

PHP et LeMonde.fr : passé, présent et avenir
Olivier Grange-Labat <grange@lemonde.fr>



Quelques mots pour présenter LeMonde.fr

Notre migration vers PHP

PHP au Monde.fr

Conclusion



Quelques mots pour présenter LeMonde.fr

Notre migration vers PHP

PHP au Monde.fr

Conclusion

Quelques mots pour présenter LeMonde.fr

Le Monde interactif :

- société créée en décembre 1998
- filiale du Monde (66%) et du groupe Lagardère (34%)
- aujourd'hui, environ cinquante salariés, parmi lesquels :
 - rédaction : central, newsletters, rendez-vous, autres (icono., etc.)
 - informatique : projets (7), production (3), informatique locale (3), études (1), autres (2)
- modèle économique : publicité, b2b et b2c (vente de contenu et abo.)
- LeMonde.fr est le premier site français d'information généraliste

Quelques mots pour présenter LeMonde.fr

Juin 1999 :

- site dynamique construit avec une technologie propriétaire
- infrastructure d'hébergement principalement Sun
- 12 millions de pages/vues par mois

Août 2002 :

- utilisation d'un CDN (“l'après 11 septembre”) pour dupliquer les pages HTML (en plus des images dupliquées elles depuis 2001) sur leurs milliers de “reverse proxies”

Quelques mots pour présenter LeMonde.fr

Avril 2002 :

- lancement de la zone abonnés (“des personnes payent pour se connecter sur nos serveurs”)
- 42 millions de pages/vues par mois

Octobre 2003 :

- mise en ligne de la version PHP de www.lemonde.fr
- frontaux Sun remplacés par frontaux Linux
- 44000 abonnés, 48 millions de pages/vues par mois

Quelques mots pour présenter LeMonde.fr

Mars 2005 :

- mise en production du SEPT (système éditorial de production et de trafic) développé en XUL/PHP
- nouvelle formule du site
- 61 000 abonnés, 67 millions de pages/vues par mois



Quelques mots pour présenter LeMonde.fr

Aujourd'hui :

- 100% de nos sites en PHP
- toute la gestion éditoriale en XUL/PHP
- ... mais la gestion technique et commerciale toujours sur la technologie propriétaire de 1999 (6 années après...)
- quelques serveurs non Linux
- 68 000 abonnés, 70 millions de pages/vues par mois
- record de consultation le 7 juillet 2005 (attentats de Londres et annonce des JO) : 5,6 millions de pages/vues et un pic de trafic à 300 Mbps !

En 2006, suite de la migration vers Linux, PHP et XUL.



Quelques mots pour présenter LeMonde.fr

Notre migration vers PHP

PHP au Monde.fr

Conclusion

Notre migration vers PHP

Les motivations (réflexion débutée courant 2002) :

- technologie propriétaire en fin de vie
- saut technologique pour sa nouvelle version (TCL vers Java)
- coûts de licence élevés (éditeur américain :))
- coûts d'infrastructure élevés :
 - si (frontal web == quadri-processeurs Sun)
 - et (1 quadri-processeurs Sun == \$\$\$\$)
 - alors (4 frontaux web = \$\$\$\$ \$\$\$\$ \$\$\$\$ \$\$) // en comptant la remise
- “faire mieux avec moins” (se replacer dans le contexte de l'après 2000)

Notre migration vers PHP

Etude menée en interne en décembre 2002 :

- n'était concerné que le front (sites web), pas la gestion de contenu
- “shortlist” : technologie propriétaire J2EE, Tomcat et PHP 4
- critères généraux (8) : année de création, pérennité, support, documentation, références...
- critères techniques (27) : plateformes supportées (Linux), respect des standards, environnements de développement, performance, stabilité, temps de développement (“time to market”)...
- critères financiers (6) : licences, infrastructure, formation...

PHP l'a emporté !

Notre migration vers PHP

Les avantages qui ont fait la différence en faveur de PHP :

- coûts faibles (licence, formation, prestation, infrastructure)
- conçu pour le web
- performances
- prise de connaissance rapide
- temps de développement réduits
- les références (FAI, Google, dans notre secteur le Times, El Pais...)

Notre migration vers PHP

Ce qui nous a fait douter :

- support XML
=> PHP 5, PECL, PEAR
- modèle de connectivité Oracle (pas de “connection pooling”)
=> SQL Load Balancer, FastCGI, bientôt Zend et PECL ?
- interfaces avec logiciels propriétaires
=> cela n'a pas été un problème ;-) et puis l'API Zend est (relativement) simple.

Notre migration vers PHP

Méthodologie pour la migration de technologie :

- maîtrise d'oeuvre et maîtrise d'ouvrage en interne
- mise en oeuvre d'un socle technique (“framework”) objet par une équipe
- puis écriture des classes et gabarits s'appuyant sur ce framework par une seconde équipe
- migrer un premier site moins complexe (aden.lemonde.fr), puis enfin le site d'information (www.lemonde.fr)
- migrer la gestion du contenu dans un second temps (les contenus dans la BD sont lus en PHP et modifiés par la technologie propriétaire)

Notre migration vers PHP

Pourquoi un framework :

- pour s'appuyer sur une base “saine” et éprouvée
- garantir une homogénéité de la conception logicielle sur chaque projet (principes de fonctionnement identiques)
- accélérer les temps de développement et réduire les coûts d'exploitation (réutilisation de code donc moins de temps passé à développer et à produire des bugs, “coding standards”)

Et cela a fonctionné ! Peu d'évolutions dessus depuis, mais...



Quelques mots pour présenter LeMonde.fr

Notre migration vers PHP

PHP au Monde.fr

Conclusion

PHP au Monde.fr

Première phase, migration des sites en 2003 :

1/ mise en place d'un environnement PHP :

- frontaux Linux, serveurs de dev., CVS, outil de livraison (Interflorae), “coding standards”, phpDocumentor, PHPUnit, mantis...

2/ mise en oeuvre de notre “framework”, Rosae V1 :

- 1 CdP, 2 architectes, 2 développeurs
- 200 j/h de développement

3/ migration des sites :

- 1 CdP, 4 développeurs
- 60 j/h pour le mini-site, 300 j/h pour le site infos

PHP au Monde.fr

Rosae V1 :

- utilisé pour générer toutes les pages de nos sites
- MVC : scripts et classes métiers (modèle), gabarits (vue), rosae.php (C)
- s'appuie sur Apache (mod_rewrite, mod_auth_lemonde)
- basé sur un système de cache de fichiers dans le "docroot" Apache : un flux marqué "statique" est généré une fois puis servi par Apache comme un fichier traditionnel. Décache == effacement du fichier.
- utilise les ESI ("Edge Side Includes", voir <http://www.esi.org/>)
- sait gérer différents sites, canaux de diffusion, versions, formats : notion de gURL.

PHP au Monde.fr

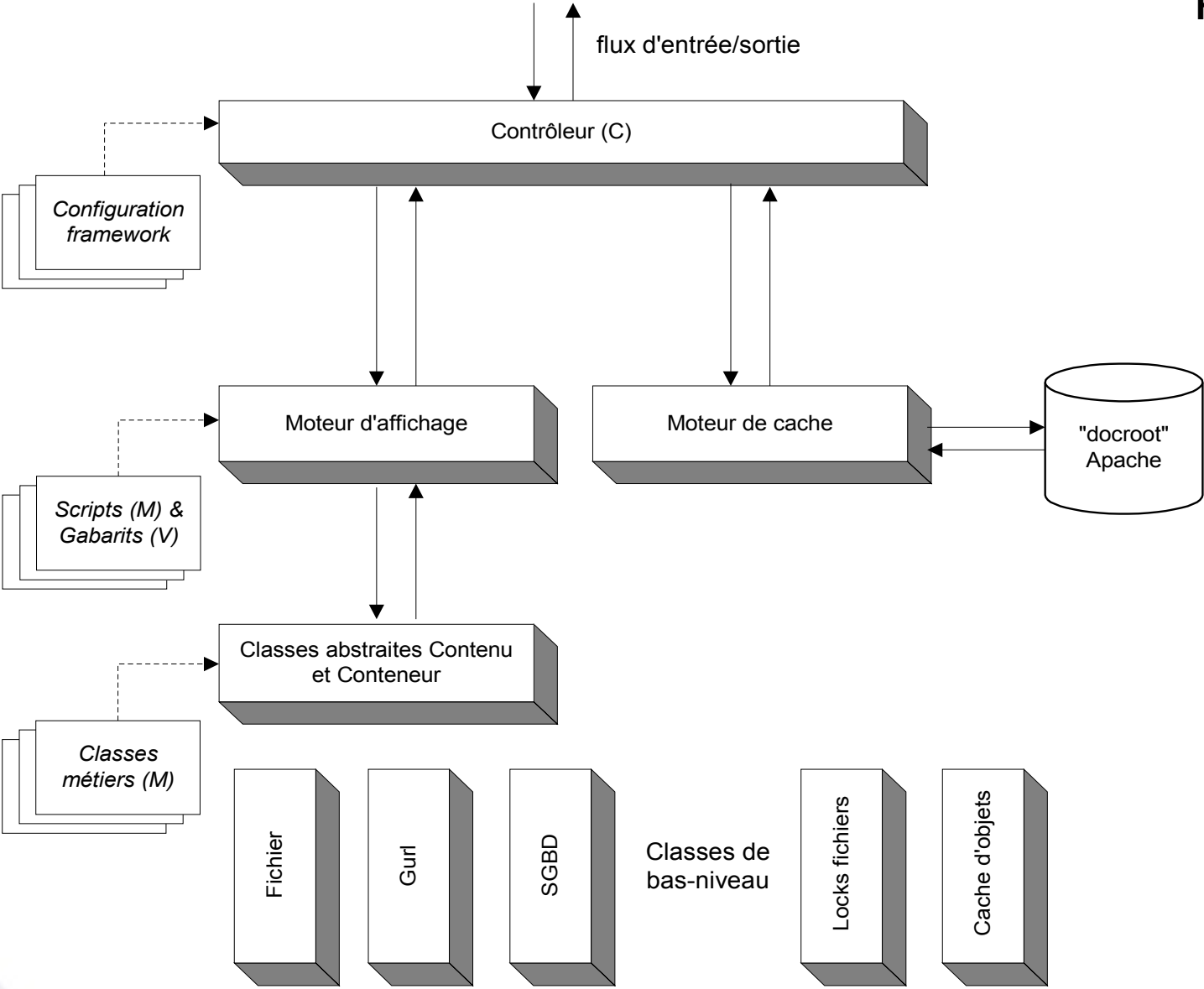
Les gURL :

- permettent de référencer un contenu dans tout le code
- sont “mappées” sur des URL publiques

par exemple :

- /info/pub/web/sequence/0,2-3208,1-0,0.html
<http://www.lemonde.fr/web/sequence/0,2-3208,1-0,0.html>
- /info/abo/web/sequence/0,2-3208,1-0,0.html
<http://abonnes.lemonde.fr/web/sequence/0,2-3208,1-0,0.html>

=> / site / version / canal / label / flag , type_conteneur1-id_conteneur1 ,
type_contenu1-id_contenu1 , id_profil . extension



PHP au Monde.fr

Deuxième phase, migration de la gestion du contenu, 2004-2005 :

1/ nouvelle étude :

- XUL côté client (application installée sur le poste local)
- PHP côté serveur (appel à des traitements en HTTP)

=> Le SEPT et Rosae V2

2/ évolution du framework : 1 CdP, 1 architecte, 2 développeurs (120 j/h)

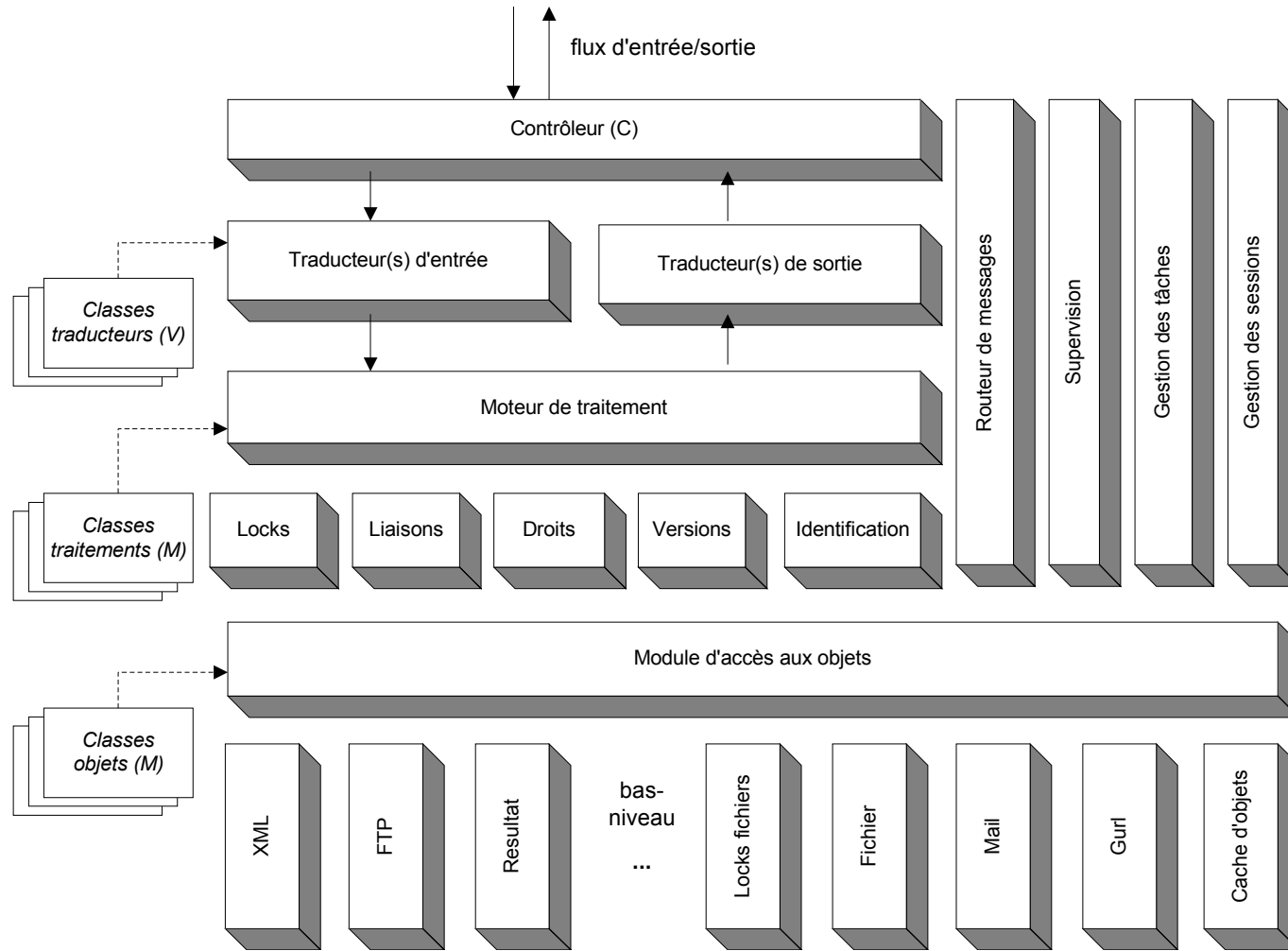
3/ développement du SEPT (PHP et XUL) : 1 DdP, 2 architectes, pratiquement toute l'informatique a participé (300 j/h) !

PHP au Monde.fr

Rosae V2 :

- ajout d'un moteur de traitement (modèle)
 - ajout de nombreux modules (Locks, Droits, Versions...)
 - ajout d'un module de persistance d'objets (“CRUD”)
 - ajout de classes de bas-niveau (Ftp, XML...)
 - notion de traducteurs d'entrée et de sortie (vues)
- => Tout est traitement (identification.loguer_utilisation, objet_editorial.lire_objet...). Une traduction en entrée et en sortie se charge des formats d'entrée et de sortie.

Rosae V2 (ajouts)

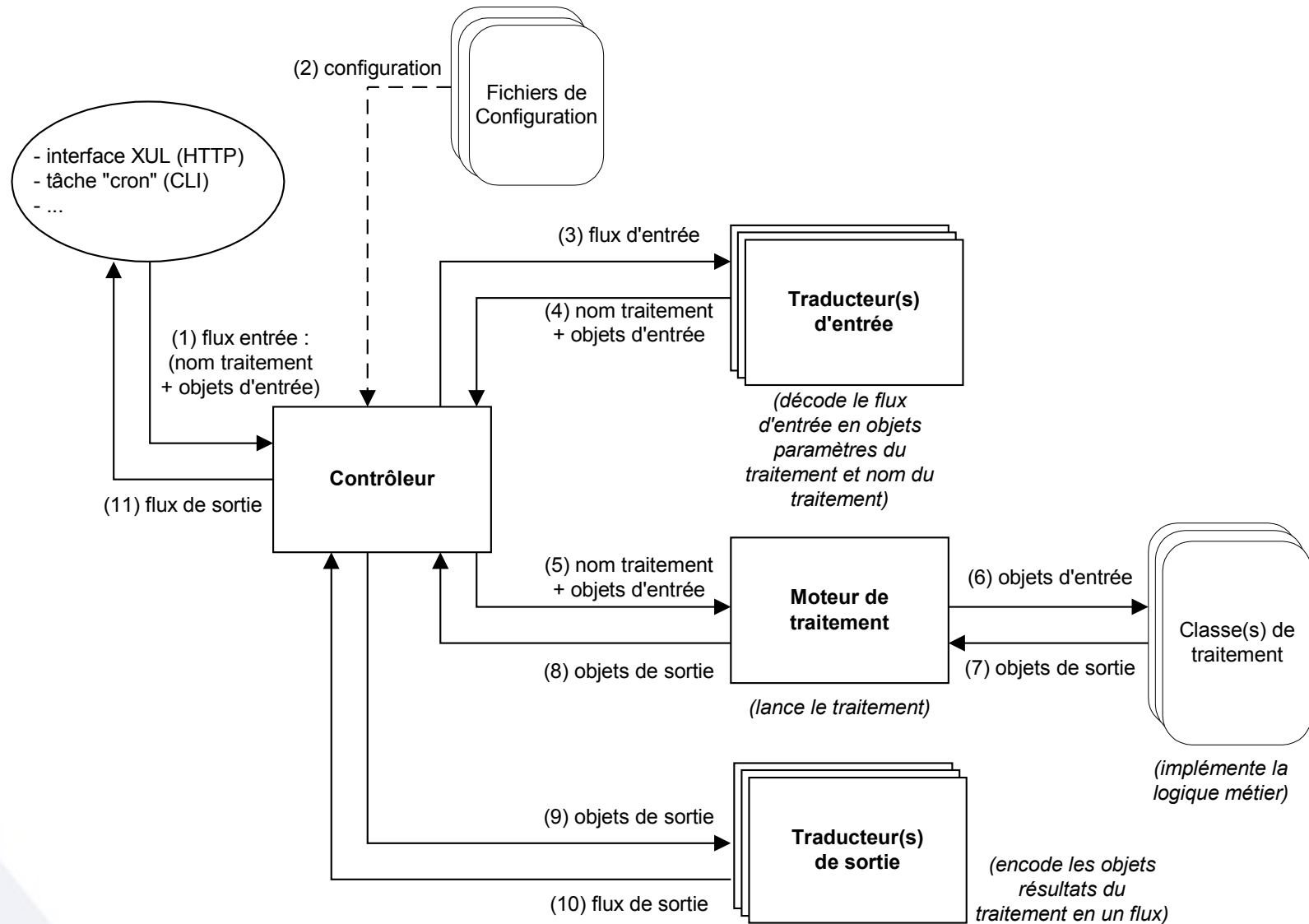


PHP au Monde.fr

Le SEPT



Exécution d'un traitement



PHP au Monde.fr

Quelques ordres de grandeur :

- Rosae : 32 000 lignes
- Gabarits (principalement HTML) : 31 000 lignes
- Scripts de présentation (principalement PHP) : 37 000 lignes
- Classes métiers : 117 000 lignes
- Le SEPT (code côté client) : 48 000 lignes
- Traitements : 24 000 lignes



Quelques mots pour présenter LeMonde.fr

Notre migration vers PHP

PHP au Monde.fr

Conclusion

Conclusion

Notre expérience de XUL

Les points forts :

- puissant (GUI ergonomique, multi-plateformes, API complète...)
- structurant (XUL/CSS pour la présentation, Javascript, bindings)
- la documentation est présente (attention FUD) : xulplanet, xulfr...
- les sources de Firefox et Thunderbird sont exploitables
- mise à jour gérées par Firefox (Le SEPT = deux extensions et la jslib)

Les points faibles :

- peu de « best practices »
- pas (encore) de CPAN/PEAR/JSAN pour XUL
- debug parfois laborieux (jslib, console Javascript, DOM Inspector)

Conclusion

Notre expérience de PHP

Les points forts :

les promesses sont tenues !

- c'est effectivement libre (licence, sources, php-dev...) :)
- 4 frontaux Sun (cpu 70%) remplacés par 4 frontaux Linux (cpu 20%)
- temps de développement divisé par 2, nombre de bugs par 3
- outils annexes (APD, eAccelerator, phpDocumentor, PHPUnit...)
- documentation de qualité

Les points faibles :

- parfois des crashes Apache

Conclusion

Le futur (2006) :

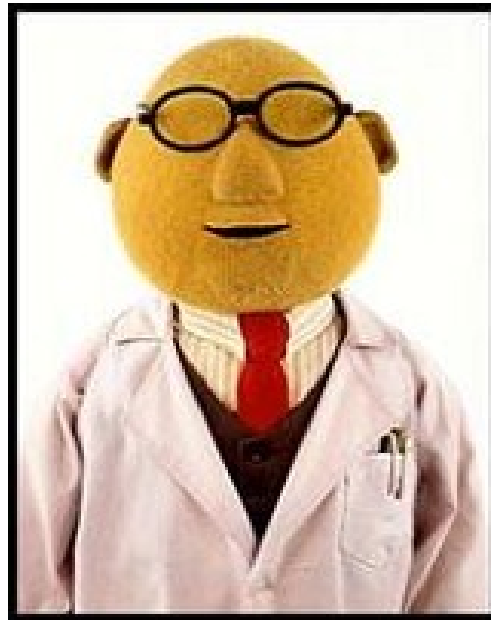
- suite de la migration vers XUL/PHP de la gestion du site
- PHP 5.1 (modèle objet, PDO, apache hooks...)
- remplacement de CVS (Subversion ?)
- partage du “docroot” entre les différents frontaux (GFS ?)
- dans le radar : PHP, Zend, eAccelerator, phing, lighttpd, memcache...

Réflexions :

- quels rôles dans le libre pour les entreprises utilisatrices ?

Questions ?





Merci !