

InriaGforge : leçons de 2 ans d'exploitation de GForge à l'INRIA

David Margery, Janet Bertot,
Christophe Demarey et Claude Inglebert

21 novembre 2007



INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE



Un peu d'histoire

Novembre 2004 : décision de démarrage d'un service de forge

Juin 2005 : le site est ouvert

Juin 2006 : migration de gforge 4.0 à gforge 4.5

Été 2007 : stats pour l'article Jres2007

Octobre 2007 : l'équipe InriaGforge fait le bilan

Attention : cette présentation reprend les chiffres présentés dans l'article, sauf mention contraire

Quelles leçons ?

C'est quoi une forge, pourquoi en déployer une et laquelle ?

Que faut-il mettre en place pour supporter 3403* utilisateurs et 949* projets ?

Qui sont les utilisateurs d'une forge et que font-ils avec ?

* chiffres de début novembre 2007

C'est quoi une forge, pourquoi en déployer une et laquelle ?

Définition d'un service de forge

Rassembler en un seul lieu des outils pour le développement collaboratif

- gestionnaire de versions
- gestionnaire de listes de diffusion
- hébergement web des pages de projet
- gestionnaire de tâches et de bugs
- gestionnaire de documents
- gestionnaire de releases
- wiki

Les exemples les plus connus :

SourceForge, Savannah, Google Code

Motivations pour un service de forge à l'INRIA

Favoriser les bonnes pratiques de développement

- dans un environnement qui peut évoluer avec la collaboration
- dans un contexte plus large que celui des projets libres
 - projets privés et publics

Support aux collaborations scientifiques

- internes à l'Inria
- avec des partenaires extérieurs
 - base de comptes ouverte

Aider à la diffusion des résultats d'un développement collaboratif

Objectifs du service

Offrir une grande facilité de collaboration

- Le processus de création de compte et de projet doit être simple

Assurer la confidentialité des données des projets hébergés

- En particulier pour les projets privés

Assurer la disponibilité du service

- Doit être utilisé au quotidien, dans plusieurs fuseaux horaires
- Ne pas dépasser
 - une heure d'interruption par mois dans les heures et jours ouvrés en France (pas d'astreintes)
 - une journée d'arrêt par an

Héberger une forge, et non en développer une

- Utiliser un logiciel éprouvé : GForge

Pourquoi GForge ?

Pour un démarrage rapide

- Un minimum de changements pour les projets
- Projets sous CVS
- Passerelle pour une migration depuis bugzilla a existé

Un logiciel ayant déjà fait ses preuves

Un logiciel libre

- Permet de faire des adaptations de son code au contexte d'utilisation

Une version adaptée pour nos besoins

Adaptation des scripts Gforge à notre installation

- projets privés dès leur création
- sécurisation des données

Personnalisation des pages d'accès au dépôt

- permettre aux projets de donner quelques informations supplémentaires
- 14 projets utilisent cette fonctionnalité

Mise en place d'accès webdav aux dépôts subversion

- 408 utilisateurs entre le 7/08/2007 et le 24/10/2007

Navigation des dépôts par les membres authentifiés

Adaptation et évolution du logiciel

C'est un logiciel libre

- les adaptations sont faciles

Quid de sa communauté ?

- intégrer nos adaptations
- maintenir sa branche

Expérience du passage à GForge 4.5

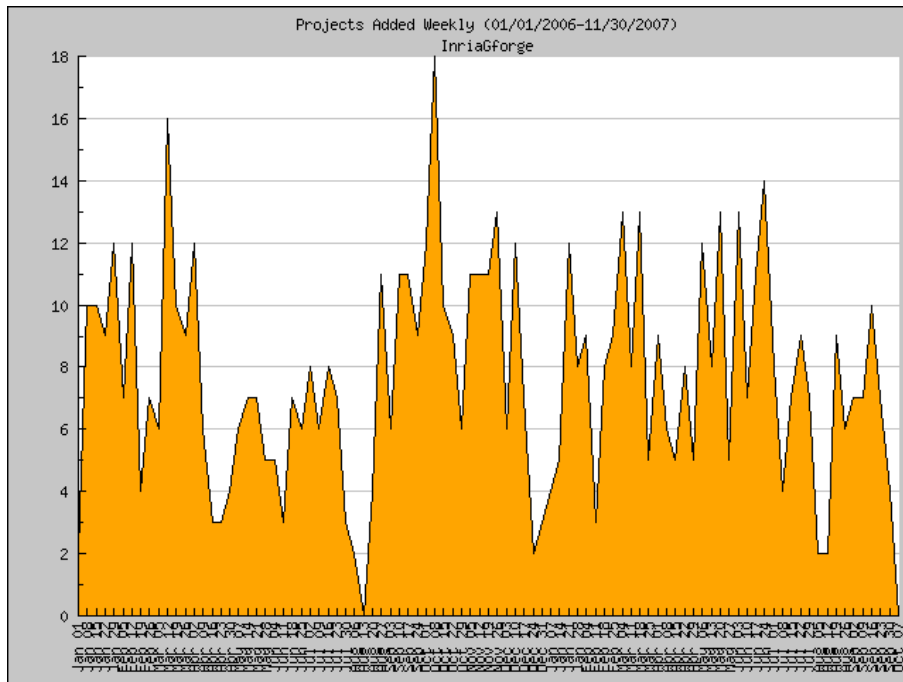
- même les petites adaptations ou corrections se révèlent coûteuses
- limiter les adaptations locales s'avère nécessaire

Et s'il n'y avait pas de Gforge 4.6 ?

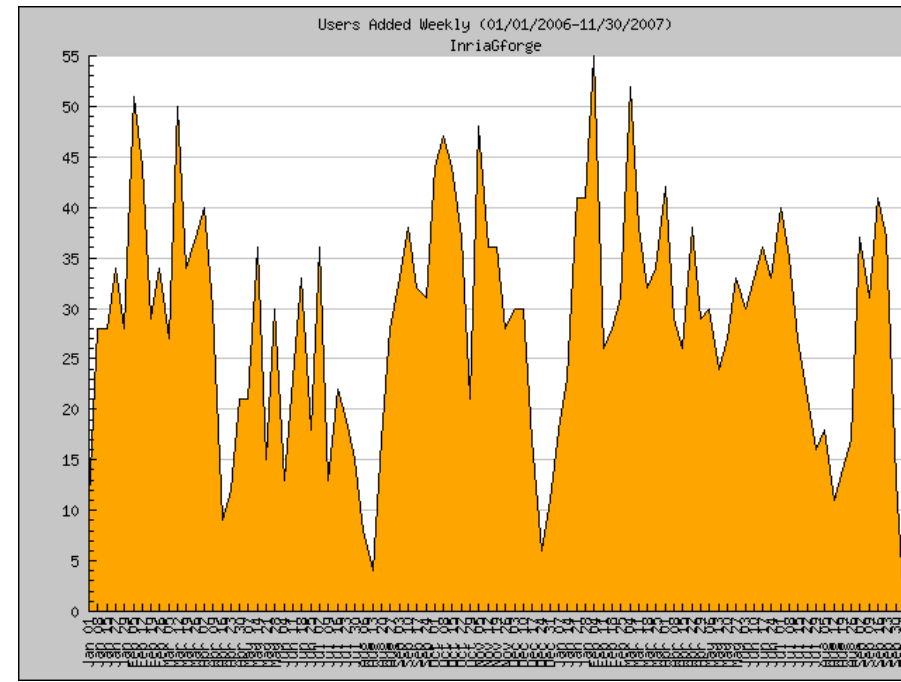
- participer à la communauté ou favoriser son émergence semble indispensable

Les résultats

C'est un service qui a trouvé son public:



Nouveaux projets par semaine



Nouveaux utilisateurs par semaine

Que faut-il mettre en place pour supporter
3403 utilisateurs et 949 projets ?

Infrastructure mise en place

Un logiciel

- GForge, version gforge.org
- Avec quelques adaptations locales

Une équipe transversale

- Plutôt un correspondant par site de l'INRIA à temps partiel qu'une ou deux personnes dédiées
- Avec ses procédures et ses processus pour partager l'administration

3 serveurs

- Pour distinguer développement, qualification et production
- Où cohabitent processus de mise à jour de sécurité et processus de déploiement de nouvelles versions de gforge

L'équipe d'administration

Equipe transversale à temps partiel

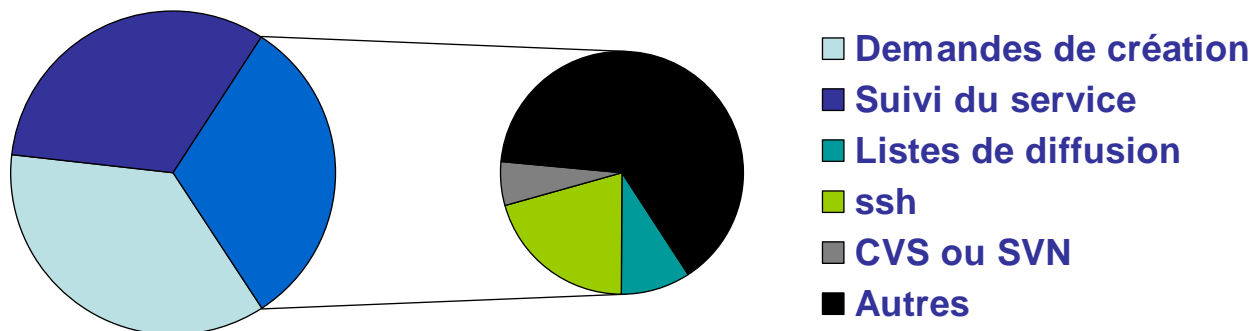
- 1 responsable opérationnel (25% d'un ingénieur à temps plein)
- 2 correspondants des directions fonctionnelles (10%)
- 5 membres (total : 25%)
- Présents sur 5 sites

Les processus de l'équipe

- Gestionnaire de tickets (RT) hors forge
- Un wiki hors forge décrivant
 - La procédure pour chaque alerte de supervision
 - La procédure pour les demandes les plus courantes des utilisateurs
 - La documentation de l'installation du serveur
- Des réunions mensuelles par visioconférence
- Une phase de validation des processus par simulation d'incidents

Le support

Sur le gestionnaire de tickets (2548 tickets au total)



- Une moyenne de plus de 3 approbations de projets par jour ouvré

Une FAQ et un document introductif

- Beaucoup de demandes initialement sur l'intégration avec Eclipse

Serveurs

La fiabilité du serveur principal suppose 2 autres serveurs

- 1 serveur de développement (devel)
- 1 serveur de qualification (qualif)
 - Identique matériellement au serveur principal
 - Pour valider les déploiements

L'installation de chaque serveur peut-être découpé en trois

- La distribution installée
- Les fichiers de configuration propres à chaque serveur
 - soit versionnés
 - soit générés depuis gforge.conf
- Le logiciel gforge, installé hors distribution,
 - Sur devel, le logiciel est un checkout de notre base de code

Sécurisation du service et des données

Par utilisation d'environnements restreints (chroot)

- Restreindre les vecteurs d'attaques
- Limiter la vision du serveur par les utilisateurs

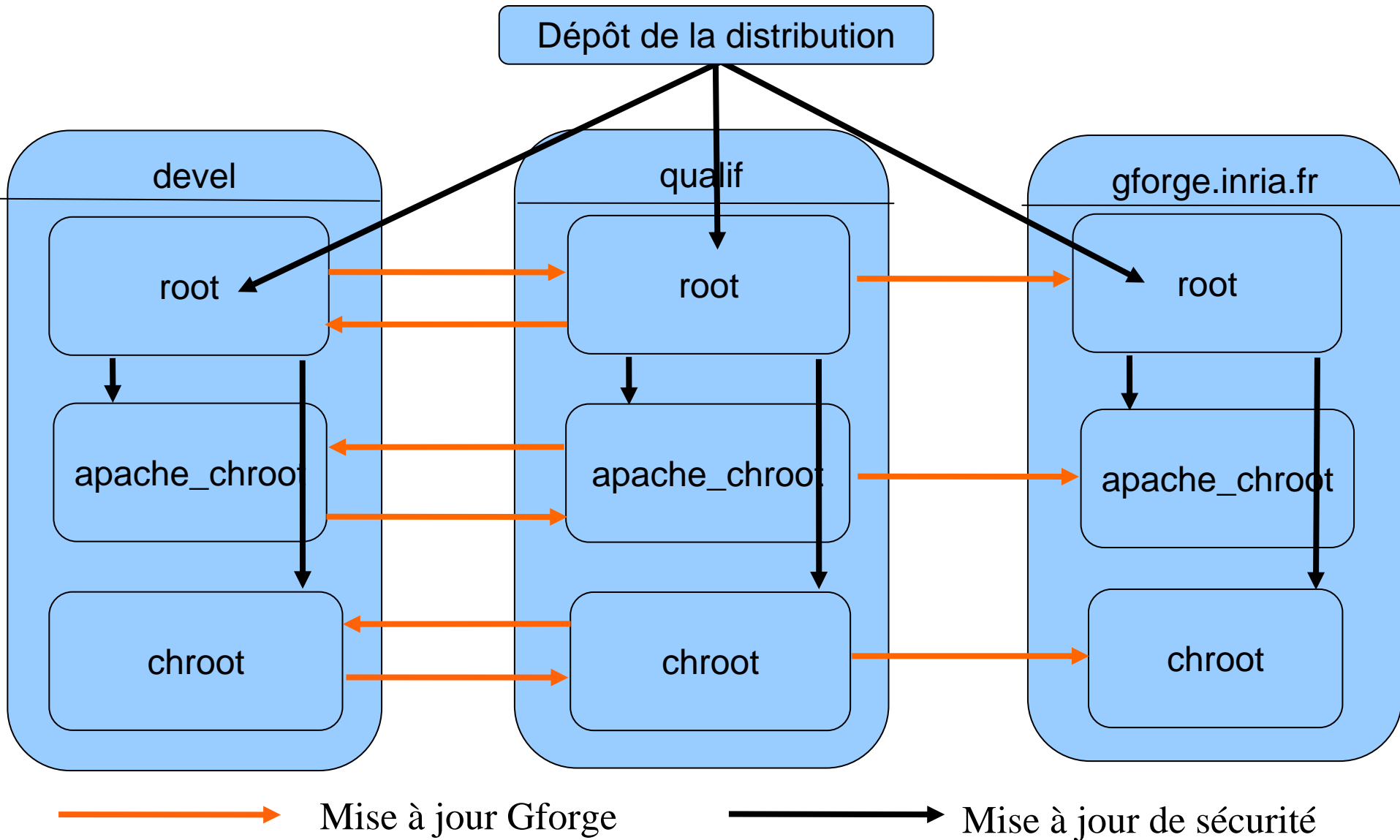
Par confinement avec suPHP

- Confiner les risques en cas de faille d'un site web d'un projet

Par test d'intégrité du serveur

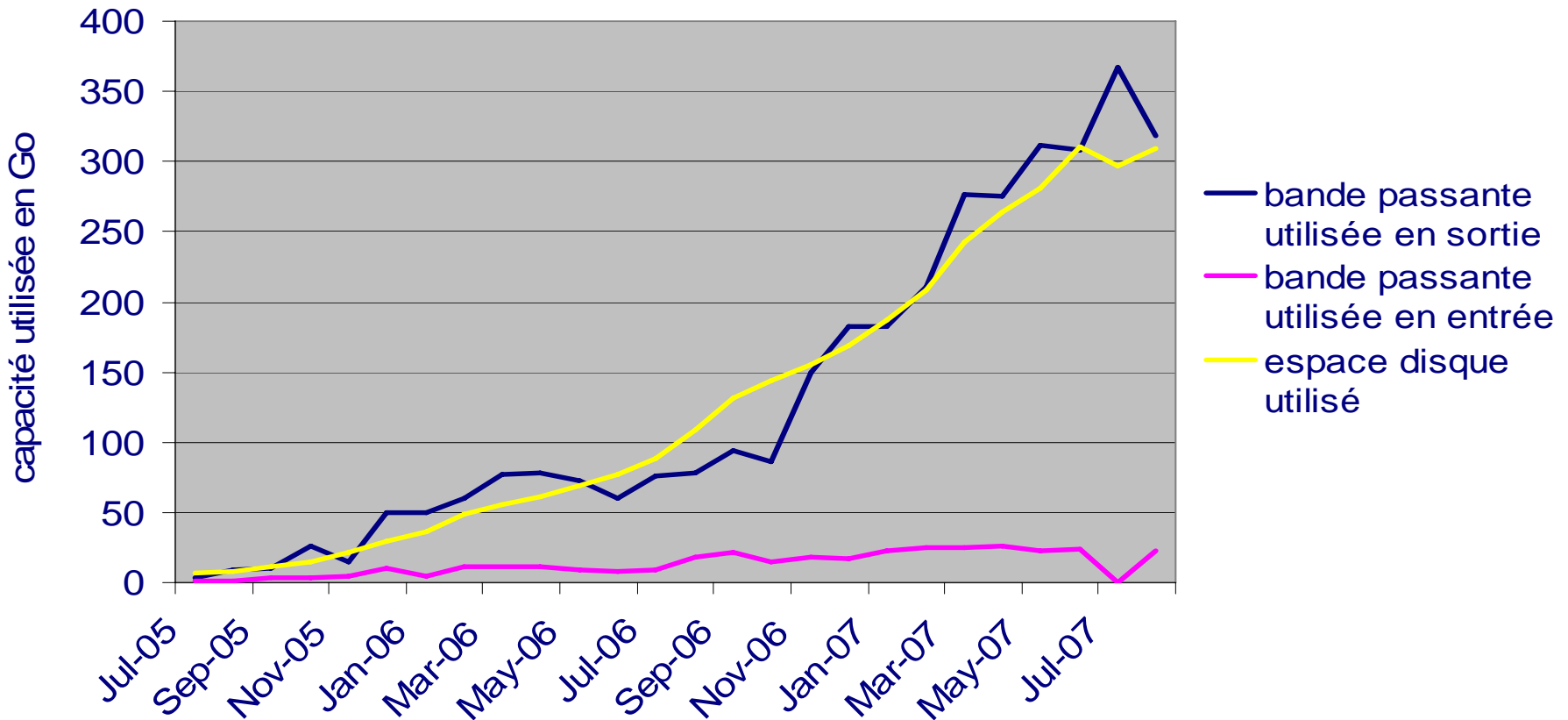
- Surveillance des binaires
- Surveillance des espaces disques utilisés par les projets

Relations et processus de mise à jour



Volumétrie du service

Evolution de la bande passante et de l'espace disque



Problèmes rencontrés

Installation d'une Debian stable (3.0) sur un DELL PowerEdge 2850

- Le noyau par défaut n'était pas suffisant

Un processus ne peut appartenir qu'à 32 groupes

- La distribution est construite sur les interfaces stables de noyau
- C'est un problème pour supporter un grand nombre de groupes privés

OpenLdap

- Configuré de manière non-optimale par défaut
 - Pas d'index
 - Taille de page de résultats trop petite
- Quelques plantages intempestifs depuis 3 mois

Les résultats

Ce qui est en place :

- 3 serveurs pour pouvoir tout tester avant la mise en production
- Une équipe avec ses procédures documentées

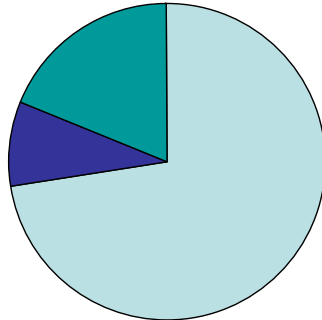
Objectifs atteints :

- Taux de disponibilité du service compatible avec les besoins du public visé
- Pas de maintenance de serveurs dédiés par les équipes scientifiques

Qui sont les utilisateurs d'une forge et que font-ils avec ?

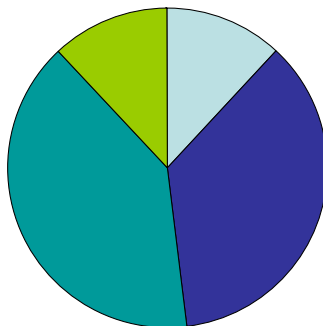
Profil des projets

Les projets ont des buts variés, d'après leur description



- Développement scientifique (633)
- Développement non-scientifique (76)
- Ecriture de documents scientifiques (164)

Taille des projets

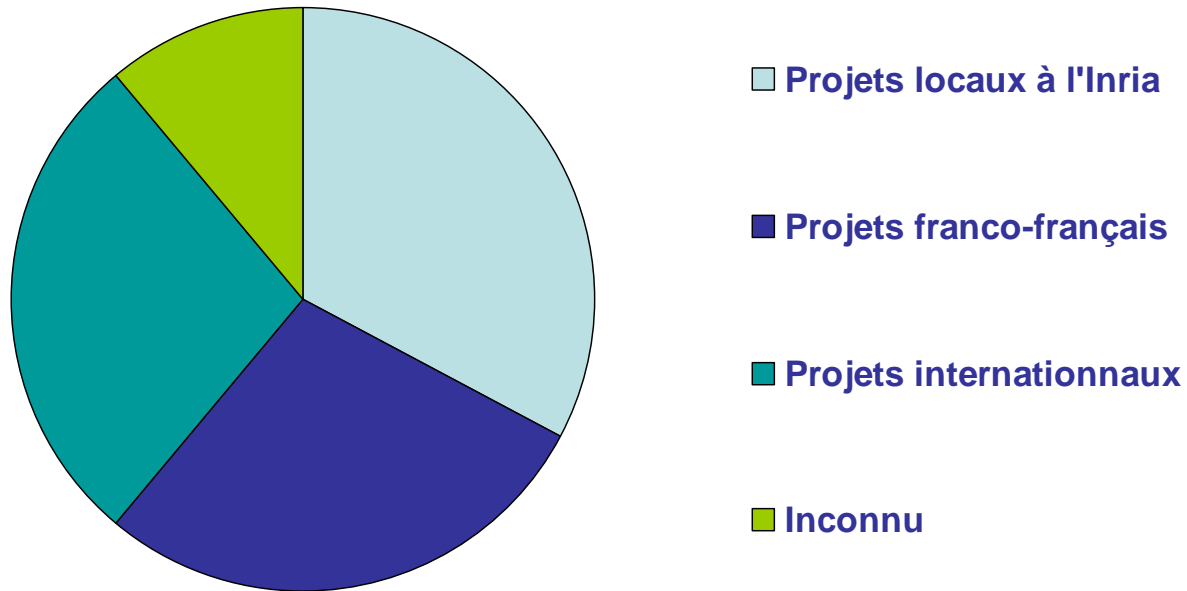


- 1 membre
- 2-3 membres
- 4-10 membres
- plus de 10 membres (112 projets)

— 62 projets ont plus de 10 contributeurs à leur dépôt au 31/10

Origine des membres des projets

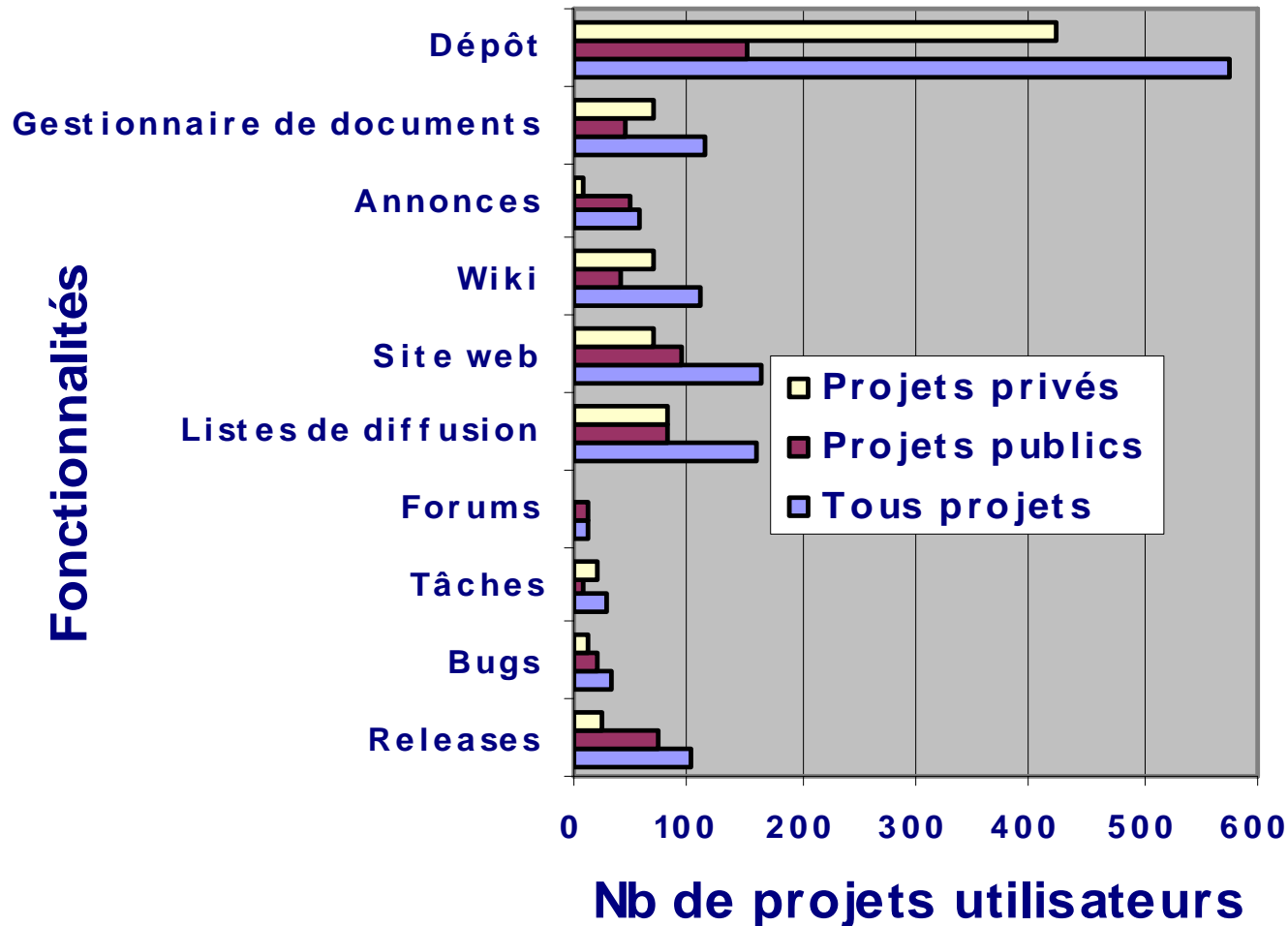
Par étude de leur adresse mail



On peut estimer que 2/3 projets impliquent des personnes non-INRIA

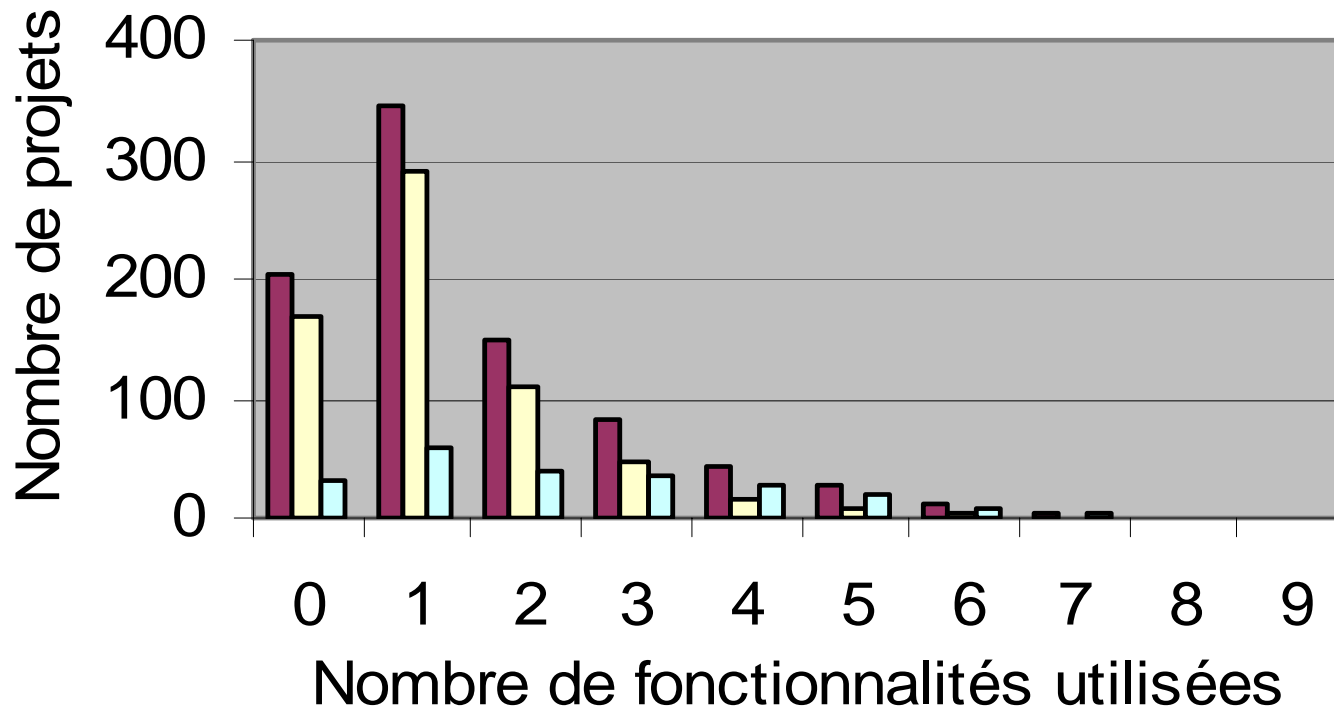
Outils de collaboration utilisés

Etude depuis début 2007



Nombre de fonctionnalités utilisées

■ Tous projets ■ Projets privés ■ Projets publics



Services de plus haut niveau

Une forge ne fait pas tout :

- pas de build ou de tests automatiques

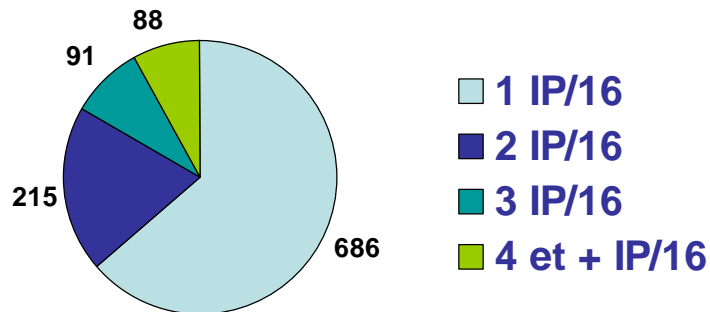
Une forge doit donc être ouverte

- Flux rss pour exporter les nouvelles sur le site web
 - 1103 IP clients
- API SOAP pour attaquer la base de données
 - 1 client
- Export d'informations générales par un mécanisme ad-hoc
 - 25 projets utilisateurs

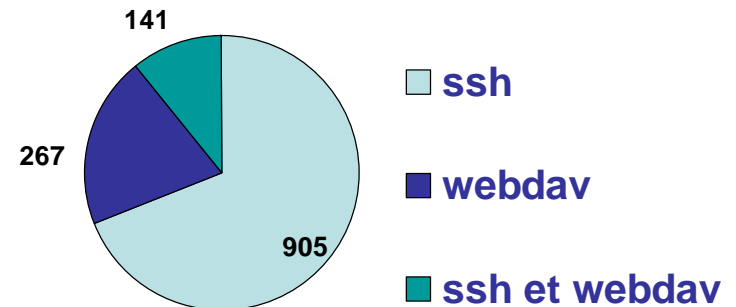
Mobilité des utilisateurs

Nous avons constaté une grande mobilité des utilisateurs

Source des connexions aux dépôts entre le 18/7/07 et le 9/9/07



Mode de connexion aux dépôts entre le 18/07/07 et le 28/10/2007



Pas les moyens de communication associés à cette mobilité

- Messagerie Instantanée
- Voix sur IP

Conclusion

Un service de forge

- répond à un réel besoin
- permet de satisfaire
 - des projets publics comme des projets privés
 - de petits comme de gros projets
- est utilisé
 - aujourd'hui pour les fonctionnalités les plus simples
 - demain ... ?
- libère les scientifiques de la gestion de leur propre serveur pour leurs collaborations via une forge INRIA & une équipe partagées

Un service de forge s'adapte à la vie d'un logiciel