science + computing

an atos company

How to Deploy a Secure, Highly-Available Hadoop Platform

Dr. Olaf Flebbe, Michael Weiser

science + computing ag

IT-Dienstleistungen und Software für anspruchsvolle Rechnernetze

Tübingen | München | Berlin | Düsseldorf

Agenda

- Environment
- Automation
- Kerberos
- Demo

Target Scenario

- Build a secure, reliable, Debian based Hadoop system for the german BSI (Bundesamt für Sicherheit in der Informationstechnik)

Federal Office
for Information Security

- Cornerstones
  - MIT Kerberos / OpenLDAP
  - Debian Jessie (8)
  - Able to rebuild from source
    - Zookeeper/HA Hadoop/Hive/Hue
  - Automation necessary
  - Upstreaming into projects
  - Offline Installation

Code donated to Apache Bigtop

- We donate the automation of Hadoop with puppet to the Apache Bigtop project
  - Advance the deployment recipes
  - Demo of automation with puppet
  - Demo of how to secure Apache Hadoop Ecosystem
- All of our code is under Apache 2 License
- Code is on https://github.com/oflebbe/inst
- Uses the Apache Bigtop convenience repositories

MIT-Kerberos

- Master/Slave Setup
- Install and configure Kerberos-KDCs
  - Kadmin Software kdc, kpropd
  - Konfiguration of kdc.conf
- Configure replication:
  - cronjob for kprop on Master und kpropd service on Slave
- Manage service- and host-Principals
  - kadmin, add_principal, modprinc, cpw, …
- Keytab-Management:
  - kadmin, ktadd

Automation needed:

- NTP
- Nagios Client Config
- Postgresql Master/Slave Replication
- Hostname Resolution
- Kerberos KDC Master/Slave Replication
- OpenLDAP Multimaster
- Ldap client config
- Saslauthd
- SSSD Config
- Locales
- Firewall

- PAM
- Puppetdb manage ssh hostkeys
- Timezone

- Zookeeper
- Hadoop nn, jn, zkfc, rm, dn
- Hive
- Hue
- Tez
- Oozie

Puppet in five Minutes:

- Declarative configuration of target state
- Stateless, without ordering but with dependencies

- puppet ecosystem:
  - puppet (engine)
  - facter (determine facts of the system, like os)
  - hiera (hierarichal lookup of target properties)
  - augeas (manage code sniplets in configuration files, in various formats)
  - puppetdb (Database of generated properties, for instance ssh host keys)
  - mcollective (massive remote commands)

Puppet Concepts

- Manifest: Puppet code
- Class:
  - logical group of Puppet code
  - Smallest entity to call
  - best practice: 1 manifest == 1 class
- Module:
  - Group of classes for a feature to handle
  - Many projects on GitHub most often with Apache 2 License
  - Deployment on Puppet Forge
  - Installation: puppet module install <author>-<module>

Puppet Concepts

- hiera
  - Configuration and instantiation of classes

- site-manifest:
  - best practice: site-Manifest almost empty, starts node classification via hiera

- catalog:
  - Assembled collective of manifests and class for a particular host.
  - Host determines deviation from target state and tries to reach target state

Puppet Modes

- masterless/apply: All manifests are local to system
  - Bigtop mode suitable for CI

- master/agent: Agent starts catalog generation on master and applies catalog on local system
  - Creates a PKI
  - SSL Connection with trust!
  - Usually used for enterprise configurations

Puppet Usage

- Typical practice:
  - Search on Puppet Forge, try out
  - Example: saz/locales
  - Example: Apache Bigtop already has some configuration classes

Automation done with puppet:

- NTP
- Nagios Client Config
- Postgresql Master/Slave Replication
- Hostname Resolution
- Kerberos KDC Master/Slave Replication
- OpenLDAP Multimaster
- Ldap client config
- Saslauthd
- SSSD Config
- Locales
- Firewall

- PAM
- Puppetdb manage ssh hostkeys
- Timezone

- Zookeeper
- Hadoop nn, jn, zkfc, rm, dn
- Hive
- Hue
- Tez
- Oozie

puppet-modules for kerberos

- Only few modules available
- Problems:
  - Some of the only implements Client-configuration
  - Some of then uses hard-coded default passwords
- Michael Weiser has improved a promising one
  - edgester/kerberos
- Unfortunately not available on PuppetForge, only on github: https://github.com/edgester/puppet-module-kerberos
- MIT License

PKINIT - Kerberos with X.509 Certificates

- PKINIT (https://tools.ietf.org/html/rfc4556) replaces user passwords with X.509-Client certificates
- Substantial increase of cryptographical strength
- Side aspect: Allows use of smartcards for Kerberos authentication
- Automation possible if a suitable PKI already exists.

PKINIT - Kerberos with X.509 Certificates

- Need for X.509-Certifikates for KDC and Kerberos Clients
- Notice the peculiarity:
  - KDC-Server-Certificate needs extendedKeyUsage with OID 1.3.6.1.5.2.3.5
  - PKINIT-Client-Certificates needs extendedKeyUsage 1.3.6.1.5.2.3.4 and the attribute subjectAltName with the value of the principal name.
- Description of PKINIT within a MIT-Kerberos-Realm using OpenSSL:

http://web.mit.edu/kerberos/krb5-1.13/doc/admin/pkinit.html

Supercharging: Use of Puppet-CA for PKINIT

- Puppet already uses SSL certificates – why not use /var/lib/puppet/certs/$fqdn.pem ?
- Puppet-CA does not support extensions for extendedKeyUsage and subjectAltName
- Developed patches for the Puppet-CA
- Unfortunately rejected by Upstream:
  - https://tickets.puppetlabs.com/browse/PUP-4014
- We will look into an alternative implementation without patching puppet

Hadoop with Puppet

- We enhanced the Bigtop templates:
  - Supporting journaling, HA namenode
  - HA Yarn resource manager
  - Configuration of Hive on Tez
  - Configuration of Hue
  - Securing the Hadoop components and web interfaces
    - Zookeeper
    - Hadoop
    - Hue
- We introduced a role concept, which is not the one which is implemented in upstream Bigtop

Hadoop with Puppet

- The bigtop puppet kerberos support is not of production quality:

➡ It uses hardcoded passwords

- Upstream github edgester/kerberos module is now a drop-in replacement for the Bigtop kerberos class

Kerberos with Hadoop

- Setup
  - The principals are named: „hdfs/fqdn“, „yarn/fqdn“ …
  - Users are „olaf“…
  - All the daemons support it for authentication
  - Kerberos works mostly out of the box.
- Authentication:
  - LGTM!

Kerberos with Hadoop:

- Authorization:
  - HDFS, a bit clumsy since user -> uid mapping is done decentralized on each node.
  - Configuration of the NSS mapping is required
    - e.g. a directory service:
    - System users hive, yarn, mapred required

Kerberos with Zookeeper

- Zookeeper:
  - Supports ACL´s, but there is no tool to set ACL´s!
  - The ZK Root is left unprotected!
    - '/'
    - '/zookeeper'
    - '/zookeeper/quota'
  - Everyone authenticated can damage HDFS journaling!
  - Hadoop, Yarn sets ACL´s (++)
  - Hive does not set ACL´s in ZK (--)
    - '/hive_zookeeper_namespace'
  - Workaround: we created a tool to set ACL´s.

Noteworthy things:

- Parallel installation of Hadoop on all nodes.
  - Synchronisation with netcat on ports
  - Formats ZK
  - Formats Namenodes
  - Starts standby HA Servers
- Trocla: Do not store passwords in manifests / configuration files, no plain passwords stored.

Upstreaming

- Apache Bigtop:
  - Fixed Debian build support: Made it in Bigtop 1.0!
  - Automation and Configuration: only partly upstreamed.

- Debian:
  - All our changes are in Debian git
  - However, only one package made it into „unstable“
    - **puppet-module-asciiduck-sssd**

- Puppet kerberos module
  - Fixes are upstream except for the use of trocla

- IT WAS A GREAT EXPERIENCE!

Upstream Fixes needed

- Hive: Must protect the hive root in the ZK with ACLS!
  /hive_zookeeper_namespace
- Zookeeper: Should secure the ZK Filesystem
- Hadoop/Bigtop: change daemon scripts to better support the systemd init replacement

- Some projects did not work with a HA Yarn RM setup:

- Sqoop/Sqoop2
- Oozie

- Hue-3.8.0 does not work with the tez jobmanager/timeline

Wrapup

- None „yet another deployment tool“ needed
  - The generic system administration tools are far more advanced with respect to enterprise grade functionality:
    - Master/Slave Kerberos
    - Multimaster OpenLDAP
    - AD Integration
  - The concepts presented can be integrated in ansible, saltstack and many more. Be opinionated!
  - Complexity can be reduced by reusing proven technology
- Kerberos Support in Hadoop is quite good
- Upstream!

Demonstration

- Life Demo

Thanks !

Dr. Olaf Flebbe / Michael Weiser
science + computing ag
www.science-computing.de

Telefon: +49 7071 9457-0

E-Mail: oflebbe@apache.org