

# Dossier : le Big Data va-t-il forcer le modèle L.A.M.P. à muter ?

Si l'on en croit, Bertrand Diard, un des entrepreneurs français qui a connu le plus de réussite ces cinq dernières années avec la création de Talend, « le 20ème siècle restera marqué comme le siècle de la mondialisation et du numérique, le 21ème sera celui de la donnée ». Mais l'explosion des données, ce phénomène baptisé «Big Data» est-il conciliable avec les actuelles infrastructures opensource ?

La priorité des sociétés actuelles – à commencer par les start-up, par définition innovantes, tournées vers l'avenir et au cœur des nouveaux usages – changent. Il ne s'agit plus simplement d'avoir une présence sur le Web au travers d'un site performant. Il faudra aussi de plus en plus être capable de résister à la vague de données qui s'annonce. Il faudra savoir surfer sur « l'augmentation de 650 % des données d'entreprise dans les prochaines années, dont 80 % seront non-structurées » prévue par Gartner pour en tirer un avantage concurrentiel.

Ce nouveau paradigme stratégique « du 21ème siècle » va-t-il emporter avec lui le modèle d'infrastructure jusqu'ici privilégié par de très nombreuses start-ups ? Possible. Mais pas sûr. Possible parce que Talend n'est en effet pas le seul à défendre cette idée d'une nouvelle ère. Tous les acteurs majeurs du secteur IT proposent leur solution maison pour séduire les jeunes pouces et leur faire choisir un modèle de serveur orienté Big Data plutôt que L.A.M.P.

Mais pas sûr, parce que la question se pose de savoir si les briques de LAMP ne pourraient pas elles-aussi s'adapter aux nouvelles exigences du Big Data. Et plus largement si le modèle lui-même ne pourrait pas évoluer (vers un L.A.N.P. par exemple : Linux, Apache, NoSQL, PHP/Python/Pearl). Autant de questions que Silicon a posé à différents acteurs – qu'ils soient de l'open-source ou non – dans ce dossier sur les conséquences que ce changement de paradigme annoncé pourrait avoir sur LAMP.

## A lire

- [I - Les Atouts d'une infrastructure L.A.M.P. \(Linux, Apache, MySQL, PHP\)](#)
- [II - LAMP à l'heure d'une nouvelle ère : vers un LANP ou vers un nouveau M ?](#)
- [Interview : Alvin Richards, 10gen : « Aujourd'hui le M de L.A.M.P., c'est MongoDB »](#)
- [Interview : Olivier Chouraki, MADGIC : « MySQL est rapidement devenu le goulot d'étranglement de notre architecture »](#)
- [Interview : Benoît Jacquemont, SMILE : « PHP est devenu un vrai langage pour l'entreprise »](#)
- [Interview : Pierre Couzy, Microsoft : « Windows Azure permet de déployer une stack LAMP sur IaaS »](#)

# I – Les Atouts d’une infrastructure L.A.M.P. (Linux, Apache, MySQL, PHP)

Avant de voir s’il a atteint ses limites, ou s’il est dépassé, il faut revenir sur les bases qui ont fait le succès de LAMP. A la fin des 90, le modèle a été une révolution dans le monde des serveurs. Il prouvait sans contestation possible que des solutions open-sources pouvaient être aussi complètes que les solutions propriétaires et les concurrencer. Depuis, cette révolution est devenue un grand classique. Un classique qui n’a pas stagné. Bien au contraire. Chacune de ses couches a évolué et s’est améliorée au fil du temps grâce à des communautés très actives. Et même si de nouvelles questions pèsent sur ce modèle – notamment l’avenir de MySQL – il conserve à bien des égards de très nombreux atouts.

Le premier est d’être simple. Composé de quatre briques – [Linux](#) pour le noyau du serveur, sur lequel se greffe le serveur Web [Apache](#), la base de données [MySQL](#), le tout complété par du [PHP/Python/Pearl](#) – la stack est robuste, éprouvée, simple à mettre en place et a montré son efficacité.

**Linux** est né en 1991 des travaux d’un étudiant devenu célèbre : Linus Torvald. En le faisant passer sous licence GNU/GPL en 1992, Torvald donne un nouvel élan à son projet qui attire de très nombreux contributeurs. D’une ébauche de noyau, Linux devient un OS qui offre tous les avantages de l’open-source : gratuit, ouvert, modifiable et adaptable, évolutif, soutenu par une communauté... autant de caractéristiques qui assureront son succès auprès des Webmestres.

Mais ce sont ses performances au sein de LAMP qui lui a permis de s’imposer face à UNIX, l’autre OS open-source. Un des éléments qui a su séduire est l’optimisation extrêmement poussée entre son Kernel et le serveur Web Apache. C’est par exemple au niveau du noyau que se gère la gestion de la concurrence (le dispatch des visiteurs en fonction de la charge de travail). Avec pour résultat des performances remarquables – voire inégalées.

**Apache** justement. D’après [Netcraft](#), plus d’un serveur HTTP sur deux (58 %) est un serveur Apache. Un succès qui s’explique pêle-mêle par sa modularité, les nombreux langages qu’il est capable d’interpréter (Perl, PHP, Python et Ruby) et surtout ses grandes possibilités de configuration.

Tout comme le Kernel Linux sont intimement liés, Apache et **PHP** forment un duo. Le mod\_php d’Apache est particulièrement rapide quand il s’agit de traiter du PHP (plus que PHP-CGI). Technologie phare sur le Web côté serveur, le langage PHP est à l’origine (en 1994) une bibliothèque logicielle en Perl créée par Rasmus Lerdorf pour enregistrer le nombre de visiteurs ayant consulté son CV sur son site web.

Là encore, c’est la facilité qui en a assuré le succès. Peu typé, souple, pragmatique, le langage est simple à prendre en main. Son installation ne pose pas de problème. L’accès aux bases de données bénéficie de nombreux modules et ses fonctionnalités ne cessent de s’enrichir. Presque vingt ans plus le langage fait partie des incontournables indétrônables du [Top 10 de l’Index TIOBE](#).

Ne restait plus qu’à trouver une base de données pour avoir une offre complète. PostgreSQL, plus ancien et lui aussi open-source, aurait pu remporter la partie. Mais c’est bien **MySQL** – créée en 1995

et dont le nom est un mélange de celui de la fille de son co-créateur Michael Widenius (My) et de SQL (Structured Query Language, son langage de requête) – qui est devenu le plus populaire.

Directement utilisable en production, avec des performances élevées en lecture, sans configuration particulière, multi-OS, multi-langage, [multi-thread](#) et multi-utilisateur, MySQL est aussi plus rapide sur les faibles volumes que PostgreSQL du fait qu'il n'effectue pas de tests d'intégrité. Depuis la version 3.23.15, MySQL permet également de gérer la réplication très facile à mettre en place (même si elle ne permet qu'une redondance limitée)... Seul défaut, bien que théoriquement non plafonné, MySQL devient moins performantes sur les gros volumes de données complexes, alors que PostgreSQL peut travailler avec des bases de plusieurs Teras. Mais la simplicité était du côté de MySQL.

Au final, et en plus des atouts liés à chacune de ses briques, ce mariage des quatre technologies au sein de LAMP est bon marché. Il ne demande pas d'achat de licences particulières (toutes les briques sont gratuites – hors frais de support éventuel) ce qui en fait une alternative à faible coût comparé aux outils propriétaires. Bref, en 1998 – année où pour la première fois le magazine allemand de référence C't utilise cet acronyme – tout était en place pour l'émergence de cet ensemble alternatif, modulaire et flexible.

Autre point crucial, la stack est très facile à essayer. A l'opposé, les solutions de traitement intensif de données propriétaire comme HANA – celles qui pourraient rendre LAMP caduque – ont un coût non négligeable. Et pas de version Freemium pour les tester. SAP le sait et propose depuis peu sa base in-memory en mode Cloud après une certification sur AWS pour réduire les coûts d'entrée (environ 4\$ de l'heure). Mais un inconvénient persiste par rapport à une particularité de l'open-source qui a fait le succès de LAMP : HANA requiert des spécificités matérielles particulières. Il n'est toujours pas possible de la télécharger et de la tester telle quelle.

A supposer qu'une solution orientée Big Data soit plus en adéquation avec les besoins futures des entreprises, celle-ci ne pourra remplacer LAMP dans de nombreux projets qu'en respectant ces deux conditions : universalité du support et simplicité de mise en place. Or une telle solution n'existe pas... Sauf à faire évoluer LAMP ?

A l'usage, LAMP s'est en effet imposé par sa légèreté. Une légèreté que la concurrence (Windows Server / IIS / SQL Server / ASP.NET ou dans un univers JAVA du JBOSS ou du Tomcat) a souvent qualifié de "limite". Mais au vu de la popularité actuelle de LAMP, celle-ci reste bien un atout.

Tout comme l'autre autre mot clef de LAMP : l'adaptabilité. Cette capacité à évoluer et à être « tout terrain » tient d'une part à la possibilité d'intégrer de nouvelles technologies dans le mix (PHP peut par exemple être remplacé par du Python ou du Perl et inversement) et d'autre part à des communautés très actives sur chaque brique.

Un simple exemple. LAMP est d'ores et déjà prêt à supporter les architectures ARM du simple fait que le noyau Linux l'est. Celles-ci ne sont pas encore très répandues dans le monde des serveurs mais la recherche continue de gains d'énergie dans les datacenters les rend de plus en plus populaires.

Les serveurs LAMP continuent d'ailleurs à s'adapter aux évolutions. Notamment au IAAS. Ou pour le dire de manière plus juste, la popularité actuelle du modèle a poussé les responsables des nouvelles offres Cloud à les rendre compatibles avec LAMP. Il est ainsi aujourd'hui parfaitement

possible de mettre la stack entière sur une machine virtuelle sur Windows Azure ou AWS pour bénéficier de la scalabilité de ses solutions ET des avantages de LAMP.

Reste que – malgré cette adaptation au Cloud – l'accélération de la volumétrie des données à traiter pourrait tout de même donner un coup de vieux aux serveurs LAMP, peu adaptés originellement au Big Data. Une crainte renforcée par la consomérisation de l'IT qui semble se faire aussi sentir dans les technologies pour serveurs. Les entreprises veulent de plus en plus aller vers ce qu'il y a de plus simple avec des solutions « out of the box » qui n'ont plus besoins d'être déployées (comme SAP HANA One ou l'alliage HADOOP-SQL Azure sur Windows Azure), et dont les parties techniques et de maintenances sont externalisées.

Bref, « un clef en main » à l'opposé de l'esprit de LAMP qui est – pour le meilleur et pour le moins bon – le représentant parfait du « faites le vous-mêmes » pour garder la maîtrise de la technologie, de ses données et de la totalité de sa stratégie IT. Des atouts qui poussent de nombreux utilisateurs de solutions open-source à appeler de leur vœux un « LAMP-like » pour le Big Data : simple, robuste, ouvert, universel et standardisé.

## **II – LAMP à l'heure d'une nouvelle ère : vers un L.A.N.P. ou vers un nouveau M ?**

Dans son article «De Nouveaux Outils pour une Nouvelle ère », Ravi Kalakota, consultant et auteur de l'ouvrage [E-Business 2.0: Roadmap for Success](#), [constatait qu'](#)« avec le Big Data, c'est la première fois en 30 ans que nous repensons les bases de données et la manière de gérer les informations ». Une ère sous le signe de Hadoop, NoSQL, BigTable et autres Appliances comme Oracle Exadata Database Machine ou SAP HANA qui pourraient rendre caduques des outils qui ne savent pas gérer ces problématiques.

Les commerciaux de SAP, Amazon, Microsoft ou Oracle ne se privent d'ailleurs pas pour faire remarquer aux Startups qu'elles auront de gros problèmes pour gérer leurs serveurs lors de la montée en charge si elles commencent sur un modèle LAMP. On l'aura compris, si ces prévisions se confirment et que le Big Data d'aujourd'hui devient le Normal Data de demain, le gros point faible de LAMP se trouvera au niveau sa brique SGBD.

L'air du temps va-t-il rendre MySQL caduque ? En tout cas Oracle a changé le discours marketing et ré-orienté de plus en plus la base de données vers les seules usages webs, laissant le marché des entreprises et des applications critiques à Oracle Database.

Si on élargit la réflexion sur cette brique, on peut également se demander si MySQL a un avenir tout court avec la défiance du monde open-source envers Oracle. Pas de roadmap claire sur le long terme. Augmentation des tarifs de support. L'arrivée de nouvelles alternatives communautaires comme MariaDB (fork du créateur de MySQL) avec de nouveaux supports comme SkySQL.

Sans oublier le fait que MySQL deviendrait moins ouvert. C'est ce qu'affirme en tout cas Sergei Globuchik, Vice Président de MariaDB, en constatant la disparition de tests unitaires et d'un certain

nombre d'historiques de révision qui permettraient aux implémentations alternatives de garder un maximum de compatibilité avec la base originale. Une évolution qui pourrait rapidement pousser LAMP à changer son M pour le P de PostgreSQL Ou pour le N de NoSQL) pour devenir le meilleur « LAMP-like » capable de traiter ces nouvelles problématiques du Big Data.

En fait, comme le rappelle Ravi Kalakota, une solution open-source existe déjà pour cette « nouvelle ère » qu'il décrit. Elle s'appelle MongoDB. « MongoDB est devenu la base de donnée leader dans le NoSQL avec plus de 100.000 téléchargement par mois ». Et cerise sur le gâteau, elle commence par un M.

Son installation au côté de Apache, de PHP/Pearl/Python (pour lesquels MongoDB est livré avec des pilotes, même si son langage natif est le JavaScript) et d'une distribution comme Ubuntu Server [reste très simple](#). Autre avantage, MongoDB possède [un connecteur](#) pour une des technologies phares les plus à même de s'imposer comme un standard dans le Big Data : HADOOP. Résultat, des voix s'élèvent pour demander son intégration pour redonner un coup de fouet au modèle LAMP.

C'est une piste intéressante que défend Guy Harrison, vice-président de la R&D de l'unité Database Management chez Quest Software (Dell). Pour lui, « MySQL n'a pas fait le succès de LAMP parce qu'il était plus rapide ou plus sûr que les alternatives d'Oracle, mais parce qu'il était plus simple à déployer, à administrer et moins coûteux. [...] De la même manière, le choix de MongoDB oblige à des compromis quand on le compare à la concurrence. Mais il est économique et developer-friendly ». Et de conclure que pour lui MongoDB est à la fois « le MySQL du monde NoSQL » et « un nouveau M potentiel pour LAMP ».

Malheureusement, tout ceci n'est qu'une éventualité. Il n'y a pas encore à proprement parler de stack « LAMP-like » pour le Big Data et l'introduction de MongoDB n'est pas à l'ordre du jour. Conséquence, la concurrence se positionne sans attendre cette mutation de LAMP. Car les éditeurs de solutions propriétaires, s'ils ne donnent pas dans l'open-source, ont en partie retenue la leçon que leur a infligé la stack. Leurs solutions ne sont pas facilement testables ? Qu'à cela ne tienne. Le Cloud permet d'essayer in-situ.

Des briques open-sources sont plus performantes que les briques maisons ou les clients ne veulent pas tout acheter chez un seul fournisseur ? Pas de problème... L'ouverture est devenue une règle commerciale parfaitement intégrée. A tel point que des observateurs de l'open-source comme Dan Woods, journaliste pour Forbes et auteur d'une vingtaine de livres dont « Open Source for the Enterprise », prédisent que le (ou les) prochain(s) modèle(s) de référence pourrai(en)t être hybride(s).

Reste une question sur le Big Data lui-même. S'il y a bien une leçon que LAMP a donné, c'est celle que pour réussir, il faut faire simple. Or aujourd'hui, aucune solution de traitement de données intensives n'est vraiment pratique. En tout cas aussi pratique qu'un LAMP ou suffisamment pour s'imposer jusque dans les PME. Hadoop par exemple n'est pas une usine à gaz. Mais pour certains ce n'est pas loin d'être le cas.

En résumé, sur le long terme, c'est la capacité de LAMP à évoluer et à répondre aussi à la combinaison des 3 V – volume, vitesse, variété – qui fera la pérennité du modèle. Mais c'est bien la simplicité que cherche les entreprises. Et aucune autre stack ne « ringardisera » LAMP sur ce point.

« C'est une opportunité pour des startups comme Cloudera qui cherchent à faire grandir l'écosystème d'Apache Hadoop en rendant son usage plus simple », note Ravi Kalakota. Pour mémoire, Cloudera est une distribution d'Hadoop (un outil graphique desktop pour gérer Hadoop).

En attendant, le Big Data n'est pas à la portée de tous. Et cette situation est aussi une chance pour LAMP. Car, entre les solutions open-source pas encore assez « user-friendly » et des solutions propriétaires payantes – et encore coûteuse pour des PME – le modèle pourrait bien rester le choix de prédilection des entreprises, notamment de celles qui débutent.

## **Conclusion :**

LAMP dépassé pour les TPE et petites PME qui ne veulent pas s'occuper de leurs infrastructures IT ? Oui. Par le Cloud et les outils clefs en main. La consomérisation de l'IT – poussé par les grands acteurs du marché- comme Microsoft qui pour la première fois propose de tester Windows Server sans serveur, directement dans Azure – peuvent mettre à mal un modèle simple, mais à déployer soi-même.

LAMP pas scalable pour le Big Data ? Pour l'instant non (il faut dire qu'à la base, il n'a pas été conçu pour cela). Mais avec une nouvelle brique comme MongoDB (ce qui répondrait au passage aussi bien aux problématiques Big Data que e-commerce), l'histoire pourrait connaître un rebondissement.

Mais en attendant l'évènement de cette nouvelle ère tant annoncée, LAMP brille toujours dans le royaume des serveurs. Et pour briller également dans le Cloud, virtualisé. C'est en tout cas le second souffle qu'on lui souhaite pour ses usages traditionnels.

## **A lire**

- [Interview : Alvin Richards, 10gen : « Aujourd'hui le M de L.A.M.P., c'est MongoDB »](#)
- [Interview : Olivier Chouraki, MADGIC : « MySQL est rapidement devenu le goulot d'étranglement de notre architecture »](#)
- [Interview : Benoît Jacquemont, SMILE : « PHP est devenu un vrai langage pour l'entreprise »](#)
- [Interview : Pierre Couzy, Microsoft : « Windows Azure permet de déployer une stack LAMP sur IaaS »](#)