

E-mails et archivage



Sommaire

■ L'archivage, au-delà du stockage et de la sauvegarde

Le stockage à la base de tout

Sauvegarde et archivage

L'archivage est-il une nécessité ?

Qui est concerné par l'archivage ?

■ Nécessité de s'appuyer sur une politique d'archivage

Pourquoi construire une politique d'archivage

De la politique d'archivage au système d'archivage électronique

■ Les solutions d'archivage d'e-mails

Les différents modes d'archivage opérationnel

Comment choisir l'outil ?

Veiller à la performance

L'archivage, au-delà du stockage et de la sauvegarde

Bien distinguer gestion de l'information, conservation, stockage, sauvegarde et ... archivage ?

Le stockage à la base de tout

Le stockage est la première étape du traitement de l'information dont l'e-mail fait évidemment partie. On peut même dire qu'il conditionne l'existence de l'information : si les données qui constituent l'information produite par une personne ou un système ne sont pas enregistrées et stockées quelque part dans une mémoire informatique ou sur un support physique, l'information n'a plus d'existence matérielle ; elle est perdue ou se limite à une information orale ou mémorielle qui ne peut donc pas être archivée.

Les supports et systèmes de stockage offrent une gamme variée d'outils face à des besoins variés de gestion de l'information dans l'espace et dans le temps :

- technologies optiques, numériques,
- sécurité,
- capacité,
- résistance à l'obsolescence,
- temps d'accès aux données,
- modalités de migration.

Sauvegarde et archivage

Certains utilisateurs et quelques acteurs de la gestion de l'information ont encore tendance à confondre sauvegarde et archivage. Il s'agit pourtant de notions bien distinctes.

La finalité de la sauvegarde est uniquement de permettre une copie des données d'origine dite copie de sécurité afin d'éviter de les perdre en cas de dysfonctionnement du dispositif sur lequel elles sont enregistrées ; la durée de conservation est déconnectée de la valeur du contenu et relative à la périodicité de la sauvegarde (journalière, hebdomadaire, mensuelle, annuelle).

À l'inverse, l'archivage doit permettre une conservation qui peut être beaucoup plus longue, voire ad vitam aeternam. De plus, l'archivage permet une interrogation aisée et fine des objets conservés. Contrairement à la sauvegarde, les données archivées sont considérées comme figées, c'est-à-dire non modifiables.

D'autre part, l'archivage doit permettre de consulter les informations indépendamment de leur système d'origine (ce qui suppose d'extraire et de conserver les informations dans des formats pérennes et si possible ouverts tels que XML) et des droits d'accès initialement affectés aux documents.

Retenons donc que si la sauvegarde est conçue pour la restauration de données et systèmes perdus, l'archivage quant à lui est conçu pour conserver des documents de référence ou probants.

Il y a encore peu, le cycle de vie des données au sein de l'entreprise pouvait se résumer à :

- la création de la donnée,
- son utilisation courante et sa gestion au quotidien, modifications, accès...,
- l'archivage des données à des fins légales ou réglementaires,
- la destruction ou l'archivage patrimonial (histoire de l'entreprise).

Actuellement l'archivage, dans sa version électronique permet de remonter très en amont dans le cycle précédent et intervient ainsi dans une utilisation courante, au quotidien dès l'instant où les données sont figées, ce qui ne modifie en rien leur emploi régulier. Quel meilleur exemple pour illustrer cela que l'e-mail dont la principale caractéristique est justement celle d'être figé (non modifiable) dès sa création.

Gestion Electronique de Documents et Archivage

L'archivage électronique est également à distinguer des outils de gestion électronique de documents. Ces derniers visent en général à répondre aux besoins de partage et d'amélioration de la productivité des services. A contrario, contrairement aux systèmes d'archivage électronique, ces outils prennent rarement en compte la gestion de la valeur probante de ces documents en garantissant leur intégrité sur le long terme.

Correspondances dématérialisées et mémoire d'entreprise

L'archivage est-il une nécessité ?

Ainsi l'archivage électronique ne doit pas être vu comme la simple « dématérialisation de l'archivage papier ». L'une des principales différences provient du fait que les documents électroniques sont archivables dès lors qu'ils ne sont plus modifiables, ce qui n'enlève rien à leur accessibilité. On est loin des interrogations fastidieuses des archives dans les arrières caves poussiéreuses et il serait dès lors plus précis d'utiliser une autre terminologie comme archivage actif (proposé par le Gartner) ou encore « archivage dynamique » afin de bien marquer la différence.

La conséquence en est une prise de conscience très en amont de la notion d'archive sans attendre la quasi fin de vie des documents. La loi elle-même est très claire sur le sujet dans la mesure où l'un des critères de recevabilité d'un écrit électronique en est la garantie d'intégrité depuis son origine et non pas uniquement pendant sa période d'archivage au sens classique du terme.

Comme vu précédemment, l'e-mail constitue un parfait exemple de ce qui précède en matière d'archivage. En effet il possède cette caractéristique de ne plus être modifiable à partir du moment où il existe. Dès lors que l'on conserve des e-mails il s'agit en fait d'un archivage au sens actif du terme. Plutôt que de parler d'archivage des e-mails il serait donc plus juste de présenter cela comme une gestion optimisée et efficace de ces derniers.

Pour être tout à fait précis, il est parfaitement possible de « gérer » ses e-mails au quotidien et au bout d'un certain temps de les archiver au sens classique du terme. Cependant une telle approche a pour principal mérite de se priver de tous les avantages liés à l'archivage dynamique des e-mails et de leur gestion au quotidien.

Qui est concerné par l'archivage ?

De plus en plus d'informations transitent par les e-mails, ce qui ne va pas sans poser un certain nombre de problèmes auxquels tout chef d'entreprise se doit de trouver des réponses.

Sans vouloir aborder ici les aspects purement sécuritaires, traités par ailleurs, nous nous attacherons à plutôt définir la mise en place de moyens destinés à simplifier la gestion des e-mails au quotidien, tant du point de vue de l'utilisateur, que du directeur informatique et du chef d'entreprise et ce tout en respectant les aspects légaux et réglementaires.

Côté utilisateur, il y a l'exigence d'un confort maximum à pouvoir retrouver ses e-mails facilement sans pour autant être systématiquement obligé de les organiser par dossier et autre sous dossier. Cette classification montre d'ailleurs très vite ses limites lorsqu'un e-mail peut être rattaché à deux dossiers différents.

Côté directeur informatique, les e-mails représentent vite un véritable cauchemar compte tenu de l'évolution à la fois de leur nombre et de leur volume moyen. La solution la plus souvent mise en place consiste à limiter radicalement la taille des boîtes de chaque utilisateur ! Par rapport à ce qui précède, la sauvegarde des données ne doit plus s'appliquer aux données figées dans la mesure où elles sont éligibles à l'archivage. Il y a là un gain potentiel considérable à réaliser par simple élimination des doublons et autre duplication d'information totalement inutile à supprimer. L'autre vrai problème consiste à pouvoir si possible éliminer les e-mails qui n'ont aucun sens à être conservés mais toutefois avec la certitude de ne pas éliminer des e-mails potentiellement utiles.

Remarque : le décret n° 2006-358 du 24 mars 2006 relatif à la conservation des données des communications électroniques impose aux opérateurs de communication de conserver la trace des échanges électroniques et donc des e-mails. Bien que cette loi ne concerne a priori que les opérateurs, elle concerne en réalité de plus en plus d'entreprises ayant mis en place des réseaux ouverts et notamment Wi-fi.

Côté archiviste, l'e-mail constitue une source d'information de plus en plus critique, notamment sur le plan juridique et patrimonial mais elle constitue également une source d'information difficile à collecter et à traiter du fait d'une part des volumes d'informations véhiculées et d'autre part de la nature

très personnelle et non-structurée de cet outil. Il est essentiel pour les archivistes d'une part d'être en mesure de sélectionner puis de collecter facilement les messages méritant d'être conservés sur le long terme, et d'autre part de disposer d'informations structurées (« méta-données ») exploitables pour retrouver facilement l'information.

Côté chef d'entreprise il s'agit d'avoir la garantie de ne pas perdre d'information stratégique ou non, qu'il s'agisse d'un aspect commercial, technique, comptable ou financier voire patrimonial. Cette garantie de conservation est aujourd'hui essentiellement dictée par des obligations légales et réglementaires sachant que de plus en plus de contraintes naissent en la matière, généralement dictées par un souci de traçabilité de l'ensemble des opérations au sein de l'entreprise.

Enfin toujours sous l'angle des contraintes il ne faut pas oublier de prendre en compte celles liées au droit à la conservation de données personnelles, d'où la problématique juridique directement liée aux exigences Informatique et Libertés pour le respect de la vie privée qui induit une durée limitée en rapport avec la finalité d'une conservation.

Par rapport aux aspects juridiques de l'e-mail voir la fiche 4.

Nécessité de s'appuyer sur une politique d'archivage

Les règles de fonctionnement définies dans la politique d'archivage des e-mails devront entre autre préciser d'une part les modalités de prise en compte des messages « privés » que l'entreprise n'est pas habilitée à archiver, et d'autre part les conditions d'accès aux messages une fois archivés (délais et conditions de restitution, droits d'accès aux archives...)

Les modalités opérationnelles devront quant à elle prévoir de gérer le dédoublement des messages et des pièces jointes associés aux messages reçus par plusieurs utilisateurs de façon notamment à limiter l'espace de stockage dédié à leur archivage.

Le lien avec la sécurité dans l'entreprise

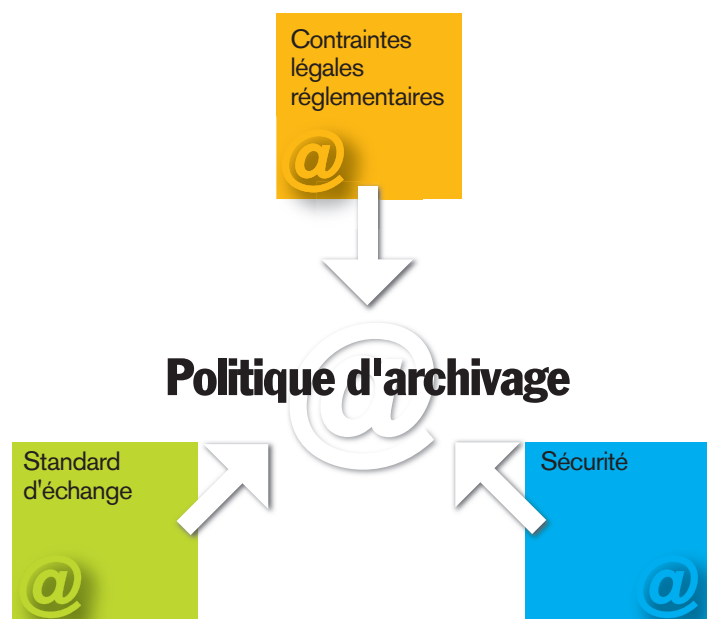
Pourquoi construire une politique d'archivage

L'archivage des e-mails doit être vu comme un projet à part entière dans l'entreprise et abordé dans ce sens avec la définition d'objectifs précis, d'identi-

fication des contraintes, du retour sur investissement attendu, d'un accompagnement éventuel auprès des utilisateurs,...

En matière de retour sur investissement la mise en place d'un système d'archivage actif des e-mails doit permettre entre autres d'économiser énormément d'espace disque du simple fait de sortir des procédures traditionnelles de sauvegarde l'ensemble des e-mails comme indiqué précédemment. Là où un e-mail et son contenu sont malheureusement aujourd'hui répliqués des dizaines de fois sans aucune utilité, un bon système d'archivage doit permettre la quasi unicité de l'information tout en la laissant accessible à un ensemble de personnes.

Afin d'aider la direction de l'entreprise à gérer ce type de projet il est fortement recommandé de s'appuyer sur une politique d'archivage (PA) qui permet avant tout de bien définir les acteurs en présence, leurs rôles et leurs responsabilités, de lister les contraintes qui pèsent sur l'entreprise (légales, réglementaires, internes), de fixer les objectifs poursuivis en particulier les durées de conservation, de décrire les fonctions attendues par le système d'archivage électronique, de préciser l'environnement de sécurité nécessaire. La politique d'archivage se présente ainsi au centre de différents éléments qui eux-mêmes évoluent dans le temps, d'où la nécessité d'actualiser cette politique régulièrement



Disposant d'une politique d'archivage adaptée aux e-mails il est alors possible d'en déduire un cahier des charges précis destiné à pouvoir choisir parmi telle ou telle offre du marché.

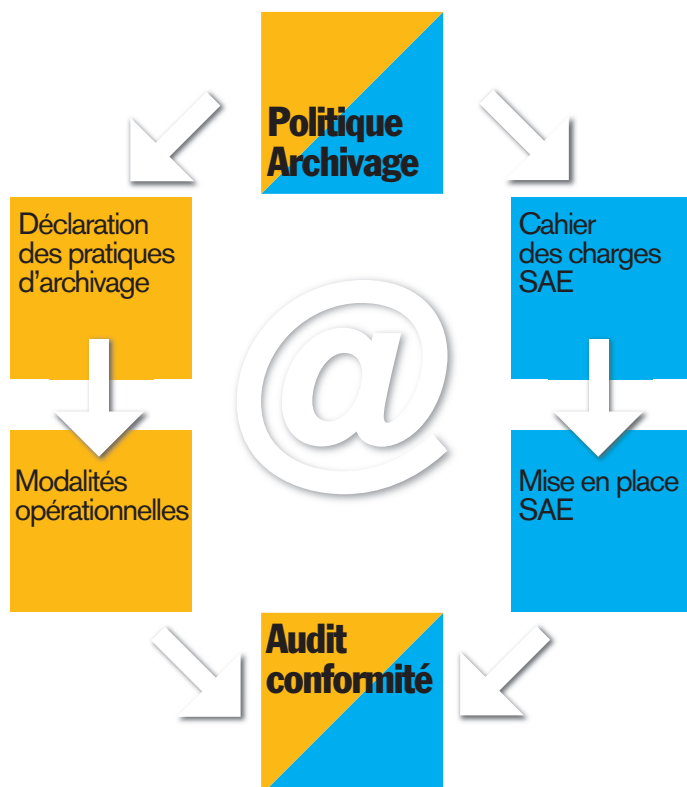
Remarque : Il est important de noter qu'une telle politique d'archivage, quoique décrite par rapport aux e-mails doit permettre également de traiter l'ensemble des données de l'entreprise, évitant par la même une gestion de l'archivage de type silo, interdisant toute mutualisation de moyens.

Enfin après avoir effectué la mise en place du système retenu il est indispensable de procéder à un audit destiné à vérifier la conformité de ce qui est installé par rapport aux objectifs et exigences définis dans la politique d'archivage.

De la politique d'archivage au système d'archivage électronique

Le schéma ci-après précise sur la partie droite la logique pratique de mise en place d'un système d'archivage électronique adapté tel que défini dans la politique d'archivage tandis que la partie gauche indique les documents indispensables à la bonne réalisation d'un tel système.

Pour plus de précisions sur le sujet voir le site : <http://www.ssi.gouv.fr/fr/confiance/archivage.html>



Les solutions d'archivage d'e-mails

Bien identifier son niveau de risque pour choisir Les différents modes d'archivage opérationnel

En simplifiant le propos, il existe à ce jour trois grandes philosophies d'archivage des e-mails.

1. Par l'utilisateur qui est responsabilisé

La première solution consiste à décrire pour chaque utilisateur l'ensemble des procédures retenues par l'entreprise et à adopter sur son poste de travail : organisation des répertoires, création de fichiers .PST, etc... Certes la plus économique mais seulement en apparence, il s'agit d'une solution dangereuse car sans la mise en place de contrôles réguliers et coûteux, le chef d'entreprise n'aura absolument aucune maîtrise de la conservation ou non des données vitales échangées par e-mail.

2. Espace d'archivage sécurisé

Afin de pallier ce dernier inconvénient majeur, une deuxième solution consiste pour l'utilisateur à déplacer ses e-mails, à la demande et/ou en fonction d'une politique prédéfinie, vers un espace mutualisé de stockage disposant de toutes les fonctionnalités de recherche et de migration nécessaires. Sur son poste de travail l'utilisateur garde ainsi la trace et la visibilité sur l'ensemble de ses e-mails mais ne subsiste en réalité qu'un lien pour ceux qui ont été déplacés.

Il existe une variante de cette méthode.

Cette solution peut également être mise en œuvre en deux temps. Dans un premier temps, l'utilisateur transfère tout ou partie des messages qu'il émet ou reçoit dans une application métier, un système de GED (Gestion Electronique de Documents) ou un espace de travail collaboratif en associant aux messages des informations complémentaires (client concerné, type de message ...). L'archivage électronique est alors réalisé à partir de ces applications et peut ainsi être mieux structuré que s'il est réalisé directement à partir des boîtes aux lettres. De nombreuses solutions de GED ou de travail collaboratif offrent aujourd'hui cette possibilité de versement en standard.

3. Archivage déporté totalement automatisé

Enfin une troisième et dernière orientation consiste à opérer une copie systématique de tous les e-mails entrant et sortant vers l'espace d'archivage (ce qui nécessite d'avoir son propre serveur de messagerie). L'utilisateur peut alors se permettre de suppri-

mer l'ensemble des e-mails de son poste de travail dès qu'il les a traités dans la mesure où il accède à l'ensemble de sa base e-mail au travers du système d'archivage, généralement disponible via un simple navigateur web.

Attention. Cette solution consistant à automatiser totalement l'archivage rend problématique l'exclusion des e-mails « privés ».

Mise en place de la solution : en interne versus en mode externalisé

En fonction de la solution retenue, certaines peuvent être mises en place en interne dans l'entreprise ou être totalement externalisées en mode ASP chez un tiers archiveur. Là encore tout dépendra du type d'architecture dont l'entreprise dispose (voir fiche 8). En effet les solutions admissibles dépendent étroitement du fait que l'entreprise possède ou non son propre serveur d'e-mails (cf. troisième solution). Enfin certaines sociétés proposent un ensemble complet logiciel et matériel tandis que d'autres offrent uniquement la partie logicielle tout en préconisant un certain type de matériel.

L'ensemble des grands acteurs informatiques (IBM, HP, EMC, HDS, CA, SUN, ...) proposent ainsi des solutions d'archivage d'e-mails tandis que d'autres présentent des solutions spécifiques comme AXSOne, SYMANTEC (plus connu pour ses antivirus) ou encore ZANTAZ, actuel leader en matière d'archivage d'e-mail.

Comment choisir l'outil ?

La première chose est de vérifier l'ensemble des caractéristiques fonctionnelles suivantes :

- sécurisation de l'ensemble des e-mails d'un point de vue de l'accessibilité (confidentialité), de l'intégrité, de la traçabilité, éventuellement de la non répudiation (impossibilité de remettre en cause l'envoi ou la réception d'un e-mail) ;
- rationalisation possible des espaces de stockage en fonction de la fréquence et de la criticité des e-mails. Il s'agit en fait de prévoir au besoin un système de migration adapté ;
- catégorisation possible des e-mails afin de permettre la rationalisation évoquée ci-dessus ;
- conservation des e-mails de façon sécurisée et surtout intègre avec vérification de l'intégrité de façon régulière ;
- indexation de l'ensemble des e-mails et des pièces jointes de telle sorte à pouvoir les retrouver facilement ;

- pérennisation des e-mails afin d'assurer une conservation à moyen voire long terme ;
- partage des e-mails au travers de la définition des droits accordés aux utilisateurs ;
- destruction des e-mails à expiration du délai de conservation.

Si la grande majorité des solutions proposées aujourd'hui offre cet ensemble de caractéristiques, la façon d'y répondre n'est pas toujours identique. Ainsi telle solution privilégiera plutôt l'utilisateur et l'ergonomie, telle autre offrira plus de facilités en matière d'administration centralisée, une troisième optimisera les espaces de stockage, etc...

Veiller à la performance

Le choix devra donc se faire en fonction de vos propres objectifs fonctionnels et de vos propres contraintes d'environnement informatique et de coût. Il faudra également veiller à l'importance en matière de performance des systèmes informatiques. En effet, archiver est une chose, mais retrouver en est une autre sachant toutefois qu'il s'agit de la base même de l'archivage. Il est clair que retrouver un e-mail parmi plusieurs milliers n'est pas la même chose que parmi plusieurs centaines de milliers voire de millions. Si le volume global des e-mails à conserver est en général tout à fait acceptable, le nombre d'items est quant à lui en général considérable et la taille des index peut même dépasser la taille de l'e-mail d'origine !

Enfin, avant de se décider pour une solution, bien veiller également à la notion d'évolutivité, de réversibilité (très important dans le cas d'un tiers archivage) et d'interopérabilité afin de ne pas devoir remettre en cause le système retenu dans quelques années ce qui aurait pour principale conséquence de coûter extrêmement cher à l'entreprise.

Pour en savoir plus :

- Livre blanc « L'archivage électronique à l'usage du dirigeant » téléchargeable sur :
 - Cigref : http://cigref.typepad.fr/cigref_publications/fedisa/index.html
 - FedISA : <http://www.fedisa.eu/index.php?pc=articles>
- Dunod, collection Management des Systèmes d'Information : « Dématérialisation et archivage électronique. (Mise en œuvre de l'ILM)