

MozDex

Mozdex.com

Mozdex.com is a project to build a whole internet search engine around Nutch technologies.

Goals

Mozdex.com aims to re-establish a 100million page index through the summer of 2005 and reach 1 billion pages by next year.

Availability

[mozdex.com search engine](#)

Features

We have enabled Clustering, Ontology, PDF/WORD/XML and other features of Nutch.

Hardware

Web Servers (two)

- Athlon64 3000
- 1 Gig ram
- 2x40 gig disk
- Java 1.4.2
- Resin Application Server

Database Servers (four)

- Athlon64 3000
- 2 Gigs Ram
- 2x300 Gig Drives
- Java 1.4.2

Query Servers (four)

- Athlon64 3000
- 2 Gigs ram
- 2x40 gig disks
- Java 1.4.2

All servers are interconnected on GigE (over copy). We utilize a variant of CFS for global storage across the servers.

Performance

Query response is incredibly fast. We will be scripting some performance benchmarks under load and making this information available here in the near future.