

Drupal's page loading performance in 2010

Wim Leers ~ <http://wimleers.com/>

Drupal.org, IRC, Twitter, LinkedIn: wimleers

Bachelor thesis

“Improving Drupal’s page loading performance”

- Promotor: Prof. dr. Wim Lamotte
- Co-promotor: dr. Peter Quax
- Mentors: Stijn Agten & Maarten Wijnants



Goal

- **Faster web sites**

- Speed = satisfaction = more & happier visitors = more revenue



Goal

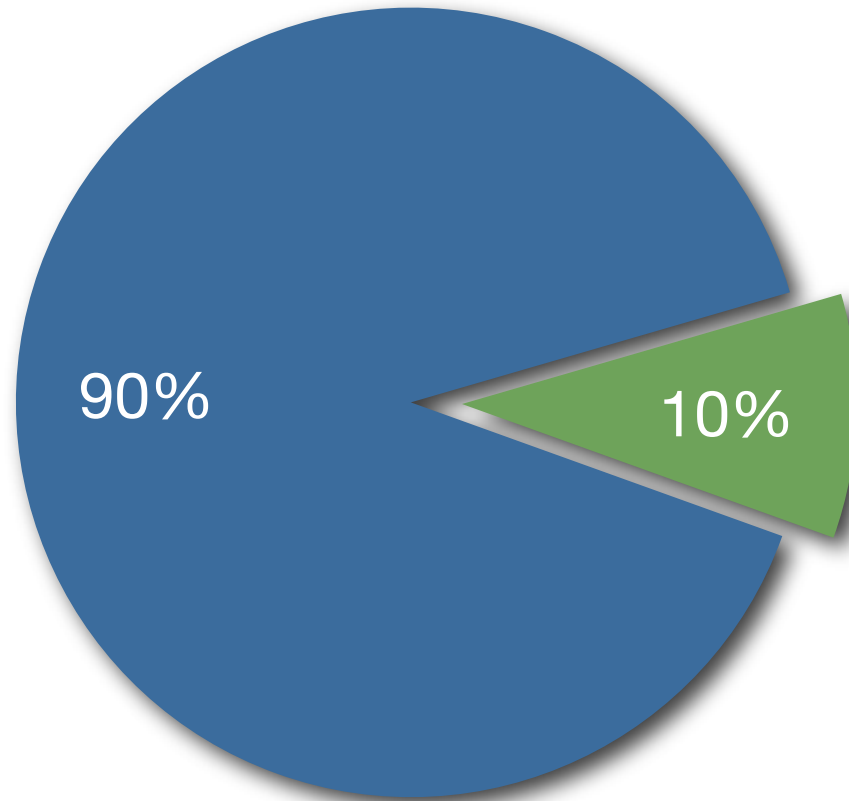
- **Faster web sites**

- Speed = satisfaction = more & happier visitors = more revenue

- Google: +0.5s → -20% searches



Terminology: page loading performance

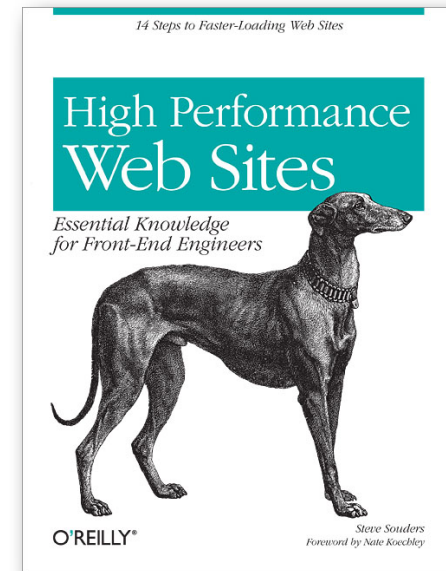


- CSS, JS, images ...
- HTML

Drupal's page loading performance

- One of the most effective measures:

“Use a CDN”

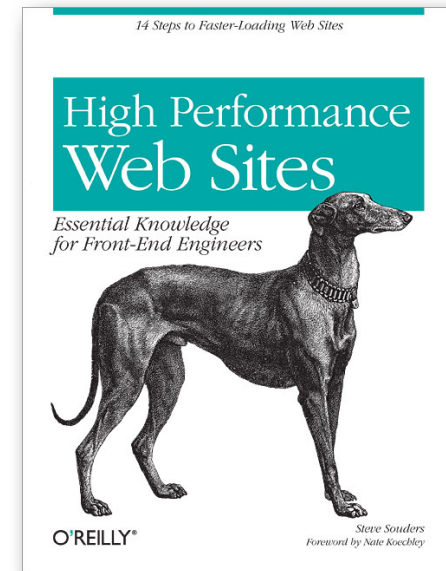


Drupal's page loading performance

- One of the most effective measures:

“Use a CDN”

- Drupal: ~~not yet possible!~~ **Possible now! :)**



Terminology: CDN



Key properties of a CDN

Key properties of a CDN

- Geographical spread

Key properties of a CDN

- Geographical spread
- Pull versus Push

	Pull
transfer protocol	automatically
+	virtually no setup
-	no flexibility

Key properties of a CDN

- Geographical spread
- Pull versus Push

	Pull	Push
transfer protocol	automatically	FTP, SFTP, rsync, WebDAV, Amazon S3 ...
+	virtually no setup	flexibility
-	no flexibility	setup

Key properties of a CDN

- Geographical spread
- Pull versus Push
- Lock-in

	Pull	Push
transfer protocol	automatically	FTP, SFTP, rsync, WebDAV, Amazon S3 ...
+	virtually no setup	flexibility
—	no flexibility	setup

Profiling: *Episodes*

Profiling: *Episodes*

- Measures “episodes” during page loading

Profiling: *Episodes*

- Measures “episodes” during page loading
- **Real measurements:** JS in browser, for *each* visitor
- No simulation!

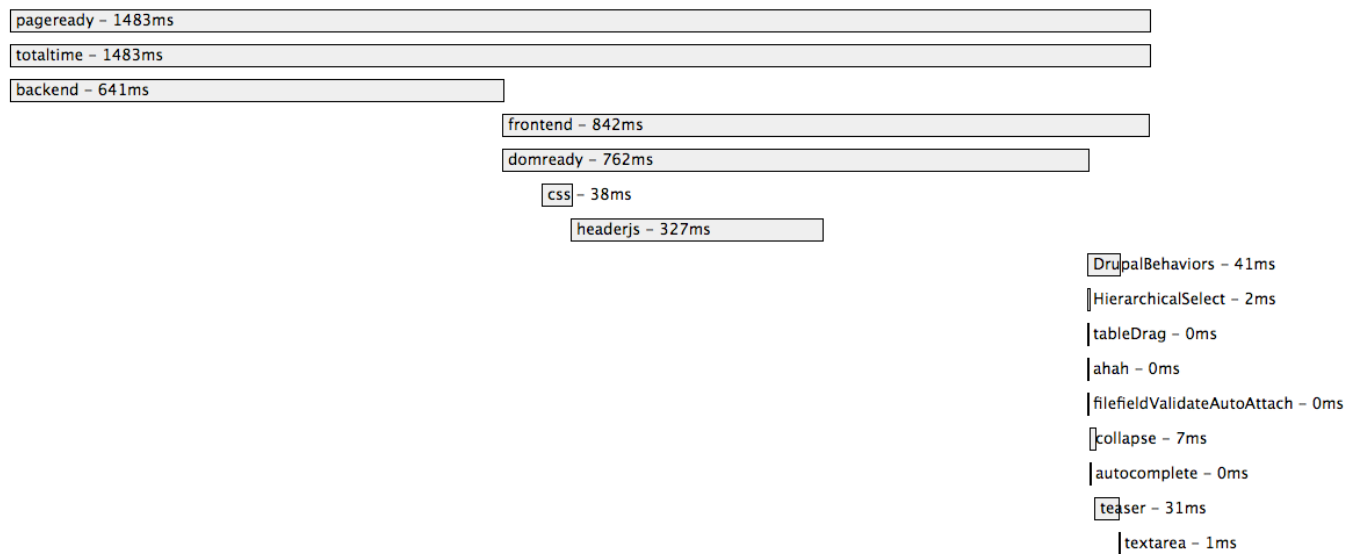
Episodes module

Episodes module

- **Drupal module** that offers Episodes integration

Episodes module

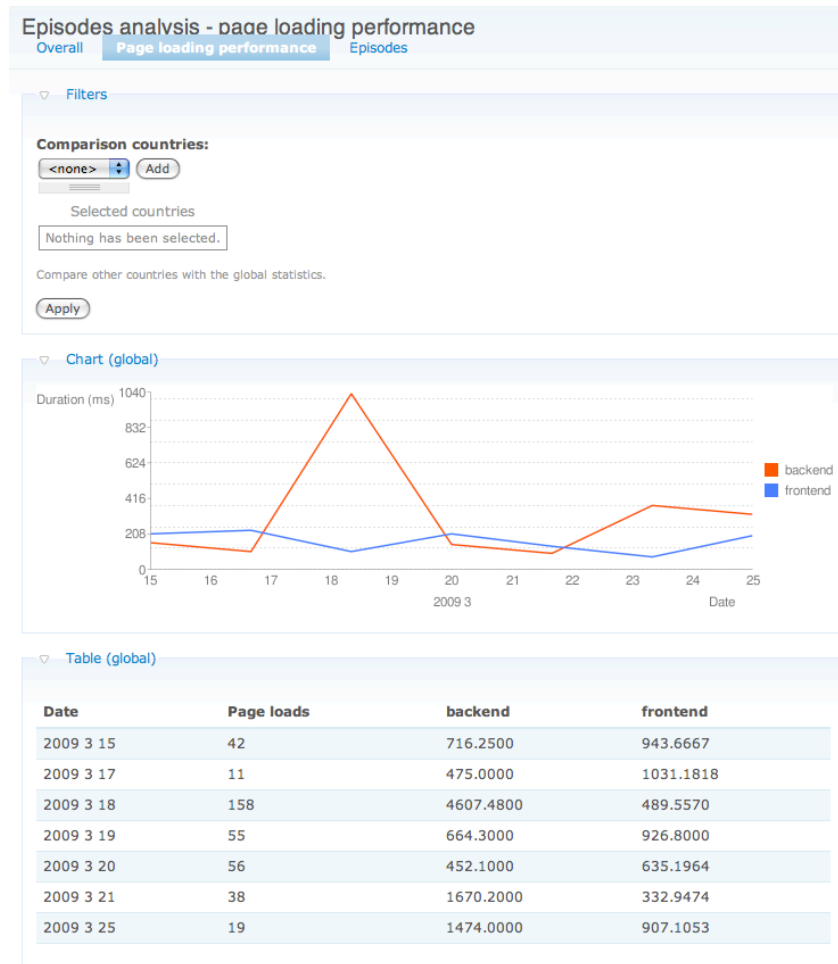
- **Drupal module** that offers Episodes integration



Episodes Server module

Episodes Server module

- **Drupal module** that visualizes the collected measurements



Daemon: *File Conveyor*

Daemon: *File Conveyor*

1. **Configuration:** simple XML file

Daemon: *File Conveyor*

1. **Configuration:** simple XML file
2. **Detection:** instantaneous

Daemon: *File Conveyor*

1. **Configuration:** simple XML file
2. **Detection:** instantaneous
3. **Processing:** store image more efficiently ... — extensible!

Daemon: *File Conveyor*

1. **Configuration:** simple XML file
2. **Detection:** instantaneous
3. **Processing:** store image more efficiently ... — extensible!
4. **Syncing:** supports many protocols (FTP, Amazon S3 ...) — extensible!

Daemon: *File Conveyor*

1. **Configuration:** simple XML file
2. **Detection:** instantaneous
3. **Processing:** store image more efficiently ... — extensible!
4. **Syncing:** supports many protocols (FTP, Amazon S3 ...) — extensible!
5. **Result:** SQLite DB with CDN URLs

Daemon: *File Conveyor*

1. **Configuration:** simple XML file
2. **Detection:** instantaneous
3. **Processing:** store image more efficiently ... — extensible!
4. **Syncing:** supports many protocols (FTP, Amazon S3 ...) — extensible!
5. **Result:** SQLite DB with CDN URLs

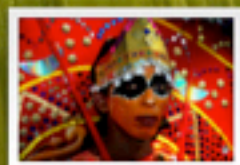
Project homepage: <http://fileconveyor.org/> — code on GitHub



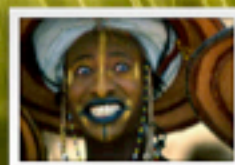
HD
249.88 GB, 167.58 GB free



monitored
No items



party.jpg
500 × 333



scary smile.jpg
500 × 333



traditional clothing.jpg
160 × 240

Daemon's capabilities: scenario 1

Company wants to switch from CDN provider X to Amazon S3.

- CDN X: FTP
- Amazon S3: custom protocol

Daemon's capabilities: scenario 1

Company wants to switch from CDN provider X to Amazon S3.

- CDN X: FTP
- Amazon S3: custom protocol

- Setup:
 - CDN X → Amazon S3
 - File Conveyor

Daemon's capabilities: scenario 1

Company wants to switch from CDN provider X to Amazon S3.

- CDN X: FTP
 - Amazon S3: custom protocol
-
- Setup:
 - CDN X → Amazon S3
 - File Conveyor
 - Alternative: write a lot of code

Daemon's capabilities: scenario 2

U.S. company expands to South-Korea

Daemon's capabilities: scenario 2

U.S. company expands to South-Korea

- Setup:
 - North-American CDN
 - Static file server in South-Korea
 - File Conveyor + language-/subdomain-based logic to pick CDN/server

Daemon's capabilities: scenario 2

U.S. company expands to South-Korea

- Setup:
 - North-American CDN
 - Static file server in South-Korea
 - File Conveyor + language-/subdomain-based logic to pick CDN/server
- Alternative: global CDN or slower web site in South-Korea

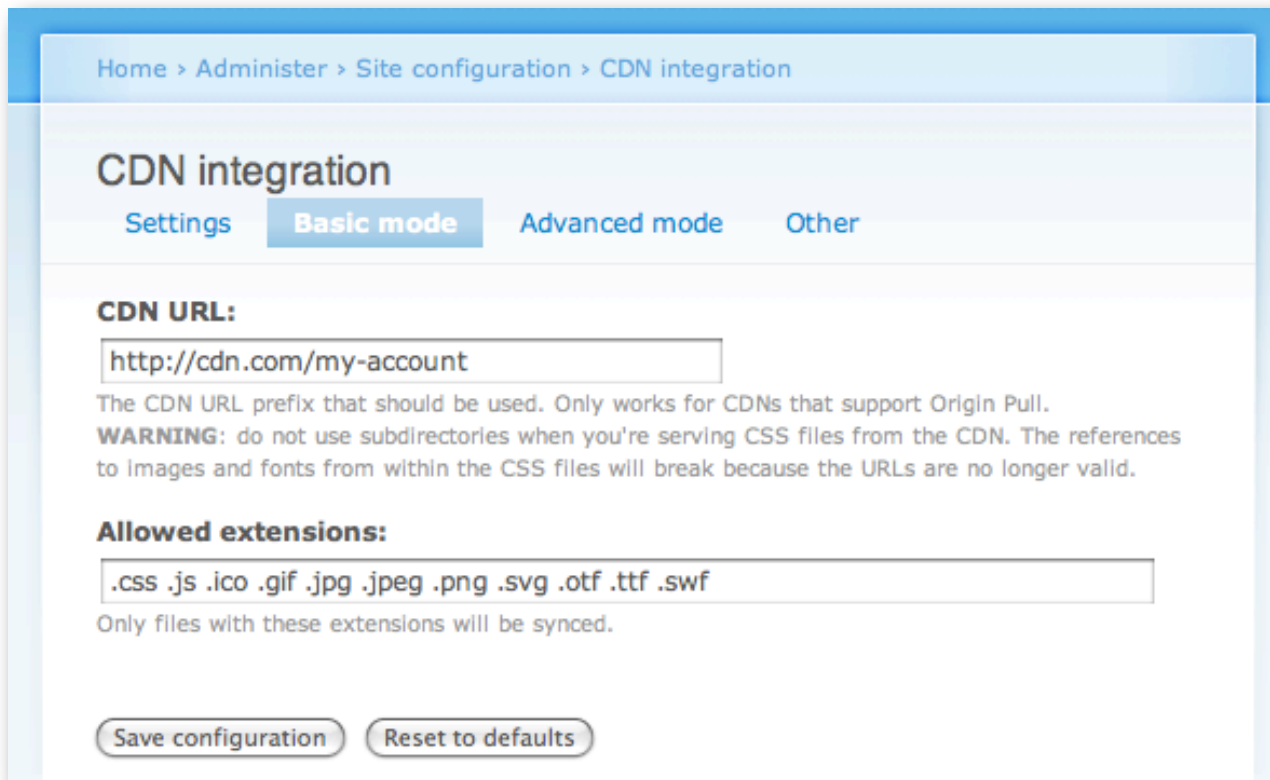
CDN integration module

CDN integration module

- Drupal core patch for Drupal 6, committed to Drupal 7 & **Pressflow** 6

CDN integration module

- Drupal core patch for Drupal 6, committed to Drupal 7 & *Pressflow* 6
- Basic mode: Origin Pull CDN



Home > Administer > Site configuration > CDN integration

CDN integration

Settings **Basic mode** Advanced mode Other

CDN URL:

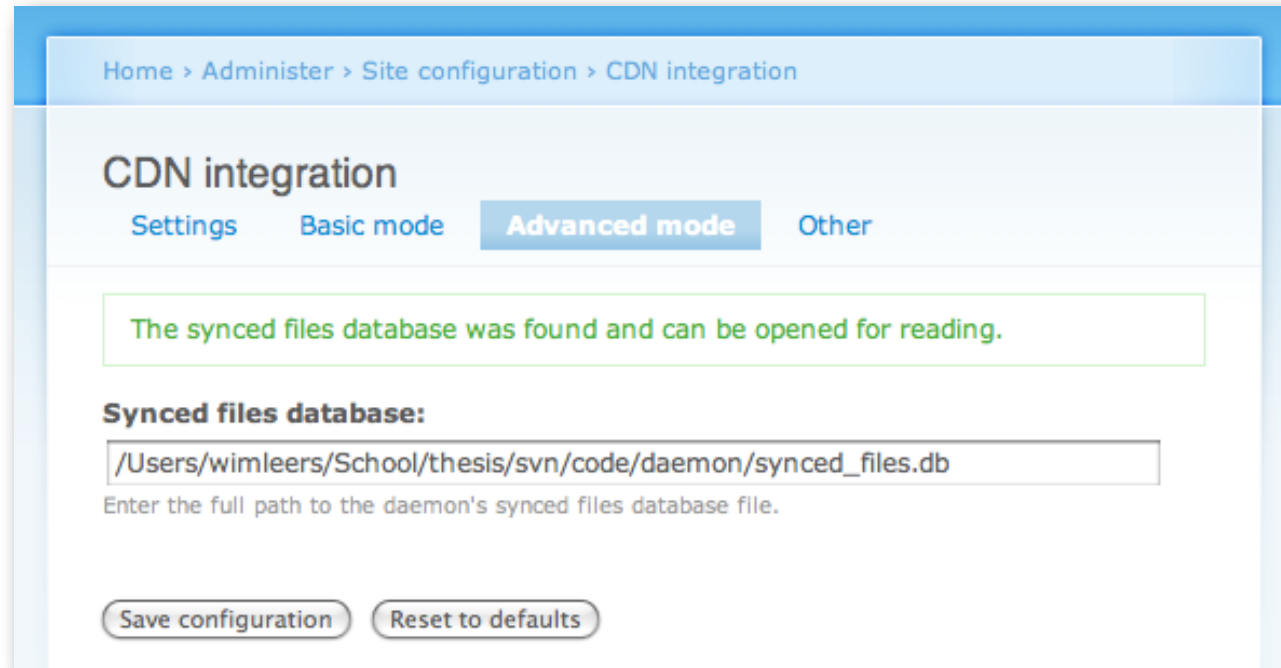
The CDN URL prefix that should be used. Only works for CDNs that support Origin Pull.
WARNING: do not use subdirectories when you're serving CSS files from the CDN. The references to Images and fonts from within the CSS files will break because the URLs are no longer valid.

Allowed extensions:

Only files with these extensions will be synced.

CDN integration module

- Drupal core patch for Drupal 6, committed to Drupal 7 & *Pressflow* 6
- Basic mode: Origin Pull CDN
- Advanced mode: **any CDN** (File Conveyor)



Home > Administer > Site configuration > CDN integration

CDN integration

Settings Basic mode **Advanced mode** Other

The synced files database was found and can be opened for reading.

Synced files database:

Enter the full path to the daemon's synced files database file.

Save configuration Reset to defaults

Test case

The screenshot shows a web browser window with the address bar displaying "News | DriverPacks.net". The website header features the "DriverPacks.net" logo on the left and a green banner on the right that reads "RÉPARER WINDOWS Prévenir les erreurs de Windows" with a "GRATUITEMENT" button and the URL "www.Quad-Cleaner.com". Below the header is a navigation menu with buttons for "News", "Downloads", "Forum", "Bugtracker", "Well-known users", "About", and "Donate".

The main content area is highlighted in yellow and contains the following text:

What are the "DriverPacks"?

Each DriverPack is a package of drivers that fit a certain category. There are for example DriverPacks for Sound devices and for Graphics devices. For each of these DriverPacks we include *every* device driver of that category. On top of that, we solve tricky problems — such as driver conflicts — for you.

They are intended to be used while creating a Windows 2000, XP, 2003 or Vista (and soon Windows 7) installation disc/image. We also provide you with the software to make that as simple as possible.

We've built this to **simplify the lives of network administrators and employees of pc shops**, but also enthusiast pc users.

Sounds cool? Then explore our [downloads](#) and come and say hi in our [forum](#)!

On the right side of the page, there is a "Latest DriverPacks" section with tabs for "XP", "Vista", and "7". Below the tabs is a list of links: "Chipset 9.02", "Mass Storage 9.01", "Graphics A 8.12.1", "Graphics B 8.12.1", and "Graphics C 8.12.1".

Below the links is a message: "Welcome from Belgium! For the fastest experience, you are downloading the images, CSS and JS from our static file server in Belgium." Below this is a status bar showing "35 online".

At the bottom right, there is a "save on Delicious" widget with tags: "windows drivers software unattended xp installation tools hardware slipstream driver" and a count of "548 others".

At the bottom left, there is a "New site live!" announcement:

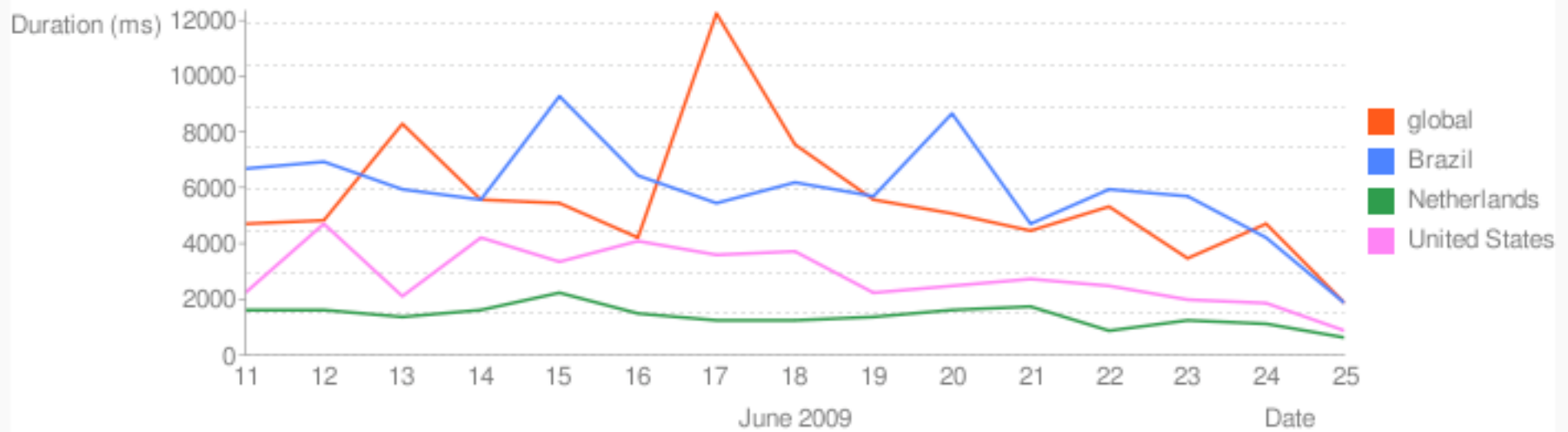
New site live!

Posted on Thu, 06/11/2009 - 03:14 by [Wim Leers](#)

Finally the time has come! The new web site went live today!

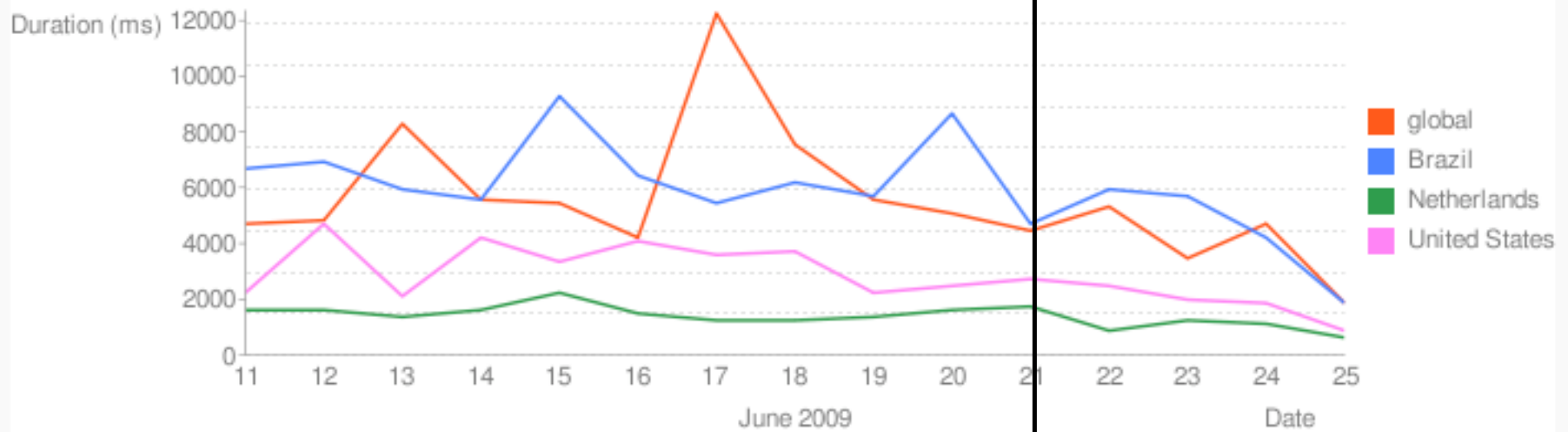
Test case

Comparison of frontend episodes



Test case

▽ Comparison of frontend episodes



Bachelor thesis: conclusion

Bachelor thesis: conclusion

1. Speed of web site very important

Bachelor thesis: conclusion

1. Speed of web site very important
2. Episodes → monitor page loading performance

Bachelor thesis: conclusion

1. Speed of web site very important
2. Episodes → monitor page loading performance
3. CDN integration module → best method to integrate Drupal with a CDN

Bachelor thesis: conclusion

1. Speed of web site very important
2. Episodes → monitor page loading performance
3. CDN integration module → best method to integrate Drupal with a CDN
4. File Conveyor → flexible CDN integration

It doesn't end here!

What else exists ...

What else exists ...

High-traffic Drupal sites:

What else exists ...

High-traffic Drupal sites:

The logo for Pressflow, featuring the word "Pressflow" in a bold, blue, italicized sans-serif font with a white outline and a grey drop shadow. A small "TM" trademark symbol is located at the bottom right of the word.

PressflowTM

What else exists ...

High-traffic Drupal sites:



“Pressflow is a distribution of Drupal with integrated performance, scalability, availability, and testing enhancements.”

What else is coming ...

What else is coming ...

1. CDN integration 2

What else is coming ...

1. CDN integration 2

2. ESI

What else is coming ...

1. CDN integration 2

2. ESI

3. Master thesis

CDN integration 2

Goals

CDN integration 2

Goals

1. Significant expansion of functionality

CDN integration 2

Goals

1. Significant expansion of functionality

Obsoleting modules:

Parallel & Simple CDN

New mode: **Media Mover**

CDN integration 2

Goals

1. Significant expansion of functionality
2. Awesome UX

Obsoleting modules:

Parallel & Simple CDN

New mode: **Media Mover**

CDN integration 2

Goals

1. Significant expansion of functionality
2. Awesome UX

Obsoleting modules:

Parallel & Simple CDN

New mode: **Media Mover**

Bojhan Somers

CDN integration 2

Goals

1. Significant expansion of functionality
2. Awesome UX
3. Excellent docs + high-quality screencasts

Obsoleting modules:

Parallel & Simple CDN

New mode: **Media Mover**

Bojhan Somers

CDN integration 2

Goals

1. Significant expansion of functionality
2. Awesome UX
3. Excellent docs + high-quality screencasts
4. Streamlined CDN-specific UI

Obsoleting modules:

Parallel & Simple CDN

New mode: **Media Mover**

Bojhan Somers

CDN integration 2

Goals

1. Significant expansion of functionality
2. Awesome UX
3. Excellent docs + high-quality screencasts
4. Streamlined CDN-specific UI

Obsoleting modules:

Parallel & Simple CDN

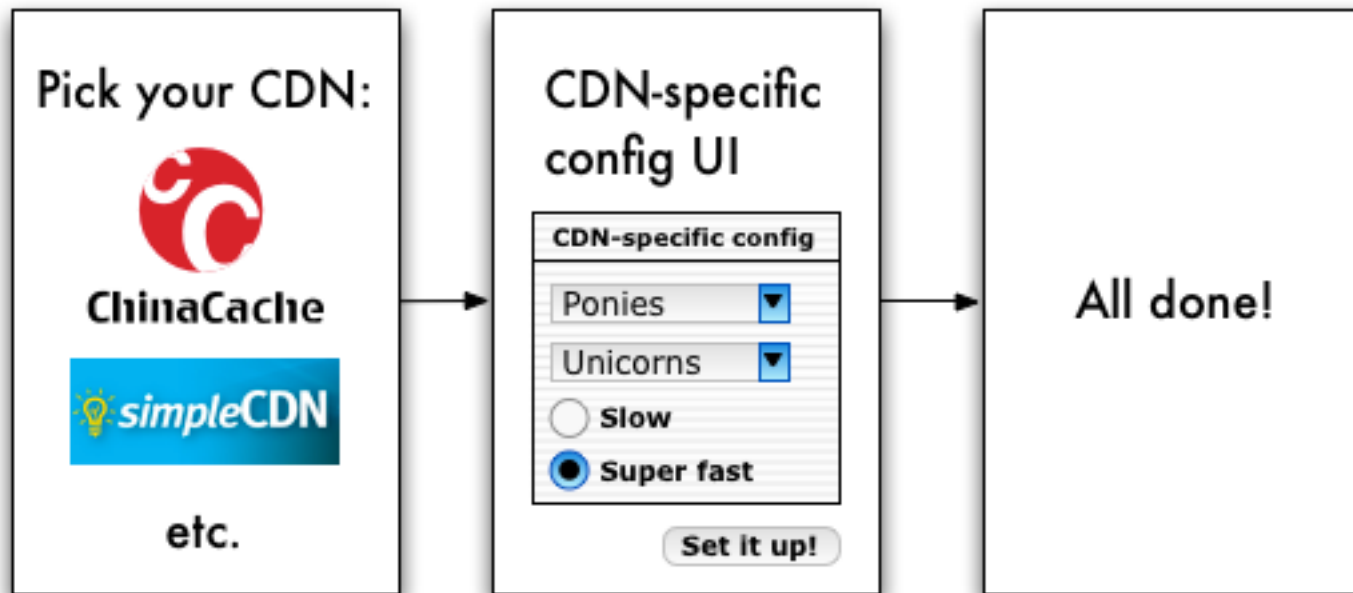
New mode: **Media Mover**

Bojhan Somers

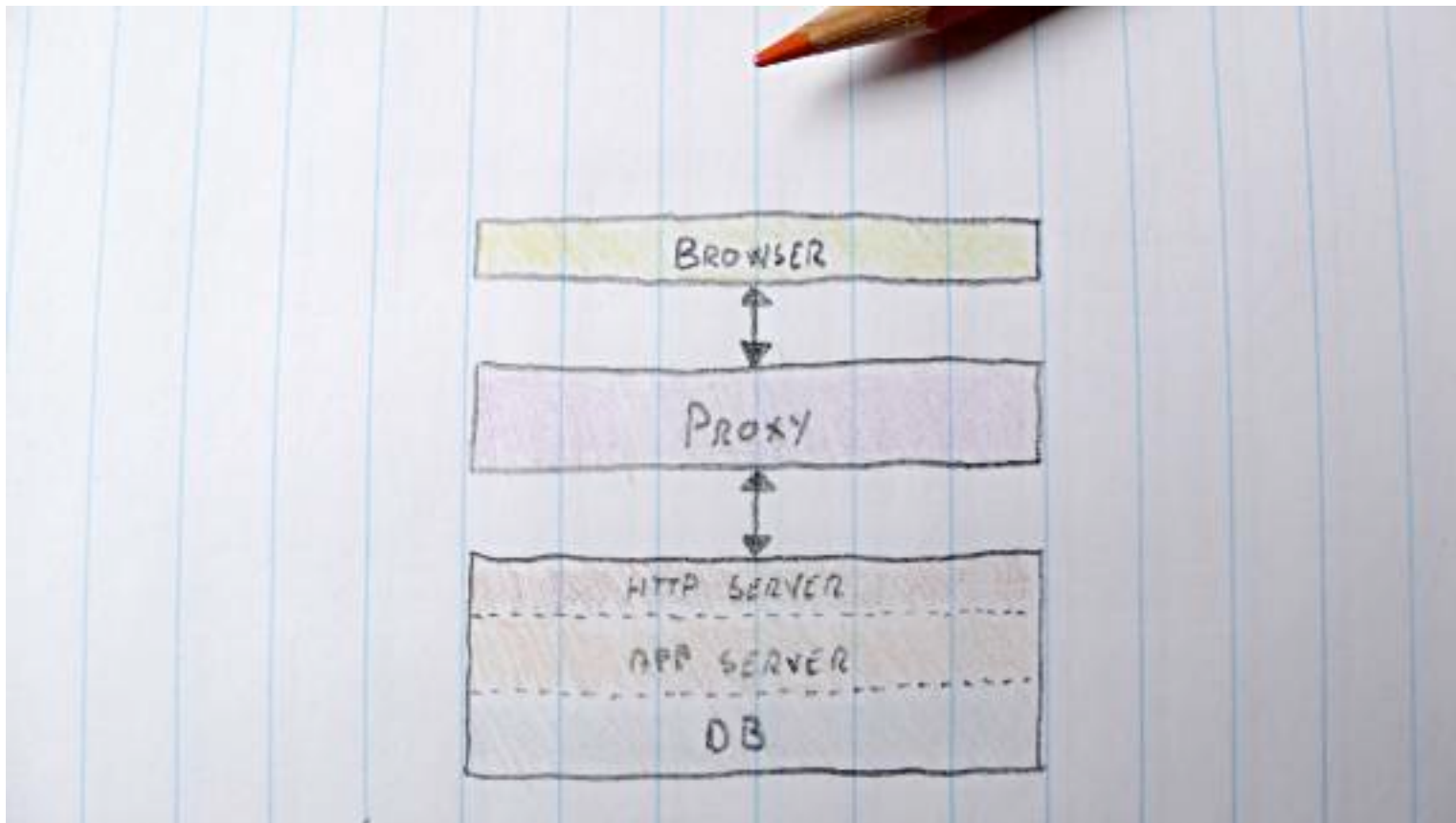


ChinaCache

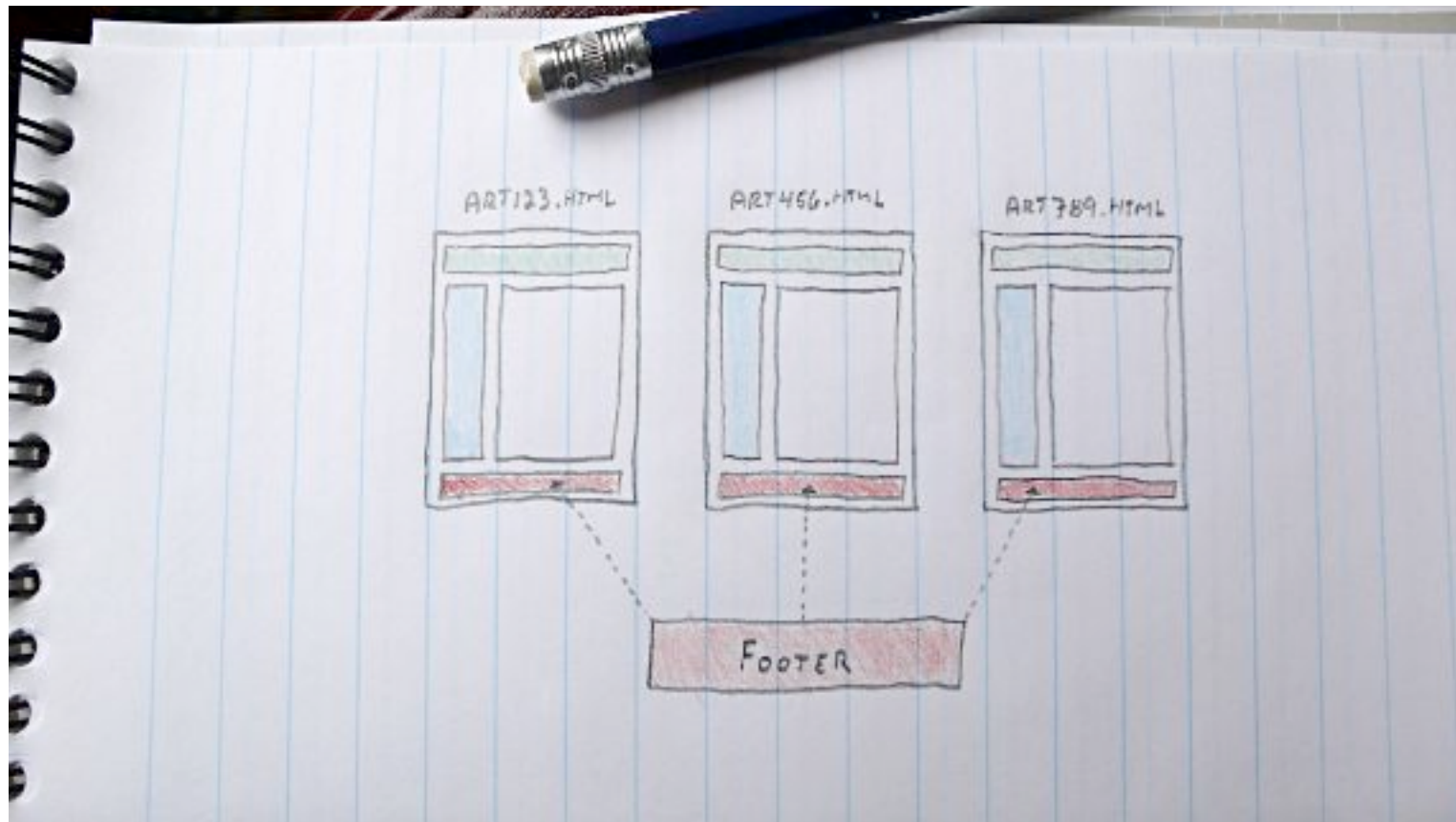
CDN integration 2



ESI



ESI



ESI

```
<body>
  <div id="header">
    <esi:include src="http://drupal.org/esi/block/header" />
  </div>
  <div id="left-sidebar">
    <esi:include src="http://drupal.org/esi/region/left-sidebar" />
  </div>
  <div id="content">
    <esi:include src="http://drupal.org/esi/content/some-path" />
  </div>
  <div id="footer">
    <esi:include src="http://drupal.org/esi/block/footer" />
  </div>
  <!--
    http://drupal.org/esi/<type>/<id>
  -->
</body>
```

ESI

ESI

- ESI module: <http://drupal.org/project/esi>

ESI

- ESI module: <http://drupal.org/project/esi>
- In development

ESI

- ESI module: <http://drupal.org/project/esi>
- In development
- **Pressflow**™ developers are actively participating

ESI

- ESI module: <http://drupal.org/project/esi>
- In development
- **Pressflow**™ developers are actively participating
 - David Strauss (Four Kitchens)

ESI

- ESI module: <http://drupal.org/project/esi>
- In development
- **Pressflow**™ developers are actively participating
 - David Strauss (Four Kitchens)
 - Josh Koenig (Chapter Three)

ESI

- ESI module: <http://drupal.org/project/esi>
- In development
- **Pressflow**™ developers are actively participating
 - David Strauss (Four Kitchens)
 - Josh Koenig (Chapter Three)
- Note: not every CDN supports ESI!

Master thesis

“Web Performance Optimization: Analytics”

- Promotor: Prof. dr. Jan van den Bussche



Master thesis

Master thesis

- An application that can automatically extract conclusions out of Episodes logs and visualize them.

Master thesis

- An application that can automatically extract conclusions out of Episodes logs and visualize them.
- e.g.:

Master thesis

- An application that can automatically extract conclusions out of Episodes logs and visualize them.
- e.g.:
 - “http://d.o/ is slow in Belgium, for users of the ISP Telenet”

Master thesis

- An application that can automatically extract conclusions out of Episodes logs and visualize them.
- e.g.:
 - “http://d.o/ is slow in Belgium, for users of the ISP Telenet”
 - “http://d.o/project/esi has slowly loading CSS”

Master thesis

- An application that can automatically extract conclusions out of Episodes logs and visualize them.
- e.g.:
 - “http://d.o/ is slow in Belgium, for users of the ISP Telenet”
 - “http://d.o/project/esi has slowly loading CSS”
 - “http://d.o/project/cdn has slowly loading JS for visitors that use the browser Internet Explorer 6 or 7”

Master thesis

- An application that can automatically extract conclusions out of Episodes logs and visualize them.
- e.g.:
 - “http://d.o/ is slow in Belgium, for users of the ISP Telenet”
 - “http://d.o/project/esi has slowly loading CSS”
 - “http://d.o/project/cdn has slowly loading JS for visitors that use the browser Internet Explorer 6 or 7”
- Challenges: complex data mining, distributed computing, storage, performance

Master thesis

- An application that can automatically extract conclusions out of Episodes logs and visualize them.
- e.g.:
 - “<http://d.o/> is slow in Belgium, for users of the ISP Telenet”
 - “<http://d.o/project/esi> has slowly loading CSS”
 - “<http://d.o/project/cdn> has slowly loading JS for visitors that use the browser Internet Explorer 6 or 7”
- Challenges: complex data mining, distributed computing, storage, performance

See <http://wimleers.com/tags/master-thesis>

Conclusion

Conclusion

1. Speed of web site very important
2. Episodes → monitor page loading performance
3. CDN integration module → best method to integrate Drupal with a CDN
4. File Conveyor → flexible CDN integration

Conclusion

1. Speed of web site very important
2. Episodes → monitor page loading performance
3. CDN integration module → best method to integrate Drupal with a CDN
4. File Conveyor → flexible CDN integration
- 5. CDN integration module 2 → more awesome!**

Conclusion

1. Speed of web site very important
2. Episodes → monitor page loading performance
3. CDN integration module → best method to integrate Drupal with a CDN
4. File Conveyor → flexible CDN integration
- 5. CDN integration module 2 → more awesome!**
- 6. ESI → big potential**

Conclusion

1. Speed of web site very important
2. Episodes → monitor page loading performance
3. CDN integration module → best method to integrate Drupal with a CDN
4. File Conveyor → flexible CDN integration
- 5. CDN integration module 2 → more awesome!**
- 6. ESI → big potential**
- 7. Master thesis → more insight**

Conclusion

1. Speed of web site very important
2. Episodes → monitor page loading performance
3. CDN integration module → best method to integrate Drupal with a CDN
4. File Conveyor → flexible CDN integration
- 5. CDN integration module 2 → more awesome!**
- 6. ESI → big potential**
- 7. Master thesis → more insight**

This talk will be available at <http://wimleers.com/talk/fosdem-2010>

Need Drupal page loading performance help?

Contact me: <http://wimleers.com/contact>

This talk will be available at <http://wimleers.com/talk/fosdem-2010>