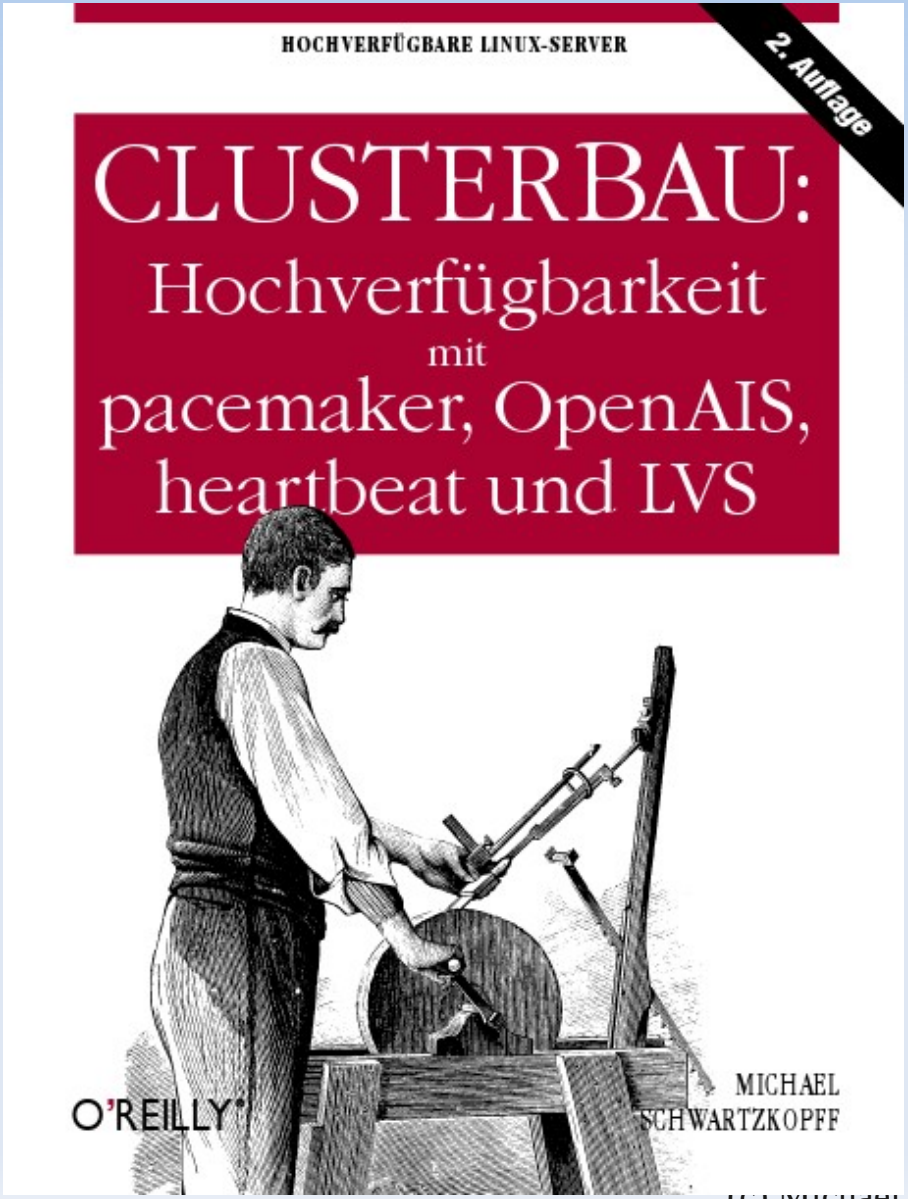
Quo vadis Linux-HA?

Developments in Linux Clustersoftware

Michael Schwartzkopff
misch@schwartzkopff.org

Background


HOCHVERFÜGBARE LINUX-SERVER
2. Auflage
CLUSTERBAU:
Hochverfügbarkeit
mit
pacemaker, OpenAIS,
heartbeat und LVS
O'REILLY
MICHAEL
SCHWARTZKOPFF

Linux Clustersoftware

- Linux Virtual Server
  - Scales nicely...
  - … but is not high available.
- Linux-HA
  - Offers high availability ...
  - … but does not scale.

  - Code is not maintained any more.
  - Is beeing replaced by a collection of other programs.

Linux-HA Version 1 (`heartbeat`)

- Two servers exchange *heartbeats*.
- If the standby server does not receive *heartbeats* from the active node any more it starts the services.
- Configuration in a plain text file *`haresources`*.
- Cons:
  - No monitoring of the services.
  - No management of the resources.

Linux-HA Version 2 (CRM)

- Mangement of the resources by a *Cluster Resource Manager* (CRM):
  - Aktive monitoring of the reosurces by the cluster software.
  - Up to 16 nodes in a cluster.
  - Resources free moveable between the nodes.
  - Communication in the cluster via `heartbeat`.
  - Constraints determine which resource should run on what node.

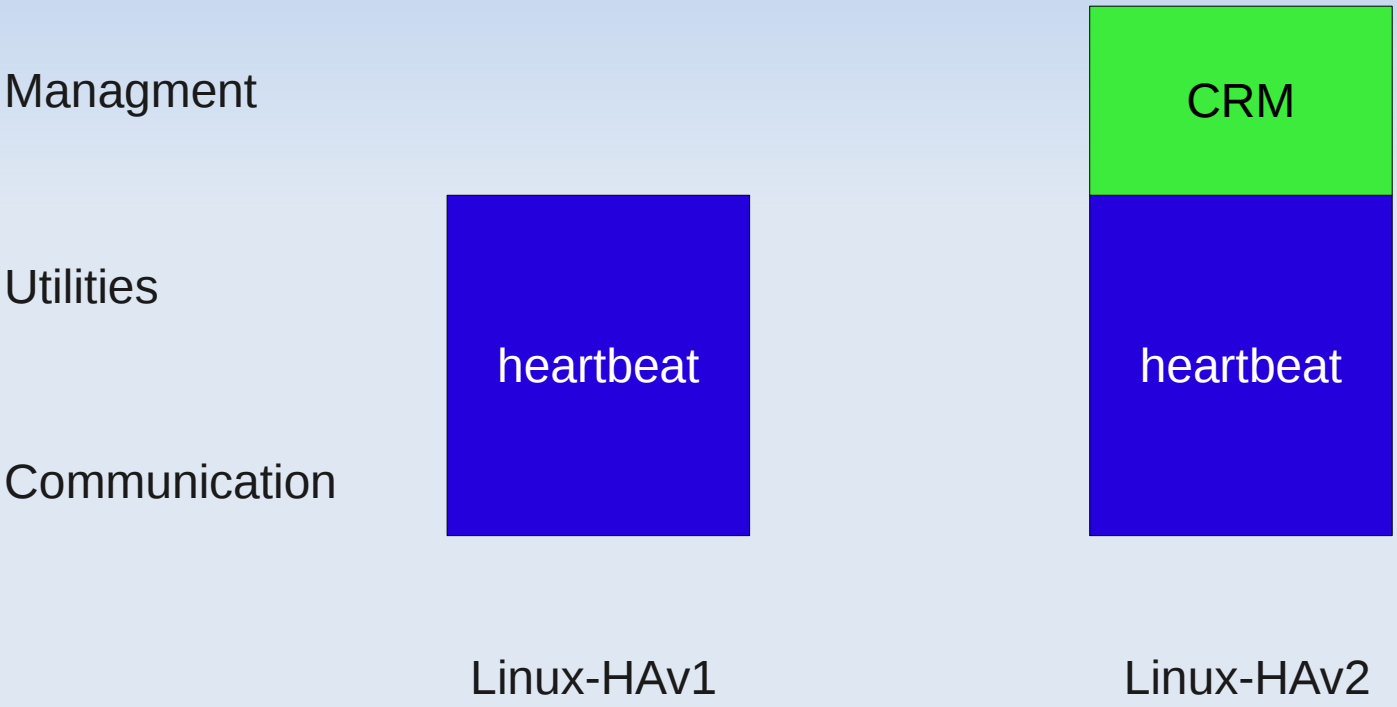
Constraints in Linux-HAv2

- **Order**
  A resource should start before or after an other resource.
- **Colocation**
  A resource should run on the same node as an other resource.
- **Location**
  A resource should run on a node with specific attributes.

Features of Linux-HAv2?

- Pros
  - Monitoring of the resources.
- Cons
  - Configuration in XML.
  - Administration via the command line.
  - GUI not really usable.
  - Not maintained any more after 2007 (version 2.1.4).

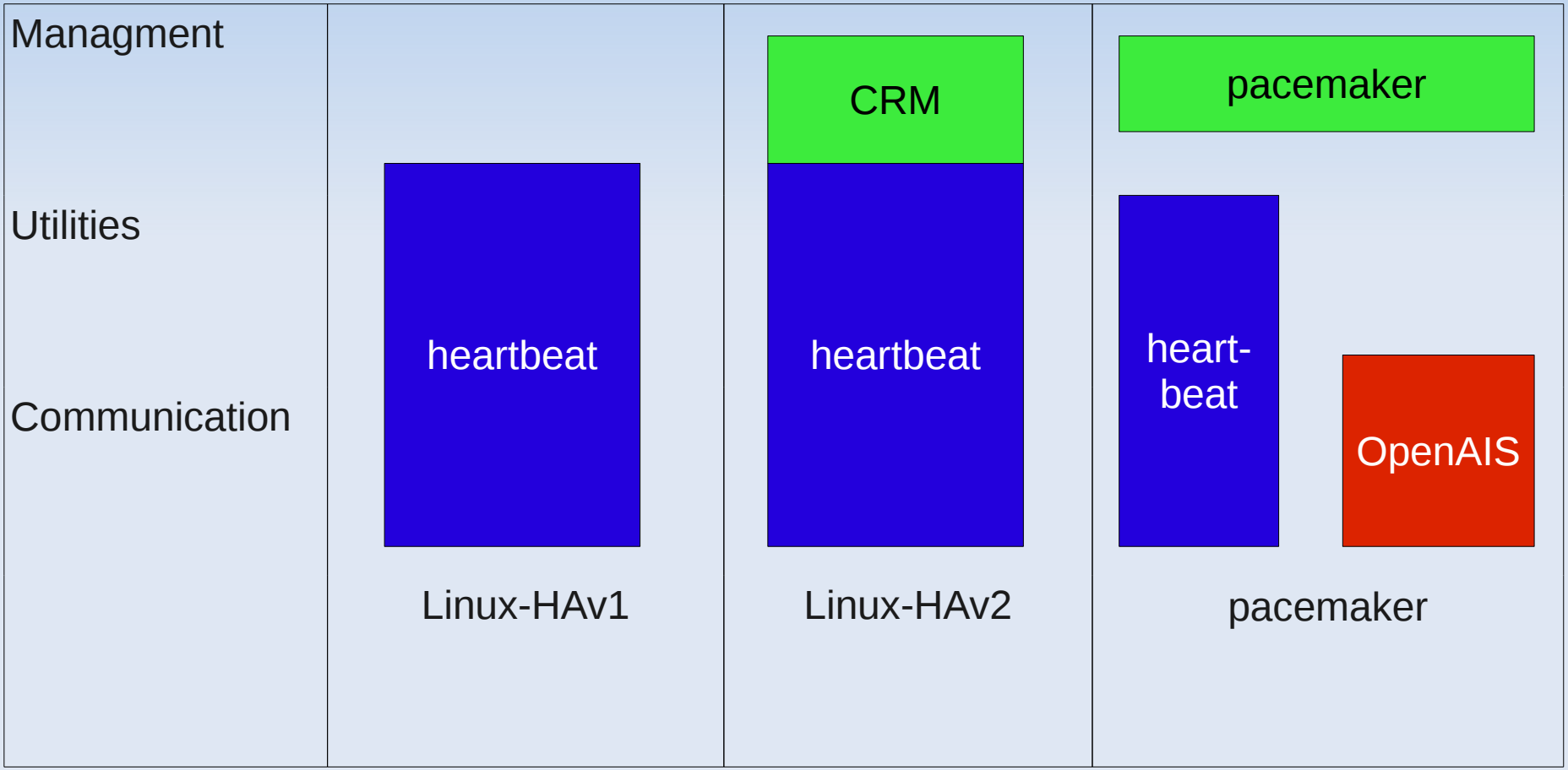
Linux-HA


Managment
Utilities
Communication
heartbeat
CRM
heartbeat
Linux-HAv1
Linux-HAv2

Restart

- If the CRM uses `heartbeat` only for the communicaton in the cluster, this could be done also by another program.
- The other software was `OpenAIS`.
- The developers organized the CRM code in a separate project called `pacemaker`. It uses either `heartbeat` or `OpenAIS`.
- Some utilities from the `heartbeat`-package were still nescessary.

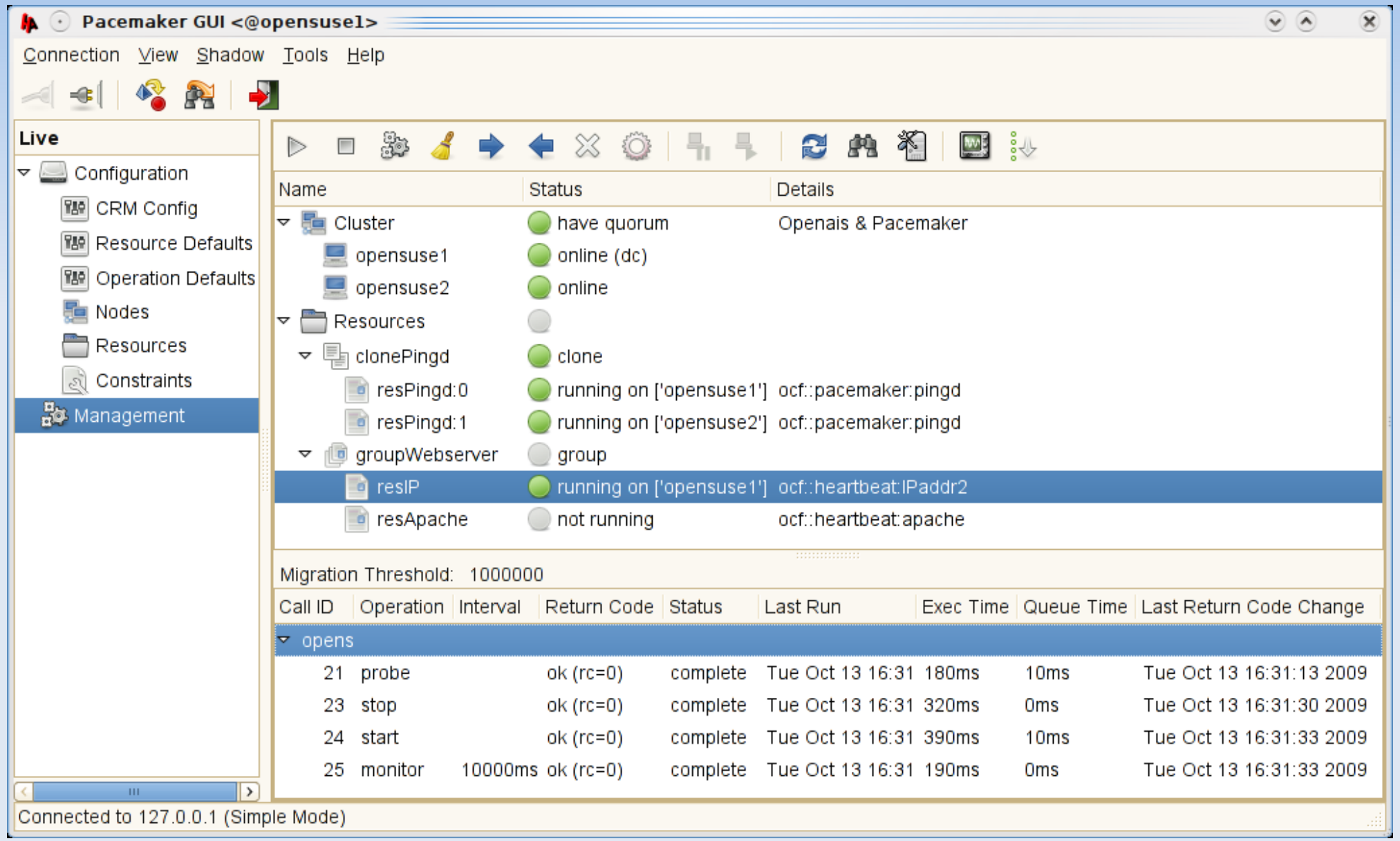
pacemaker


Managment
Utilities
Communication
heartbeat
Linux-HAv1
CRM
heartbeat
Linux-HAv2
pacemaker
heart-
beat
OpenAIS
pacemaker

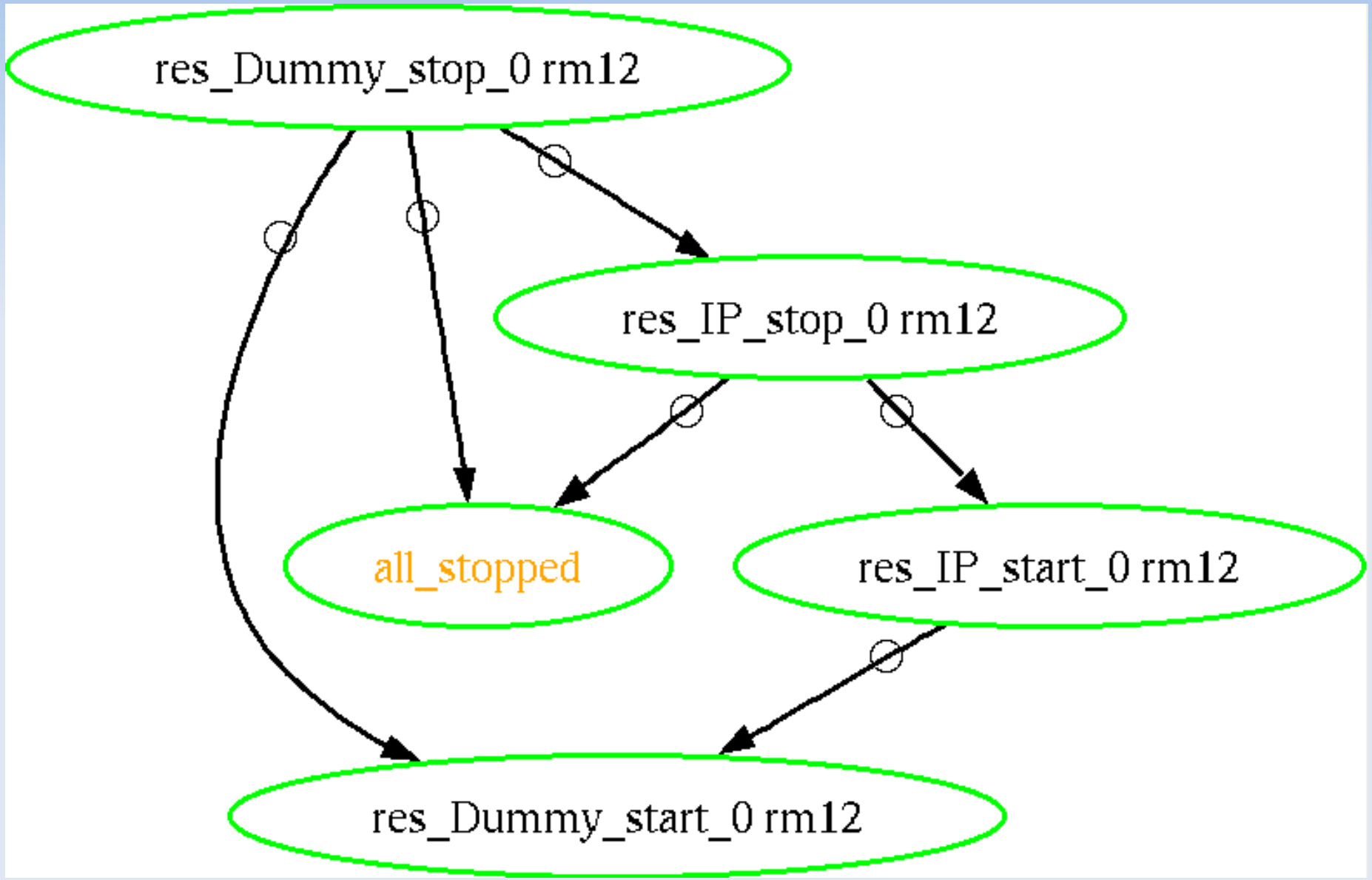
pacemaker

- Uses `heartbeat` or `OpenAIS`.
- Suitable GUI.
- Own subshell additional to the CLI.
- Sandboxes for testing.
  - First test what would happen, if ...
    → only afterwards activate the new configuration.
- Exact history, what happened why.

pacemaker: The GUI



Sandboxes with graphics!



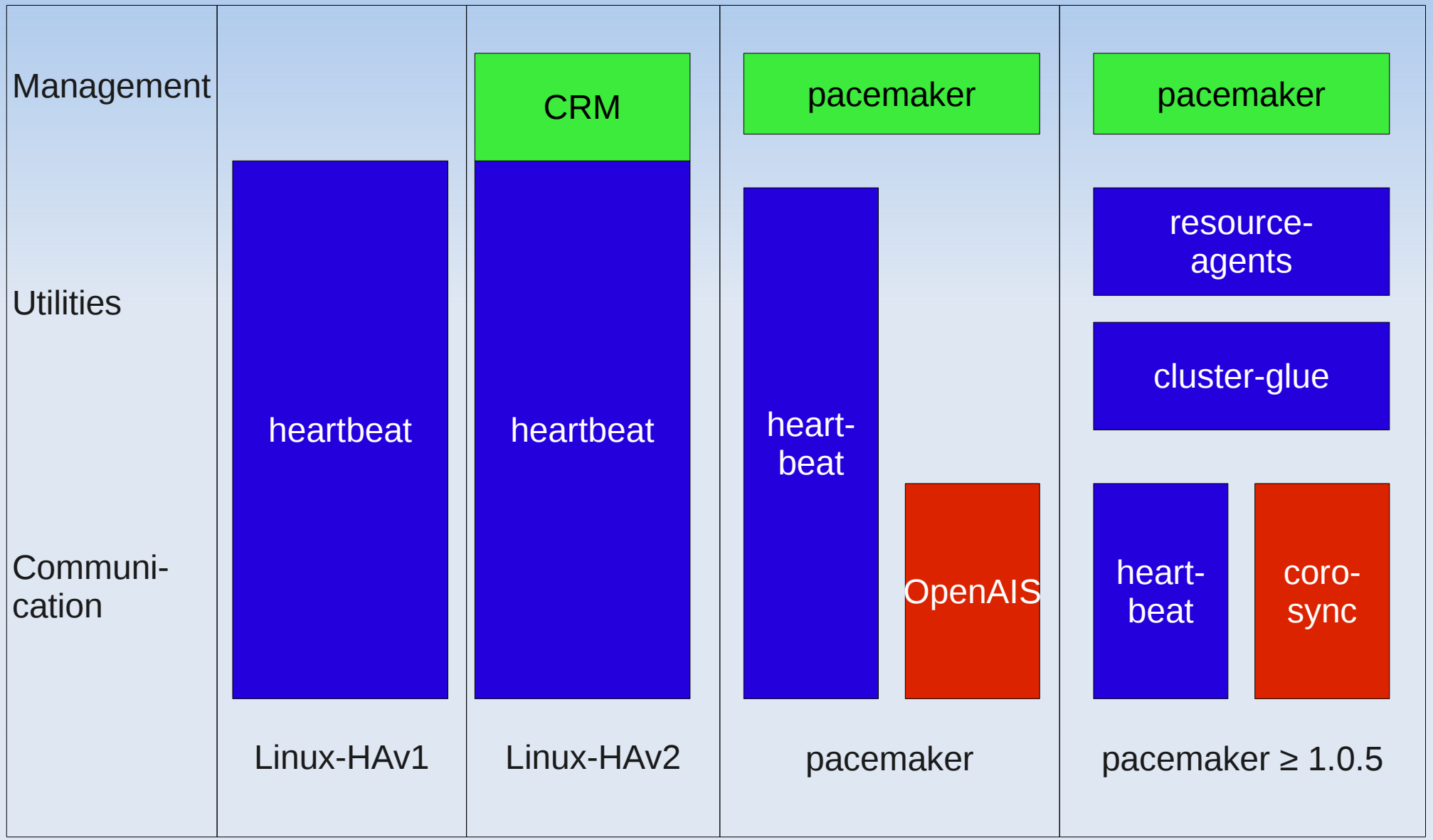
The development goes on

- OpenAIS offers much more as needed by pacemaker. The communication stack is forked into a separate project called corosync.
- corosync is responsible for the communication in the cluster.
- OpenAIS takes care for all ”higher” services in a cluster.
- Simple clusters with pacemaker only need corosync.

Decomposition of heartbeat

- With pacemaker 1.0.5, heartbeat is decomposed into three projects:
  - cluster-glue includes all nescessary utilities.
  - The package resource-agents combines all resource agents, the interface between pacemaker and the application's binaries.
  - In heartbeat (version ≥ 3.0.2) remains, what is left from the project.

HA-cluster with Linux



Master Control Process

- `pacemaker` in version 1.1 got a Master Control Process. So it can be started outside from `corosync` though the `init`-system.
- Of course, `corosync` has to be running.
- Causes less trouble.
- Is more stable.

DRBD Managment Console

- From company Linbit.
- Written in Java.
- Originally ment to manage DRBDs
- Now full blown cluster management tool:
  - Nodes, resources, constraints
  - Graphical representation of the configuration
  - Emphasis on virtual machines: Config & console.

Utilizations in pacemaker 1.1

- Every node has utilizations like CPUs, RAM configured.
- Every resource has utilizations configured.
- The cluster manager takes care that resources only run on nodes with enough utilizations.

- TODO: Make resource utilizations dynamic.

Role Based Access Model

- From version 1.1 administrators access to resources can be limited:
  - You can only destroy your own resources.
- Complete RBAC model implemented.
- Think about a provider offering HA services.

External Quorum Providers

- Quorum descisions are quite simple up to now.
- `corosync` can use external quorum providers.
- `cman` is a possible external provider.
  - For now it is the only one.

Beering

- Lavandevil
  Schusterusstrasse 3

  U Richard Wagner Platz: U7 / M45

Thank you very much for your attention!

Questions?