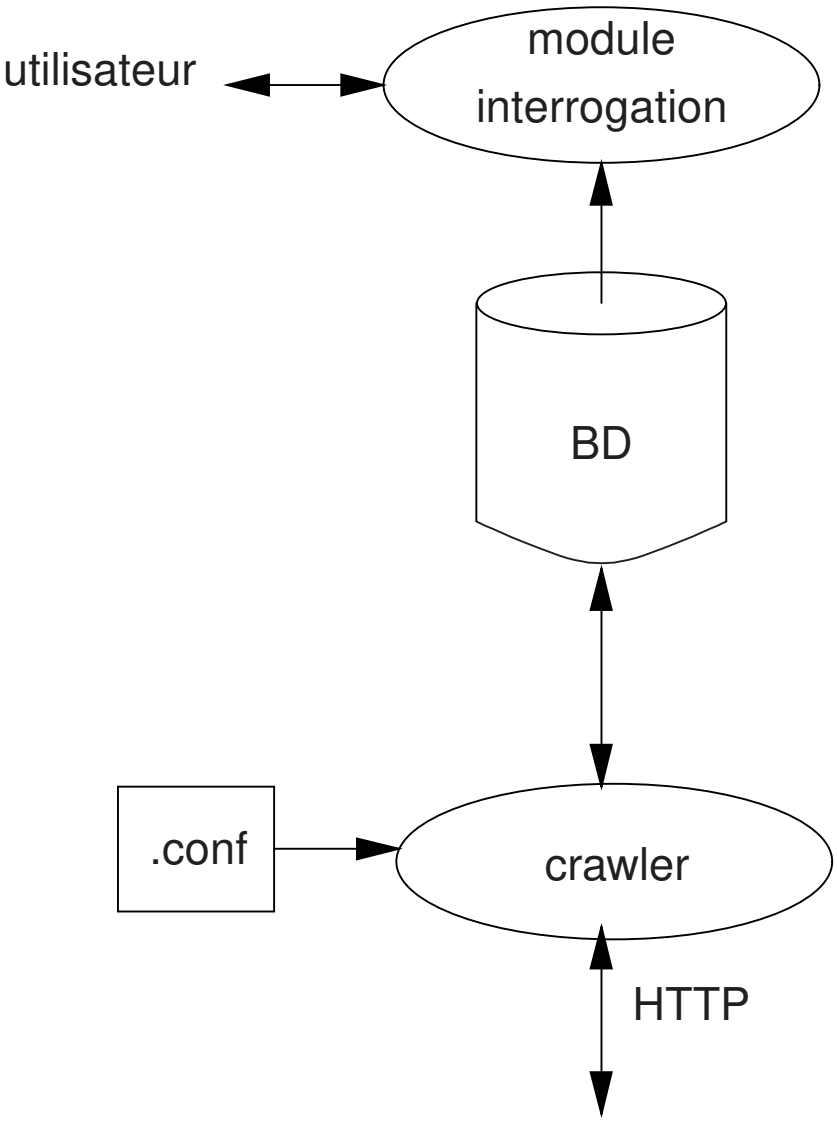
Programmation d'un mini-Google

Mihaela Sighireanu

Mihaela.Sighireanu@liafa.jussieu.fr
http://www.liafa.jussieu.fr/~sighirea/cours/mini-google/

Objectif

Implémentation d'un moteur de recherche de type GOOGLE.



Point de départ

Une implémentation simplifiée :

- *crawler* en Java
- BD en Postgres
- module d'interrogation en Perl

... et 6h de cours sur le Web, HTTP, CGI, etc.

Point(s) d'arrivée

Amélioration de l'implémentation de base sur au moins deux des points suivants :

- *crawler* concurrent
- *crawler* incrémental
- BD compacte (dictionnaire, compression de texte)
- classification des pages (*PageRank*)
- classification des résultats d'une recherche

Entrées/sorties

Pré-requis :

- programmation Java ou C
- BD relationnelles
- systèmes d'exploitation
- algorithmique des graphes

Acquis du projet :

- programmation Web : CGI et HTML
- programmation concurrente
- algorithmique avancée