

L'optimisation de sites web en Flash pour un meilleur référencement (1ère partie)

[Retour au sommaire de la lettre](#)

Le format Flash de Macromedia a toujours été considéré, depuis qu'il existe, comme une source de problèmes pour le référencement d'un site qui l'utilise. Si l'ajout de petites animations Flash à l'intérieur d'un site bâti en HTML ne pose généralement aucun souci, un site web conçu 100% en Flash reste aujourd'hui encore quasi invisible pour les moteurs. Deux ingénieurs canadiens nous explicitent pourtant dans cet article une méthodologie qu'ils ont mise en oeuvre pour obtenir une visibilité beaucoup plus forte sur Google et consorts à ce type de site. Explications...

Une technologie en forte progression

Avec une progression de plus de 98 % (donnée provenant des statistiques d'Adobe), la technologie "Flash" est devenue l'un des formats les plus répandus dans la création d'applications interactives, de jeux ainsi que de diffusion vidéo sur Internet.

Or, de nos jours, plutôt que de taper l'URL directement dans la barre d'adresse du navigateur, l'un des premiers réflexes de l'internaute est de trouver l'information ou le site désiré *via* un moteur de recherche.



Cette habitude incite donc à soulever une question : par le biais d'un moteur de recherche, est-il aussi aisé de repérer, dans les résultats affichés, les sites qui utilisent la technologie "Flash" que ceux en format HTML ?

Comprendre les résultats

Google étant, sans contredit possible, le moteur de recherche le plus fréquenté dans de très nombreux pays, nous l'avons choisi (NDLR : dans sa version canadienne, les auteurs de cet article étant basés dans ce pays) comme référence afin d'effectuer quelques recherches et analyses pour mieux comprendre l'impact du "Flash" dans les résultats. Dans le graphique ci-dessous, nous retrouvons donc une comparaison des résultats de recherche entre un site "Flash" et un site développé en langage HTML.



Le nombre renvoyé par la zone "résultats" est celui fourni par le moteur de recherche lorsqu'on saisit comme requête son nom de domaine. En bleu, le site web est développé en HTML. En vert, un autre site est, lui, conçu en Flash.

De par cette simple expérience, il est évident que nous pouvons obtenir davantage d'URL distinctes (i.e. de pages indexées) pour les sites HTML, soit une variété et un grand nombre de portes d'entrée, qui nous permettent d'accéder à un contenu pertinent concernant notre recherche.

Explication technique de ces résultats

Afin de mieux analyser les résultats, nous avons comparé les deux versions de codes de ces sites et essayé d'en comprendre l'interprétation qu'en font les "spiders" de Google.

Code HTML

```
<table border="0" cellpadding="0" cellspacing="0">
<tr>
<td height="10"></td></tr></table> <br> <br>
<table border="0" cellpadding="0" cellspacing="0">
<tr>
<td colspan="2" class="petittexte10"> Agence | 07-05-2007 </td></tr>
<tr>
<td colspan="2" class="souttitre4"> BBDO organise un bed-in </td></tr>
<tr>
<td colspan="2">
<font size="2" face="Verdana, Arial, Helvetica, sans-serif"
class="textbeginpale13"> L'agence BBDO Montréal
a pris le contrôle du site Dailywom.tv dans
le cadre d'un bed-in destiné à obtenir le
compte du Musée des beaux-arts de
Montréal (MBAM).
<a href="article.aspx?id=22006" target="_top">
</a></font> <br>
<table border="0" cellpadding="0" cellspacing="0"> <tr>
<td height="10"></td>
</tr> </table>
</tr> </table>
```

VS

Code FLASH

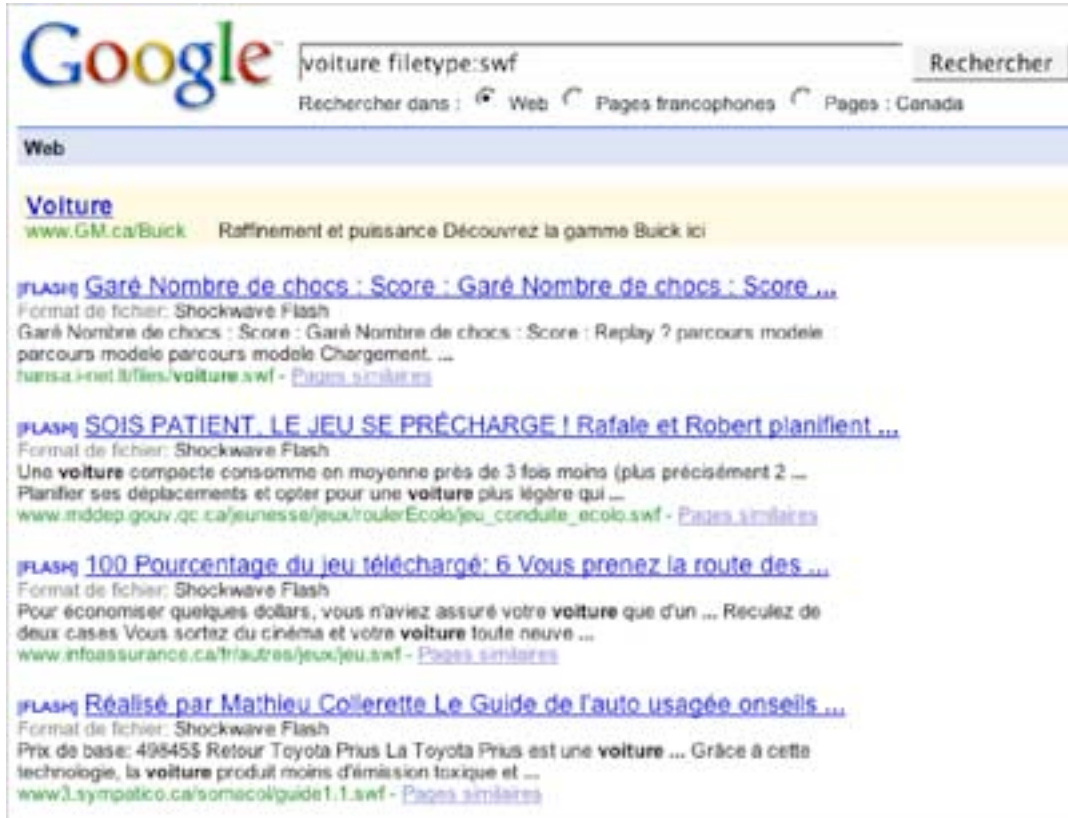
```
<div id="Layer2"></div>
<div id="Layer3">
<object classid="clsid:D27CDB6E-4E8D-11dF-9E88-
444553400000"
codebase="http://download.macromedia.com/pub/shockwave
/flash/swflash.cab#version=7,0,19,0" width="800"
height="550" title="Main_site">
<param name="movie"
value="seven25f.swf" />
<param name="quality" value="high" /> <embed
src="seven25f.swf" quality="high"
pluginspage="http://www.macromedia.com/go/getflashplayer"
type="application/x-shockwave-flash" width="800"
height="550">
</embed>
</object>
</div>
```

On peut constater, dans les illustrations ci-dessus, qu'à l'inverse d'un site HTML (où le contenu est pris en considération), le moteur de recherche n'est pas en mesure d'analyser le contenu à l'intérieur d'un site conçu en "Flash". En effet, en première approximation, ce format est perçu

comme une "boîte noire" dans laquelle il est impossible d'en lire le contenu. Il s'agit en réalité d'un média à part entière (fichier SWF), tout comme l'est un fichier vidéo.

Indexation des fichiers SWF par Google

Google peut cependant lire le contenu de certains fichiers SWF. Pour s'en persuader, il est possible de spécifier au moteur de recherche le type de fichier ciblé dans sa requête (ex : "voiture filetype:swf"). Les résultats s'afficheront en liste de médias (SWF), pertinents à cette recherche, mais, hors contexte :



*Google sait lire certaines animations Flash.
Le terme [FLASH] est alors affiché à gauche du titre de la page.
En revanche, les résultats sont loin d'être pertinents...*

Lorsqu'un site conçu en "Flash" est construit avec l'aide de plusieurs fichiers SWF, une seule partie du site sera retrouvée, mais sans son contexte global (HTML, Javascript, autres animations SWF, etc..). Bref, en excluant toutes les composantes normales d'un site. Il est alors impossible de sélectionner et d'ouvrir un de ces fichiers, puisque le reste des éléments est manquant.

Actuellement, aucune solution valable n'est proposée par les différents moteurs de recherche sur ce point. Google n'a proposé que certaines "guidelines" (<http://www.google.com/support/webmasters/bin/answer.py?answer=72746&query=flash&topic=&type=>), afin de respecter les règles du SEO (Search Engine Optimization), mais sans plus.

Des techniques à éviter

Il existe plusieurs techniques afin d'améliorer le positionnement de sites produits en "Flash", mais une mise en garde s'impose : certaines d'entre elles sont considérées comme illégales ou très "limites" par les moteurs de recherche. La liste suivante répertorie certaines solutions ou pistes, mais ces dernières correspondent, pour la plupart, à des techniques de positionnement dans les moteurs de recherche illicites ou tout du moins à déconseiller :

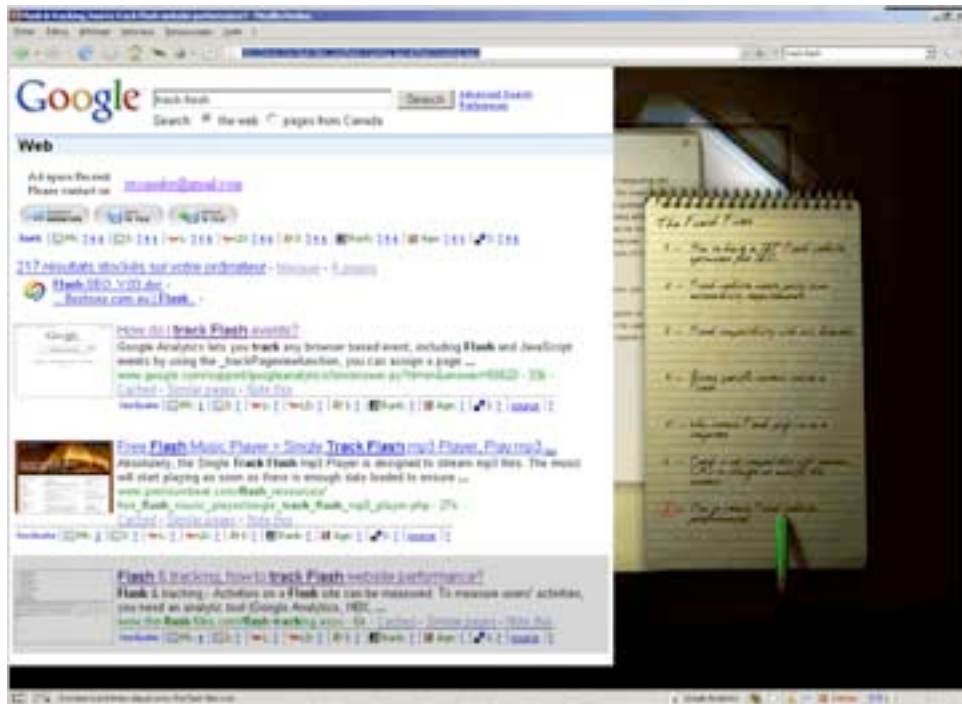
- Utilisation de balises "NOEMBED" : procédé trop utilisé, rarement à des fins honnêtes, donc peu considéré à ce jour.
- Cloaking : ce procédé, qui propose une version de page web adaptée aux spiders des moteurs de recherche lorsque l'un d'entre eux explore votre site, est très mal considéré par les moteurs et pénalisé. A exclure.
- *Doorway pages* ou pages satellites : ce procédé présente des pages cachées avec redirection (même constat que pour le cloaking).

Une solution innovatrice

Une solution pour obtenir le référencement d'un site web conçu en Flash peut, bien entendu, être de créer un autre site, en HTML cette fois. Mais cela peut s'avérer long et coûter relativement cher.

Nous avons donc mis en place une architecture différente, une stratégie nouvelle de création de sites en Flash, sur un site entièrement développé grâce à cette technologie, disponible à l'adresse <http://www.the-flash-files.com/>, et nous avons identifié différentes requêtes dans le but de vérifier si le site apparaît dans les résultats d'une recherche effectuée sur Google.ca.

- Recherche sur la requête "track Flash" : 9 400 000 résultats
the-flash-files.com est le 3e résultat avec l'URL www.the-flash-files.com/flash-tracking.aspx.
Lorsqu'on clique sur le lien, on arrive bel et bien sur la section en rapport avec le site, en "Flash".



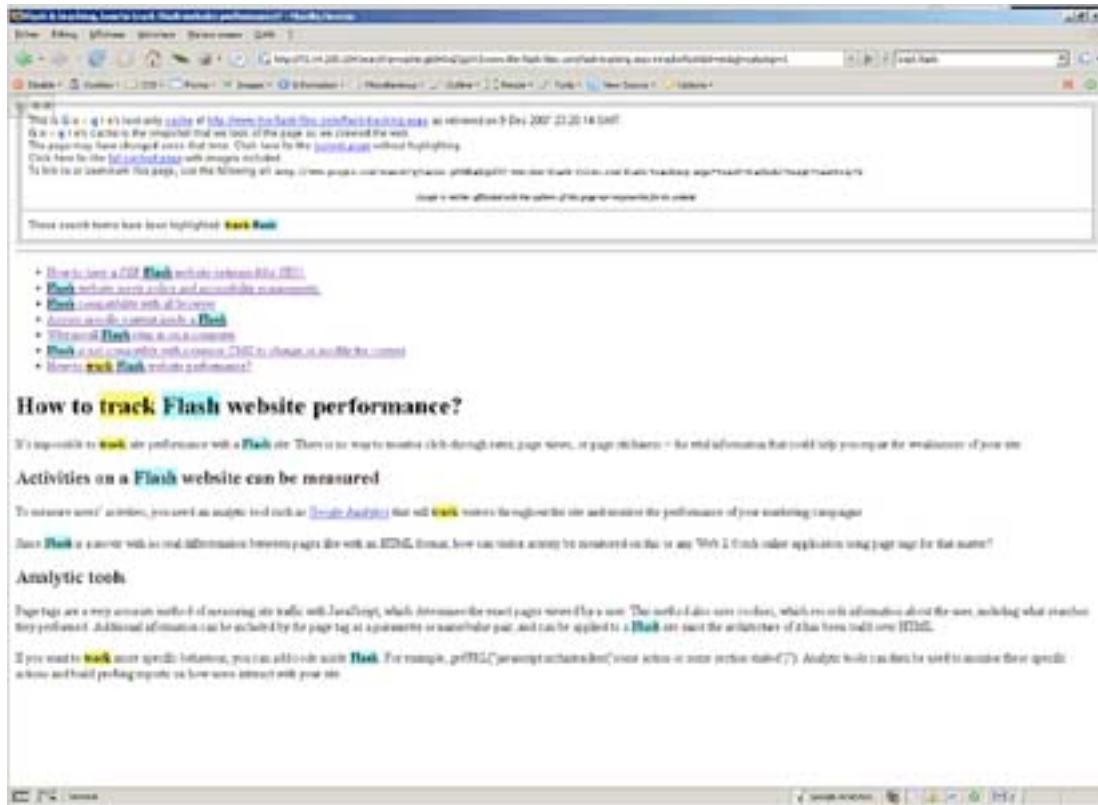
- Recherche sur "sef flash website" - 168,000 résultats
the-flash-files.com – 1er et 2ème résultat avec les URL
www.the-flash-files.com/flash-seo.aspx (1)
www.the-flash-files.com/flash-tracking.aspx (2)



Dans l'image ci-dessous on peut voir un exemple de la version du site qui est en cache chez Google.

Nous pouvons ainsi constater que la version texte en cache correspond bien au contenu de la page identifiée dans le "Flash".

<http://72.14.205.104/search?q=cache:g60HXaZQpXYJ:www.the-flash-files.com/flash-tracking.aspx+site:the-flash-files.com&hl=fr&client=firefox-a&gl=ca&strip=1>



Un Pagerank de qualité

Suite à l'application de cette technique, Google a attribué à la page d'accueil du site the-flash-files.com un Pagerank de 6. Ceci démontre donc qu'il accorde une importance aux liens qui pointent vers un site entièrement en Flash. D'ailleurs, l'ensemble des pages du site possèdent maintenant leur propre Pagerank.

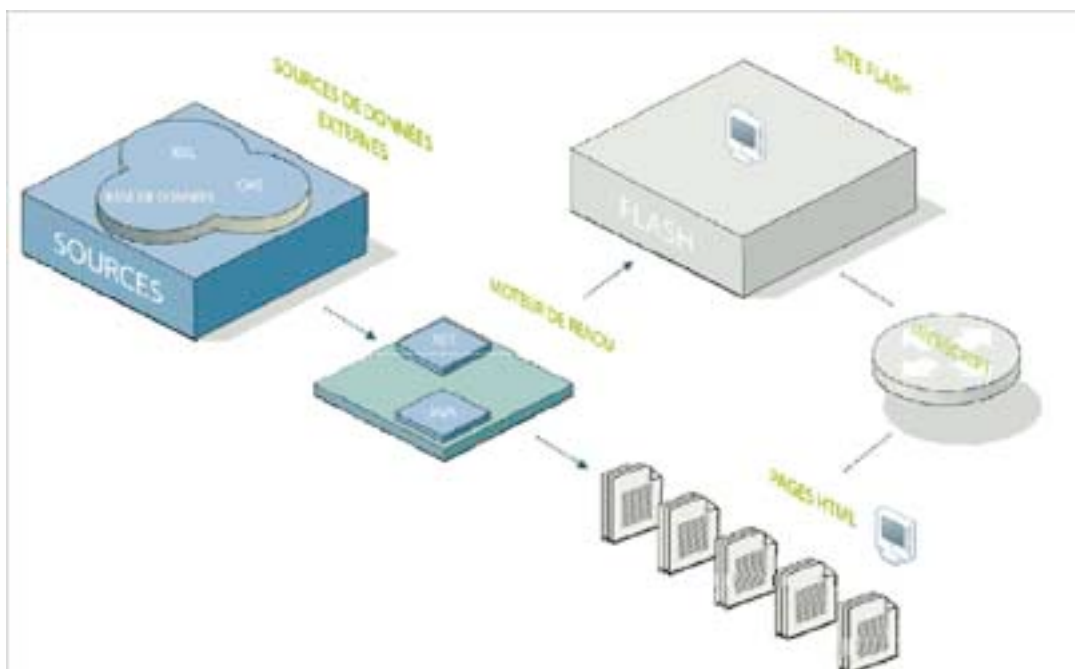
Nous sommes en mesure de dire que cela confirme un certain gage de qualité et de reconnaissance accordé par Google, et ce, malgré le fait que le site soit entièrement en Flash.

Analyse technique de la solution proposée

Contrairement aux pages satellites, la solution proposée ici n'attire pas l'utilisateur sur un contenu qu'il croit être le bon, pour finalement le rediriger vers un contenu alternatif. On propose à l'utilisateur et aux moteurs de recherche exactement le même contenu, mais en appliquant le principe de "dégradabilité", c'est-à-dire en superposant les couches, le "Flash" étant une couche de présentation ou d'habillage au même titre que les *Cascading StyleSheet* (CSS).



Pour ce faire, chacune des pages HTML, devra inclure la totalité du contenu texte dans un "DIV" et celui-ci pourra être remplacé par l'objet Flash si la détection du plug-in est concluante. Cette dernière sera effectuée à partir d'un JavaScript, plusieurs outils "open source" sont disponibles tel que SWFObject (<http://blog.deconcept.com/swfobject/>).



Une règle fondamentale qu'il est nécessaire de respecter afin de mettre en application cette solution, est d'avoir un contenu central qui sera distribué sur les différentes couches de présentation, tout en respectant méticuleusement le contenu. Ainsi, on assure l'intégrité du contenu présenté aux utilisateurs des moteurs de recherche, tout en s'assurant qu'il sera toujours exactement le même en tout temps, et ce, pour toute la durée de vie du site.

Comparaison des résultats

Nous pouvons maintenant comparer les résultats analysés au début de ce présent article (résultats non optimisés avec la technique présentée) avec ceux du site en "Flash" (optimisés avec la technique proposée):



On constate que nous avons maintenant dix (10) pages indexées avec autant de liens répertoriés. Ceci signifie donc qu'il existe dix (10) portes d'entrées distinctes, contenant dix (10) liens directs vers un contenu pertinent, même si conçu avec une structure HTML classique contenant la sémantique nécessaire à une bonne optimisation.

La zone grise

Afin de s'assurer que notre technique cadre bien avec les recommandations des principaux moteurs de recherche, nous nous sommes attardés à démontrer que nous respectons les principaux "guidelines"

- Créer une hiérarchie claire entre les pages et créer des liens de texte efficaces ;
- Tester et s'assurer du bon fonctionnement du site avec un "browser" texte ;
- Réaliser les pages pour les utilisateurs et non pour les moteurs de recherche ;
- Ne pas présenter un contenu différent de celui qui est destiné aux moteurs de recherche ;
- Éviter les textes ou les liens cachés ;
- Éviter les pages satellites "doorway" créées exclusivement pour les moteurs de recherche.

Malgré le fait que nous respectons toutes ces règles à la lettre et que nous n'entrons en contradiction avec aucune autre, nous sommes conscients qu'une zone grise persistera, tant qu'une charte de pratique "officielle" du référencement ne sera pas créée.

Un regard vers l'avenir

Il demeure primordial de comprendre dans notre démarche, que nous ne voulons pas remplacer les méthodes de référencement traditionnelles de sites HTML, riches en contenu. Nous désirons plutôt présenter une solution pour les sites qui offrent une expérience plus immersive grâce aux technologies multimédias et qui doivent eux aussi être répertoriés dans les moteurs de recherche.

L'arrivée massive sur le web de ces contenus multimédias, qui ne sont certainement pas moins pertinents, est avantageuse si les moteurs sont en mesure d'augmenter leur vision des résultats de recherches. La toile n'en sera que plus riche et la situation bénéficiera à tous.

Dans la deuxième partie de cet article, qui sera publiée dans le numéro d'avril 2008 de la lettre "Recherche et Référencement", nous nous attacherons plus particulièrement à expliciter notre méthodologie au travers d'un exemple pratique. Pour les plus impatients, vous pouvez explorer notre site (<http://www.the-flash-files.com/>) pour comprendre plus en profondeur comment tout cela fonctionne...

Nicolas Guillemot (*Spécialiste SEO Nurun Montréal*)
Frédéric Caron (*Chef d'équipe Flash Nurun Montréal*)

Réagissez à cet article sur le blog des abonnés d'Abondance :

<http://abonnes.abondance.com/blogpro/2008/02/optimisation-de-sites-web-en-flash.html>