



Pièges et erreurs à éviter

pour garder le cap dans la tempête



Dans un contexte macro-économique de ce début 2025 chargé d'**injonctions contradictoires** pour les entreprises, il est très souvent difficile de concilier impératifs financiers de **croissance économique pour les Métiers**, leviers d'Innovation, accélération des cas d'usages Métiers augmentés par l'IA avec **frugalité budgétaire, durabilité environnementale** et respect des **contraintes sociales et réglementaires ESG**.

Les organisations ayant adopté le FinOps ces dernières années connaissent l'**importance stratégique du « Shift Left »**, en intégrant les bonnes pratiques FinOps à l'échelle et le plus tôt possible dans la chaîne de valeur, comme on l'a vécu sur la qualité du code applicatif, le DevOps ou la cybersécurité, le SecOps ces dernières années.

Une tentation de « **l'urgence court termiste** » pourrait, chez d'autres organisations moins avancées, avoir comme conséquences négatives une **approche « Cost Killer »** créant des tensions profondes, contradictoires avec la bonne compréhension des fondamentaux du FinOps, impactant l'innovation, la compétitivité et risquant de créer des situations de blocage, de crise voire de décrochage.

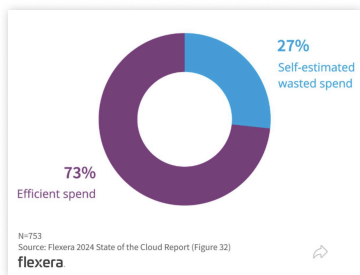
Dans cette Newsletter, on repose le décor de la **Réduction des coûts et les limites associées**, avant de rentrer dans le vif du sujet sur la **Refacturation (Chargeback)**, le concept d'**Unité de Valeur Métier** pour le Business (**Unit Economics**) pour se projeter sur ce qui pourrait être le **Graal du FinOps** pour aligner le Cloud sur la Stratégie de l'entreprise.

1

Réduction des coûts et limites



Le FinOps, concaténation de “Finance” et “DevOps”, désigne un cadre opérationnel et une pratique culturelle qui **optimise la valeur Métiers (Business)** du cloud (source : Fondation FinOps [FinOps Framework Overview](#)).



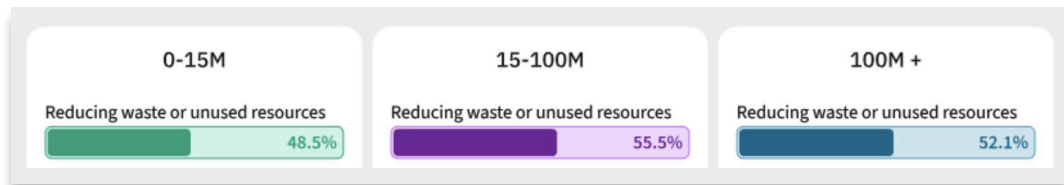
◀ What's your estimated wasted cloud on IaaS and PaaS ?

Dans l'inconscient collectif, le terme est encore trop largement dévoyé au profit de “réduction des coûts”. En effet, plusieurs études, dont [State of the Cloud Report 2024](#) de Flexera, évalue à **27%** auprès de 753 répondants le gâchis (“Waste”) de dépenses Cloud non utilisées / mal utilisées par rapport aux besoins.

Ainsi, pour se rassurer, on pourrait dire que le gâchis réduit de **32%** (en 2022) à **28%** (en 2023) jusqu'à **27%** (en 2024).

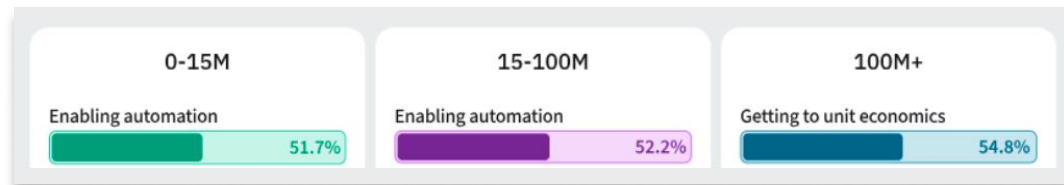
► On pourrait le corroborer également aux priorités remontées par l'étude [“State of FinOps 2024”](#) de la FinOps Foundation :

- Key Priorities for Practitioners by Cloud Spend -



► Lorsque cette fois, on rentre dans le détail des priorités ayant augmenté, classées par dépenses Cloud : **Increased priorities by cloud spend...**

- When looking at increased priorities (not key priorities) -



...on comprend que la priorité de faire le lien avec la **Valeur dégagée pour les Métiers** (mesurée notamment à travers les Unit Economics) reste la priorité des plus grands consommateurs de Cloud à l'échelle mondiale mais que “Pour la première fois (en 2024), la réduction du gaspillage (est devenue) la **priorité absolue des FinOps Practitioners**, tous niveaux de dépenses confondus. Cela (a pu être) influencé par les tendances macroéconomiques, les entreprises cherchant à **réduire leurs dépenses** sans réduire la valeur de leurs investissements dans le cloud.”



Comment garder le cap dans la tempête ?

Car si on y réfléchit à deux fois, les efforts, la charge, le ROI...vont très vite se heurter à un **plafond de verre** sur les Optimisations de base (Right-Sizing, décommissionnement de ressources non utilisées et sous-utilisées, Reserved Instances (RIs), Savings Plans (SPs)...les limites rencontrées étant de plus en plus importantes sur la complexité des infrastructures, l'interdépendances des workloads avec les couches applicatives, les arbitrages de priorisations...

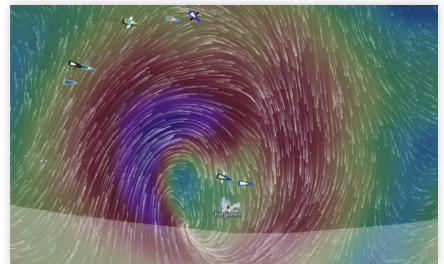


En effet, plus que jamais **prioriser** - comme on l'a toujours fait - à l'époque avec les concepts de classes de services, de **SLAs**, les **applications Métiers les plus critiques et porteuses de valeur**, devraient nous occuper avec un niveau de granularité plus fin, une capacité d'évolutivité temporelle (plus ou moins critiques selon les jours, mois, périodes de l'année...), d'évolutivité géographique et selon le sous-jacent d'infrastructure.

En un mot, c'est bien dans la **"tempête d'incertitude économique"** qu'on va reconnaître celles et ceux qui savent **naviguer en "bons marins"**, capables de réduire la voilure quand et où il le faut, tout en sachant accélérer dans les dépressions pour bénéficier de leviers de vitesse extraordinaires, **conférant des avantages compétitifs décisifs pour la victoire.**



C'est ce qu'a fait avec un formidable talent **Charlie Dalin**, à plusieurs reprises, notamment en Atlantique au **Pot-Au-Noir**, et également et surtout dans **les mers du Sud**, au niveau des îles Kerguelen, pour bénéficier de conditions d'accélération offertes par les vents du front et battre le record des **64 jours de l'Everest de la Voile**: le Tour du Monde en solitaire du Vendée Globe.



Crédit : L'équipe



"C'est comme le train, il ne faut pas le louper !"
"Chaque mètre gagné compte et vaut de l'or."

2

Chargeback et Unit Economics : leviers d'alignement

► Quelle forme prend le Chargeback dans le FinOps au sein de 5 sociétés de renom ?



● Bpifrance

En 2020, Bpifrance a démarré une stratégie d'hybridation, dite « **Cloud First** », accompagnée d'un enjeu de maîtrise des budgets. La mise en place progressive de la stratégie FinOps Multi Cloud a mis en avant la nécessité de **passer du showback au chargeback** tout en étendant le périmètre des usages supervisés dans le Cloud. Un des premiers challenges a été la **mobilisation de toutes les parties prenantes** sur ces sujets.

En 2023, le contrôle de gestion et les Achats se sont impliqués pour aboutir, avec la Gouvernance Cloud et FinOps, pilotée par Hela ben Farhat, à la mise en place d'un processus de **chargeback** et d'une culture axée sur la **frugalité** au niveau des **équipes**.

L'acculturation des équipes via la **refacturation** a permis de responsabiliser les utilisateurs sur leurs consommations et leurs dépenses. Les coûts sont remontés aux utilisateurs de ce qu'ils utilisent réellement en comparaison des coûts de la facturation (du billing). La mise en place des KPIs autour de la vue capacitaire du cloud privé, combinée à la définition d'un modèle de coût commun, a permis de disposer des **coûts unitaires et par application**.

Si ces coûts incluent les ressources provisionnées par les équipes et non utilisées, c'est dans le but de les inciter à **optimiser leurs usages et réduire l'empreinte carbone**. Ce type d'initiative FinOps a permis de créer des **synergies** avec l'équipe Green de Bpifrance, permettant de mutualiser les efforts autour du **Green IT**.

● Hermès

La refacturation tire son origine depuis le **on Premise**, par le besoin d'avoir un modèle de coût cohérent. Cela a été étayé dans une **documentation d'explication de la refacturation**.

Pour le Cloud, il est rapidement apparu la nécessité de formaliser (en s'appuyant sur le Tagging et la mise en place de **Centres de Coûts**) une démarche similaire. La documentation de refacturation a donc été amendée. Parmi les besoins : identifier les entités portant l'Opex.

Trois étapes clés : le Design (avec calculatrice, estimation d'une architecture initiale, première estimation coût par environnement) puis l'**Enveloppe budgétaire** (E05 en Mai et E09 en Septembre), le **Ré-estimé** pour le **Build** puis pour le **Run**. Alexandre Fedi.

La gestion fine des dépassements et des mécaniques de **Start-Stop** et de **Stop** en cas d'anomalies permettent **un retour à l'équilibre budgétaire**.

Des **KPIs** offrant la possibilité de **s'auto-évaluer**, permettent d'objectiver l'ensemble sous l'angle de la gouvernance (nombre de personnes formées...), le **taux d'applications refacturées**, le **taux de couverture des Savings Plans (SPs)** et bien sûr les **économies réalisées...**





● EDF

EDF, de par ses activités stratégiques énergétiques et nucléaires s'appuie sur du **cloud hybride** sous trois formes : **public (40%), privé (30%), de confiance (30% à horizon 2030)**. Les activités de vente, de production et industrielles n'ayant pas du tout, et on le comprend, la même sensibilité.

Comme l'explique Florian Beaurain, deux piliers sont clés : **le suivi du contrat et des savings**, raison pour laquelle une campagne de Tagging a été mise en œuvre avec l'objectif d'identifier les coûts directs sur chaque ressource au niveau

des projets. S'agissant des coûts indirects (Sécurité, Réseau...) **l'allocation** est effectuée **proportionnellement aux coûts directs**.

S'agissant du mécanisme de **refacturation**, il y a un **mark-up** entre les coûts des consoles des Cloud Providers et les coûts de refacturation. **Les budgets** sont gérés au niveau du **contrôle de gestion**, tout comme les **budgets prévisionnels**.

La reventilation est effectuée à la maille des **Business Units (BUs)** et des **projets**, la répartition des discounts est effectuée au prorata de leur consommation. Les coûts des Savings Plans sont amortis sur les applications ayant bénéficié des discounts, de même sur HPC (qui présente un niveau de complexité plus fort au niveau des shared costs et allocations), **le taux de couverture des Savings plans engagés** étant suivi méticuleusement. L'allocation est effectuée par **BU/Projet/Environnement** ce qui facilite la reventilation des coûts.



● ST Microelectronics

L'allocation des coûts chez ST s'effectue au niveau des **souscriptions** et des **resources groups**. En effet, Pierre-Emmanuel Nuiry a eu l'occasion de présenter son approche à Paris chez Microsoft et à Barcelone pour la Fondation FinOps (cf. FinOpsXE).

Plusieurs caractéristiques intéressantes sont résumées dans la planche ci-dessous :

- une clé d'**allocation unique**;
- l'utilisation de **Tags** pour des cas bien précis (HPC) mais pas holistique
- l'association d'une clé d'**allocation** à **un seul et unique centre de coûts (Cost Center)**;
- Et surtout, « **pas un centime laissé de côté** ».

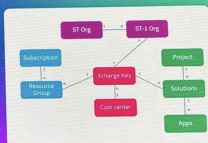


Costs Allocated at Subscriptions + Resources Groups

- A single Allocation Key
- Azure Subscription & Resources Groups
- Tag not used in allocation key, however, used
- for tracking single CAD HPC project executions
- An allocation key has a unique Cost center

Why: ~20% of running resources are not Taggable

Best Practice:
An Allocation key for 100% cost charge back done outside the CSP platform to enable yearly company changes



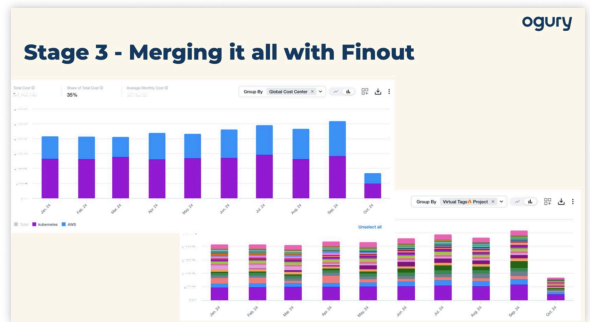
La combinaison de ces bonnes pratiques permet une **refacturation à 100%**, effectuée en dehors de la plateforme Cloud (en l'occurrence Microsoft Azure), permettant de gérer plus facilement les changements organisationnels éventuels de l'organisation.

ogury

● Ogury

Et dans le monde Kubernetes, qu'en est-il ? Ogury avait besoin d'une **allocation de coûts automatisée** et supportant le Multi-Cloud, au-delà d'AWS, sous un concept de « **Tag Project** » avec une hiérarchie flexible et dynamique, du reporting et de l'alerting pour du **Showback**.

Cette approche, outillée par **Finout** a permis à **Adrien de Castelnau** et ses équipes d'obtenir une vue détaillée, à jour et demain de se projeter sur un « **unit cost** » par impression, par requête et de gérer l'**allocation des shared resources** (DBs, stockage...).



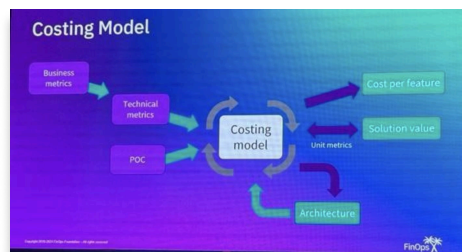
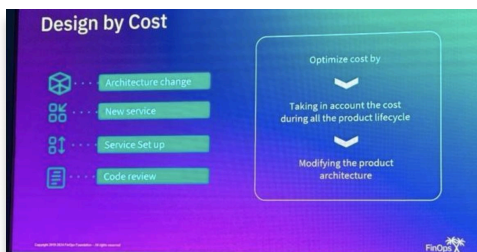
► Quels sont les exemples parlants de Unit Economics au sein de grandes multinationales françaises?

Une fois ancrée cette mécanique de ré-allocation et de refacturation, l'étape ultérieure est clairement d'associer des "Business Metrics", des éléments de Valeur Métiers, aussi appelés **Unit Economics** dans le FinOps, afin de mieux aligner les dépenses Cloud sur la Valeur Stratégique dégagée pour l'entreprise.

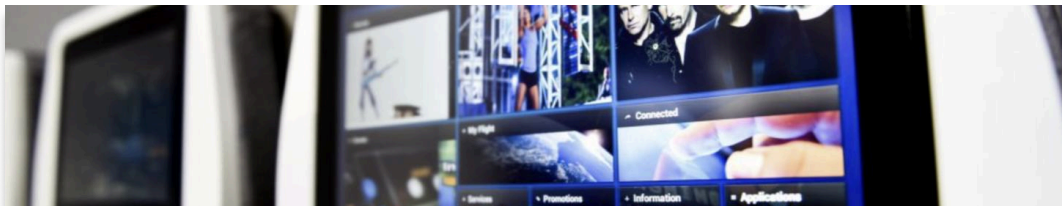
THALES

● Thales

Ainsi Thalès Avionics, au travers de son approche "Design by Cost" a décidé de prendre l'Unité de Mesure la plus fiable à remonter et la plus commune : le "**Cost per CPU**". En effet, comme l'expliquait fin 2023 à la conférence FinOps Foundation de Paris et dernièrement fin 2024 à la conférence FinOpsX de Barelone, **Philippe Galeano**, "**La métrique n'est pas parfaite, on préférerait le Coût par UE, mais ça donne un point de départ**".



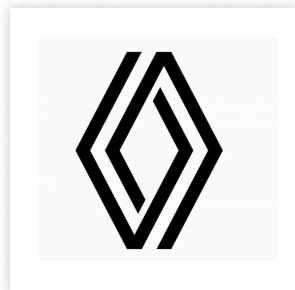
De cette approche exposée dès 2023, on voit que cela a permis d'évoluer vers un Costing Model plus complet de **"Cost per feature"** et d'introduire les Unit metrics pour construire une valeur orientée solution et faciliter le **"Shift Left"** pour intégrer la dimension **"Coûts"** le plus tôt possible dans le chaîne de valeur de la conception <https://www.finops.org/insights/cost-aware-product-decisions/>, en l'occurrence ici les écrans, les programmes et les appliances des contenus, jeux et applications des millions d'équipements embarqués dans les avions utilisés par 1,6 millions de passagers chaque jour.



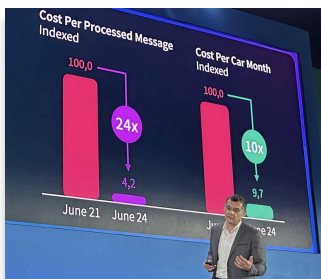
● Renault

L'approche Unit Economics chez Renault, qui elle aussi avait été évoquée dans la Newsletter de Novembre 2024 **"FinOps Europe Maturité"**, est révélatrice du besoin des industriels, en particulier dans l'automobile.

Sous injonction contradictoire de l'utilisation de l'innovation dont l'IA et sous **pression croissante de réduction des coûts**, de l'**empreinte carbone** de disposer d'un tableau de bord pour la DSI et les Métiers en suivant le **coût par message** (messages en milliards issus des télémetries, des commandes à distance, des mises à jour...) et de le décliner par **un ratio du Métiers** en suivant le **coût par message** (messages en milliards issus des **Nombre de messages** par les **coûts totaux du cloud**..



Atteindre cette réduction par facteur **x10 en 3 ans** est phénoménal compte tenu de l'évolution de la croissance de la production **x3 de 80k à 200k véhicules** produits chaque mois durant cette période.



Ainsi la combinaison **"Ratio Coûts totaux des messages générés vs. Coûts totaux du Cloud"** est un excellent moyen de mesure de l'efficacité Valeur du Cloud pour les Métiers pour ce cas d'usage, alors que les coûts du Cloud ont augmenté de 2021 à 2024, mais le Ratio de la Valeur dégagée prend ici tout son intérêt, comme l'expliquait fin 2024 **Frédéric Mingot**.



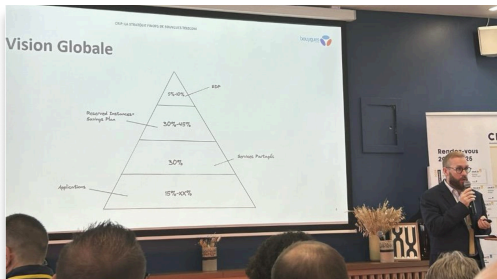


● Bouygues Telecom

Maël Louvet, dans son intervention à la Matinale FinOps du CRIP, mettait au coeur de son intervention **“Le meilleur service au meilleur coût”**, et dans un contexte Applications-driven sur Kubernetes revenait sur la nécessité impérieuse :

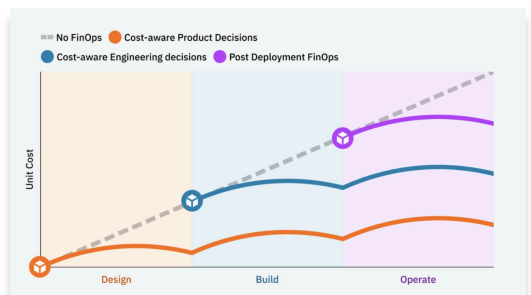
- d'avoir une **vision globale** de là où est généré la Valeur et donc l'impact d'une maîtrise au **plus près des applications**, car de rester très / trop “Infrastructure”-driven (comme le sommet d'une pyramide ou la partie émergée d'un iceberg) ne fait qu'effleurer le sujet et perd en impact.

Pour mettre en perspective, le site <https://www.bouyguetelecom.fr/> encaisse en pic jusqu'à 1 million de clients / heure. Bâti sur AWS (EKS), l'approche de BPU (Rentabilité des Opérateurs Télécoms) est chevillée au corps. En effet une interaction téléphonique ou chat d'un conseiller coûtant 1€, on réalise que sur un forfait de 10€/mois, les approches “chasse aux coûts” sont d'emblée limitées et que le **Design to Cost** est essentiel.



Cette approche, d'ailleurs corroborée par de nombreux retours d'expérience et largement couverte au sein de la FinOps Foundation a même été documentée dans cette optique à travers le Cost Awareness Product Design pour permettre un vrai **“Shift Left”** à l'image de l'éco-conception pour le **GreenOps**.

Plus l'approche est pensée et mise en œuvre en amont plus les **bénéfices Valeurs/Coûts** sont importants. C'est ce qui va nous occuper dans les années à venir.

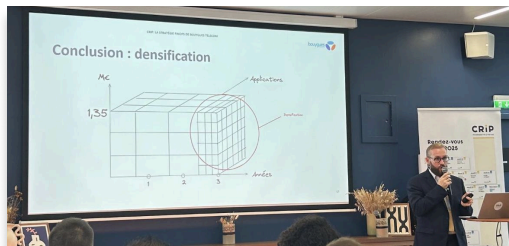
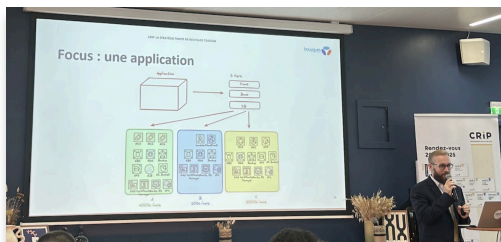


Ainsi, si on poursuit l'exposé de Bouygues Telecom dans cette approche, on comprend aisément qu'une **application 3 tiers** : Front-end, Back-end, Base de données (Front, Back, DB), peut - dans le monde des containers kubernetes et microservices - se retrouver (selon les **classes de services sélectionnés**, les **SLAs choisis**, les **options activées...**) pour simplifier sous une offre de service **Standard (A)**, **Basique (B)** ou **Premium (C)**.

Les coûts associés **4 000€/mois**, **200€/mois**, **15 000€/mois** pris pour l'exemple, doivent permettre de faire des **arbitrages éclairés: quels environnements?** Prod/hors-Prod, Run/Build? ARN vs x86? Refactorisation ou non? TTM/HELM? intégration d'outils...?

Une fois cette **introspection sur les cas d'usage**, finalement dépendant des **SLAs** et des **enjeux Métiers**, donc de la **Valeur attendue**, il sera ainsi plus facile de choisir les applications candidates éligibles à quelle classe de service.

S'il on garde notre angle "**Unit Economics**", cette approche cible démontre, s'il en était besoin, que plus l'on va vers des architectures modernes cloud-natives, plus la connaissance, la maîtrise et la cap et la capacité de projection à l'échelle dans la durée des Unit Economics sont plus que jamais indispensable pour faire les bons choix du sous-jacent Cloud et permettre, ce que Maël illustre dans son schéma vers la **densification**.



L'automatisation peut être poussée sous l'angle d'un "**quota financier**": déclenchement via API à atteintes de seuils et les Développeurs sont "**sensibilisés**" en voyant les coûts avant chaque Deploy, ce qui aide également le **Green**.

3

Le Graal : Valeur et Performance (EVM & CPI)



● AXA

Le cas d'usage Axa présenté par [Anderson Oliveira](#) et, lui aussi, détaillée dans la Newsletter de Novembre 2024 "FinOps Maturité" est une référence en matière de calcul de Unit Economics. Pourquoi?

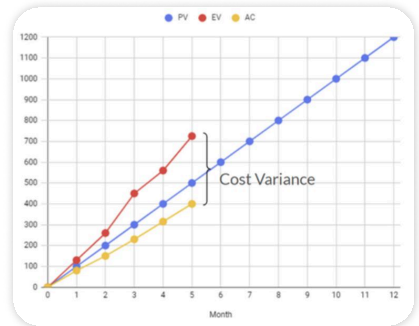
D'abord, il s'agit de partir de 3 variables de données : 1) La Valeur budgétée (ou planifiée) attendue du travail planifié (PV: Planned Value), 2) Le coût actuel du travail réalisé (AC : Actual Cost), 3) La Valeur Obtenue du travail réalisé (EV : Earned Value)

Puis, on cherche à définir la Valeur Obtenue (EV : Earned Value)? 1) Déterminer l'Unit Economics actuel d'une période donnée, 2) Comparer à l'Unit Economics planifié sur la même période, 3) Multiplier le pourcentage par le Montant Budgété de la période.

Ainsi, $EV = PV \times \% \text{ Complete}$, on en déduit :

$$\%Complete = \frac{Unit\ Economic_{Target}}{Unit\ Economic_{Month}}$$

PV, pour simplifier dans cet exemple, prend en compte une ventilation budgétaire mensuelle linéaire **Planned Value (en bleu)**, l'**Earned Value (en rouge)**, l'**Actual Cost (en jaune)**.



L'objectif est évidemment de maximiser l'**Earned Value** tout en maîtrisant l'**Actual Cost**, car l'un des effets pervers peut être qu'en diminuant les Coûts, on **impacte négativement la valeur Métiers (Business Value)** à cause d'une variance négative.

CPI Formula

$$CPI = \frac{Earned\ Value}{Actual\ Cost}$$

Earned Value = Planned Value X % Unit Economics

Budgeting + Value Planning = Unit Economics

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Cost	100	100	100	100	100	100	100	100	100	100	100	100
Value	200	200	200	200	200	200	200	200	200	200	200	200
Unit Econ. (cost/value)	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

Comme précédemment indiqué, Anderson Oliveira a écrit un article complet sur son blog, son implication au sein de la Fondation, en particulier dans le groupe de travail KPIs, a permis d'ajouter la définition de CPI - Cost Performance Indicator. Cette méthodologie devrait devenir un standard pour mieux communiquer sur la Valeur Métiers dégagée par le FinOps pour sortir du « préjugé » de réduction des coûts.

Réduire les coûts est réducteur, Aligner le Cloud sur la Stratégie est l'accélérateur !

Dans une période de tempête, quelle est la meilleure stratégie FinOps à adopter? Est-ce de réduire drastiquement tous les coûts partout, en mode Optimist? pour revenir sur les fondamentaux?



La réponse est **Oui**, pour beaucoup de cas d'usage où l'on sait que réserver des **ressources (RIs/SPs)** et suivre le taux d'utilisation (>80%), le **Right-Sizing**, le **décommissionnement** des ressources non utilisées, sous-utilisées, le **nettoyage (Storage, Back-Ups, Network, Logs...)** sont des bonnes pratiques extrêmement bien documentées et indispensables dont on estime que s'il y a **30%** de gâchis, d'over-spending, alors ces pratiques peuvent largement contribuer à hauteur de 15 points, ce qui sur des budgets en Millions et dizaines de Millions peut rapidement être conséquent.

Mais on a vu aussi que cette approche peut rapidement présenter des limites (plafond de verre et lassitude des équipes à mener ce type de campagnes à des rythmes de plus en plus rapprochés, risques et peurs d'impacts sur les SLAs, les données sauvegardées en cas de nécessité de Restore, en cas d'attaque Cyber...); il y a donc des **pièges et erreurs à éviter pour garder le cap dans la tempête**, en particulier dans **cette ère de l'IA**.

De nombreux témoignages montrent donc la voie à travers la **refacturation** et la mise en œuvre d'**Unités de Valeur Métiers (Unit Economics)** en lien avec le Business, pour aller chercher les 15 points supplémentaires. Cela demandera certainement du Replatforming, qui a un coût et une latence avant d'avoir un ROI, qui nécessite une connaissance et un **Mapping des Applications** les plus stratégiques et porteuses de business pour être en capacité de faire les bons choix, pour le Fin et le Green. Mais aligner une technologie de pointe sur **les enjeux les plus essentiels** pour **gagner la course**, n'est ce pas là tout l'Art du FinOps?



Cloud FinOps



Cloud FinOps est un cabinet **100% dédié FinOps indépendant** intervenant en France et à l'international qui aide ses clients à « **booster la Valeur du Cloud** » et à vraiment « **contrôler leurs dépenses Cloud** ».



Cloud FinOps



cloud-finops.io