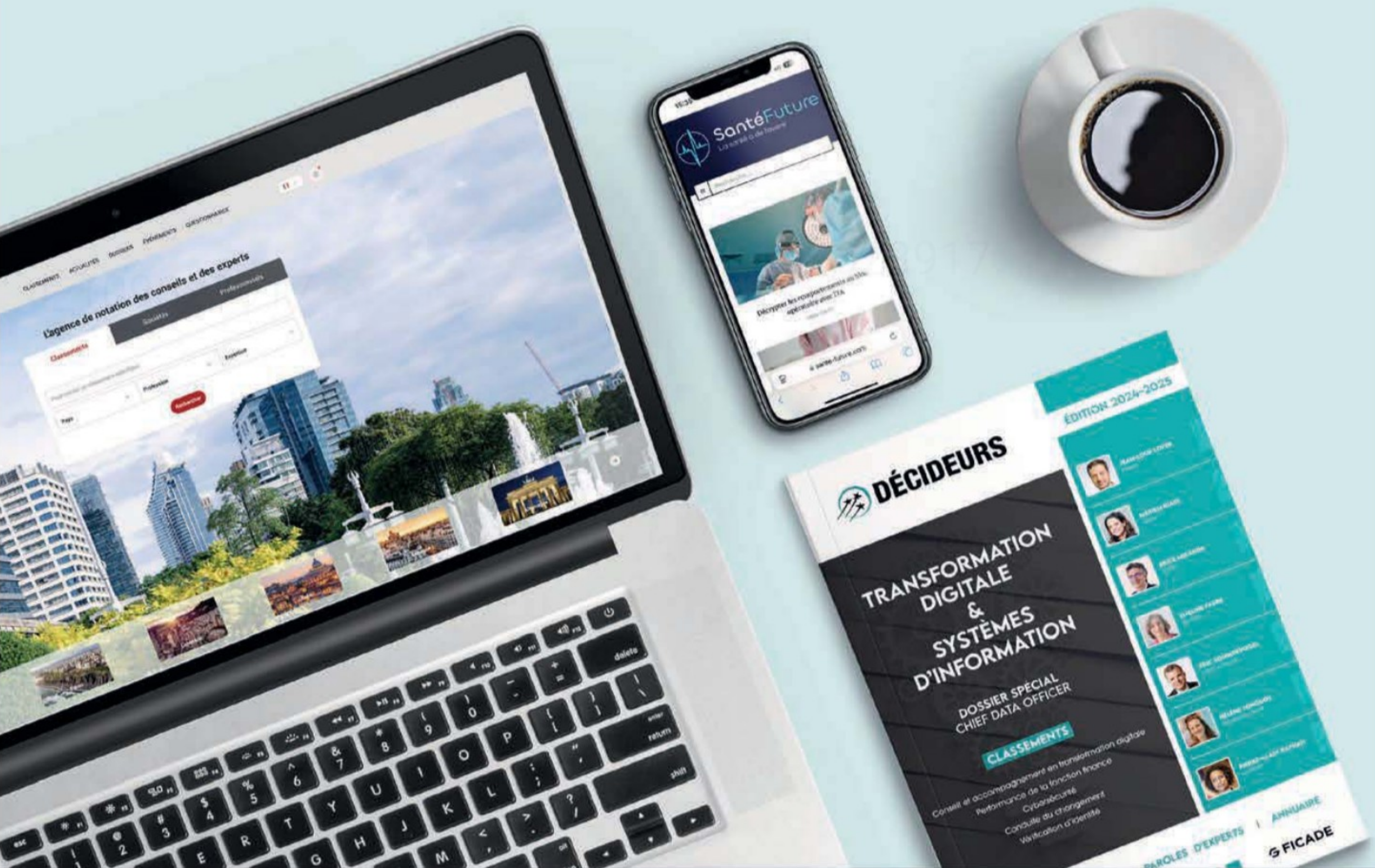
Joséphine Delas est membre du CEFYCS (Cercle des femmes de la cybersécurité), une association française qui vise à promouvoir la place des femmes dans les métiers de la cybersécurité.



DÉCIDEURS

TRANSFORMATION DIGITALE & SYSTÈMES D'INFORMATION

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



L'INFORMATION STRATÉGIQUE POUR BIEN CHOISIR VOS PARTENAIRES
COMMANDER LE GUIDE

Évoluez à votre rythme.

Gérez les charges à haute densité grâce au refroidissement liquide, tout en conservant votre infrastructure existante refroidie par air.

Vertiv™ Liebert® XDU070 offre un refroidissement liquide à haute densité sans raccordement au réseau d'eau de rejet de chaleur, en exploitant les systèmes de refroidissement par air existants, évitant ainsi des changements d'infrastructure majeurs et coûteux.



En savoir plus :

[Vertiv.com/XDU070-FR](https://www.vertiv.com/XDU070-FR)



Catégorie Hardware
Infrastructures Critiques
Data Center

