

NOTE DE SYNTHÈSE

Énergie et Numérique : des défis réciproques

anRT

ASSOCIATION NATIONALE
RECHERCHE TECHNOLOGIE



**LE POUVOIR DE
L'INTELLIGENCE
COLLECTIVE**

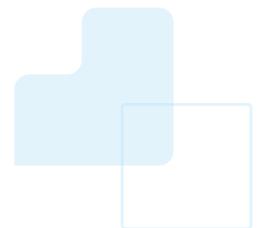
MAI / 2022
LES CAHIERS FUTURIS

Sous la présidence d'Olivier Appert, membre de l'Académie des technologies
et de Gérard Roucairol, Président honoraire de l'Académie des technologies
Richard Lavergne, co-président
Denis Randet, co-président
Mihaja Razafimbelo, rapporteur
Clarisse Angelier, ANRT, directrice de publication

Ces travaux sont soutenus financièrement par les souscripteurs FutuRIS :



AI CARNOT, AIR LIQUIDE, AMPIRIC - AIX-MARSEILLE UNIVERSITE, ANR, BERGER-LEVRAULT, BOUYGUES, BRGM, CEA, CNRS, CPU, EDF, ENGIE, FACEBOOK, GE HAELTHCARE, INSERM, INSTITUT MINES TELECOM, INRIA, INSTITUT PASTEUR, IRIS SERVIER, MINISTERE DE L'EDUCATION NATIONALE ET DE LA JEUNESSE, MINISTERE DE L'ENSEIGNEMENT SUPERIEUR, DE LA RECHERCHE ET DE L'INNOVATION, NOKIA, REGION PAYS DE LA LOIRE, SHNEIDER ELECTRIC, SNCF, TOTAL ENERGIES



Le contenu n'engage que la responsabilité de l'ANRT en tant qu'auteur et non celle des institutions qui lui apportent leur soutien.

MEMBRES ACTIFS DU GROUPE DE TRAVAIL



ALAIS	Jean-Christophe	Air Liquide
ALVAREZ-HERAULT	Marie-Cecile	Ancre/INP Grenoble
ASTIC	Jean-Yves	RTE
AUVERLOT	Dominique	France Stratégie
AZAIS	Philippe	CEA
AZARMAHD	Telman	EdF
BACHA	Seddik	INP Grenoble
BAREUX	Gabriel	RTE
BERTHOMIEU	Nadine	Ademe
BERTIERE	François	Académie des Technologies
BERTRAND	Arnaud	ATOS
BONIFACE	François	GRDF
BOUYER	Etienne	CEA
BULLIER	Guillaume	CRE
BURTIN	Alain	EDF
BUSATO	Guillaume	RTE
CHAILLOT	Christophe	Orange
CLAUSSE	Marc	Ancre-Allistene/INSA Lyon
DAUVE	Sebastien	LETI-CEA
de WARREN	Nicolas	Uniden
DUPONT DE DINECHIN	Benoit	Kalray
FAUCHEUX	Ivan	CRE
FERRASSE	Jean-Henry	Aix-Marseille Université
FERREBOEUF	Hugues	The Shift Project
FOURNEL	Guillaume	CRE
GAME	David	RTE
GELENBE	Erol	Académie des Technologies
GEOFRON	Patrice	Université Paris Dauphine
GILLET	Guillaume	Engie
GIRAUD	Guillaume	RTE
GRELLIER	Christian	Bouygues
GUILLOT	Françoise	Safran
HARTMANN	Joël	STMicroelectronics
HAUET	Jean-Pierre	Comité scientifique d'Equilibre des Energies
KAMIYA	George	IEA/EXD/SIO
KHERROUF	Samira	Ademe
LABRY	Delphine	Ministère des Armées
LATROCHE	Michel	CNRS
LEBLANC	Elvire	CEA
LECAILLE	Aurélien	GRTgaz
LEMOINE	Fabrice	Ancre/Université de Lorraine
L'HELGUEN	Eric	EMBIX
LUCCHESE	Paul	CEA
MASSINES	Françoise	CNRS
MAZAURIC	Vincent	Shneider Electric
MAZINGAND	Hervé	Technip Energies
MENTING	Jos	Engie
MONCOMBLE	Jean Eudes	Conseil Français de l'énergie/WEC
MONTAGNE	Xavier	Ministère de l'Enseignement supérieur et de la recherche
MORTELETTE	Tiphaine	Leyton
ORGERIE	Anne-Cecile	Ancre-Allistene/CNRS
POSTEL-VINAY	Grégoire	Ministère de l'Economie et des Finances
PRIEM	Thierry	CEA
PROULT	David	CEA
QUEHEN	Audrey	Engie
ROCHE	Nicolas	Enedis
SCHMITT	Michel	CGE/Ministère de l'Economie et des Finances
SLAOUI	Abdelilah	CNRS
THIBAUT	Muriel	IFPEN
TISSERAND	Isabelle	Laposte
VANDELAER	Olivier	Engie

SOMMAIRE

INTRODUCTION	6
1. ESTIMATION DE L'ÉVOLUTION DE LA CONSOMMATION D'ÉNERGIE DU NUMÉRIQUE	8
2. LIMITES PHYSIQUES ET TECHNOLOGIQUES	12
1 Le ralentissement de la loi de Moore	14
2 L'ordinateur quantique	14
3. INFRASTRUCTURES NUMÉRIQUES. ÉTAT DE L'ART ET ÉVOLUTION	16
1 Paramètres des systèmes informatiques	16
2 Nouvelles architectures	18
2-1/ Le FOG computing : l'architecture nouvelle de télécommunications de la 5G	18
2-2/ L'edge computing et les IOT, le continuum computing	18
2-3/ La fin du processeur universel : le recours à la spécialisation et au parallélisme	19
3 L'optimisation des systèmes	20
4 Sécurité et résilience	22
4. ANALYSE SECTORIELLE	24
1 Électronique, informatique, télécommunications	24
1-1/ Composants ; réalisations du CEA-LETI et de ST Microelectronics	24
1-2/ Les processeurs parallèles Kalray	26
1-3/ Les data centers ; les réalisations d'OVH	27
1-4/ Les super calculateurs, les fabrications d'Atos	28
1-5/ Le programme quantique	28
1-6/ Télécommunications	29
1-7/ Le contrôle des supply chains	29
2 Production et transport d'électricité	30
2-1/ Production	30
2-2/ Optimisation amont-aval	30
2-3/ Commercialisation et services énergétiques	31
2-4/ Le transport d'électricité	31
3 Automobile	33
4 Aéronautique	34
5 Bâtiment et aménagement	35
5. CONCLUSION ET RECOMMANDATIONS	38

INTRODUCTION

Cette note récapitule le travail 2021-2022 d'un groupe rassemblant entreprises, organismes de recherche et pouvoirs publics. Ce groupe est réuni depuis 2017 par l'ANRT, sous la présidence d'Olivier Appert, en appui à la Stratégie Nationale de Recherche en Energie. Il se consacre à l'analyse des enjeux technologiques liés à la transition énergétique. Il le fait d'un point de vue pragmatique, concret, correspondant à l'expérience de ses membres.

Les travaux du groupe avaient montré l'importance des technologies du numérique, en particulier pour l'organisation et le fonctionnement du réseau électrique modifié par l'introduction des énergies renouvelables intermittentes. De façon globale, ces technologies constituent l'un des leviers de la lutte contre le changement climatique. Toutefois, dès lors que l'Accord de Paris sur le climat implique une diminution des émissions de CO₂ de -7,6% par an d'ici à 2050, il est impératif d'analyser l'empreinte environnementale de tous les secteurs d'activité, tout au long du cycle de vie. Et le numérique n'échappe pas à cette obligation : l'extraction et la transformation des minerais en amont, les consommations électriques et les émissions liées à la fabrication et au fonctionnement des équipements et des traitements de l'information, l'efficacité et la complétude du recyclage des matériaux en aval du cycle sont autant de sujets de vigilance. Il est essentiel que les gains induits par le numérique ne soient pas annulés par ses impacts environnementaux. Le numérique est comme le cholestérol : il y a le bon et le mauvais. Cette année, le groupe s'est donc intéressé à la consommation d'énergie du numérique, problème incontestable, puisque la croissance du besoin en capacité de calcul et d'échanges semble sans limite.

Certaines études ont annoncé des catastrophes, mais les prophètes de malheur exagèrent : ils ne tiennent pas compte des progrès du numérique, y compris dans la manière de s'en servir. Le problème demeure, mais nos travaux montrent qu'il l'est moins dans les secteurs applicatifs que dans celui du numérique lui-même. Nous avons vu que la production d'électricité, l'automobile, l'aéronautique, le bâtiment, sont des secteurs très dépendants des questions énergétiques. Partout, le numérique est un moyen pour améliorer l'organisation du travail, la réduction de la consommation d'énergie et le fonctionnement des produits. Et un moyen nécessaire, en particulier dans le secteur primordial de l'électricité.

Certes, si la consommation électrique est toujours un élément du prix de revient, en fabrication et fonctionnement, il est intéressant de la réduire, mais ce n'est pas un point de blocage. En revanche, c'en est un pour l'industrie du numérique : désormais, la dissipation de chaleur limite les performances, depuis les circuits intégrés jusqu'aux superordinateurs. Et comme on ne peut plus compter sur la loi de Moore et la poursuite de la miniaturisation, il faut trouver d'autres moyens.

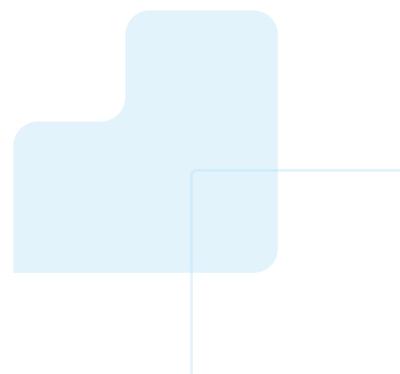
Or, ces moyens impliquent une véritable révolution. En effet, il va falloir mettre en cause la séparation traditionnelle entre la fabrication de processeurs universels, qu'on peut acheter partout, et celle des systèmes. Par exemple, l'automobile du futur sera conçue par des constructeurs maîtrisant la conception des processeurs embarqués, et ayant accès - tout en maîtrisant leurs informations confidentielles - à des superordinateurs capables de mettre au point les modèles d'intelligence artificielle les concernant.

Ce qui s'annonce, c'est un bouleversement des chaînes de valeur. Notre travail en donne une idée pour quelques secteurs industriels, mais cela va être vrai pour tous.

Cette révolution peut ouvrir des occasions, mais risque, si nous ne réagissons pas tout de suite, de donner la maîtrise de notre industrie à des pays qui maîtriseront l'interaction nouvelle entre l'électronique et les produits.

La France et l'Europe font face à un double problème :

- Maîtriser l'ensemble de la chaîne du numérique ;
- Que les secteurs applicatifs disposent du personnel compétent pour concevoir - en liaison avec l'industrie européenne du numérique - le numérique utilisé dans la conception, le fonctionnement et la maintenance des matériels qu'ils produisent.



01

ESTIMATION DE L'ÉVOLUTION DE LA CONSOMMATION D'ÉNERGIE DU NUMÉRIQUE

Le trafic Internet mondial a été multiplié par plus de 1000 entre 2000 et 2019. Cependant, une analyse récente de l'AIE montre que la consommation d'énergie et l'empreinte des émissions des technologies de l'information et de la communication sont restées stables, car l'efficacité énergétique de ces technologies s'est beaucoup améliorée. Cela a démenti les prévisions catastrophistes, comme celle parue dans Forbes en 1999, selon laquelle en 2010 la moitié de la production américaine d'électricité serait absorbée par l'économie numérique et Internet.

En réalité, les prévisions sont difficiles, en raison de la rapidité avec laquelle la demande, l'efficacité des TIC, et la manière dont on les utilise évoluent. Les méthodologies et les hypothèses diffèrent selon les organisations. Les résultats extrêmes, sur lesquels on communique beaucoup, résultent de modèles simplistes. Par exemple, pour la consommation des centres de données (data centers), ces modèles donnent des résultats plus de deux fois supérieurs à la fourchette probable (200-350 TWh).

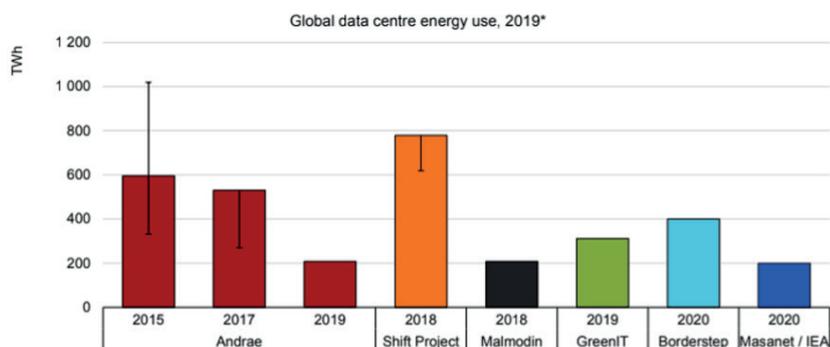
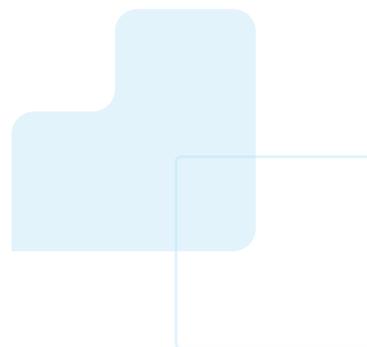


Figure 1 : Estimations de la consommation énergétique des data centers
Source : Présentation IEA, G. Kamiya

La donnée va être modifiée par la poursuite de l'accroissement des besoins et par l'arrivée des technologies émergentes (5G, Internet des objets, intelligence artificielle et machine learning, block chain,...). Pour contenir la consommation et les émissions, il faudra investir de façon considérable dans la recherche et le développement. Et s'il faut réduire autant que possible leur empreinte directe, les plus grands impacts des technologies numériques sur les émissions proviendront de leurs applications dans d'autres secteurs (électricité, transport, bâtiment,...), bien que leur emploi puisse en fin de compte réduire la consommation et les émissions totales de ces secteurs.



En France, selon un rapport du Sénat, le numérique aurait eu en 2019 une empreinte carbone de 15Mt, soit 2% de l'empreinte carbone du pays, les terminaux représentant à eux seuls 81% de cette empreinte.

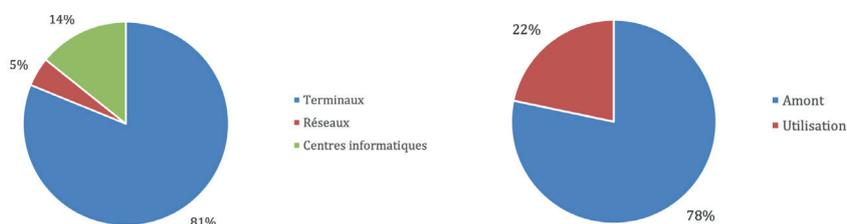


Figure 2 : Empreinte carbone 2019 du numérique en France. Source Sénat, mission d'information sur l'empreinte environnementale du numérique - 2020

Des scénarios volontaristes, proposés par exemple par l'ADEME ou le Ministère de la transition écologique envisagent une réduction sensible de la consommation d'électricité du numérique d'ici 2030. Cela dépendra en fait de plusieurs facteurs, comme les choix faits pour la couverture de la 5G, le développement de l'Internet des objets et celui de la localisation des calculs (edge computing).

Le Conseil général de l'économie a fait une autre prévision pour 2030, en extrapolant les données de l'ARCEP (Autorité de régulation des communications électroniques et des Postes).

	Consommation 2018 (GWh)	Variation du parc	Variation du temps d'utilisation	Efficacité énergétique	Consommation 2030 (GWh)
Téléviseurs	10 000	-0,6%	-3,2%	-5%	4 996
Smartphones	240	+1,9%	+4,5%	-5%	255
Ordinateurs-tablettes	7 510	0%	+4,5%	-7%	4 924
Box	5 030	0%	0%	0%	5 030
Data centers	7 870	+10,1% (pour moitié)		-4%	10 096
Réseaux	5 100	+10,1% (pour moitié)		-4%	6 543
Total	35 750				31 843

Tableau 1 : Prévision 2030 : hypothèses et résultats.
Source : « Réduire la consommation énergétique du numérique », Ministère de l'économie et des finances – Décembre 2019

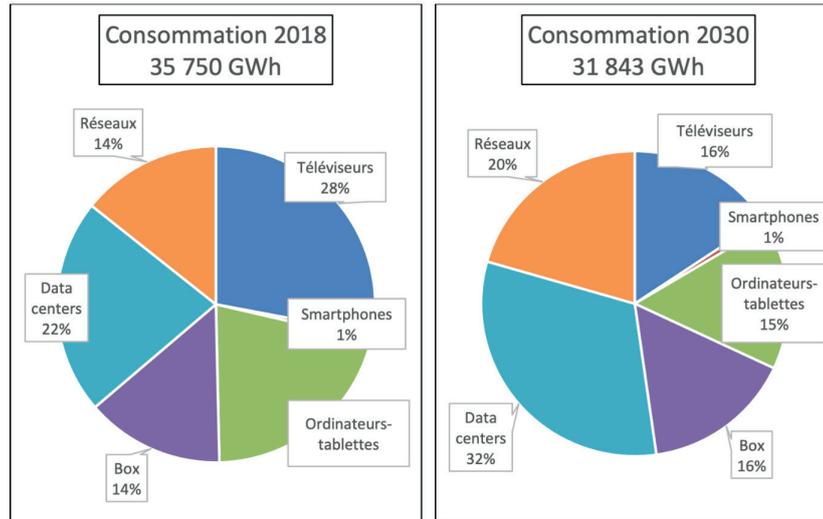


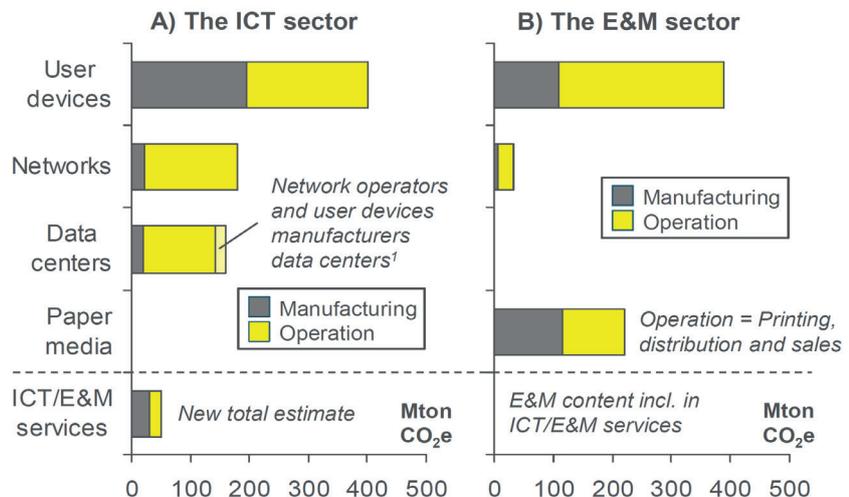
Figure 3 : Répartition de la consommation électrique
 Source : « Réduire la consommation énergétique du numérique », Ministère de l'économie et des finances – Décembre 2019

Une baisse globale est attendue, mais la consommation des data centers et des réseaux augmente sensiblement.

Un point particulier concerne la vidéo, pour laquelle le passage à la résolution 4K augmenterait la consommation de 10%.

Cependant, les mesures de la consommation énergétique sont indirectes, car les technologies de l'information n'ont pas été équipées pour être mesurées. Il faudrait y remédier.

Un autre point est la fabrication des équipements. Elle représenterait aujourd'hui 40% de l'impact global du numérique. Il existe malheureusement peu d'études approfondies sur ce sujet, si ce n'est une étude publiée en 2018 par Jens Malmödin (Ericsson) et Dag Lundén (Telia Company) intitulée « The Energy and Carbon Footprint of the Global ICT and E&M Sectors 2010–2015 » d'où est extrait le graphique ci-après donnant des estimations de l'empreinte carbone des secteurs ICT (Information and Communication Technology) et E&M (Entertainment & Media) en 2015 :



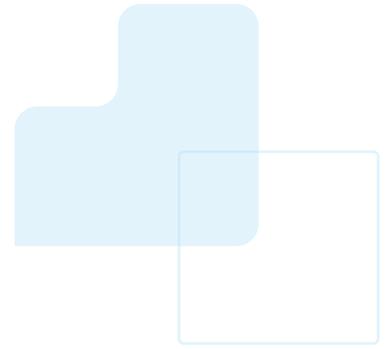


Figure 4 : Estimations de l'empreinte carbone des secteurs ICT (Information and Communication Technology) et E&M (Entertainment & Media) en 2015

Les équipements (téléphones, tablettes, ordinateurs portables, télévision) et plus généralement les objets intelligents (IoT) se comptent en milliards et ce nombre croît rapidement. Comme ils sont fabriqués en grande partie hors d'Europe, nous n'avons guère de moyens d'action directs, si ce n'est la diffusion de labels indiquant l'empreinte carbone.

Les progrès en durabilité des équipements réduisent l'impact de leur fabrication. De même il convient de favoriser la réparabilité et le recyclage. Des réglementations et normes existent, notamment sur l'obsolescence programmée². Une autre partie concerne la limitation des ressources consommées grâce à une loi sur l'économie circulaire.

Le point à relever est que les données précises concernant les équipements et leur recyclage semblent manquer dans le domaine du numérique. Il s'agit d'ailleurs d'une recommandation du Conseil Général de l'Économie (CGE) que les organismes œuvrant dans la transition écologique (Ademe) définissent et publient, en lien avec les autorités de régulation des communications électroniques (Arcep), un corpus de données minimales communes (parc d'équipements et de services numériques) par secteur et le fassent figurer dans les bilans RSE des entreprises.

À ce stade, nous pouvons retenir les points suivants :

- le trafic va continuer à augmenter de façon importante ;
- on mesure mal la consommation énergétique du numérique en raison d'une méthodologie évolutive et par manque de données fiables et exhaustives ;
- la fabrication des équipements représenterait aujourd'hui 40% de l'impact global du numérique ;
- il y a une grande divergence des projections à cause d'estimations différentes des gains de productivité.

2- Article L. 213-4-1.-I. du code de la consommation : La réglementation s'applique à travers le refus de l'inamovibilité des batteries, qui a fait l'objet d'une directive européenne retranscrite en droit français par un décret du 10 juillet 2015, et l'obligation d'information sur la disponibilité des pièces détachées (article L. 111-4 du code de la consommation).

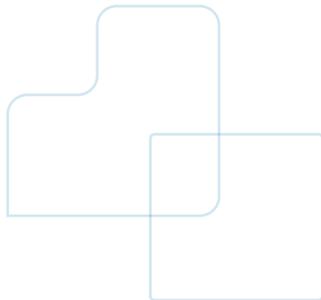
02

LIMITES PHYSIQUES ET TECHNOLOGIQUES

POINT DE VUE THERMODYNAMIQUE

Le lien entre énergie et information a été pour la première fois remarqué dans le paradoxe du Démon de Maxwell (1867). La levée de ce paradoxe a permis de clarifier le rôle thermodynamique des technologies de l'information et d'établir l'équivalence entre information manquante et entropie. En effet, le rôle des technologies de l'information est d'accroître la connaissance dont on dispose sur un système donné. D'un point de vue thermodynamique, cette ambition contrarie l'évolution naturelle d'un système isolé vers une augmentation de son information manquante si bien que l'on doit admettre, pour respecter le second principe de la thermodynamique, que le système sur lequel on souhaite acquérir de l'information est couplé à une source d'énergie noble de telle sorte que l'entropie évacuée sous forme de chaleur vers le thermostat est supérieure à l'information acquise sur le système. Le processeur est la machine thermodynamique qui réalise cette acquisition : pour accroître la connaissance d'un système donné ou de manière équivalente baisser son entropie, elle consomme de l'énergie noble (d'origine électrique), et la dissipe dans un thermostat en conservant globalement les transferts d'énergie, selon le premier principe de la thermodynamique. Sa machine duale étant le moteur de Szillard (1929), la classification thermodynamique suggère de voir le processeur comme une machine frigorifique dont l'efficacité est définie par un coefficient de performance. Ce dernier est utile pour calculer le service thermodynamique rendu par la digitalisation au terme d'un cycle élémentaire de fonctionnement. Afin de disposer du processeur pour le calcul élémentaire suivant, son cycle de fonctionnement se « ferme » sur un mécanisme d'effacement (paradigme de Landauer³) qui permet de rapporter la valeur de l'information acquise à l'énergie qui lui a été allouée par la polarisation des circuits d'électronique logique :

- actuellement, le coefficient de performance atteint par la technologie CMOS pour affecter une mémoire tourne autour de 10^{-5} ;
- avec l'électronique de spin ou spintronique⁴, ce coefficient de performance gagnerait 2 voire 3 ordres de grandeur, même si cette technologie n'est pas encore capable de proposer un ordinateur complet.



3- Formulé pour la première fois en 1961 par R. Landauer d'IBM, le niveau d'énergie minimal qu'il faut consentir pour effacer un bit d'information s'établit à $k_B T \ln 2$ où k_B est la constante de Boltzmann et T la température du système physique considéré. Cette valeur postulée empiriquement a été confirmée expérimentalement en 2012 par une équipe de l'Ecole Normale Supérieure de Lyon.

4- La spintronique ou électronique de spin propose un contrôle de l'information sur le spin de l'électron (par application d'un champ magnétique) et une lecture de l'information grâce au courant de charge de l'électron.

LOIS DE MOORE ET KOOMEY

Depuis l'avènement de la technologie CMOS et du microprocesseur (1960), les progrès réalisés dans la capacité d'intégration des transistors et l'augmentation de la finesse de gravure ont amélioré les performances des processeurs de telle manière que :

- le nombre de transistors des microprocesseurs sur une puce de silicium double tous les deux ans (loi de Moore, 1971) ;
- le nombre de calculs par Joule (unité d'énergie) dépensé double tous les 18 mois environ (loi de Koomey, 1965).

Bien qu'il ne s'agisse pas de lois physiques, mais seulement d'extrapolations empiriques, ces prédictions se sont révélées exactes :

- entre 1971 et 2001, la densité des transistors a doublé chaque 1,96 année ;
- entre 1946 et 2009, le nombre de calculs par joule dépensé a doublé environ tous les 1,57 ans.

En conséquence, les ordinateurs sont devenus de plus en plus petits et de moins en moins coûteux tout en étant de plus en plus rapides et puissants.

Néanmoins, le paradigme de Landauer qui fixe l'énergie minimale qu'il faut consentir pour effacer une information élémentaire (2,80 zJ à la température ambiante) constitue une limite physique de la loi de Koomey dont l'extrapolation devrait être atteinte à la fin de la décennie 2020 avec l'architecture actuelle des ordinateurs. Les limites de la loi de Moore sont décrites ci-après. Dans les faits, les premiers signes d'essoufflement des lois de Moore et Koomey se manifestent depuis le milieu de la décennie 2000 – 2010.

Afin de repousser ces limites, d'autres paradigmes de calcul ont été proposés, notamment en 1973 par C. Bennett d'IBM, qui s'appuient sur une distinction entre entropie physique et entropie computationnelle :

- l'entropie physique est associée aux fluctuations d'un système macroscopique pour atteindre l'équiprobabilité de ses micro-états ; la chaleur dégagée est liée aux variations de ces fluctuations quand le système évolue et l'irréversibilité est due à une évolution imposée en un temps fini ;
- l'entropie informationnelle est associée à la perte d'information ; la moitié des pertes est due à l'irréversibilité (effacement de la mémoire des entrées), l'autre moitié à la dissipation des circuits de commutation des transistors polarisés par des échelons de tension.

Dans ce contexte, on pourrait envisager :

- de ralentir les processeurs en les polarisant par des rampes de tension suffisamment lentes ;
- sous réserve de disposer d'une mémoire suffisante, de conserver les étapes intermédiaires du calcul, pour pouvoir « rembobiner » de manière réversible l'exécution du calcul complet et récupérer l'énergie associée aux mécanismes d'effacement.

Afin de disposer du résultat en un temps comparable, on adopterait alors la parallélisation. Alors que la possibilité de réaliser un calcul réversible résulte classiquement d'un compromis entre les empreintes matérielle (capacité mémoire et processeurs) et énergétique, elle constitue une qualité intrinsèque au calcul quantique (voir 2.2).

1/ LE RALENTISSEMENT DE LA LOI DE MOORE

Depuis plus de quarante ans, la loi dite de Moore a caractérisé la croissance exceptionnelle de la puissance des microprocesseurs. Cette loi empirique résulte de la capacité des fabricants de semi-conducteurs à doubler tous les dix-huit mois la densité de transistors gravés sur le silicium au cours des années passées. Cependant depuis quelques années, le maintien du rythme de cette évolution est confronté à des difficultés considérables. Les limites imposées par la puissance dissipée et la densité d'intégration complexifient de plus en plus les progrès possibles en matière de miniaturisation des transistors.

Sur le plan de l'intégration, des problèmes nouveaux apparaissent. L'extrême densité des transistors (plusieurs dizaines de millions par mm^2 en technologie 5 nanomètres) impose d'empiler de plus en plus de niveaux de métallisation pour les interconnecter, ce qui a tendance à faire perdre en partie l'avantage de l'augmentation de leur performance par l'augmentation des délais de propagation des signaux dans les circuits. L'impossibilité de continuer à réduire régulièrement en 2D les dimensions critiques de la lithographie nécessite de passer à des solutions innovantes d'intégration en trois dimensions des transistors. Par ailleurs, l'industrie des semi-conducteurs se trouve aujourd'hui confrontée à une autre limite physique : les transistors nanoscopiques approchent de la taille de l'atome (0,1 nm). À cette échelle, le comportement des particules est décrit par la physique quantique, dont les lois non déterministes mettent en défaut le fonctionnement attendu des transistors. Bien que, pendant plusieurs décennies, la miniaturisation seule ait suffi à stimuler l'innovation, en permettant d'inventer de nouveaux usages et en assurant ainsi la croissance, l'avenir de l'industrie des semi-conducteurs doit désormais se construire par-delà la loi de Moore.

2/ L'ORDINATEUR QUANTIQUE

Une des raisons poussant à l'ordinateur quantique est le risque que le numérique se heurte au mur de l'énergie.

Un système quantique est caractérisé par son « état ». Celui-ci est représenté par un vecteur unitaire $|E\rangle$ dans un espace vectoriel de Hilbert (sur le corps des complexes). Ce vecteur est initialement immobile mais sous l'influence d'actions extérieures il évolue sur une sphère unitaire de façon déterministe. Il s'agit de la solution de l'équation de Schrödinger, valable tant qu'on ne mesure rien.

Une quantité mesurable est représentée par un opérateur (dit une « observable ») qui a sur cet espace des vecteurs propres et des valeurs propres réelles.

On peut prévoir les résultats possibles de sa mesure : ce sont précisément ces valeurs propres, avec pour probabilité les modules carrés des composantes du vecteur sur la base de ces vecteurs propres. Pour que cette mesure ait un sens, il faudra si l'on refait la même obtenir le même résultat ; ceci est garanti par le fait que lorsqu'on fait la mesure, le vecteur E tel qu'il était disparaît et devient le vecteur propre correspondant à la valeur propre qu'on a mesurée .

Un ordinateur quantique est fabriqué à partir de qbits. Un qbit est un système quantique à deux états de base, représenté par un vecteur dans un espace à deux dimensions. Son état peut être écrit $a|+\rangle + b|-\rangle$, on l'appelle état « cohérent ». Deux qbits sont représentés par un vecteur d'un espace à 4 dimensions : $a|++\rangle + b|+-\rangle + c|-+\rangle + d|--\rangle$, et cet état est appelé « intriqué ». Si on passe de 2 qbits à 10 qbits par exemple on évolue dans un espace à $2^{10}=1024$ dimensions.

Un algorithme va faire évoluer ce vecteur d'état. Il est constitué d'une série de « portes quantiques » qui le modifient pas à pas. Une porte quantique est une impulsion radiofréquence du type de celles de la résonance magnétique nucléaire, sur un ou plusieurs qbits.

À la fin de son exécution, on « va au résultat » en mesurant quelque chose.

L'avantage du quantique est que chaque opération se fait sur le vecteur d'état dans son ensemble, donc affecte tous les états de base à la fois. C'est équivalent à un calcul massivement parallèle. Dans l'exemple de la recherche d'un nom d'abonné à partir de son numéro de téléphone, on montre comment, après avoir créé un état qui combine les noms de tous les abonnés, un opérateur appelé « Oracle » permet de changer le signe du coefficient de l'abonné qui a le numéro de téléphone donné au départ. Il reste alors à appliquer une série de portes quantiques pour concentrer au maximum l'état du système sur celui dont le coefficient vient d'être inversé (suite de l'algorithme de Grover). L'état du système est alors "presque à coup sûr" l'état recherché. Pour avoir une certitude, il convient de refaire une seconde fois la même opération ou de la faire en parallèle sur deux ordinateurs quantiques identiques.

En pratique, un qbit peut être un électron isolé avec un spin ou un noyau de spin $1/2$, une jonction Josephson, des ions ou des atomes de spin $1/2$ piégés à très basse température, des condensats de Bose-Einstein, des quantum dots, des spins de noyaux d'atomes au sein d'une molécule, pourvu qu'on ait deux états de base...

L'ordinateur quantique existe réellement jusqu'à 5 ou 10 qbits. Au-delà, la moindre interaction entre l'ordinateur quantique et l'extérieur peut être considérée comme une mesure et crée une décohérence qui entraîne des erreurs. En réalité, pour la plupart des spécialistes, il faudrait au moins une centaine de qbits pour que l'ordinateur quantique devienne intéressant.

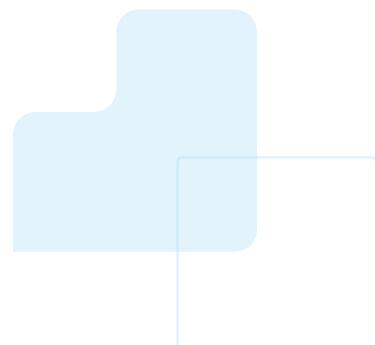
Les pistes pour réduire les erreurs seraient de travailler à très basse température et si possible avec le minimum de connexions entre l'intérieur et l'extérieur du cryostat, ou de calculer plus vite que le bruit thermique, ou finalement de corriger les erreurs comme pour les ordinateurs classiques.

Concernant la consommation énergétique du quantique, deux raisons principales peuvent confirmer que le calcul quantique serait moins énergivore que le calcul classique. La première est que le calcul demande nettement moins d'étapes. La deuxième raison serait que les « portes » sont réversibles et ne transfèrent qu'une énergie récupérable. Ces avantages restent théoriques à ce jour.

Le calcul quantique permettrait théoriquement de :

- factoriser de très grands nombres (cryptographie : algorithme de Shor) ;
- trouver la configuration d'énergie minimum d'une molécule complexe en chimie ou en biologie ;
- optimiser des systèmes très complexes : logistique, finance, réseaux de toutes sortes, physique des particules, ... ;
- accélérer le fonctionnement de l'intelligence artificielle.

Du fait de son mode de fonctionnement, un ordinateur quantique est souvent optimisé pour un seul type d'utilisation (par exemple le recuit simulé). Dans quelques années, on pourra probablement utiliser l'ordinateur quantique pour sa performance dans certains calculs très spécialisés, mais même cela n'est pas pour demain.



03

INFRASTRUCTURES NUMÉRIQUES. ÉTAT DE L'ART ET ÉVOLUTION

1/ PARAMÈTRES DES SYSTÈMES INFORMATIQUES

LES PRINCIPAUX PARAMÈTRES D'USAGE DES SYSTÈMES INFORMATIQUES SONT :

1. La puissance de traitement, exprimée en nombre d'opérations par seconde. Elle-même dépend de :

- la fréquence ;
historiquement, on a pu jusqu'ici l'augmenter constamment, en poussant la miniaturisation des circuits intégrés (loi de Moore). Mais on plafonne désormais à 4-5 GHz ;
- le nombre de circuits intégrés par processeur, et le nombre de processeurs ;
- le mode de calcul, avec aujourd'hui l'introduction de processeurs spécialisés pour accroître la puissance de traitement sans augmenter la fréquence (dont la mise en œuvre de traitements en parallèle).

Toutefois, la spécialisation des processeurs oblige à concevoir en même temps les circuits et les logiciels, alors que les processeurs universels pouvaient être fabriqués séparément, en très grande série. D'autre part, la spécialisation, et plus particulièrement la parallélisation, périclète jusqu'à 90 % du logiciel existant.

2. L'intensité du trafic, ou débit, exprimée en nombre de messages simultanés.

Celle-ci est liée au mode de calcul et à l'organisation du système.

Les opérations qui demandent une grande puissance ne peuvent être traitées que sur des supercalculateurs, ou au moins des calculateurs groupés :

- une part du trafic est la communication de ces calculateurs avec les terminaux, téléphones, objets communicants... ;
- une autre est la conséquence de l'usage d'une algorithmique distribuée comme celle de la blockchain, qui veut substituer aux validations centrales des validations par les usagers eux-mêmes ;
- une autre est le streaming des vidéos. Elle croît avec la définition des images. Faut-il dépasser 4K ?

3. La latence, ou temps de transmission d'une information d'un point à l'autre.

La latence dépend directement de la distance parcourue par l'information.

4. La consommation d'énergie

Il y en a de deux sortes :

- la fabrication et le recyclage des matériels. L'Europe s'est donné des normes, dont les fabricants s'inspirent, encore plus pour leurs objectifs à 10 ans ;
- le fonctionnement.

Les deux sont difficiles à mesurer :

- la première parce que les cycles complets sont mal traçables, d'autant que le gros de la fabrication est en Asie, avec une production d'énergie fortement émettrice de CO₂ ;
- la consommation en fonctionnement dépend des configurations de calcul, qui varie en permanence et de façon non décelable. C'est un vrai sujet de recherche.

5. La complexité algorithmique

En fonction de la quantité de données à traiter, cette complexité s'exprime en termes de taille de la mémoire ou du nombre d'opérations qu'utilise un programme d'ordinateur qui met en œuvre un algorithme. Lorsque l'algorithme est distribué, c'est-à-dire lorsqu'il résulte de l'exécution par des acteurs indépendants et asynchrones communiquant par messages, alors le nombre de messages échangés constitue un critère supplémentaire de complexité.

6. La plateforme

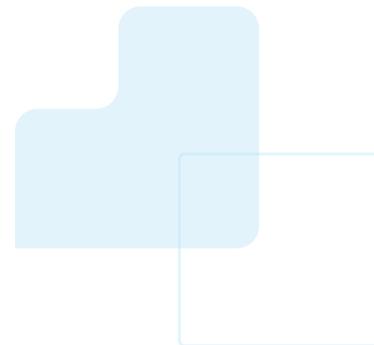
L'hébergement des données dans des calculateurs relativement puissants (data centers) permet aux usagers d'accéder à des puissances de traitement qu'ils ne pourraient pas se payer. Et la combinaison d'un grand nombre de données ouvre la voie à des services nouveaux. Mais la plateforme a un envers : la propriété des données, ou au moins les droits d'accès. Pour échapper à la mainmise des Gafa sur leurs données, les Européens ont décidé de constituer le système de labellisation GAIA-X. Les Gafa en font partie ; heureusement, car vouloir se passer de leur base coûterait trop cher. Mais ce sont les grands industriels européens qui seront les garants de notre indépendance. Les Allemands l'ont bien compris.

7. Les architectures, les compromis centralisation-décentralisation, l'optimisation des systèmes

Puissance de calcul et plateforme poussent à la centralisation, réduction de la latence et des échanges à la décentralisation.

Le compromis se traduit par des choix architecturaux, avec étagement et mise en réseau des centraux et des terminaux :

- le FOG, qui est l'architecture de La 5 G, avec des clouds étagés ;
- l'edge, qui consiste à faire le plus possible de traitement en local (en association avec des capteurs-actionneurs, notamment dans les voitures autonomes ; pour celles-ci, l'edge est d'ailleurs indispensable pour réagir assez vite, mais les modèles d'intelligence artificielle doivent être chargés depuis des supercalculateurs).



2/ NOUVELLES ARCHITECTURES



2-1 / LE FOG COMPUTING : L'ARCHITECTURE NOUVELLE DE TÉLÉCOMMUNICATIONS DE LA 5G

L'architecture courante est en transition : on constate la mise en place d'une couche intermédiaire, le FOG computing (des millions de processeurs avec des réseaux locaux, connectés avec les data centers, mais aussi qui gèrent les connexions en bout du réseau avec les mobiles, les objets connectés, les capteurs de surveillance, les capteurs du smart-grid, les capteurs/actionneurs industriels, demain les véhicules connectés, etc.). Le FOG computing est réalisé avec des ordinateurs compacts et fiables, comme les NUC d'Intel qui peuvent aussi agir en tant que « routeurs SDN », interconnectés par des réseaux locaux avec des câbles Ethernet ou optiques qui font partie de la structure en dur, et des petits serveurs pour stocker des données et offrir des services. Son rôle est d'alimenter les logiciels soutenant les applications (mobile, domotique, sécurité et surveillance, gestion du trafic automobile, jeux, etc.), qui au bout du réseau sont connectées aux « terminaux » (capteurs, actionneurs ou mobiles), tout en assurant un accès assez rapide au système.

Les services du FOG sont très profitables pour les opérateurs. Actuellement le FOG se spécialise : il est différent pour chaque type de service, par exemple selon qu'on cherche la sécurité ou la rapidité.

Chaque FOG particulier est un ensemble de machines : d'une part les clients (qui représentent des fins de chaînes, comme les applications) et d'autre part les services. Cette structure a aussi un réseau interne (en dehors de l'Internet public) dans lequel il y a alors un contrôleur SDN et plusieurs NUC. Le contrôleur SDN crée les connexions de façon dynamique pour traiter les demandes qui sont adressées au FOG.

La consommation d'énergie se mesure notamment dans un NUC, en mesurant le courant.

Le contrôleur SDN joue le rôle de stabilisateur du système. Pendant les surcharges, le temps de réponse augmente et le contrôleur SDN adapte les connexions, ce qui permet au temps de réponse de se stabiliser.

Il existe toujours un compromis entre la qualité de service et la consommation d'énergie. On peut optimiser en choisissant le meilleur chemin entre les différents serveurs et réseaux.

En conclusion, le FOG computing apporte des progrès considérables (IoT, voiture connectée) en permettant de spécialiser le support. De plus, la complexité du réseau informatique a augmenté. Le défi est alors de mettre en place des outils de mesure pour comprendre la consommation énergétique du système entier. Il faudra expérimenter et pousser la connexion des outils de mesure. Et enfin, les algorithmes de contrôle dans les contrôleurs SDN permettent de faire des choix raisonnés et partagés pour choisir la performance (soit en consommation d'énergie, soit en temps de réponse).

Par ailleurs il est fondamental de constater que cette architecture de télécommunications est avant tout celle d'un système informatique distribué entre plusieurs clouds interconnectés. La capacité des acteurs traditionnels des télécoms (opérateurs et équipementiers) à maîtriser ce changement de paradigme est une vraie question, de même que leur compétitivité future vis-à-vis des GAFAs

2-2 / L'EDGE COMPUTING ET LES IOT, LE CONTINUUM COMPUTING

Les calculs ne sont plus effectués uniquement dans un point central (cloud ou data center central) mais de manière éclatée, en tenant compte de la localisation de la production des données. Le traitement - ou prétraitement - peut être effectué au sein d'edge computers, au plus proche de là où la donnée est créée. S'il s'agit d'utiliser des données qui diffèrent notamment par la localisation de leur production, le traitement doit être remonté vers un cloud plus centralisé. Pour l'intelligence artificielle et l'apprentissage automatique, l'élaboration (l'entraînement) du modèle du phénomène à étudier se fait de manière centralisée dans le cloud (souvent via un supercalculateur installé à cet effet), car il s'agit de bénéficier à la fois d'une puissance de calcul considérable et de pouvoir croiser des données d'origines très différentes. Une fois le modèle élaboré il peut être chargé dans les dispositifs dont le comportement seront soumis au modèle. C'est en particulier le cas pour la mise en œuvre de voitures autonomes.

2-3 / LA FIN DU PROCESSEUR UNIVERSEL : LE RECOURS À LA SPÉCIALISATION ET AU PARALLÉLISME

Jusqu'à environ 2005, on a pu, grâce à la miniaturisation des transistors selon la loi de Moore, augmenter la performance d'un micro-processeur tout en diminuant sa consommation d'énergie. Ensuite, il a fallu arrêter l'augmentation de fréquence, car au-delà de 4 à 5 GHz la chaleur à dissiper devenait trop importante.

Pour continuer cependant d'accroître la performance selon les besoins, il y a deux voies possibles : la spécialisation des processeurs et le calcul parallèle, les deux pouvant d'ailleurs être combinées.

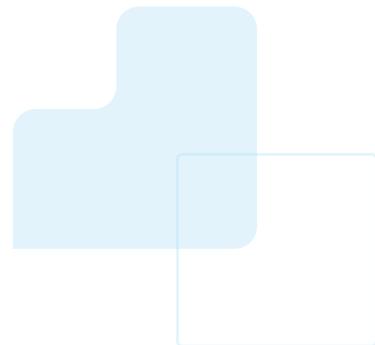
Le problème, c'est la programmabilité. Comme une partie du code applicatif est exécutée dans le composant spécialisé, et que par ailleurs l'exploitation du parallélisme massif nécessite de revoir les algorithmes existants, il faut redévelopper le patrimoine applicatif pour pouvoir bénéficier de l'ensemble des nouvelles capacités de traitement.

Plus un composant est spécialisé et plus il est rapide et efficace énergétiquement, mais il devient moins facile à programmer ; par exemple, les codes Fortran écrits il y a 30 ans ne sont plus utilisables.

Le parallélisme peut être mis en œuvre aux différents niveaux :

- les processeurs ;
- les circuits intégrés, qui peuvent maintenant regrouper jusqu'à quelques centaines de processeurs (les Américains appellent sockets de tels circuits) ;
- les serveurs, où plusieurs « sockets » peuvent coexister ;
- les grands systèmes, comme ceux des supercalculateurs ou du cloud, qui accueillent des milliers de serveurs en réseau local.

Une très grande partie du patrimoine applicatif est fait de programmes séquentiels ou faiblement parallèles. Pour que les applications puissent tirer parti du parallélisme, il va falloir réécrire les programmes, et reconcevoir les méthodes de résolution de problèmes (algorithmes) qu'ils mettent en œuvre. 90% au moins des logiciels existants sont à revoir.



3/ L'OPTIMISATION DES SYSTÈMES

Lors du dimensionnement d'un système énergétique, on obtient des données temporelles (demande en énergie, prix d'achat/vente de l'énergie, profil climatique du site) et des données technologiques (catalogue de convertisseurs d'énergie, catalogue de moyens de stockage à court et long terme). Connaissant ces données, on détermine l'installation optimale, notamment les unités de production, les moyens de stockage, la taille des unités et la planification des usages. Il s'agit de ne pas surdimensionner et d'éviter les conflits entre systèmes, mais en même temps de répondre à la demande. C'est un enjeu majeur pour les entreprises en énergie.

Pour se fixer un objectif, le critère classique est le compromis entre coût du système et performance. Aujourd'hui, on y ajoute la prise en compte de critères environnementaux, notamment l'empreinte carbone. L'idée est alors de fédérer les sciences de l'énergie et celles de l'environnement, pour trouver comment mettre en musique tous les critères environnementaux. La mise en équation des contraintes environnementales et sociales est difficile. Et la multiplicité des critères va nécessiter une puissance de calcul plus élevée. La solution est de rechercher un optimum raisonnable (proche de l'optimum réel) dans un temps raisonnable (dépendant des acteurs et du secteur industriel visé). Il faut aussi intégrer les aléas dans la prise de décision (aléa sur la ressource, la demande, les conditions économiques, etc...). Cette intégration peut demander des temps de calcul prohibitifs. L'enjeu est de trouver des méthodes alternatives. Modéliser à tous les pas de temps (par exemple 10 min pour la chaleur ou 10 s pour l'électricité) conduit à une taille de problème excessive. Il faut réduire le volume des données en créant des données types (cluster type).

Pour modéliser un système physique, tenir compte des usages de tous les composants permet d'ajuster le choix de la précision physique. Plus elle sera élevée, mieux le modèle rendra compte du fonctionnement (par exemple, la dégradation du système, etc.), mais plus le temps de résolution sera long.

Il est alors primordial de communiquer avec les experts du numérique pour mettre en place soit des modèles de données en faisant de l'apprentissage, soit des modèles dits hybrides qui sont des modèles physiques plus simples couplé à des modèles de données.

Un premier axe de progrès est d'arriver à *mesurer la consommation énergétique*.

Les chiffres que l'on a actuellement sur la consommation générale du numérique sont à la hausse. La plupart des études montrent que c'est vrai depuis les débuts du numérique. Malgré toutes les améliorations de l'efficacité énergétique, l'accroissement du trafic et du nombre d'équipements l'emportent, surtout avec le déploiement de l'internet des objets (IoT). Cependant, il est compliqué d'avoir une idée précise de la consommation du numérique. Plusieurs idées fausses doivent être revues, comme dire qu'un équipement informatique non utilisé consomme peu. Un serveur qui n'est pas utilisé continue de consommer de l'énergie à presque 50% de sa consommation pleine. Et c'est le cas de la plupart des infrastructures internet, qui ne sont pas souvent utilisées à saturation.

Des serveurs identiques ne consomment pas la même énergie. Ils ont la même performance mais ne consomment pas de la même manière ; la différence peut être de 10 à 15%. Faire le profil énergétique d'un serveur en se basant sur les performances du CPU est également faux. Le taux d'utilisation du CPU n'est pas un bon indicateur.

Obtenir la consommation énergétique d'une application ou d'un service donné est compliqué, car plusieurs sous-services s'exécutent en même temps, communiquent entre eux et utilisent des serveurs différents et ont donc des consommations différentes. Sachant aussi que certains serveurs ne sont guère disponibles à la mesure.

Pour optimiser cette consommation, il est alors nécessaire d'avoir des modèles déployés sur des clouds grand public mais aussi privés pour essayer d'estimer la consommation des services. Cela se fait sur des architectures de plus en plus distribuées. Les calculs théoriques sont erronés du fait de l'empilement des couches matérielles et logicielles. Il s'agit de chercher à comprendre cette consommation énergétique ainsi que sa distribution pour pouvoir concevoir des systèmes d'informations adaptés au besoin de l'utilisateur.

Le deuxième axe est de *concevoir des systèmes informatiques « juste assez »*. Par exemple, sur la digitalisation de l'énergie, quels objectifs doivent être pris en compte : performance, sécurité des données, efficacité en temps réel, tolérance aux pannes (redondance, surprovisionnement, comment optimiser une infrastructure déjà construite ou neuve). Pour cela, il est important de développer des outils d'analyse et d'aide à la décision en prenant en compte toutes les contraintes. Cela nécessite beaucoup d'interdisciplinarité et d'échanges. Les systèmes d'information seront contraints d'un point de vue énergie, en termes de pic de puissance, ou d'alimentation en énergie renouvelable intermittente.

Le troisième axe concerne *l'amélioration de l'efficacité énergétique des infrastructures*. Sachant que les nouvelles générations de mobiles s'empilent, il faudra compenser leurs consommations par les gains sur les usages dans d'autres secteurs.

Le dernier axe concerne *la sobriété* pour réduire la consommation énergétique. Cette sobriété peut consister à rendre le numérique plus "vert", à analyser le cycle de vie des matériaux et améliorer le recyclage, à se poser la question de l'internet en priorisant les capteurs et intelligences numériques les plus utiles, et finalement à poser les questions d'acceptabilité sociétale de cette sobriété.

On peut emprunter le cadrage global à la Commission européenne. En juillet 2020, celle-ci a publié sa stratégie "intégration du système énergétique", définie comme la planification et le fonctionnement coordonnés du système énergétique considéré « comme un tout », tous vecteurs énergétiques, infrastructures et secteurs de consommation confondus. Son objectif est de permettre une décarbonation efficace, abordable et en profondeur de l'économie européenne.

Les trois piliers de l'intégration sectorielle sont :

- un système énergétique « circulaire », avec l'efficacité énergétique au centre (ex : installations de production combinée de chaleur et d'électricité ou via l'utilisation de certains déchets et résidus, chaleur fatale provenant de procédés industriels ou de centres de données, énergie produite à partir de biodéchets) ;
- une électrification accrue des secteurs d'utilisation qui s'y prêtent (ex : recours aux pompes à chaleur pour le chauffage ou les

procédés industriels à basse température, électromobilité, fours électriques dans certaines industries etc.) ;

- l'utilisation de carburants renouvelables et de carburants peu carbonés, y compris l'hydrogène, pour des applications dans lesquelles le chauffage direct ou l'électrification ne sont pas la meilleure solution ou pas possibles, pas efficaces ou trop onéreux. Par exemple, le gaz naturel dans une phase de transition entre charbon et solutions décarbonées ; l'hydrogène renouvelable et bas carbone dans des procédés industriels, voire dans les transports routiers et ferroviaires lourds ; les carburants de synthèse dans les transports aérien et maritime (biomasse, biogaz etc.).

Concrètement, les caractéristiques de l'intégration sectorielle (ISE) sont :

- approche énergétique multi-vectorielle, pour garantir décarbonation, sécurité d'approvisionnement, résilience et maîtrise des coûts ;
- le principe de l'efficacité énergétique au cœur ;
- la demande énergétique devra être décarbonée, notamment via l'électrification (nucléaire pour ceux qui en font le choix, électricité et gaz renouvelables) ;
- l'ISE est également clé pour faciliter la pénétration de l'électricité renouvelable dans le système énergétique (via P2G et G2P notamment). Elle prévoit aussi l'optimisation des infrastructures existantes (acceptabilité, coûts etc.) pour assurer la flexibilité et le back-up ;
- recours au gaz et aux gaz renouvelables, notamment quand l'électrification s'avère trop difficile ou trop coûteuse ;
- modèle énergétique en mutation, plus décentralisé, apparition de systèmes énergétiques locaux, de « communautés énergétiques », coexistant avec le modèle centralisé traditionnel ;
- accroissement de l'interfaçage / appel à la digitalisation notamment, qui est aussi facteur de risques (cybersécurité).

4/ SÉCURITÉ ET RÉSILIENCE

Le premier besoin, en particulier pour le réseau électrique, est de restaurer le plus vite possible l'état des systèmes qui tombent en panne.

L'autre est de résister aux cyber-attaques. La cybersécurité est un enjeu majeur pour le numérique. Il existe divers groupes de discussion à une échelle internationale. À l'échelle européenne, la question de la cybersécurité vise à partager les meilleures pratiques et standards. Aujourd'hui, le système énergétique n'est plus linéaire, mais présente des complexités sur les flux de données passant par différents acteurs. On parle aussi des couplages entre différentes énergies : gaz/électricité etc...

Cette interdépendance propage le risque, depuis l'utilisateur jusqu'au système. Le projet européen apporte une cohérence d'approche globale dans cette gestion de risque. La cybersécurité amène certes des coûts supplémentaires, notamment énergétiques, mais est essentielle pour assurer une continuité de fonctionnement du système entier.

Avec différents systèmes qui coexistent, la communication intersectorielle est indispensable pour contrôler et optimiser les interdépendances. Que ce soient les données d'échange ou les matériels, la gestion d'une communication numérique intense et en temps réel est nécessaire pour équilibrer les systèmes. Un système de contrôle décentralisé (par exemple, Edge-to-cloud) apportera de la valeur mais aussi de nouveaux points potentiels de cyberdéfaillance. Aujourd'hui, compte tenu de la grande diversité des transcriptions réglementaires dans les États membres, l'UE cherche à mettre en place des contraintes régulatrices auprès des industriels. La directive européenne NIS 1.0 sur la résilience des entités critiques révèle de grandes différences de mise en œuvre entre les États membres. Dans le cadre de la version NIS 2.0, la Commission a lancé un cycle de consultations qui devrait se terminer début 2022.

En 2019, l'attaque de SolarWinds a démontré que la vulnérabilité de la chaîne d'approvisionnement était désormais un enjeu majeur de la cybersécurité industrielle. Un système supposé de confiance peut propager des vulnérabilités à d'autres systèmes. L'attaque de la chaîne d'approvisionnement de SolarWinds a également mis en évidence l'existence de failles majeures dans le secteur de l'énergie : le manque de préparation et la faible

résilience systémique.

Il existe différentes recommandations à l'intention des «clients» et des «fournisseurs», mais elles ne sont souvent pas mises en œuvre (réglementation) ni appliquées (manque de sensibilisation, coûts et efforts, etc.).

Les principaux constats et recommandations liés à la cybersécurité sont les suivants :

- le rythme des changements induit de nouvelles vulnérabilités dont il faut savoir se protéger rapidement. Les dispositifs de sécurité sont toujours en retard sur les menaces ;
- les systèmes de surveillance dotés de capacités de visibilité totale font défaut ;
- les mesures de cybersécurité représentent un coût, et il est absolument nécessaire que tous les acteurs/opérateurs du système agissent en conséquence. Les entreprises devraient consacrer 10 à 15 % de leur budget informatique à la protection contre les violations de données et les attaques. Les dépenses supplémentaires liées à la cybersécurité industrielle pourraient atteindre une valeur égale ;
- les collaborations intersectorielles sont indispensables pour gérer le cyber-défi de l'ensemble du système énergétique ;
- il est indispensable d'avoir une approche réglementaire harmonisée dans l'UE ; le NIS 2.0 de l'UE est un pas dans la bonne direction ;
- le risque cybernétique de la chaîne d'approvisionnement doit être géré ;
- un système de cyber-certification pour les technologies critiques est indispensable ;
- les personnes qualifiées restent insuffisantes sur notre marché de l'emploi. La cybersécurité industrielle n'est encore que trop souvent effleurée dans les formations académiques. Un investissement important dans la formation théorique et pratique est nécessaire.

O4

ANALYSE SECTORIELLE

1/ ÉLECTRONIQUE, INFORMATIQUE, TÉLÉCOMMUNICATIONS

Comprendre la digitalisation commence par les composants matériels d'un système numérique.

La France a le potentiel de remplir les chaînes de valeur de la transformation numérique :

- la fabrication de composants réalisée par le CEA-LETI et ST Microelectronics ;
- Kalray, spécialiste dans les processeurs parallèles ;
- les data-centers d'OVH ;
- les supercalculateurs ainsi que le programme quantique d'Atos ;
- la filière télécommunication, du point de vue d'Orange ;
- et le contrôle des supply chains, par l'exemple d'Atos et de Kalray.

1-1 / COMPOSANTS ; RÉALISATIONS DU CEA-LETI ET DE ST MICROELECTRONICS

La fabrication des composants électronique a deux axes principaux :

- le cycle de vie, avec tout ce que cela implique dans les matériaux, l'approvisionnement et la capacité à produire de la façon la moins consommatrice ;
- les données et l'apport de l'intelligence artificielle pour démultiplier les traitements fondés sur leur analyse.

Les industriels sont mobilisés sur les enjeux de consommation des objets électroniques (analyses de cycle de vie). Le CEA-LETI est extrêmement motivé pour travailler sur ces sujets. L'enjeu proposé est de réduire la consommation d'un facteur 1000 d'ici 10 à 15 ans. Pour ce faire, il propose un plan d'action avec 9 axes. Chaque axe est défini aussi sur la technologie : le premier concerne les semi-conducteurs, le deuxième niveau l'architecture et le dernier niveau le logiciel embarqué en lien étroit avec l'application visée.

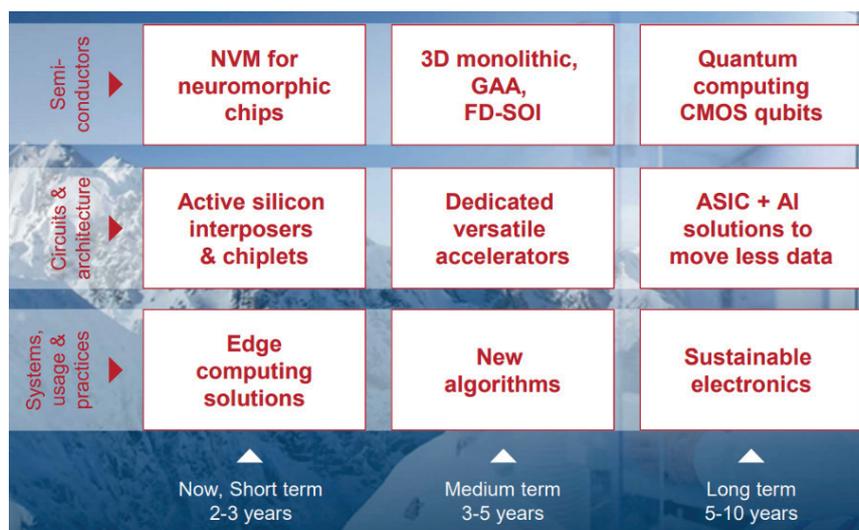


Tableau 2 : Les 9 axes du plan d'action proposé. Source : Présentation CEA-LETI, Sebastien Dauve

Aujourd'hui, pour adresser les problématiques définies, il est nécessaire d'avoir une approche système sur ces trois niveaux.

Dans le premier niveau, qui est le court terme, on peut parler de calcul neuromorphique. On traite massivement l'intelligence artificielle sur le cloud. L'enjeu est d'être capable de monter des circuits d'intelligence embarquée au niveau d'un objet lié à l'IoT. En performance de calcul et en performance énergétique, on est encore loin des capacités du cerveau d'une abeille, et beaucoup de travail est nécessaire. On peut chercher à modifier les architectures et rapprocher le plus possible l'unité de calcul et l'unité de mémoire, car d'un point de vue énergétique, le mouvement des données représente 90% de la consommation des processeurs.

Aujourd'hui on peut utiliser des mémoires non volatiles, cohérentes avec la technologie CMOS. Plusieurs critères de comparaison sont définis, comme la puissance de programmation, l'endurance, la vitesse d'écriture etc... On explore une mémoire résistive et une mémoire 0- électrique qui offrent des critères largement pertinents comparés à la mémoire flash très utilisée. On peut alors déjà avoir une consommation très réduite.

Toujours à court terme, on a d'autres moyens de réduire la consommation. Par exemple de développer un substrat de type silicium on insulator (SOI) et d'utiliser une technologie appelée "fully depleted SOI (FD SOI)" pour faire des microcontrôleurs ; elle offre intrinsèquement une alternative intéressante en termes de consommation, en particulier pour des applications embarquées.

A cinq ans, le gain dans le rapport entre la puissance de calcul et la consommation d'énergie pourrait être d'un facteur 100.

Aujourd'hui, STMicroelectronics travaille à 28 nanomètres ; le CEA-LETI prépare le 10 nanomètres qui sera mis sur le marché en 2025-2026. Sous 10 nanomètres, l'intérêt du SOI est difficile à démontrer. TSMC a déjà commercialisé les 7 nanomètres. Le CEA-LETI explore des concepts de nanosheets empilés qui pourraient permettre d'aller vers les 5 nanomètres et un peu plus loin.

STMicroelectronics est très présent sur le marché des objets intelligents (IoT), où la consommation d'énergie est déjà un vrai problème. Le trafic internet va croître très fortement avec l'arrivée de la 5G, qui permettra de multiplier les IoTs. Les échanges d'informations pourraient atteindre 25 à 50 ZB en 2025 (1 ZB = 10²¹ octets), et 500 ZB en 2030. L'intelligence artificielle utilisée par les IoTs se réalise dans le cloud ou les data centers, via des logiciels qui tournent sur de gros calculateurs. ST Microelectronics estime que les data centers pourraient consommer de 25 à 50 fois plus d'énergie à l'horizon 2030. Cela représenterait 100 à 300 Twh. Il existe des architectures qui essaient de maîtriser cette augmentation. Mais le moyen le plus efficace, c'est de maximiser les traitements locaux sur les données produites localement, ce qu'on appelle l'edge computing.

Un modèle de réseau de neurones est entraîné par apprentissage en central pour bénéficier de toutes les données de circulation accumulées par le plus grand nombre de véhicules possibles et sur une durée la plus longue possible (ce qui nécessite une très grande capacité d'archivage et de puissance de calcul). Ce modèle est ensuite chargé dans chaque véhicule pour pouvoir réagir en temps réel par des processeurs spécifiques embarqués aux événements de circulation. Cette phase dite d'inférence neuronale est donc effectuée localement et relève donc de l'edge computing.

ST Microelectronics vise le traitement de l'IA en edge. L'objectif, c'est 80 % d'edge en 2030. Dans une voiture, il y a par ailleurs beaucoup de tâches d'inférence à effectuer : la reconnaissance vocale, la détection d'objets, la détection des anomalies... Actuellement, les capteurs ne font pas de traitement. Le but est donc d'agir au niveau de chaque composant :

- associer aux capteurs un traitement local d'informations ;
- doter les microcontrôleurs de systèmes de communication intégrés et de moyens de sécurité embarqués ; implanter l'inférence neuronale en local (réseaux neuronaux, apprentissage et traitement d'informations).

Pour optimiser la consommation, STMicroelectronics utilise de nouvelles technologies, notamment les circuits intégrés FDSOI et les mémoires à changement de phase (pcm), qui permettent en association avec les réseaux de neurones un traitement partiellement analogique et non juste binaire.

Un défi est aussi *d'augmenter la capacité de la mémoire* pour y effectuer des opérations. Aujourd'hui, la consommation d'énergie est principalement liée à la communication entre processeurs et mémoire. Or on a observé que les processeurs graphiques ont la capacité de traiter de l'information plus rapidement et avec moins de consommation que les CPU classiques en la traitant directement dans la mémoire.

Les traitements en local demandent que les logiciels soient intégrés dans les microprocesseurs, puis qu'on leur associe les réseaux de neurones, et enfin qu'on ajoute de la mémoire au niveau des microcontrôleurs. Le but ultime est un système autoalimenté.

1-2 / LES PROCESSEURS PARALLÈLES KALRAY

Kalray fait des processeurs pluri-cœurs. Avec l'IA, les calculs se font près des données avec des accélérateurs: l'edge computing. Il y a deux principaux acteurs du sujet: Intel et Nvidia (focalisée sur l'accélération de l'IA). Il existe un continuum qui va de l'IoT aux data centers et c'est de plus en plus sensible avec les problématiques de la 5G.

La motivation générale de l'edge est la consommation d'énergie, car on calcule localement. Aussi, le fait de travailler près des sources de données réduit nettement la latence.

Kalray connaît bien l'edge, notamment pour la conduite autonome : gestion des données avec les capteurs (caméra, capteur de mouvement, position satellite) (étape appelée perception, dans laquelle il faut calculer vite), localisation, prédiction, décision (avec une technique classique de contrôle-commande). Beaucoup de calcul se fait en local et donc nécessite beaucoup de communication (avec fréquence ultra-basse).

Au niveau des machines informatiques, on travaille avec des multicores homogènes (data centers), processeurs avec une multiplicité de cores. C'est beaucoup plus facile à programmer, mais ne passe pas bien à l'échelle à cause des partages en dehors du core. Les duplications se font avec les groupes de cores appelés manycores. La différence entre multicore et manycore est la qualité. Les multicores ne sont pas forcément spécialisés, les plus efficaces sont alors les manycores.

Kalray construit des processeurs manycores basés sur des cores programmables non spécialisés comme les gpgpu. Les machines Kalray s'utilisent en particulier dans l'edge, beaucoup en inférence neuronale, avec des programmations classiques (type java, c++) mais aussi dans les modèles d'apprentissage automatique (langage de haut niveau).

Ce type d'évolution est inclus dans le projet européen EPI (European Processor Initiative). La deuxième étape sera de viser des calculs plus complexes, par gpgpu puis sur des applications modernes. Ensuite, il s'agira de donner une accélération faisable aux applications de la deuxième étape. Enfin, Kalray souhaite offrir une alternative aux gpgpu et passer au-delà des limites de la technologie actuelle.

1-3 / LES DATA CENTERS ; LES RÉALISATIONS D'OVH

Les data centers sont des équipements particulièrement électro-intensifs. Fin 2020, l'UE a publié un rapport sur la consommation énergétique du cloud. On observe avec intérêt que parmi les scénarios présentés, le "worst case" est celui dont la dynamique de croissance de la consommation électrique est la plus prononcée ... indépendamment des activités déployées (bonnes ou mauvaises pour l'action climatique).

Or, toujours selon la Commission européenne (dans un autre rapport), le marché des données en Europe pourrait peser jusqu'à 6% du PIB durant la décennie (avec, mécaniquement, un besoin important de data centers). Autrement dit, le débat est complexe, et le déploiement de ces équipements (et la croissance des consommations induites) sont vues ou comme une menace, ou comme un facteur de transformation de l'économie européenne.

On constate que le problème se pose à l'échelle locale tout autant que globale, avec une problématique relative aux conditions d'insertion des data centers dans des systèmes locaux (systèmes électriques, mais également réseaux de chaleur).

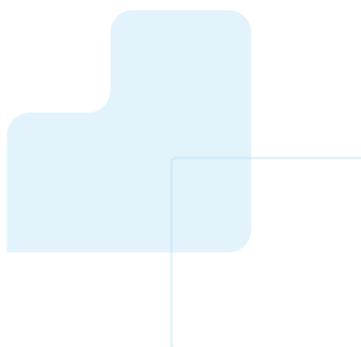
Une illustration peut se faire en comparant l'électronique à l'éclairage : on est parti des lampes à huile puis des bougies, avec des rendements lumineux très médiocres, puis on a vu arriver les lampes à incandescence avec des rendements de 1 à 2 % , puis enfin les leds avec un rendement lumineux allant jusqu'à 25%. Cette dernière mutation s'est opérée en 30 ans. Comparativement, les data centers sont à un niveau encore plus bas que la lampe à huile et il faut espérer que, dans les décennies qui viennent, le même progrès que pour l'éclairage va pouvoir se développer.

Le constat de base n'est pas nécessairement alarmant. Pour la décennie 2010, on a remarqué une stabilisation de la consommation énergétique des data centers. Ce constat est expliqué par le développement de data centers « hyperscales » (i.e. « géants »), nettement plus efficaces que les précédentes générations.

L'enjeu est de déterminer si, dans la décennie 2020, il sera possible de poursuivre cette stabilisation. Sans doute, les gains d'efficacité des équipements qui constituent les data centers ne suffiront pas : l'enjeu est notamment de mieux insérer, selon une vision de « sector coupling », les data centers dans le réseau électrique (en fournissant différents services pour l'équilibre du système électrique). Il est aussi important de considérer les conditions d'alimentation électrique, notamment par des contrats de PPA (Power Purchase Agreement), de façon à garantir un approvisionnement en électricité renouvelable (avec, en retour, une contribution au financement des renouvelables).

Malgré ces différents leviers, la concentration des data centers met en forte tension certains systèmes électriques et demande des investissements supplémentaires en infrastructures (tant dans le domaine de l'énergie que des télécommunications, notamment en ce qui concerne les réseaux de fibres optiques). Diverses projections témoignent de ce phénomène en Europe : en 2030, les data centers pourraient représenter de l'ordre de 20 à 30% de l'électricité consommée au Danemark et en Irlande (selon les opérateurs de réseaux électriques nationaux), pays qui ont développé des stratégies très agressives pour les attirer. En 2019, la Hollande a décidé un moratoire sur l'implantation de nouveaux data centers.

Il est essentiel que ce débat progresse, compte tenu de la volonté européenne d'une neutralité climatique des data centers en 2030. On observe avec intérêt qu'une large coalition d'acteurs se mobilise autour d'un « Climate Neutral Data Centre Pact » qui vise à atteindre cet objectif.



OVHCloud, fournisseur de data centers, a défini ses objectifs sur la base de cinq piliers :

1. l'efficacité énergétique (sujet du PUE). L'objectif est d'atteindre un PUE de 1,3 pour tous les nouveaux data centers en 2025 et 1,3 pour 2030 entre tous les signataires (y compris les anciens data centers, avec pour les zones chaudes un PUE de 1,4) ;
2. l'énergie décarbonée, pour qu'en 2030, 100% de l'énergie utilisée soit décarbonée, avec une liberté sur le choix du mix énergétique. Et déjà atteindre 75% en 2025 (pour les opérateurs avec une puissance supérieure à 50 kW) ; une des voies est de lier la consommation aux moyens de production verts et de créer une relation entre consommateurs et producteurs d'énergie. C'est ce qu'OVH fait pour les serveurs qu'il installe. Le problème n'est pas spécifiquement français, et OVH travaille aussi avec l'Italie, l'Inde, l'Amérique du Nord.
3. l'utilisation de l'eau. Le débat est inévitable : si en France l'accès à l'eau est encore facile, ailleurs cela commence à être contraignant, et il est important de le prendre en considération dès aujourd'hui ;
4. la réutilisation (le recyclage) de tous les équipements, (serveurs etc...) ;
5. la réutilisation de la chaleur, même si cet engagement est limité. En effet, l'utilisation de cette chaleur dépend de l'environnement du data center ; l'intérêt local est varié, sans oublier le lobbying des chauffages urbains face à la gratuité ; et nous rejetons plus de chaleur en été qu'en hiver.

1-4 / LES SUPER CALCULATEURS, LES FABRICATIONS D'ATOS

Comme exemple des cas où l'on a impérativement besoin de supercalculateurs, on peut avoir les prévisions météo, la modélisation du climat, les prospections du sous-sol, et toutes les modélisations numériques (physique des hautes énergies, crash test, séquençage de génomes, Santé et conception de nouveaux médicaments, analyse de risques, Défense...). Au-delà, on a aussi besoin de supercalculateurs pour calculer les modèles d'apprentissage s'appuyant sur des océans de données, comme dans le cas de la voiture autonome avant de télécharger ces modèles dans les véhicules eux-mêmes.

Le défi des années à venir est l'exascale (10^{18} opérations en virgule flottante par seconde, cent fois plus que ce que l'on délivre actuellement).

Les supercalculateurs du groupe ATOS sont au plus haut niveau de performance mondial.

L'iPhone 4 de 2013 a la même puissance de calcul de 1,6 giga flops (floating point operation per second) que le Cray 2 de 1985. Un MacBook Air, c'est 230 giga flops. Le supercalculateur Atos Sequana, c'est 44 millions de giga flops.

Un supercalculateur, c'est un gros radiateur. La consommation énergétique va jusqu'à 20 MW, la masse jusqu'à 200 tonnes ; chaque composant produit une chaleur extrême, jusqu'à 300 watts aujourd'hui et 800 watts d'ici 2 ans. Il peut avoir jusqu'à 10 millions de cœurs, permettant d'atteindre aujourd'hui des puissances de calcul jusqu'à 400 petaflops (1 petaflops = 10^{15} opérations par seconde). L'efficacité énergétique se dégrade en fonction de la taille du supercalculateur. Au-delà d'un certain nombre de nœuds, on ne progresse plus.

A nouveau, l'alimentation des supercalculateurs peut être mise en place dans le cadre de contrats de PPA (Power Purchase Agreement), de façon à garantir un approvisionnement en électricité renouvelable (avec, en retour, une contribution au financement des renouvelables). De même Atos a monté des partenariats de recherche pour élaborer des dispositifs d'alimentation par hydrogène vert.

1-5 / LE PROGRAMME QUANTIQUE

Une plateforme appelée Quantum Learning Machine (QLM) est développée chez Atos. Elle utilise un serveur à grande mémoire partagée (jusqu'à 48 téraoctets de mémoire vive) et permet de mettre tous les calculs en mémoire vive. Elle intègre aussi les travaux réalisés avec les instituts de recherche dans la technologie quantique. Il s'agit de modéliser le bruit quantique des dispositifs physiques et d'arriver à simuler un véritable processeur quantique avec son bruit, ou de retirer ce bruit pour simuler un calcul parfait. Il y a aussi un travail sur les algorithmes quantiques pour les lier directement au supercalculateur.

Le Quantum Safe Program étudie la suprématie quantique qui peut mettre à mal les méthodes de déchiffrement existant aujourd'hui. Ce programme vise à résister aux futures attaques quantiques.

Les enjeux tournent alors autour de l'acquisition de la connaissance sur le quantique, l'apprentissage de la programmation quantique et le test des programmes. Le but est d'optimiser les algorithmes quantiques sur les calculateurs d'aujourd'hui.

1-6 / TÉLÉCOMMUNICATIONS

L'impact actuel des technologies de l'information et de communication dans les émissions de carbone est de 3.7 % de l'impact global. La croissance de la demande de services est exponentielle, mais les émissions carbone restent stables.

Les acteurs de ces technologies sont tous engagés dans la réduction d'émissions, surtout dans le secteur des télécoms. Environ 50 % des achats privés dans le cadre de power purchase agreements (ppa) sont faits par les acteurs du numérique. Les analystes ne prennent pas en compte ces investissements dans leur impact carbone, non plus que les efforts faits dans les nouveaux data centers avec du free cooling. La demande énergétique est beaucoup plus sobre que les prévisions. Tout le secteur est en train d'être décarboné.

L'engagement des opérateurs est le suivant :

- en 2025, 30% de CO₂ en moins par rapport à 2015, cela principalement grâce aux EnR (produites ou achetées). On peut ainsi conjuguer croissance de la consommation d'énergie et décarbonation. À cela s'ajoutent les programmes d'amélioration de l'efficacité énergétique, avec les progrès du hardware et de l'algorithmie ;
- en 2040, neutralité carbone. Les émissions résiduelles incompressibles seront compensées par des puits de carbone.

Aujourd'hui les data centers d'Orange représentent 9% de la consommation totale du secteur des télécommunications.

Orange consomme près de 5,3 TWh dont 2 en France. 85% correspond au réseau et au traitement de l'information. Les data centers d'Orange représentent 9 % de la consommation totale du secteur des télécommunications.

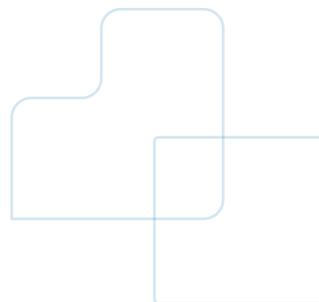
D'année en année, Orange a accumulé des puits de carbone. Il envisage aussi de s'approvisionner chez les acteurs spécialisés. Il y a deux façons d'acheter : tel que produit (y compris l'intermittence), et tel que chez l'acteur, qui peut équilibrer avec d'autres sources. Cela pallie le risque d'intermittence, mais coûte plus cher.

On cherche à faire basculer les utilisateurs vers les systèmes fixes. Il faudra savoir lier l'usage et l'architecture à la consommation. L'important est d'animer de façon globale un écosystème entier.

1-7 / LE CONTRÔLE DES SUPPLY CHAINS

Aujourd'hui 75% des composants utilisés par les fabricants européens sont produits hors d'Europe. Le programme européen European Processor Initiative (EPI) auquel Atos et Kalray participent vise à donner à l'Europe la capacité de créer ses propres processeurs. L'objectif 2030 serait d'arriver à entre 65% et 70% d'autonomie. 27 partenaires travaillent sur un processeur qui sera disponible en 2023. Il intégrera des éléments de sécurité et permettra un traitement efficace de l'entraînement, pour l'intelligence artificielle ; la consommation énergétique sera l'un de ses enjeux et devra être contenue.

Pour les circuits intégrés, à moins d'aides massives, ST ne peut suivre la course à la miniaturisation ultime, menée par deux champions asiatiques : Samsung et TSMC. Intel, qui avait semblé renoncer, a décidé de se remettre à niveau, et de mener une politique plus ouverte : davantage de dialogue avec ses clients, et sans doute la construction d'une usine en Europe.



2/ PRODUCTION ET TRANSPORT D'ÉLECTRICITÉ



Dans le domaine de l'énergie électrique, le numérique a déjà une place pérenne. Il est fondamental au niveau électrique, et il a un fort rôle dans la flexibilité. Pour EDF, les objectifs de la digitalisation sont d'améliorer les processus métier, d'innover en lien avec la stratégie d'entreprise (CAP 2030 pour EDF SA), de créer de nouveaux services et lignes de business, d'améliorer le niveau des employés (développement de compétences, expertises), et de construire un socle technologique robuste.

Autour de la donnée, on cherche à développer les IoTs, la digitalisation en interne, les smartphones pour les différentes activités, les drones pour la surveillance, et les objets Linky. En termes de partage, l'IA est omniprésente, avec l'automatisation, et le concept de dataspace. Enfin, on développe les jumeaux numériques et la réalité virtuelle (dans le cadre de formations).

De plus, on tient compte des valeurs et exigences centrales du numérique : éthique, cybersécurité, problématiques de stockage.

En général, le numérique facilite le fonctionnement de l'entreprise étendue ; il a permis de passer la crise covid-19 en poursuivant les activités à distance.

2-1 / PRODUCTION

L'ensemble des fonctions nécessaires à la production est impacté par le numérique : la sécurisation des intervenants (4G sécurisée dans les centrales), le déploiement des IoTs (écoute, visualisation), l'aide aux intervenants, la certification des tests et productions à l'ensemble des organismes de contrôle...

De plus, l'edge computing va dans le sens de la responsabilisation, et apporte une capacité de production et de fonctionnement en local. Il permet de raccourcir les temps de décision, et d'agir sur l'engorgement du réseau.

Concernant le nucléaire, l'objectif est d'augmenter la durée de vie du parc en toute sûreté, de construire les centrales de demain et de démanteler les sites. Sur les nouveaux réacteurs nucléaires à eau pressurisée (EPRs), on utilise les jumeaux numériques dans la

phase de construction, mais aussi des maquettes de surveillance de vieillissement des structures.

Quant à l'hydraulique, on veut augmenter la performance économique et industrielle, améliorer le pilotage des ouvrages et développer de nouvelles capacités.

Pour le renouvelable (solaire et éolien), les opérations de supervision et de maintenance sont centralisées dans un cloud sécurisé, et l'ensemble des données est envoyé dans un data lake renouvelable qui permet de comparer la production réalisée et celle estimée, d'avoir des indicateurs de performance en temps réel, et ainsi d'optimiser en permanence. Les VPP (Virtual power plants) permettent de piloter de manière virtuelle un ensemble de sites de production. La blockchain permet de tracer l'origine de l'énergie et donne au consommateur la garantie de traçabilité de son énergie (verte ou non verte).

2-2 / OPTIMISATION AMONT-AVAL

On fait de plus en plus face à une logique en temps réel, avec une intégration plus forte au niveau européen, en cherchant à sécuriser les transactions. Un moyen est de développer des systèmes d'IA pour réduire le pas de temps et traiter l'intermittence de production.

Pour la distribution, EDF cherche à être le tiers de confiance du marché, en adaptant le réseau aux nouveaux usages et à la production décentralisée. L'apport du numérique se trouve d'abord sur le volet industriel (supervision, maintenance, optimisation de l'ensemble du réseau, IoTs, 5G), mais aussi dans la création d'une plateforme de diffusion et de fournitures de données de référence sur la production et de la consommation d'énergie. On peut ainsi devenir réellement l'opérateur des données au service de la décarbonation de l'ensemble du secteur, à la fois industrielle, personnelle et publique

2-3 / COMMERCIALISATION ET SERVICES ÉNERGÉTIQUES

L'enjeu est de fournir une qualité d'offre et de performance commerciale, mais également une réactivité en matière d'offre (avec l'explosion des prix de l'énergie imposée sur le marché), ensuite de fournir des services accompagnant la transition écologique et le réchauffement climatique pour maîtriser la consommation et l'efficacité énergétique, et enfin de développer des services de maîtrise de l'énergie (autour des données par l'IA, ou de développer des jumeaux numériques ou des objets connectés), et les services propres à la réduction de la consommation énergétique.

Tout cela s'accompagne d'exigences fortes:

- un service de confiance : avec l'exemple des clouds, les exigences sont portées au niveau européen, au travers de la politique de labellisation de la démarche GAIA-X ;
- concernant la cybersécurité : on cherche à établir un échange transparent avec les acteurs, mais aussi on applique le principe de privacy by design dans le RGPD (Règlement Général pour la Protection des Données) ;
- la problématique du numérique responsable est au cœur des exigences : que ce soit l'IT for green ou le green-IT, on cherche à minimiser le stockage de données, avec une approche «éthique » de l'utilisation de l'IA .

2-4 / LE TRANSPORT D'ÉLECTRICITÉ

RTE, le gestionnaire français du réseau de transport d'électricité entre 400 kV et 63 kV, gère le réseau classique, les raccordements avec les industriels, ainsi que les raccordements avec les plus gros producteurs d'EnR.

Le réseau de transport d'électricité se comporte comme une infrastructure à trois piliers : la tension, le courant sur chacune des lignes et la fréquence globale du réseau.

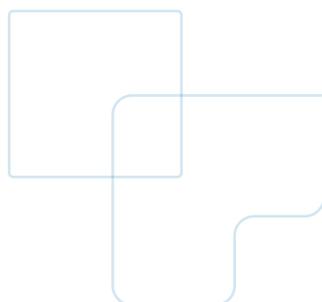
Les phénomènes que le système veut principalement éviter sont :

- l'écroulement de fréquence, quand on a un déséquilibre entre la production et la consommation ;
- l'écroulement de tension, principalement lié à l'éloignement entre la production et la consommation ;
- la surcharge en cascade de lignes électriques.

La répartition des flux de courant dépend de la répartition géographique des équipements de production et des consommateurs. Elle dépend aussi de l'impédance des éléments du réseau, liée aux caractéristiques des conducteurs. Et enfin, de la topologie du réseau, qui représente son intelligence même (réseau bouclé par exemple).

Le gestionnaire du réseau garantit que le flux reste stable en régime normal et dégradé par des pertes d'ouvrage. La simulation d'une perte d'un ouvrage (ligne, poste ou groupe de production) permet d'anticiper les réponses du système. Il est indéniable que les EnRs augmentent significativement la variabilité et l'incertitude des flux.

La transition énergétique donne un système électrique en pleine évolution, voire en pleine révolution. Les moyens de production EnRs croissent rapidement : or ils sont peu pilotables à cause des fortes incertitudes sur les prévisions d'ensoleillement et de vent. Dans ce contexte, RTE doit gérer et adapter le réseau en temps réel. Le numérique permet la mise en œuvre de diverses mesures. Ainsi RTE peut procéder à des écrêtements ponctuels d'EnRs pendant un temps de l'ordre de quelques heures dans l'année. Cela permet un meilleur dimensionnement du réseau : en écrétant 0.3 % du volume global d'EnR, on estime obtenir d'ici 2035 une économie de 7 milliards d'euros sur les investissements réseau.



Une autre solution consiste à changer l'architecture de pilotage du réseau électrique. Jusqu'à maintenant, le pilotage se faisait sur 2 couches : le poste électrique (protection locale, temps en secondes) puis le niveau central (sur une durée de 10 min). La masse d'informations devient difficile à gérer en temps réel ; on propose donc l'installation d'une couche intermédiaire, qui sera une couche de contrôle. Elle gèrera des sous-parties du réseau, entre 5 à 50 postes pour une gestion en boucle fermée. Cela résoudra des problèmes de type congestion, par exemple. L'opérateur n'est plus pilote mais navigateur : on passe au niveau du prévisionnel.

L'implémentation de cette couche de contrôle se concrétise par de nouveaux automates adaptatifs sur des zones. Ils permettent de gérer les écarts de temps, de l'ordre de la minute, on peut mettre en place des actions automatiques. Ces automates permettent aussi de créer une interface avec les producteurs. L'objectif est de près de 200 automates en 2035.

Les leviers à disposition des automates sont les suivants :

- ils peuvent changer la topologie du réseau, en ouvrant par exemple les disjoncteurs pour déboucler le réseau. Cette action est gratuite, mais augmente le risque de coupure générale ;
- les automates peuvent aussi utiliser l'effacement de production d'EnR et permettre une modulation plus fine de la production, notamment en pilotant la puissance maximale que chacun des postes peut intégrer. Cette action n'est pas gratuite, car les producteurs sont indemnisés à hauteur de la production qu'ils auraient eue, mais le risque de déséquilibre sur le réseau est faible ;
- le troisième levier des automates est la solution flexible, en s'aidant des batteries smartwire pour modifier légèrement l'impédance d'un ouvrage dans la distance électrique et l'orientation des flux. Cette option a un coût très élevé à l'investissement, mais elle est moins chère à l'usage.

Avec ces leviers, l'automate intègre ses actions à un modèle du réseau électrique qui lui permet de faire des simulations. On lui fournit le coût de ses leviers et les contraintes des zones (comme les limites de courant de transit dans chacun des

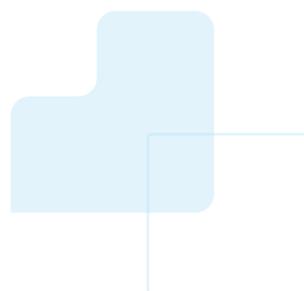
ouvrages, les capacités maximales des batteries, et les contraintes d'effacement). Tous ces éléments sont les paramètres de l'optimisation que fait le simulateur. On applique le modèle prédiction-contrôle, dont le principe est de contrôler en boucle fermée. On analyse l'état du réseau en fonction des différentes actions.

Pour installer ces automates, on utilise une architecture classique centralisée : avec les capteurs et les moyens d'actions dans les postes et chez les producteurs, les données sont envoyées sur un calculateur centralisé dans les data centers. On cherche actuellement à concevoir un mode hybride edge to cloud.

On fait face alors à des défis sur ces automates cyber-physiques :

1. on cherche des méthodes et des moyens de test adaptés à ces contrôles à boucle fermée en passant par la co-simulation, comme associer un simulateur de réseau électrique avec un simulateur de réseau télécom ;
2. on veut aussi déterminer le périmètre de ces différentes zones d'action d'automates pour avoir des critères robustes et évolutifs, et gérer les interactions entre automates (chevauchement ou influence) ;
3. on met en évidence le potentiel des systèmes auto-adaptatifs cloud to edge pour en tirer le maximum d'avantages : économie de bande passante, latence de fonctionnement, résilience, robustesse face aux pertes de postes.

On cherche aussi à élargir à d'autres problématiques : la gestion de la tension permet de dimensionner le réseau d'une manière différente par rapport à la gestion des transits, et donc d'intégrer cette problématique dans les modélisations.



3/ AUTOMOBILE

La numérisation a permis de réduire la consommation des véhicules. De 1990 à 2000, la consommation moyenne d'une voiture par km a diminué de 6,5 %. Pour une période plus longue (1990-2010), les chiffres correspondants sont - 17% sur la consommation moyenne. La diésélisation a apporté une amélioration supplémentaire de 15 %. En comparant deux voitures à moteur diesel avant et après l'introduction de l'injection directe pilotée (1991), on observe un gain de 26% dans sa consommation. Ce qui a permis cela est le pilotage très fin de l'injection avec des calculateurs de plus en plus complexes et une pression plus forte au niveau du moteur. Il faut contrôler et optimiser un grand nombre de variables. L'une est la consommation et les autres les émissions de polluants (oxyde d'azote et particules).

Sur l'essence, l'injection électronique a commencé bien avant. En 1984, on utilisait déjà 2Ko de mémoire avec un gain de 15%. Puis l'injection directe est arrivée, avec un gain supplémentaire de 10%.

D'autre part, les transmissions automatiques consomment 20 à 25% de moins que les boîtes mécaniques.

La dernière étape de numérisation concerne l'optimisation à partir de la quinzaine de combinaisons possibles entre le moteur thermique, la boîte à crabot, le moteur électrique, ce qu'on appelle la motorisation hybride

La numérisation a bénéficié à l'ensemble de la flotte automobile. Cela n'a pas empêché pour le transport routier une hausse en termes de proportion d'émission de CO₂ en Europe.

La sévérisation des normes d'émission des véhicules neufs imposera une part croissante de véhicules électriques ou hybrides rechargeables, au détriment des véhicules thermiques. Leur part de marché devrait atteindre 40 % en 2030. On n'envisage toutefois guère de changement au niveau de la consommation d'énergie.

Par contre, l'impact des véhicules auto-partagés peut être important : le nombre total de km parcourus pourrait baisser de 30 à 40%. Les scénarios extrêmes des simulations (ex Lisbonne), montrent qu'en gardant les trains et métros et en remplaçant les bus par des robots-taxis, partagés

ou non, on pourrait avoir les mêmes mobilités qu'aujourd'hui avec une flotte divisée par 10, un nombre de véhicules qui roulent divisé par 3 et un nombre de véhicules stationnés divisé par 15. Les temps de transports seraient aussi réduits.

Mais la perspective de véhicules autonomes partagés reste lointaine, probablement au-delà de 10 ans, comme l'expose le rapport *Pichereau*. Le problème est de pouvoir garantir l'absence d'accidents. Comme techniquement le pilotage a recours à l'intelligence artificielle, il faut avoir accumulé suffisamment d'expérience des situations que peut rencontrer un véhicule. Des centaines de milliers de voitures équipées d'enregistreurs roulent, avec des équipements de repérage et des niveaux de pilotage autonome variables. Pour construire un apprentissage permettant d'éviter les accidents, on fait des simulations massives à partir des données des scènes de conduites enregistrées, ou fabriquées à partir de "feuilles blanches". Ces simulations testent aussi les réactions du système à des événements rares.

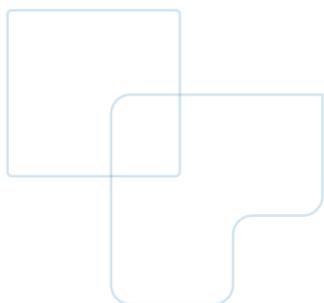
Les processeurs qui assurent le pilotage des voitures sont du type neuromorphique. Ils sont embarqués, seul moyen pour prendre assez rapidement les décisions de conduite. Mais les modèles statistiques dont ils disposent doivent être rafraîchis périodiquement, et seuls des ordinateurs très puissants peuvent les leur fournir.

Les plus avancés sont des constructeurs américains, Tesla et Waymo. Tesla a choisi de développer lui-même ses processeurs et ses ordinateurs. Waymo prévoit une voiture autonome équipée de 5 lidars (méthode de télédétection et de télémétrie semblable au radar), 27 caméras, 6 radars et d'une énorme puissance de calcul.

En conclusion, la numérisation, qui a révolutionné les moteurs, aura désormais un impact direct secondaire sur la consommation. La grande transformation, c'est l'assistance à la conduite. Au niveau extrême du véhicule autonome auto-partagé, on pourrait réduire considérablement le nombre de véhicules et de kilomètres parcourus. C'est une perspective encore lointaine, qui nécessitera des progrès techniques et des changements sociaux et politiques. Mais le bouleversement est amorcé, et cela oblige dès maintenant les constructeurs de voitures à établir des rapports beaucoup plus étroits avec la conception des composants électroniques et des ordinateurs.

En matière de R&D, on peut retenir les recommandations du rapport Pichereau⁵ de renforcer les efforts. Notamment sur :

- l'IA (données, connaissances, algorithmes et systèmes) des véhicules automatisés et connectés, ainsi que des véhicules et systèmes de transport autonomes ;
- les interactions entre le conducteur et le système automatisé (CNRS notamment) ;
- la mise au point d'une intelligence artificielle sûre, de confiance, vérifiable et explicable ;
- la cybersécurité et la sécurisation des échanges des données des véhicules automatisés ;
- l'architecture électrique et électronique du véhicule ;
- les logiciels de reconnaissance de forme.



4/ AÉRONAUTIQUE

Les principaux usages du numérique sont la simulation, les équipements de pilotage, les analyses de données et la maintenance prédictive.

La simulation va jusqu'à faire « voler au sol » les avions que l'on conçoit. Les gains en temps et coût de conception, et en sobriété des avions, sont considérables. Mais les besoins de calcul sont énormes, et Airbus utilise des ordinateurs parmi les plus puissants. Il est important de rappeler que cela a un impact environnemental positif, en permettant d'aller plus loin dans la frugalité des aéronefs.

Les équipements de pilotage comportent un nombre croissant de moteurs électriques, qui agissent sur des gouvernes de plus en plus fines.

L'analyse de données permet à Airbus d'offrir des services de maintenance prédictive qui comptent de plus en plus dans son activité. C'est la plate-forme Skyways, utilisée par beaucoup de compagnies aériennes. La puissance d'Airbus est suffisante pour qu'il ait pu garder le contrôle de cette plate-forme, alors qu'il l'a mise au point avec Amazon.

L'apport que peuvent avoir les technologies numériques à la décarbonation du trafic aérien ne sont pas négligeables notamment sur l'optimisation des trajectoires de vol : un des leviers de décarbonation du secteur aérien est d'utiliser finement la modélisation météorologique pour profiter au mieux des vents afin de réaliser des économies importantes en vol. Cette optimisation nécessite aussi de grandes puissances de calcul.

5- Le déploiement européen du véhicule autonome : Pour un renforcement des projets européens, par Damien Pichereau

5/ BÂTIMENT ET AMÉNAGEMENT



L'association du numérique, de la construction et de l'urbanisme comporte plusieurs aspects :

- le BIM est un système permettant de rassembler tous les intervenants d'une construction : conception, exécution, contrôle des travaux. Il peut être défini comme une maquette numérique permettant de partager les informations en temps réel afin d'assurer la conformité entre la conception et la réalisation. Cette maquette permet aussi de gérer de façon optimale le bâtiment dans la durée grâce à cette bibliothèque de données. De plus, on a un jumeau numérique pour modéliser le comportement du bâtiment et sa consommation, ou pour tester des configurations architecturales différentes (dans le but de mieux concevoir en fonction du climat et des apports solaires, mais aussi des contraintes de dimensionnement comme les contraintes sismiques). De plus, avec le BIM, des contraintes techniques, juridiques, de formation sont prises en compte ;
- une plateforme de service urbain pour la gestion des bâtiments donne une ouverture aux nouveaux opérateurs urbains. Avec les compteurs communicants, on peut gérer et faire de l'effacement à distance pour une production et distribution d'électricité à l'échelle du bâtiment (câblages etc.) ;
- le dernier aspect est d'intégrer le bâtiment à l'échelle du quartier.

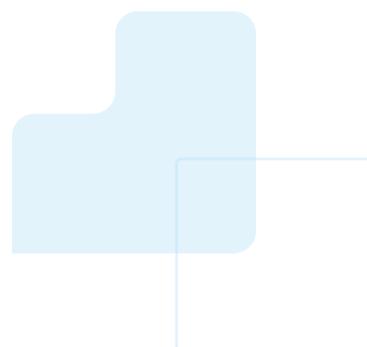
Le smart grid a changé d'échelle et tend vers le quartier. Les retours d'expérience montrent qu'un travail d'analyse et d'optimisation est nécessaire. En 2021, un bâtiment a pu être piloté avec un outil qui récupère en temps réel la consommation et la production énergétique.

Concrètement, on a trois niveaux à traiter : l'infrastructure réseau, la gestion de tous les systèmes d'information et la conception de l'aménagement urbain.

Avec les acteurs du numérique, de l'énergie et des services de l'État, trois objectifs sont poursuivis :

- un objectif technologique, touchant l'interconnexion du réseau, pour faire de l'effacement par exemple ;
- un objectif économique, pour obtenir un modèle économique viable des smart grids ;
- un objectif réglementaire pour, avec les partenaires, expliquer aux services de l'État comment faire évoluer les modèles économiques de l'énergie. Concernant ce point, les premiers travaux ont débuté en 2015, et finalement en 2021 un décret sur les réglementations a été adopté.

Dans le but de piloter en temps réel et de visualiser la consommation d'un quartier, la première étape s'est faite sur une parcelle privée : un îlot urbain comprenant un bâtiment de bureaux, un bâtiment de logement et un logement élevé. Cette étape du projet a été livrée en 2015 et a montré une baisse de 60% de la consommation globale. La deuxième étape était à l'échelle du quartier ; la géothermie a été utilisée, et la consommation globale énergétique a encore diminué. . Le troisième exemple, dans un quartier de Lyon, a permis de concrétiser une plateforme de service urbain au niveau d'un îlot. Cette plateforme est garantie dans le temps par un contrat urbain.



Les défis relevés avec l'exemple de Batignolles sont :

- on a rassemblé les différents acteurs numérique, énergétique et immobilier ;
- le terrain de jeu s'est élargi, avec une partie résidentielle et tertiaire sur une surface près de 300 000 m² ;
- le projet a été investi sur fonds propres, avec un budget très bas pour le prix d'un démonstrateur, montrant qu'on peut mener des opérations importantes avec un budget raisonnable.

Pour Issygrid, même si les plus vieux bâtiments datent de moins de 10 ans, la remontée des informations globales (du réseau électrique, logement, résidentiel) a rencontré certaines difficultés. Par exemple, pas moins de 14 systèmes étaient à interconnecter, alors que ces systèmes n'étaient pas tous standardisés par le même protocole.

Cette difficulté a été dépassée, malgré les problèmes techniques et la méfiance de la population pour utiliser le monitoring énergétique (environ 50% des habitants). Les premières logiques d'effacement ont pu être testées grâce au pilotage avec la standardisation du système numérique. Concernant le profil énergétique, la chaleur renouvelable a atteint 85% des besoins, à comparer à 50% à Paris

Dans le but d'augmenter la performance, des capteurs ont été ajoutés avec l'aide des opérateurs de réseau : CPCU (réseau urbain), eau de Paris (géothermie profonde), Enedis (réseau électrique) et les bailleurs sociaux. Avec ce travail, la remontée de l'information est devenue possible, même à l'échelle des particuliers (avec des contrats de consentement, on atteint 90% d'adhésion). Puis une maquette numérique du quartier a été construite pour comprendre les consommations annuelles et mensuelles, et on a pu montrer que la géothermie fonctionne bien ; même constat pour le réseau de chaleur, par contre la globalisation ne fonctionne pas aussi parfaitement.

On a aussi constaté que l'autoconsommation collective est plus efficace que l'individuelle. Ceci requiert toutefois la discipline d'utiliser ces modalités à des heures précises et demande une traçabilité des habitants (grâce ici aux blockchains déjà mis en place).

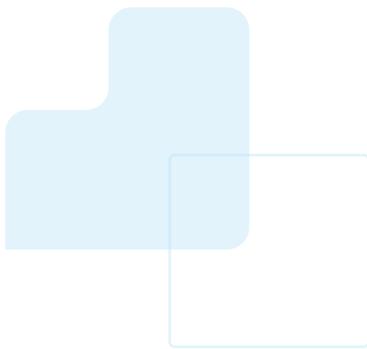
En conclusion, trois sujets importants résument les défis et les difficultés sur le nouvel écosystème du quartier: *la standardisation des protocoles informatiques, le jumeau numérique et la flexibilité.*

Hors électronique-informatique, ces secteurs applicatifs ne voient que les bénéfices du numérique :

- aide à la conception, permettant de gagner du temps et de l'argent (ex : en aéronautique, le remplacement des souffleries par des simulateurs) ;
- jumeaux numériques et maintenance ;
- organisation de l'entreprise (EDF).

Du point de vue énergétique, les principales conséquences positives sont le fait des usagers :

- réduction de la consommation (voiture, avion, bâtiment) ;
- efficacité des usages : si le concept de voiture autonome se généralisait on économiserait 30 % à 40 % de l'énergie.



05

CONCLUSION ET RECOMMANDATIONS

La place du numérique dans la transition énergétique vers une économie décarbonée est considérable. Même s'il ne doit pas se soustraire à l'obligation d'efficacité énergétique, sa consommation d'une électricité qui est de plus en plus décarbonée ne doit pas être un obstacle. Cependant, l'accroissement considérable des besoins de calcul et du nombre d'équipements - auquel rien ne semble devoir mettre fin - tend à l'augmenter, pour la fabrication et le fonctionnement.

L'impact de la fabrication ne se limite pas à l'électricité consommée. Toute la chaîne est en cause, depuis l'extraction des matériaux jusqu'au recyclage. Malheureusement, la plus grande partie de cette fabrication se fait hors d'Europe, surtout en Asie, mais on peut au moins prescrire le recyclage et interdire l'obsolescence programmée.

L'introduction du numérique est associée à la fois à la conception de nouveaux produits, à de nouveaux services, à une organisation plus efficace des entreprises et à une modification des usages. Tout cela aide à réduire les impacts énergétiques, même si ce n'est qu'une partie des objectifs des entreprises. Celles-ci respectent les prescriptions environnementales, et cherchent parfois à les devancer, mais elles sont d'abord préoccupées par les prix de revient (dont le coût de l'énergie) et la clientèle.

ÉLECTRONIQUE, INFORMATIQUE, TÉLÉCOMMUNICATIONS

Ici, la consommation d'énergie devient un obstacle à la réalisation et au fonctionnement des systèmes, alors que les besoins de calcul et de transmission ne cesseront de croître, et qu'on approche des limites de miniaturisation des circuits intégrés (loi de Moore).

La loi de Moore, en proposant des dispositifs universels de calcul de plus en plus performants à des coûts de plus en plus réduits, et la concentration industrielle qu'elle a provoquée en matière de microprocesseurs, ont privilégié le développement séparé du logiciel et du matériel. Cette séparation a provoqué la formidable dissémination du

numérique dans la plupart des activités, et la protection du patrimoine applicatif des utilisateurs. Il apparaît cependant que le recours à des dispositifs de calcul spécialisés mis en œuvre par de l'électronique classique ou du quantique, forme la seule voie envisageable actuellement pour continuer à gagner en performance, donc en quantité d'applications possibles avec une dépense énergétique maîtrisée (voir rapport de l'AdT : Les technologies matérielles supports du numérique futur, Février 2022). La contrepartie est alors que le logiciel devient dépendant du matériel et doit prendre en compte explicitement ses contraintes physiques.

Cela impose une véritable révolution aux fabricants de circuits intégrés, aux architectes de systèmes et aux programmeurs. L'introduction du parallélisme oblige à revoir 90 % des logiciels. En même temps, le développement de l'intelligence artificielle, ainsi que le recours à des accélérateurs de traitement spécialisés à la place des processeurs d'usage général, obligent à des interactions beaucoup plus étroites entre les constructeurs et leurs clients.

C'est une transformation radicale, mais c'est probablement à ce prix que numérique et énergie pourront se compléter pour offrir à la société dans son ensemble un progrès durable. Un des aspects les plus importants est le changement demandé aux programmeurs, en compétences, localisation, méthodes de travail. Cet ensemble de défis va mettre à l'épreuve les ressources et les compétences européennes dans tous les compartiments du jeu, avec de forts appels à la R&D. Pour l'Europe, qui avait pris du retard, c'est une chance nouvelle, à condition de savoir l'exploiter.

Pour les circuits intégrés, STMicroelectronics, premier constructeur européen, qui s'appuie beaucoup en recherche sur le CEA/LETI, est en bonne position sur l'Internet des objets. Son objectif est 80 % d'edge en 2030, en intégrant les logiciels dans les microprocesseurs, en leur associant des réseaux de neurones, et en ajoutant de la mémoire. Cependant, à moins d'aides massives, ST ne peut suivre la course à la miniaturisation ultime,

menée par deux champions asiatiques : Samsung et TSMC. Intel, qui avait semblé renoncer, a décidé de se remettre à niveau, et de mener une politique plus ouverte : davantage de dialogue avec ses clients, et sans doute la construction d'une usine en Europe. Pour le reste de la chaîne du numérique (processeurs parallèles, serveurs, supercalculateurs, opérateurs quantiques), l'Europe doit maintenir et amplifier un fort niveau de soutien.

L'European Processor Initiative se donne pour 2030 l'objectif de 65% à 70 % d'approvisionnement européen sur l'ensemble des matériels ; cet objectif prend en compte la consommation d'énergie. Le programme EuroHPC vise à maintenir l'Europe et la France dans le groupe des 4 leaders mondiaux de la puissance de calcul. Par ailleurs en Europe la France est le seul pays européen actif sur la totalité de la chaîne, de ST à Atos en passant par Kalray et OVH. Cela lui donne la responsabilité d'une prise de leadership dans la construction d'une position commune.

En télécommunications, l'événement actuel est le passage à la 5G, qui va ouvrir l'Internet des objets. Même si elle est énergétiquement plus efficace, il va falloir assumer un accroissement considérable du trafic sans que cela se répercute sur les émissions de CO₂. Le recours aux achats d'électricité bas carbone est un des moyens, mais il faut aussi agir, comme pour l'ensemble de l'électronique, sur la consommation électrique des systèmes.

ÉLECTRICITÉ

C'est en grande partie grâce à l'introduction du numérique que la production et la distribution d'électricité font face à l'introduction des énergies renouvelables intermittentes et à la décentralisation qui va avec. Ces modifications profondes, qui touchent à la fois la technologie, l'organisation des grands opérateurs et leurs rapports avec le gouvernement, les règles du marché, demandent beaucoup de recherche, chez les opérateurs et dans la recherche publique.

EDF, avec ses filiales, veut être l'opérateur des données, au service de la décarbonation de l'ensemble du secteur.

AUTOMOBILE

Il y a une trentaine d'années que l'électronique s'est introduite dans les voitures, pour améliorer

le rendement des moteurs et accroître le confort. Cela n'a cessé de se développer. Mais on arrive à une rupture, qui va demander une électronique nouvelle, fondée sur l'intelligence artificielle et des infrastructures de clouds connectés, du edge local dans la voiture au cloud central : la voiture autonome. Celle-ci devrait changer profondément les usages et permettre enfin une utilisation rationnelle d'un matériel qui passe aujourd'hui trop de temps à ne pas servir. Avec les voitures autonomes partagées, on pourrait gagner 30 % à 40 % d'énergie de fabrication. Les changements dans la fabrication et l'exploitation seront aussi considérables ; c'est l'entrée dans l'Internet des objets : davantage d'interactions avec les fabricants de circuits intégrés et de capteurs, orientation vers le traitement local (edge computing). Cela va demander des compétences nouvelles.

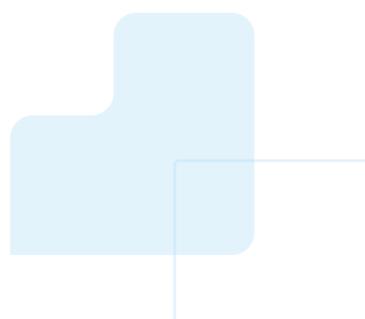
AÉRONAUTIQUE

La simulation va jusqu'à faire « voler au sol » les avions que l'on conçoit. Les gains en temps et coût de conception, et en sobriété des avions, sont considérables. Mais les besoins de calcul sont énormes, et Airbus utilise des ordinateurs parmi les plus puissants.

L'autre grand usage de l'électronique, c'est l'analyse de données pour la maintenance prédictive. La puissance d'Airbus est suffisante pour qu'il ait pu garder le contrôle de la plate-forme de service « Skyways » utilisée par beaucoup de compagnies aériennes, alors qu'il l'a mise au point avec Amazon.

BÂTIMENT

Le passage au jumeau numérique (BIM) est un facteur d'organisation d'une profession dispersée. En même temps, cela pose des problèmes de standardisation et de propriété des données. Un enjeu est de concevoir un ensemble énergétique à l'échelle du quartier : on peut réduire la consommation de 60 %. Et l'autoconsommation est beaucoup plus efficace qu'au niveau individuel. Cependant, il faut prendre garde au risque que les majors de la donnée monopolisent la relation avec les collectivités locales et imposent une standardisation fondée sur le profit plutôt que sur l'urbanisme.



RECOMMANDATIONS

1/ Prendre en compte l'ensemble des causes de consommation des systèmes numériques, de la fabrication au fonctionnement :

- les fortes divergences entre les études viennent des lacunes de l'analyse. Pour que ces études soient de vrais outils d'aide à la décision, il faut y remédier. Cela implique de standardiser ce que sont les points de mesure d'un système numérique et comment il en résulte une mesure globale. De la recherche est nécessaire par exemple dans le domaine peu exploré des relations entre algorithmes, notamment algorithmes distribués, et consommations (où il est déjà évident que la quantité d'échanges peer to peer comme ceux intervenant dans des algorithmes de bitcoin ou de blockchain n'est pas généralisable, cette quantité croissant comme le carré du nombre de participants) ;
- cette connaissance globale permettra d'optimiser l'ensemble coût-performance-qualité environnementale des systèmes, y compris les règles et protocoles d'échange de données, et en y intégrant notamment les data centers ;
- et de concevoir des systèmes « juste assez », tenant compte au plus juste de l'ensemble des contraintes.

2/ Anticiper l'interaction nouvelle entre matériels et logiciels

Les besoins en performance et en volume du traitement de l'information ne vont cesser de croître, alors que la technologie des circuits intégrés arrive aux limites de la miniaturisation. La voie d'avenir est celle des opérateurs spécialisés, donc une interaction étroite entre fabricants d'appareils et de composants. La France doit savoir tirer parti de cette nouvelle donne.

- Investir en R&D sur les technologies informatiques de nouvelle génération ;
- former des programmeurs adaptés à une interaction étroite avec les concepteurs de matériel et capables de rénover le patrimoine logiciel, en assurant aussi la cybersécurité de l'ensemble.
- Redévelopper le patrimoine applicatif des logiciels pour bénéficier de l'apport des composants spécialisés et du parallélisme massif ;
- pousser les réseaux de neurones, le calcul neuromorphique, et l'interprétation de l'intelligence artificielle, en visant le véhicule à conduite assistée ou partagée (rapport Pichereau), sans négliger l'accompagnement socio-économique.

3/ Vis-à-vis de l'Europe, construire la solidarité et faire valoir nos atouts

Asiatiques et Américains engagent des forces dont l'objectif est la domination, ou au moins l'autonomie. En énergie, la question est aujourd'hui clairement posée, mais il faudra des années pour faire valoir nos atouts. En numérique, avec la nouvelle interaction matériel-logiciel, c'est désormais l'avenir d'une grande partie de nos industries et services qui est en jeu. La France a des atouts, notamment sa présence sur l'ensemble de la chaîne du numérique.

- Pousser l'European processor initiative, en se donnant pour 2030 l'objectif de 65–70 % d'approvisionnement européen ;
- ainsi que l'initiative Euro HPC ;
- partage européen des pratiques et standards en cyber-sécurité. Encadrer la diversité des transcriptions nationales des règles européennes ;
- pousser le Climate Neutral Data Center Pact ;
- Imposer la recyclabilité et interdire l'obsolescence programmée.





anRT

ASSOCIATION NATIONALE
RECHERCHE TECHNOLOGIE

33 rue Rennequin
75017 - PARIS
Tél : +33(0)1 55 35 25 50
com@anrt.asso.fr
www.anrt.asso.fr