

4. Gestion des données urbaines dans les nuages informatiques

- Brève histoire des nuages informatiques
- Modèles de service et de déploiement
- Technologie clé : la virtualisation
- **IaaS : les points de vue utilisateur et fournisseur**
- PaaS : programmation et déploiement des applications
- Stockage de données
- Traitement de données
- Traitement de flux de données

Christine Morin

VILLES INTELLIGENTES : DÉFIS TECHNOLOGIQUES ET SOCIÉTAUX

Dans cette séquence, nous nous intéressons au niveau le plus bas dans la pyramide des services de cloud, le niveau infrastructures sur lequel s'appuient les autres couches offrant des services de plus haut niveau.

Le point du vue du fournisseur d'infrastructure

- Choix de l'emplacement des centres de données pour l'hébergement des serveurs



2

Plaçons-nous tout d'abord du point de vue du fournisseur d'infrastructures. C'est lui qui met en place et qui gère les centres de données qui hébergent les serveurs sur lesquels s'exécutent les services et applications du cloud.

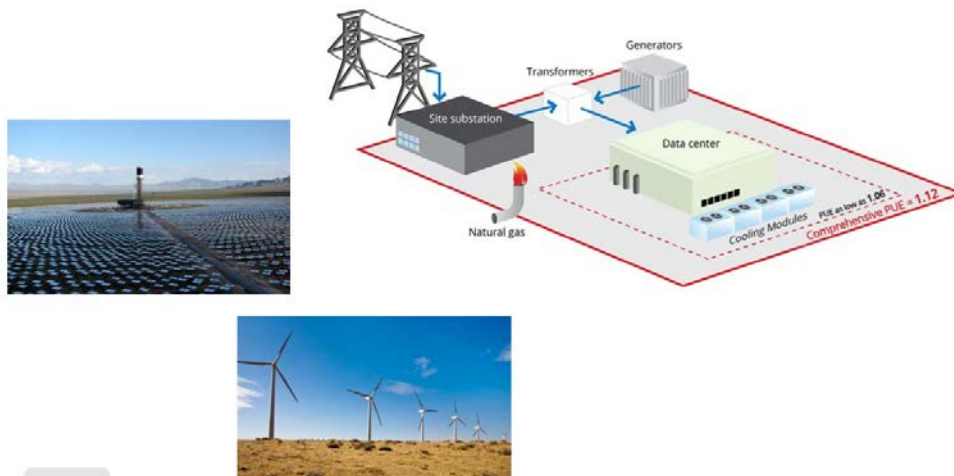
Une décision importante pour le fournisseur de cloud est tout d'abord l'emplacement des centres de données. Un centre de données sera placé dans un lieu où :

- il pourra être approvisionné en électricité, les centres de données étant très gourmands en énergie,
- il existe des solutions pour le refroidissement à moindre coût. De ce point de vue en Europe, les pays scandinaves sont très attractifs, comme le montre cette carte qui représente la localisation des centres de données de Google.

Les fournisseurs de cloud répartissent également leurs centres de données sur la planète pour les situer à proximité des bassins de population, de sorte à être proches de leurs clients et pouvoir leur offrir une bonne qualité de service lorsqu'ils servent leurs requêtes.

Le point du vue du fournisseur d'infrastructure

- Sources d'énergie et solution de refroidissement



3

Pour ce qui concerne le choix de la source d'énergie, les opérateurs de cloud utilisent non seulement l'électricité des centrales électriques traditionnelles mais ils se tournent également de plus en plus vers les énergies renouvelables produites par exemple par des éoliennes ou par des panneaux solaires.

Il faut noter que le refroidissement des machines compte pour une très large part dans la facture d'électricité des opérateurs et donc un soin tout particulier est apporté au système de refroidissement.

Le point du vue du fournisseur d'infrastructure

- Choix, installation, maintenance, évolution du parc de serveurs (calcul, stockage, réseau)



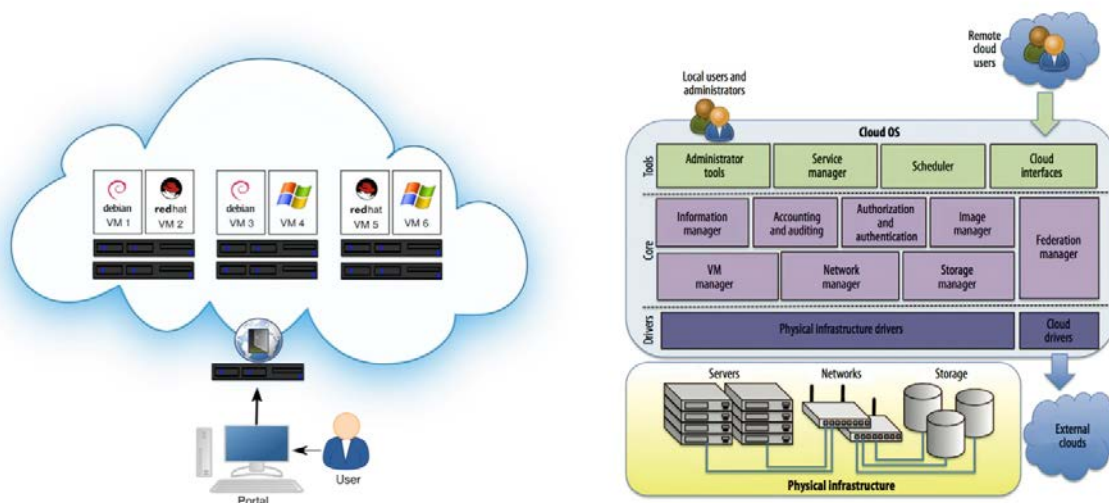
4

Les clouds d'infrastructure fournissent des ressources de calcul de stockage, de réseaux à leurs utilisateurs. Donc le fournisseur de cloud d'infrastructures est en charge du choix, de l'installation et de la maintenance du parc informatique.

Les centres de données sont constitués de grappes de calculateurs qui comportent des dizaines, voire des centaines de milliers de serveurs interconnectés.

Donc au fil du temps, les constructeurs ont optimisé le packaging des grappes pour assurer une bonne dissipation de la chaleur. Les aspects thermiques doivent en effet être pris en compte car ils ont un impact important sur les performances du data center, sur l'efficacité énergétique et aussi sur le vieillissement des composants.

Système de gestion des ressources, VM et utilisateurs



5

Tournons-nous vers les aspects logiciels dans les clouds IaaS.

Dans un cloud IaaS, les ressources sont offertes aux utilisateurs sous la forme de machines virtuelles auxquelles ils accèdent depuis leurs ordinateurs personnels à travers le réseau Internet.

La gestion des serveurs, la gestion des machines virtuelles et des utilisateurs est gérée par un système qui traite un certain nombre d'aspects :

- la sécurité avec les techniques d'authentification des utilisateurs, d'autorisation d'accès aux ressources, de détection d'intrusion.
- l'allocation des ressources aux utilisateurs à travers des politiques de placement de machines virtuelles et d'ordonnancement qui vont permettre d'optimiser la consommation d'énergie du data center comme on l'a déjà évoqué précédemment,
- le traitement des défaillances qui peuvent affecter les réseaux comme les serveurs,
- la gestion des réseaux virtuels donc le système de gestion d'un cloud IaaS est en charge de gérer les adresses Mac des machines virtuelles et également l'attribution d'adresses IP, notamment d'adresses IP publiques.
- La gestion de la facturation est aussi un autre service important dans un cloud public.
- gérer les images de machines virtuelles, donc les images qui figurent au catalogue du fournisseur de cloud IaaS et qui sont proposées aux utilisateurs,
- services de fédération qui permettent d'interfacer le cloud avec d'autres clouds, qu'ils soient privés comme dans un cloud communautaire ou qu'ils soient publics par exemple pour faire du cloud bursting lorsqu'on a à faire face à une surcharge ponctuelle.

Systèmes de gestion de clouds IaaS en open source



6

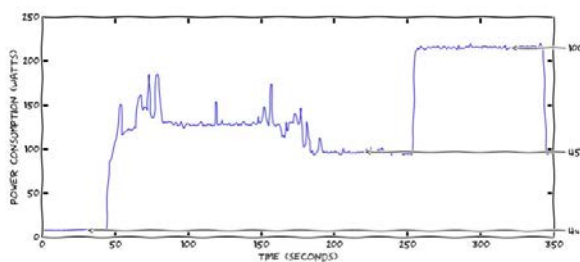
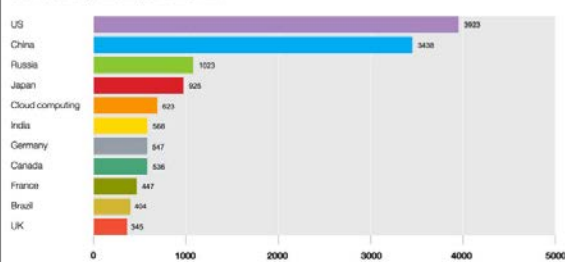
Mais il existe de nombreuses solutions qui sont en open source et qui permettent de gérer les clouds d'infrastructures. On peut citer plusieurs exemples :

- Nimbus qui est un pionnier et qui a été développé à Argonne National Lab aux États-Unis et qui vise surtout les clouds communautaires dans le monde scientifique,
- OpenNebula qui a été développé dans le contexte de l'espace européen de la recherche à l'initiative d'une université madrilène,
- Eucalyptus, un autre pionnier aux États-Unis,
- OpenStack qui est aujourd'hui le logiciel open source de référence pour la gestion des clouds IaaS et qui est très utilisé dans les clouds commerciaux mais également pour gérer des clouds privés et ou communautaires. Il est également utilisé comme support de recherche dans le monde académique. Ce logiciel a été développé à l'initiative d'un consortium constitué des grandes entreprises du domaine.

Maîtrise de la consommation énergétique

- Choix de serveurs énergétiquement efficace
- Gestion de la température du centre de données
- Extinction/allumage automatique des serveurs inoccupés
- Stratégies de placement et de consolidation

2007 electricity consumption. Billion kWh



7

Comme je l'ai déjà souligné, la maîtrise de la consommation d'énergie est un enjeu tout à fait crucial pour les clouds d'infrastructures.

Si les clouds étaient un pays, ils figureraient à la **cinquième position au plan mondial** pour ce qui est de la **consommation d'énergie** selon une étude menée par Greenpeace en 2007.

Plusieurs leviers vont permettre de contribuer à réduire la consommation d'énergie dans les clouds d'infrastructures et donc réduire aussi l'impact écologique des centres de données.

Plusieurs facteurs peuvent impacter, notamment le

- Choix de serveurs énergétiquement efficaces, les constructeurs ne cessent d'améliorer leurs serveurs,
- Gestion de la température dans le centre de données est tout à fait essentielle pour réduire la consommation d'énergie,
- Mise en place de stratégies de gestion de ressources qui vont viser à éteindre des serveurs quand ils sont inutilisés dans des périodes de charges faibles ou modérées,
- Et puis bien sûr, les stratégies de placement et de consolidation de machines virtuelles qui vont favoriser le nombre de serveurs inactifs de sorte à les éteindre.

Il est quand même important d'avoir à l'esprit cette courbe représentant la puissance consommée par un serveur.

À l'instant 50, ce serveur est allumé donc il boote son système d'exploitation.

On voit que pendant toute une période, il consomme beaucoup d'électricité.

Ensuite, on charge ce serveur à pleine puissance et on voit que la consommation électrique monte à environ 200 Watts.

On voit ainsi que la différence entre un serveur inutilisé et un serveur à pleine charge n'est pas si grande que ça, d'où l'intérêt d'éteindre les serveurs inutilisés.

Cependant, il faut faire très attention à la phase extrêmement consommatrice d'énergie, phase pendant laquelle on ne peut pas utiliser le serveur, qui correspond au redémarrage dans toutes les stratégies qui justement visent à éteindre des serveurs inutilisés. Sans quoi on risque de consommer avec ces stratégies plus d'énergie que ce que l'on va gagner.

Offre de services – Instances

Amazon Web Services
PRODUITS ET SERVICES

Instances dédiées à la demande
Les instances à la demande vous permettent de payer les capacités de calcul utilisées à l'heure, sans engagement à long terme.

Linux RHEL SLES Windows Windows avec SQL Standard Windows avec SQL Web
Windows avec SQL Enterprise

Région: **US (Frankfurt)**

| | vCPU | ECU | Mémoire (GiB) | Stockage des instances (GiB) | Utilisation de Linux/UNIX |
|--|------|-------|---------------|------------------------------|---------------------------|
| Usage général – Génération actuelle | | | | | |
| m4.large | 2 | 6.5 | 8 | EBS uniquement | \$0.105 par heure |
| m4.xlarge | 4 | 13 | 16 | EBS uniquement | \$0.33 par heure |
| m4.2xlarge | 8 | 26 | 32 | EBS uniquement | \$0.66 par heure |
| m4.4xlarge | 16 | 53.5 | 64 | EBS uniquement | \$1.32 par heure |
| m4.10xlarge | 40 | 124.5 | 160 | EBS uniquement | \$3 par heure |
| m3.medium | 1 | 3 | 3.75 | 1 x 4 SSD | \$0.087 par heure |
| m3.large | 2 | 6.5 | 7.5 | 1 x 32 SSD | \$0.174 par heure |
| m3.xlarge | 4 | 13 | 15 | 2 x 40 SSD | \$0.347 par heure |
| m3.2xlarge | 8 | 26 | 30 | 2 x 80 SSD | \$0.695 par heure |
| Calcul optimisé – Génération actuelle | | | | | |
| c4.large | 2 | 8 | 3.75 | EBS uniquement | \$0.165 par heure |
| c4.xlarge | 4 | 16 | 7.5 | EBS uniquement | \$0.329 par heure |
| c4.2xlarge | 8 | 31 | 15 | EBS uniquement | \$0.616 par heure |
| c4.4xlarge | 16 | 62 | 30 | EBS uniquement | \$1.238 par heure |
| c4.8xlarge | 36 | 132 | 60 | EBS uniquement | \$2.25 par heure |
| c3.large | 2 | 7 | 3.75 | 2 x 16 SSD | \$0.141 par heure |
| c3.xlarge | 4 | 14 | 7.5 | 2 x 40 SSD | \$0.283 par heure |
| c3.2xlarge | 8 | 15 | 15 | 2 x 80 SSD | \$0.567 par heure |

Mise en route gratuite
[Créer un compte gratuit](#)

À titre d'illustration, je vous montre des copies d'écran de l'offre IaaS du fournisseur de cloud Amazon, c'est le fournisseur le plus populaire de services d'infrastructures. On y voit un large choix d'instances, les utilisateurs pouvant choisir la capacité de leur machine virtuelle : small, médium, large et le type de stockage. Le coût à l'heure est indiqué pour chaque type d'instance, il s'agit ici d'instances dédiées, c'est-à-dire obtenues en libre-service et à la demande.

Offre de services – Instances

Amazon Web Services
PRODUITS ET SERVICES

- Amazon EC2
- Description détaillée
- Instances
- Tarification
- Instances de la génération précédente
- Options d'achat
 - Instances ponctuelles Amazon EC2
 - Instances réservées Amazon EC2
 - Instances dédiées Amazon EC2**
 - Ressources pour développeurs
- Amazon EC2 S3A
- AWS Management Portal for vCenter
- Liens connexes
 - Instances Windows
 - VM Import/Export
 - Console de gestion
 - Documentation
 - Notes de mise à jour
 - Forum de discussion

Mise en route gratuite
[Créer un compte gratuit](#)

Instances dédiées à la demande
Les instances à la demande vous permettent de payer les capacités de calcul utilisées à l'heure, sans engagement à long terme.

Linux RHEL SLES Windows Windows avec SQL Standard Windows avec SQL Web
Windows avec SQL Enterprise

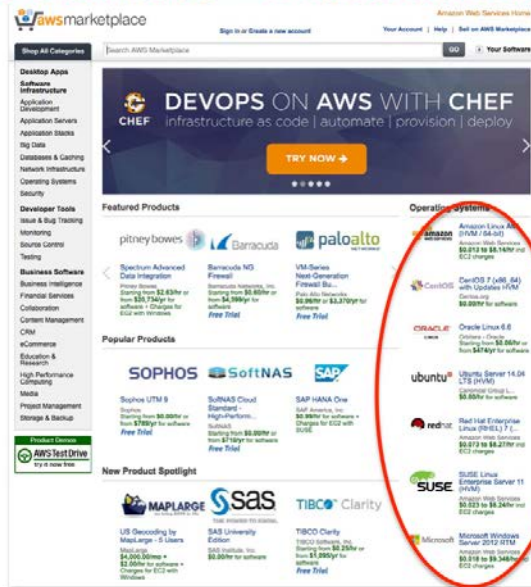
Région: **US (Frankfurt)**

| | vCPU | ECU | Mémoire (GiB) | Stockage des instances (GiB) | Utilisation de Linux/UNIX |
|--|------|-------|---------------|------------------------------|---------------------------|
| Usage général – Génération actuelle | | | | | |
| m4.large | 2 | 6.5 | 8 | EBS uniquement | \$0.185 par heure |
| m4.xlarge | 4 | 13 | 16 | EBS uniquement | \$0.33 par heure |
| m4.2xlarge | 8 | 26 | 32 | EBS uniquement | \$0.66 par heure |
| m4.4xlarge | 16 | 53.5 | 64 | EBS uniquement | \$1.32 par heure |
| m4.10xlarge | 40 | 124.5 | 160 | EBS uniquement | \$3 par heure |
| m3.medium | 1 | 3 | 3.75 | 1 x 4 SSD | \$0.087 par heure |
| m3.large | 2 | 6.5 | 7.5 | 1 x 32 SSD | \$0.174 par heure |
| m3.xlarge | 4 | 13 | 15 | 2 x 40 SSD | \$0.347 par heure |
| m3.2xlarge | 8 | 26 | 30 | 2 x 80 SSD | \$0.695 par heure |
| Calcul optimisé – Génération actuelle | | | | | |
| c4.large | 2 | 8 | 3.75 | EBS uniquement | \$0.165 par heure |
| c4.xlarge | 4 | 16 | 7.5 | EBS uniquement | \$0.309 par heure |
| c4.2xlarge | 8 | 31 | 15 | EBS uniquement | \$0.618 par heure |
| c4.4xlarge | 16 | 62 | 30 | EBS uniquement | \$1.238 par heure |
| c4.8xlarge | 36 | 132 | 60 | EBS uniquement | \$2.25 par heure |
| c3.large | 2 | 7 | 3.75 | 2 x 16 SSD | \$0.141 par heure |
| c3.xlarge | 4 | 14 | 7.5 | 2 x 40 SSD | \$0.283 par heure |
| c3.2xlarge | 8 | 28 | 15 | 2 x 80 SSD | \$0.567 par heure |

D'autres politiques tarifaires figurent au catalogue comme indiqué à gauche de la capture d'écran. Par exemple, on a les instances ponctuelles qui sont moins onéreuses mais qui peuvent être réquisitionnées à tout moment sans prévenir. Et on a également les instances réservées qui peuvent être demandées à l'avance par les utilisateurs pour une période donnée.

Amazon propose également une offre en termes de système d'exploitation que l'on va pouvoir utiliser au sein des machines virtuelles, donc différentes distributions Linux et systèmes Windows.

Offre de services – Place de marché



10

Enfin, les fournisseurs de clouds IaaS offrent de plus en plus ce qu'on appelle des places de marché où on a des machines virtuelles dans lesquelles sont pré installés des logiciels divers et variés et qui sont des machines virtuelles prêtes à être déployées pour des usages bien précis.

Offre de services – Place de marché

The screenshot displays the AWS Marketplace homepage. At the top, there's a navigation bar with 'AWS Marketplace', 'Sign in or Create a new account', 'Your Account', 'Help', and 'Get on AWS Marketplace'. Below the navigation bar is a large banner for 'DEVOPS ON AWS WITH CHEF' with a 'TRY NOW' button. The main content area is divided into several sections: 'Featured Products', 'Operating Systems', 'Popular Products', and 'New Product Spotlight'. A red circle highlights the 'Featured Products' section, which includes products like 'pitneybowes', 'Barracuda', 'VM-Series Next-Generation Firewall', and 'Amazon Advanced Data Integration'. Other products listed include 'SOPHOS', 'SoftNAS', 'SAP HANA One', 'SUSE Linux Enterprise Server 11', 'Ubuntu Server 14.04', 'Red Hat Enterprise Linux (RHEL) 7.1', 'Microsoft Windows Server 2012 R2', and 'TIBCO Clarity'.

Politique tarifaire

- Facturation en fonction de la consommation
 - Puissance de calcul
 - Tarif horaire selon le type de VM
 - Stockage
 - Tarif mensuel en fonction du type de stockage et du volume de données
 - Transfert de données entre le client et le cloud
 - Prix en fonction volume de données transférées (trafic entrant / trafic sortant)
 - Services additionnels : gestion automatique de l'élasticité, équilibrage de charge ...



12

Du point de vue de la politique tarifaire, comme j'ai déjà eu l'occasion de le dire, la facturation se fait en fonction de la consommation.

Pour ce qui concerne la puissance de calcul, le tarif est en général horaire et dépend du type d'EVM choisi par l'utilisateur.

Pour ce qui est du stockage, il s'agit bien souvent d'un tarif mensuel qui varie en fonction du type de stockage choisi et du volume de données.

Les transferts de données entre le client et le cloud sont également facturés, ceci en fonction du volume et potentiellement dans les 2 sens pour le trafic entrant et sortant du data center.

Et puis, les fournisseurs de clouds IaaS peuvent facturer des services additionnels comme la gestion automatique de l'élasticité ou de l'équilibrage de charges par exemple.

Contrat de service (SLA)



13

Entre un client et un fournisseur de services, un contrat de service ou Service Level Agreement, SLA, selon le terme consacré, est signé.

Ce contrat stipule les engagements du fournisseur vis-à-vis de son client.

Les termes du service doivent être mesurables et vérifiables puisqu'ils donnent lieu à facturation.

Dans les clouds IaaS, le contrat de service comporte des termes sur la capacité des machines virtuelles, sur le prix des machines virtuelles, également des termes sur la qualité de service comme par exemple la disponibilité des ressources.

Le système de gestion de cloud comporte :

- un service d'audit qui va permettre la traçabilité dans le cloud,
- un service de monitoring qui permet de détecter d'éventuelles violations des termes du SLA et donc de réagir en conséquence. Ce qu'il faut savoir, c'est qu'en général, il y a des pénalités qui sont prévues dans le SLA pour dédommager les clients dans le cas où le contrat de service n'est pas respecté.

En général, il y a aussi beaucoup de clauses restrictives. La définition des termes de SLA, la gestion automatisée des SLA est à l'heure actuelle un sujet de recherche extrêmement actif dans le domaine du cloud.

Le point de vue de l'utilisateur



Interface EC2 : standard de facto

- Choisir une image de machine virtuelle et la personnaliser
- Créer, démarrer, suspendre, redémarrer, sauvegarder, arrêter des machines virtuelles
- Créer un réseau privé entre des machines virtuelles
- Créer des adresses IP publiques
- Stocker des données permanentes (EBS, S3 ...)
- Utiliser des services de redimensionnement automatique, d'observation de métriques, d'équilibrage de charge

14

Enfin pour terminer, abordons le point de vue de l'utilisateur d'un cloud IaaS.

Il y a ici 2 messages importants.

Le premier message, c'est que l'interface du service Amazon EC2 est le standard de facto des clouds IaaS et en fait a été implémenté par toutes les solutions Open Source.

L'interface EC2 permet à son utilisateur de :

- choisir une image de machine virtuelle et de la personnaliser,
- créer, démarrer, suspendre, redémarrer, sauvegarder l'état d'une machine virtuelle,
- créer un réseau privé entre des machines virtuelles pour former un réseau virtuel ou un cluster virtuel,
- créer des adresses IP publiques,
- stocker des données persistantes,
- utiliser des services de redimensionnement automatique d'un cluster virtuel, d'observation de métriques, d'équilibrage de charge Et il y a également quelques services avancés comme le redimensionnement automatique,

Deuxième message, l'IaaS est vraiment l'assembleur du cloud.

L'utilisateur d'un cloud IaaS doit disposer de compétences administrateur réseau et système.

Dans le contexte de la ville intelligente, le cloud IaaS est presque invisible car les data centers ne sont en général pas déployés dans les villes et les clouds IaaS ne sont pas directement utilisés par les utilisateurs finaux. Néanmoins, ils constituent les fondations sur lesquelles s'appuient les services de plus haut niveau utilisés par les applications au service des différents acteurs de la vie intelligente.

Pour une ville durable, les clouds IaaS doivent se montrer économes en énergie et les choix de conception effectués par les opérateurs de clouds IaaS sont donc extrêmement importants dans ce contexte.

Illustrations & photos : crédits

p. 2 : droits réservés Google, localisation des centres de données Google :

<http://www.google.com/about/datacenters/inside/locations/index.html> et <http://www.google.com/about/datacenters/gallery/#/places/11>

p. 3 : droits réservés Google, <http://www.google.com/about/datacenters/efficiency/internal/> et <http://www.google.com/green/energy/>

p. 4 : droits réservés Google, <http://www.google.com/about/datacenters/gallery/#/tech/5> et <http://www.google.com/about/datacenters/gallery/#/tech/2>

p. 5 : illustration droite - CC-BY-NC-SA,

<http://opennebula.org/publication-of-the-opennebula-cloud-os-architecture-in-ieee-computer/>

p. 7 : illustration gauche :

<http://www.greenpeace.org/international/Global/international/publications/climate/2012/iCoal/HowCleanisYourCloud.pdf> ,

Graphique droite : rapport de stage Master 2 (2015 - David Guyon, Université de Rennes 1)

p. 8-9 : droits réservés, Amazon, <https://aws.amazon.com/fr/ec2/purchasing-options/dedicated-instances/>

p. 10 : Pixabay, domaine public, <https://pixabay.com/fr/billets-50-euros-argent-europe-301742/>

p. 11 : Pixabay, domaine public : <https://pixabay.com/fr/document-accord-documents-signe-428334/>

p. 11 : © z_amir, Fotolia

p. 12 : Pixabay, domaine public, <https://pixabay.com/fr/engrenages-fonction-ensemble-818456/>