DSPACE CLUSTERING

VIA PUPPET, HAPROXY AND CEPHFS

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

TABLE OF CONTENTS

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

WHAT WE WANT

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

Automatic, customizable and secure *all-in-one* configuration

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

DSpace-Clustering
OR2019

Automatic, customizable *cluster* configuration which allows for Horizontal scaling, improved availability and improved maintainability

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

WHAT DSPACE NEEDS

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

SOFTWARE STACK

- Apache web server
- Tomcat application server
- PostgreSQL database server
- Solr index server
- DSpace web application
- File system

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

HOW WE GET WHAT WE WANT

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

DSpace-Clustering
OR2019

SOFTWARE STACK

- GitLab: source code management & CI & CD
- Puppet: configuration management & orchestration
- Custom Puppet module for DSpace

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

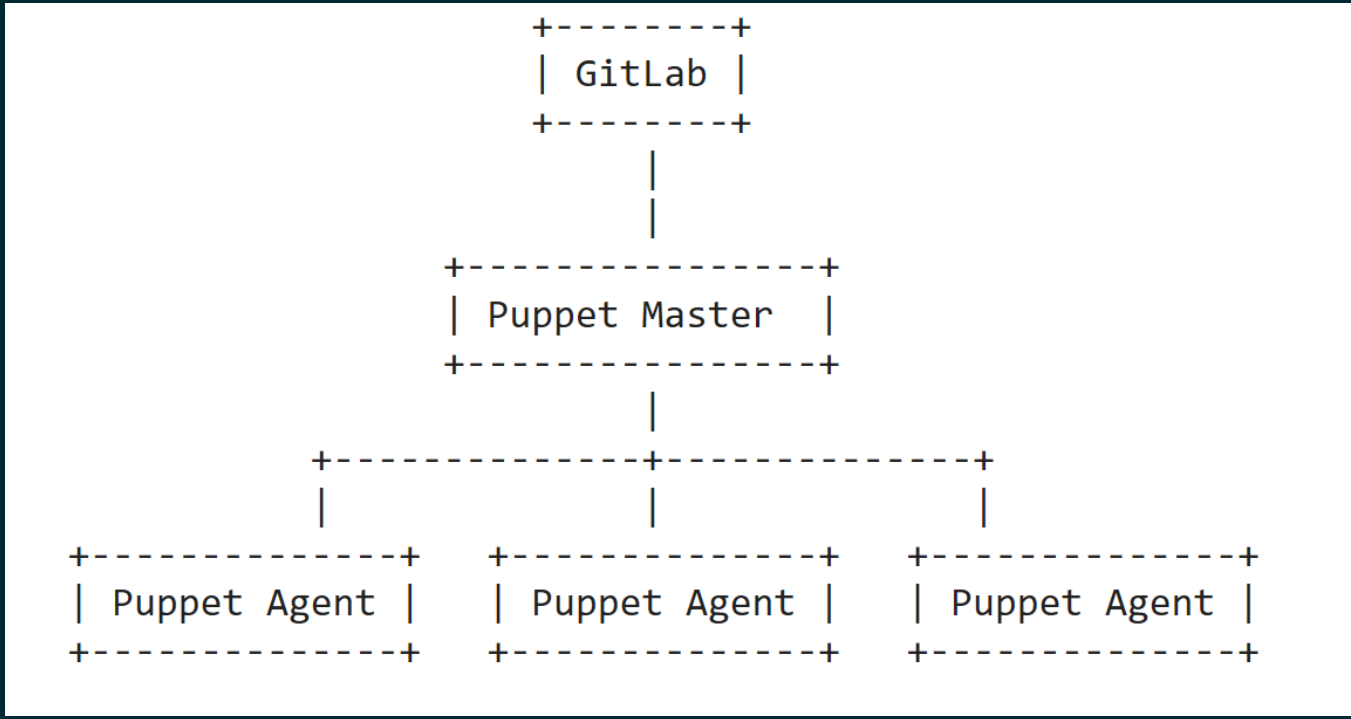
CUSTOM PUPPET MODULE FOR DSPACE

- Handles installation & configuration of all DSpace components
- Handles firewall, orchestration and communication in cluster setting

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

GITLAB + PUPPET



Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

DSpace-Clustering
OR2019

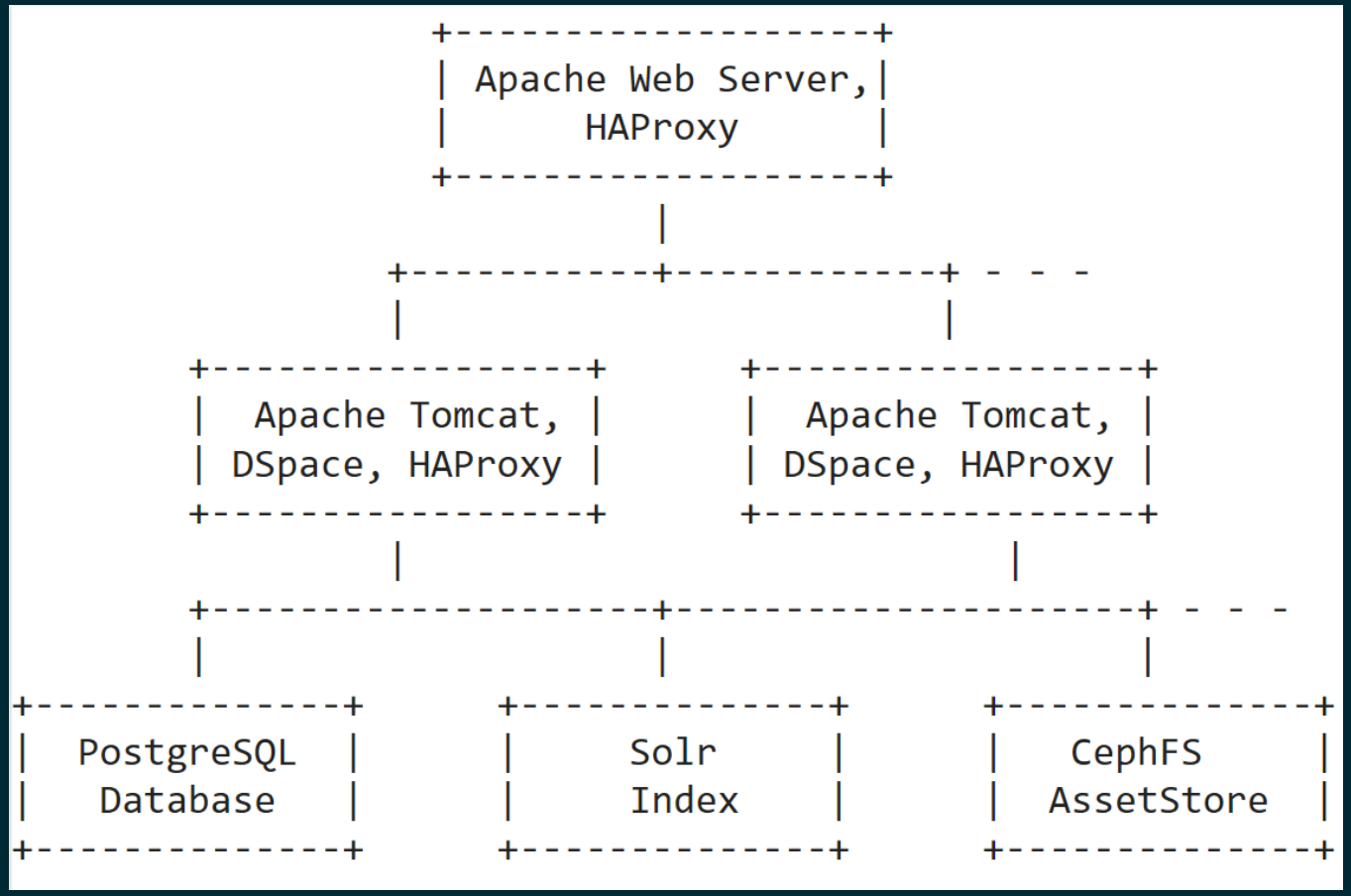
PUPPET AGENT ON A DSPACE APPLICATION NODE

- Gets configuration from Puppet Master
- Clones/pulls the source code
- Fills and places template files
- Runs Maven and Ant
- Restarts Tomcat

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

DSPACE COMPONENTS



Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

SHARED RESOURCES IN A CLUSTER SETTING

- Database
- Index
- AssetStore directory (we use *CephFS*)
- Other shared directories (e.g. *exports* or *oai*; we use *CephFS*)
- Session data? No, actually we just pin the client on a worker

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CHARACTERISTICS OF OUR CLUSTERING

- The components portal, application, database and index are installed and configured on designated nodes
- All DSpace-specific server components listen to local host
- Web server listens on the public interface
- In cluster mode all components are transparently connected via HAProxy
- Secure firewall and monitoring configuration

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CONFIGURATION AND CUSTOMIZATION VIA PUPPET MODULE

- General Configuration of System/Java/Tomcat environment variables
- Configuration of general DSpace-related settings
- Git repository with slightly modified basic DSpace source code, e.g. hot fixes
- Git repository with different branches of customized code (maven overlay mechanics)

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CODE EXAMPLE

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CONFIGURATION CLASS

```
class { 'dspace::configuration':
  dspace_name                => 'Institutional DSpace',
  git_repository             => 'https://github.com/DSpace/DSpace.git',
  git_revision               => 'dspace-6.3',
  separated_code_repositories => true,
  dspace_custom_repository   => 'https://gitlab.my-institution.edu/dspace/custom.git',
  git_revision               => 'master',
  java_memory_share          => 0.4,
  database_name              => 'dspace',
  database_user              => 'dspace',
  database_passwd            => 'dspace',
  mail_user                  => 'dspace',
  mail_pass                  => 'dspace',
  mail_sender                => 'noreply@my-institution.edu',
  mail_feedback_address      => 'feedback@my-institution.edu',
  mail_admin_address         => 'admin@my-institution.edu',
  authentication_method      => 'org.dspace.authenticate.PasswordAuthentication',
}
```

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

ALL-IN-ONE INSTALLATION

```
# Install DSpace application
include 'dspace::application::source'

include 'dspace::application::server'

# Install database server
include 'dspace::database::server'

# Install index server
include 'dspace::index::source'

include 'dspace::index::server'

# Install web server
include 'dspace::portal::server'
```

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CLUSTER INSTALLATION

```
# Base node configuration for each node

firewallchain { 'INPUT:filter:IPv4':
  ensure => present,
  policy => drop,
  before => undef,
}

firewall { '100 Allow inbound SSH':
  dport    => 22,
  proto    => tcp,
  action   => accept,
}

#...
```

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CLUSTER INSTALLATION

```
# Portal (Web Server) node configuration

class { 'dspace::portal::node':
  cluster => 'myCluster'
}

class { 'dspace::portal::server':
  # ...
}
```

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CLUSTER INSTALLATION

```
# Application (Tomcat/DSpace) node configuration

class { 'dspace::application::node':
  cluster => 'myCluster'
}

class { 'dspace::application::source':
  # ...
}

class { 'dspace::application::server':
  # ...
}
```

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CLUSTER INSTALLATION

```
# Database (PostgreSQL) node configuration

class { 'dspace::database::node':
  cluster => 'myCluster'
}

class { 'dspace::database::server':
  # ...
}
```

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

CLUSTER INSTALLATION

```
# Index (Tomcat/Solr) node configuration

class { 'dspace::index::node':
  cluster => 'myCluster'
}

class { 'dspace::index::source':
  # ...
}

class { 'dspace::index::server':
  # ...
}
```

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg

THANK YOU!

Visit us on GitHub.

Soon on Puppet Forge.

Bernd Nicklas, Paul Münch
Kontakt: dspace@uni-marburg.de

Philipps Universität Marburg