

The Tool Labs



Your \$HOME
away from
home.

What are the Tool Labs?

- It's a reliable, scalable hosting environment for community developers.
- It's “close to the metal”, sharing datacenters with the WMF production environment.
- It's supported both by WMF ops, and by dedicated staff and volunteers.
- It's designed around the needs of our projects.

Why the Tool Labs?

- The Wikimedia projects are increasingly dependent on external tools, we rely on them so we need to support them.
- Tool *maintainers* need support too! We make it easy to share responsibility, or to adopt orphaned tools.
- Sharing resources mean MOAR POWAR!!1!
(and more scalability).

Design objectives

1. Reliability

- Tools must *stay* available!

2. Scalability

- We need room to grow.

3. Simplicity

- It should be simple for a new developer to hit the ground and start coding.

Tool Labs Features

- Supports web services, continuous bots, and scheduled tasks.
- Self-serve management of your tools.
- Access to replicated production databases.
- Web server is fully featured; deploy and run Mediawiki in minutes.
- Redundancy up the wazoo!

Tool Labs Features

- Tool maintenance can be shared.
- Gridengine for dispatching jobs.
- Web proxy provides anonymized logs.
- Shared filesystems within the project with *timetravel*.
- Designed to make maintainers' lives easy: mosh, SSH, SFTP are all supported without complicated proxy setup.

... the rules?

- Open Source only.
- Open Data only.
- Don't be a dick.
- Don't break things for others.
- Private information needs to be handled carefully, if at all.

The Tool Labs



Cool!

How do I join
the fun?

Wikitech

<http://wikitech.wikimedia.org/>

- Wikitech is the portal to all things Labs
- Create an account, and put your SSH keys in your account preferences
- Request for shell access is automatic, request access to the Tool Labs at:
[[Special:FormEdit/Tools Access Request]]

The big picture

- There are, essentially, four components to the Tool Labs:
 - The bastion hosts
 - The grid
 - The web cluster
 - The databases

Pretty picture: [\[\[File:Labs Tools topology.png\]\]](#)

The bastion hosts

`ssh://tools-login.wmflabs.org`

`ssh://tools-dev.wmflabs.org`

- This is where users log in and do their thing.
- They are functionally identical, but we request that heavy processing not take place on -login (compiles, etc) to keep interactive performance snappy.

The grid

- Implemented with Open Gridengine, the Open Source fork of SGE (~~Sun~~ Oracle Grid Engine)
- You submit jobs to it from the bastions or web services, it finds a nice cozy place to run your job on one of the execution hosts.
- Synchronously or asynchronously
- One-shot, or continuous (the grid will restart it if it fails, and keep it running)

The web cluster

- Proxy in front open to the 'net
- <http://tools.wmflabs.org/>
- Supports SSL
- Load is distributed between identical backends (statically), and they can all serve as cold spares to each other.
- Uses suphp to run scripts and CGI
- WSGI coming soon to a Theatre Near You!

The databases

- Two sets: replicas and project-local
- All of the project databases replicated, with access level comparable to registered user.
- User and tools can create databases.
- Project-local database is perfect when you don't need joins with the replicas.

Tool accounts

- Also known as “service groups” in the Labs parlance.
- Separate UID and GID, an account for the tools themselves.
- All the maintainers are members of the group, and can manage it.
- Any number of services “under one banner”

When to use tool accounts?

- You should never run continuous processes, or publicly accessible tools from your own account (what if you get hit by the proverbial lorry?) so you need at least one.
- But when to use more than one?
 - When you expect they might have different maintainers someday
 - When you prefer to keep separate tools separate.

Think of the orphans!

- But, also importantly, tools can be lost to the mists of time when a maintainer drifts away.
- This way, important tools with no maintainers can be taken over by other community members!
- (No, we don't yet know the rules for that – but we know it's not going to be done against maintainers' wills, or unreasonably fast).

It's all open source anyways

- In the end, another community member could always just make a new tool running the same code.
- ... but this makes it possible to have a backup plan so that we never lose tools the projects have come to rely upon.

What next?

- We'll be (and have been!) holding workshops to help you around the Tool Labs.
- Joiiiiin uuuus!

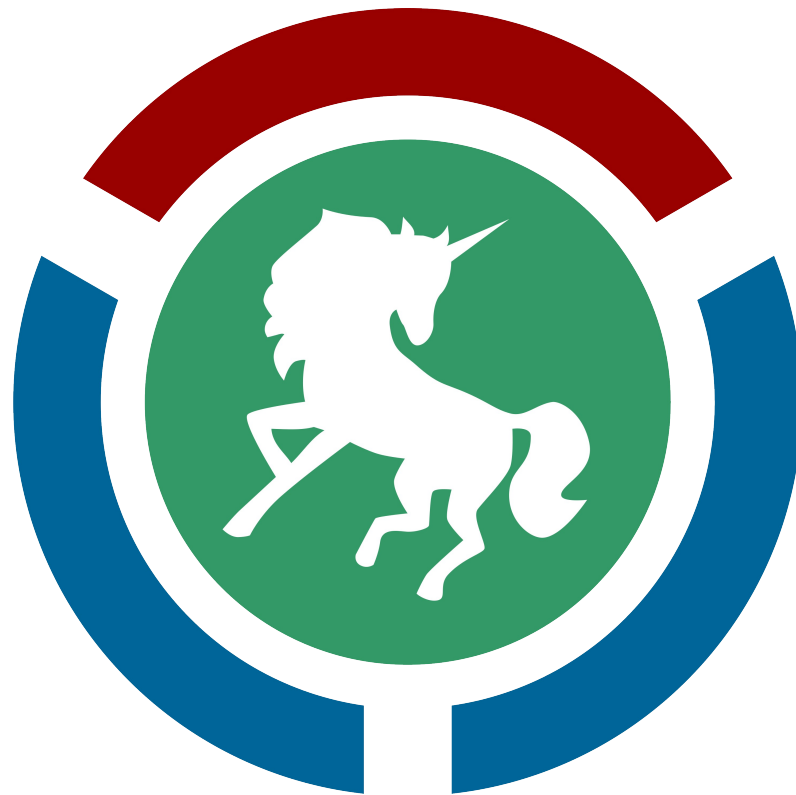
https://wikitech.wikimedia.org/wiki/Nova_Resource:Tools

It's too complicated!

- We're here to help!

#wikimedia-labs on freenode
labs-l@lists.wikimedia.org

Questions?



WIKIMEDIA
LABS