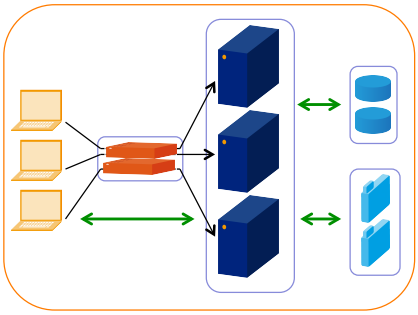
IdP Clustering

High Availability and Load Balancing



SWITCH

SWITCHaai Team
aai@switch.ch

You want to prevent service outages

Possible problems:

- HW failures
  - Server component failure
  - power failure.
  - Network failure
- Software failures
- Service overload
- Downtimes due to maintenance (major upgrades)
- …

What you usually do

- Take one box
  - Harden it through redundant components (power, network, disk, memory, CPU's, backplane (?))
- Or take another box
  - Organize failover (cold standby)
- Or take a couple of boxes
  - Organize load balancing
- ...

Challenges with the IdP

- The setup of the IdP and the whole environment is more complex than with a single-server IdP.
- Special configuration of the IdP is required.
- Load balancing requires special hardware or software.
- IdPs in SWITCHaai store some data in a database. Therefore, clustered IdPs need some kind of clustered database or some replication mechanism.

Stateful or not?

For stateless systems, building a cluster "is easy".
But: The IdP is stateful, in general.

- *Conversational state:* Short-term session during login process
  - Managed outside of the IdP software
  - Requires sticky sessions on load balancer
  - At present, there is no solution provided to replicate this data
- *Non-conversational state:* Data the IdP stores
  - Managed by the IdP software
  - Examples: Persistent ID, User consent data, IdP User Session
  - The IdPv3 provides flexible mechanisms to store such data, e.g. in the client or in a common database, so that the data is available to all nodes.

Storage Recommendations

Storage Entity	Recommended Storage	Scope
Persistent ID	***Common Database***	**Cluster**
User consent	***Common Database***	**Cluster**
IdP User Session	**Client**	**Per Client**
Transient ID	*Common Database*	Cluster
SAML artifact	*Common Database*	Cluster
Conversation Session	Memory	Per Node
Message replay cache	Memory	Per Node

Remarks:

- "Common Database" means some central/clustered database or a database replicated between nodes.
- SAML artifact:
  Irrelevant if SAML 2.0 artifacts are not used/required at all
- Alternatives for Message replay cache:
  Common Database or memcached (depending on security requirements)

Secret key management for cookie encryption

- The IdP User Session is stored in an encrypted cookie in the browser. The key to encrypt/decrypt this cookie should regularly be rotated. In a clustered setup, all nodes need to share the same key. It's recommended that one node generates a new key and copies it to the other nodes.

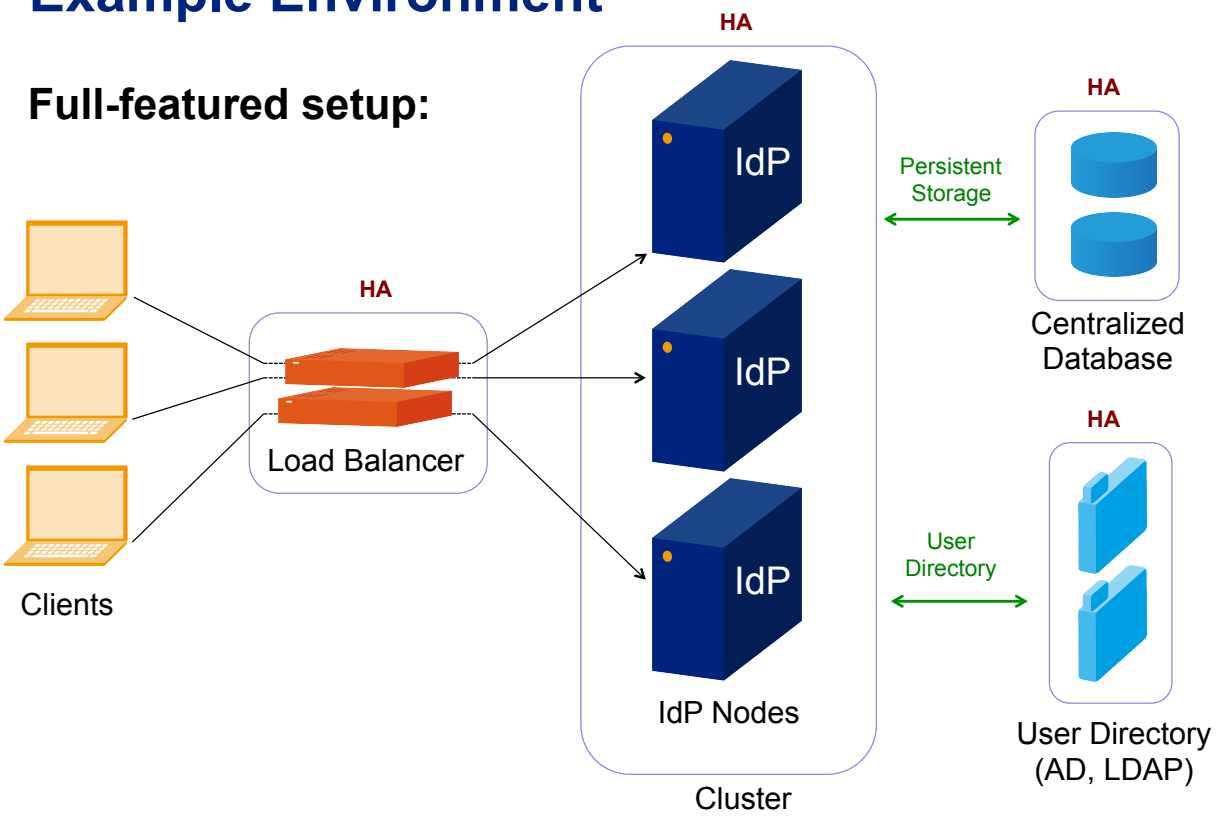
- Setup:
  - Decide for a node that is responsible for generating the secret keys and copying them to the other nodes.
  - Install an appropriate cronjob.
  - Our guide "Shibboleth Identity Provider Clustering" describes the details and shows an example cronjob script: https://www.switch.ch/aai/guides/idp/clustering/

Examples

Who	Network	Processing	Persistent storage
Uni Bern (IdPv3)	NGINX (active-active) HTTP Loadbalancer	2 IdPs	Use of central MSSQL-cluster
Uni Genève (IdPv2)	F5 BIG-IP Loadbalancer (sticky)		MySQL DB Cluster
Uni Lausanne (IdPv2)	HW load balancer (active-passive)	2 IdPs (active-passive)	external MySQL-DB (also HA: Heartbeat + DRBD)
Uni Zürich (IdPv2)		3 IdPs	external MySQL database
HES-SO Fr (IdPv2)		2 IdPs active-active	
Uni Marburg (IdPv2)	NGINX Loadbalancer	2 IdPs, memcached,	1 external PostgresDB server
SWITCH (IdPv2)	Anycast address	2 IdPs active-passive	Local MySQL-DB, replicated by cron

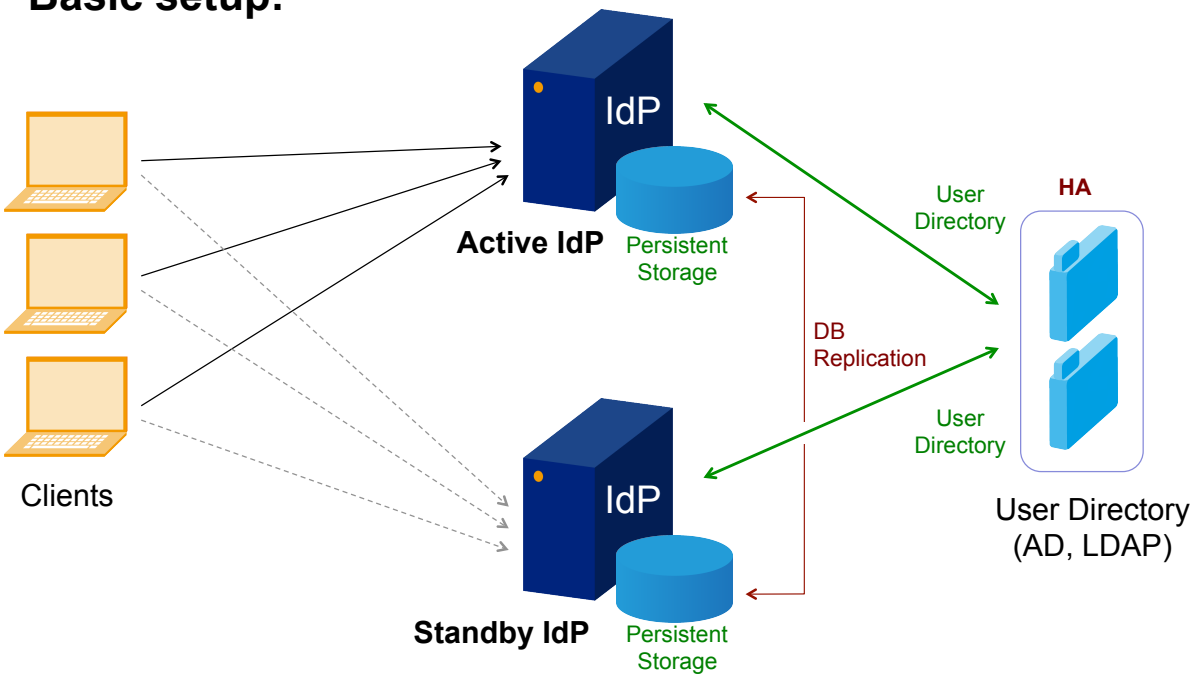
Example Environment

Full-featured setup:



Example Environment

Basic setup:



Considerations for planning an IdP cluster

You need to think about

- Which type of setup do you need?
- What kind of database do you need?
- Which additional hardware or software is required?
- Which further considerations are relevant for your organisation?

There are many mechanisms and options available to setup a suitable environment. The setup to choose depends on the requirements and the possibilities of your organisation.

References

Documentation from SWITCH:

- **Shibboleth IdP Clustering**
  https://www.switch.ch/aai/guides/idp/clustering/

Documentation from the Shibboleth Consortium:

- **Clustering**
  https://wiki.shibboleth.net/confluence/display/IDP30/Clustering
- **Storage**
  https://wiki.shibboleth.net/confluence/display/IDP30/Storage
- **Secret Key Management**
  https://wiki.shibboleth.net/confluence/display/IDP30/SecretKeyManagement