NDGF Site Report

2017-04-24, HEPiX Spring 2017, Budapest

Mattias Wadenstein

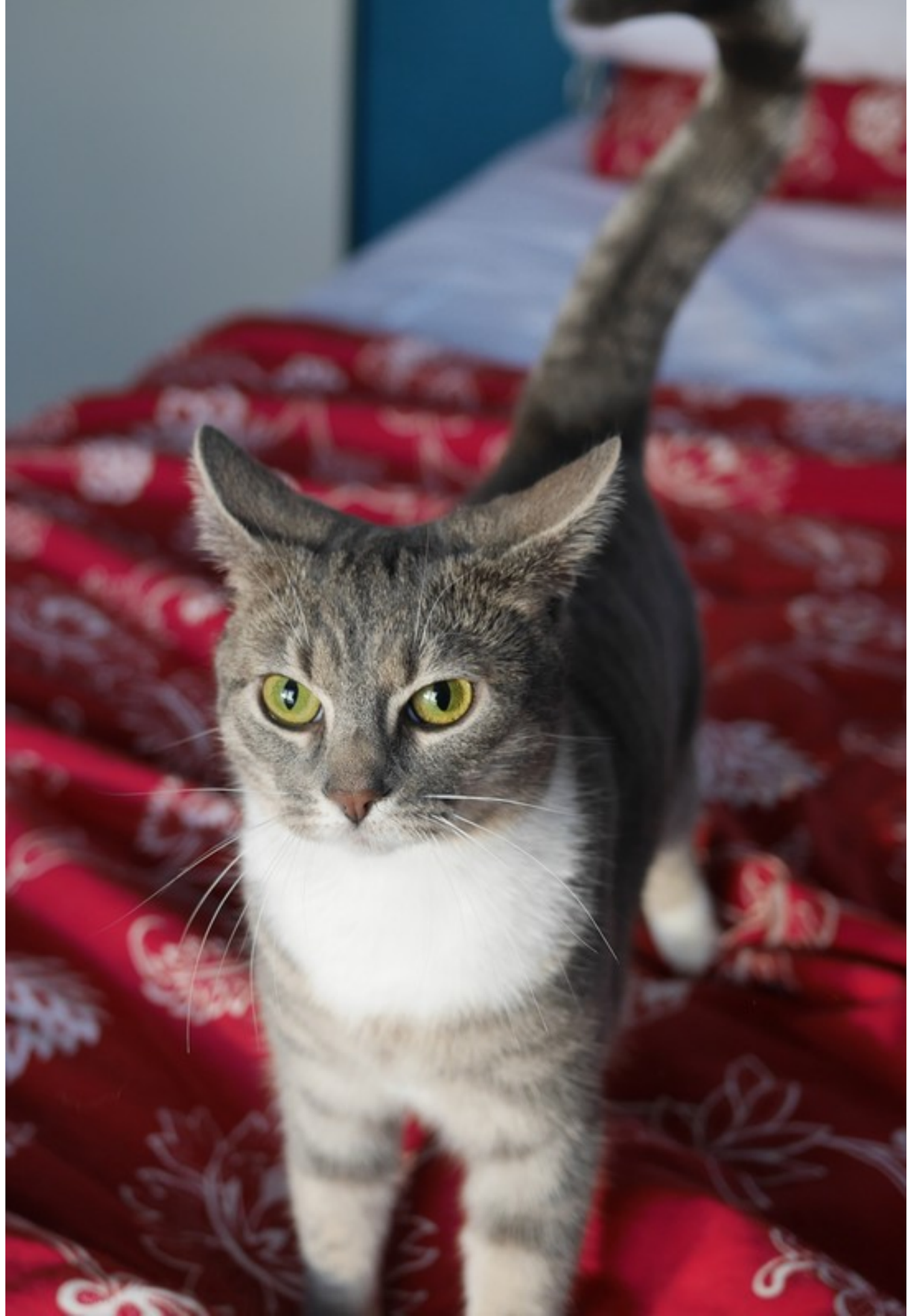
norden

NordForsk

Nordic e-Infrastructure Collaboration

Overview

- NDGF overview
- Staff changes
- HA dCache updates
- Site reports
- NeIC 2017



NDGF Overview

- Distributed WLCG tier-1 site
  - 6 Nordic academic HPC sites with dCache pools and ARC-CEs
  - And disk from IJSs T2 in Slovenia
- Supports ATLAS and ALICE
  - Targets: 6% of ATLAS tier-1 & 9% of ALICE tier-1 resources
- Thanks to the diskspace consolidated into the tier-1 from the tier-2s SI-SIGNET-T2 and SE-SNIC-T2 we currently have second largest ATLASDATADISK among all tier-1 sites
  - Could consolidate more disk in here, if you want. :)

Staff changes as NDGF-T1

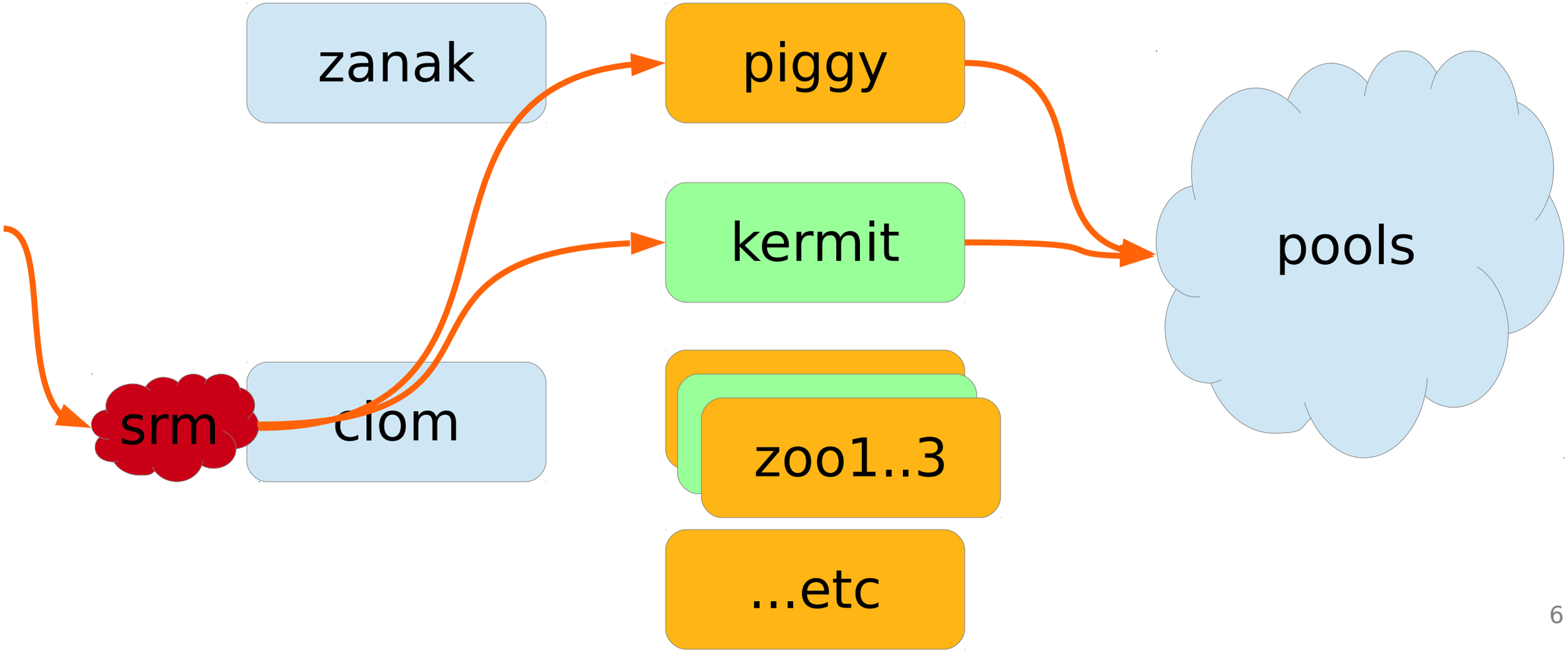
- New dCache developer starting next week, Vincent Garonne, who some of you might know from ATLAS DDM (rucio etc).

HA dCache upgrades

- With the new HA features in dCache we can do system updates including reboot into new kernels with no downtime
- Can typically be done in a day, but takes a bit of watching to make sure we don't interrupt any client accesses
- Can also do dCache upgrades of headnodes without anyone noticing, over a couple of days
  - Unless something goes wrong, of course
- Hardware and headnode upgrades on different days
  - Headnode upgrades depends on haproxy draining state – this is reset by a reboot of the hardware that runs haproxy

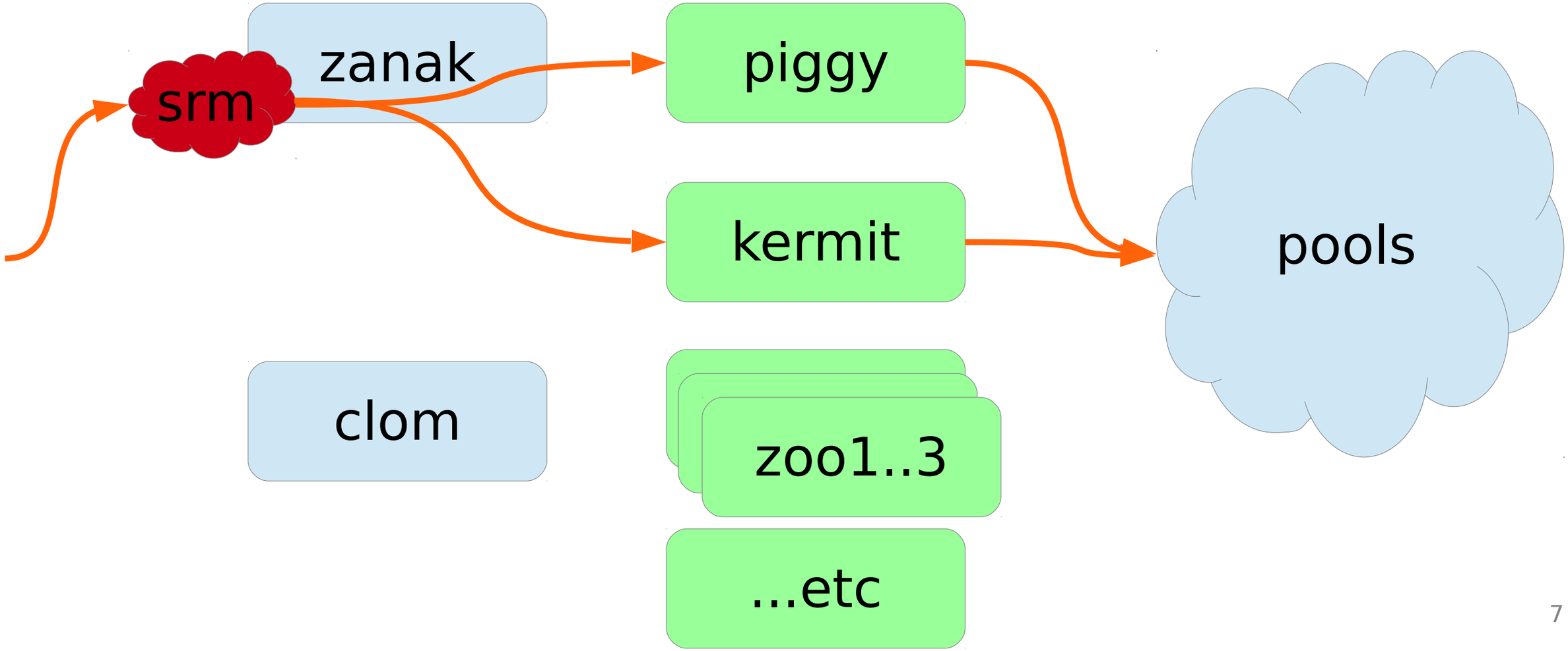
HA dCache setup at NDGF

- Yellow are VMs on clom, green on zanak
- “srm” refers to a virtual IP handled by haproxy, srm.ndgf.org



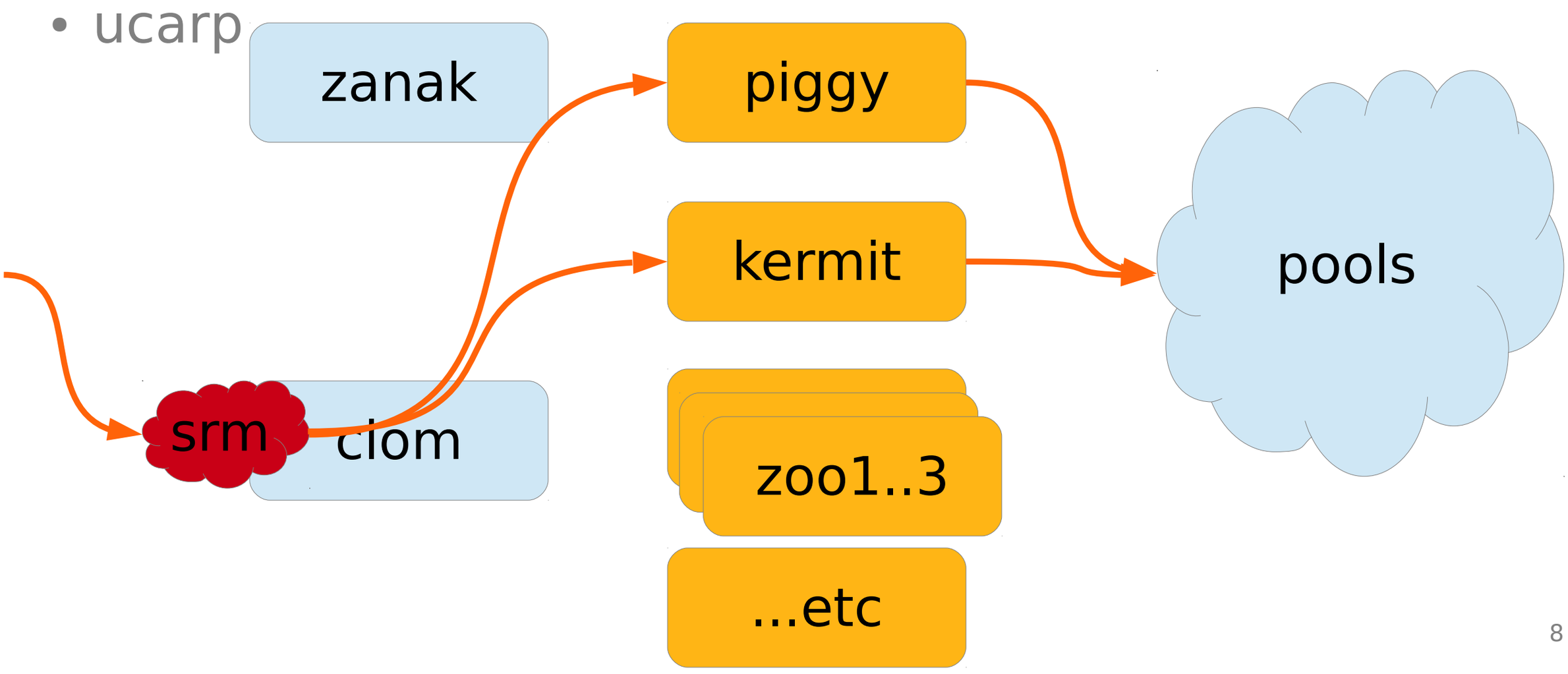
Rebooting hardnodes

- Migrate all the VMs to zanak
- Do repmgr failover of postgres to zanak
- Do ucarp failover by means of signals



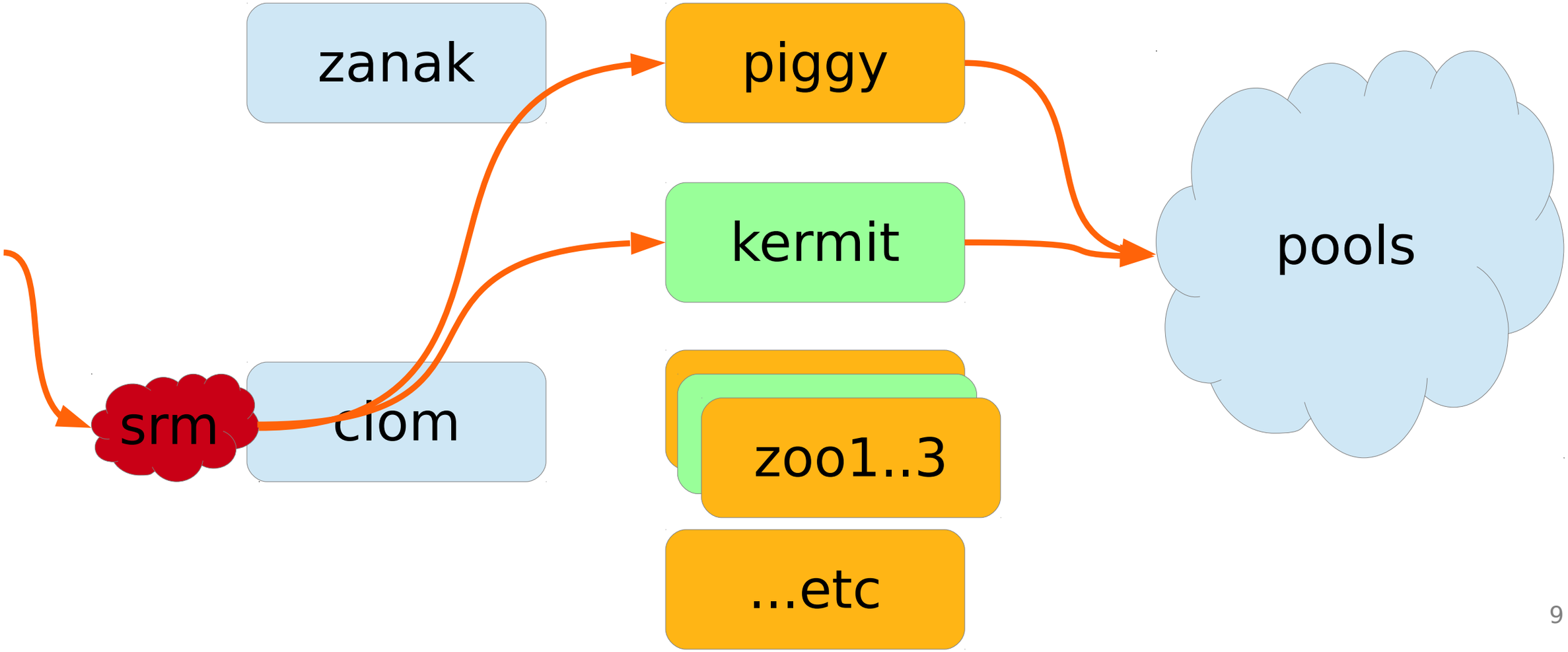
Rebooting hardnodes

- Reboot clom
- Migrate VMs to zanak
- Postgres repmgr failover to clom
- ucarp



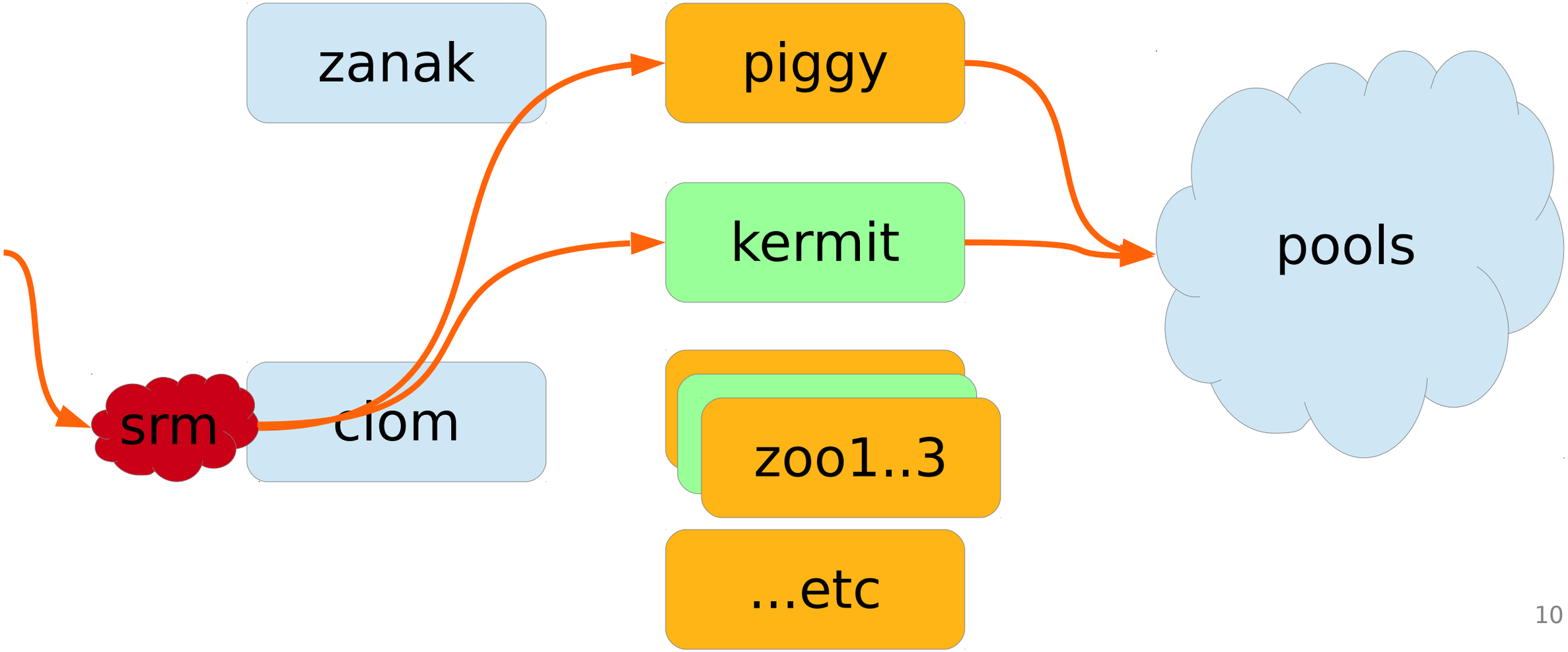
Rebooting hardnodes

- Migrate back to our standard distribution of VMs
- All done for the day, with no user-visible disruption



Upgrading headnodes

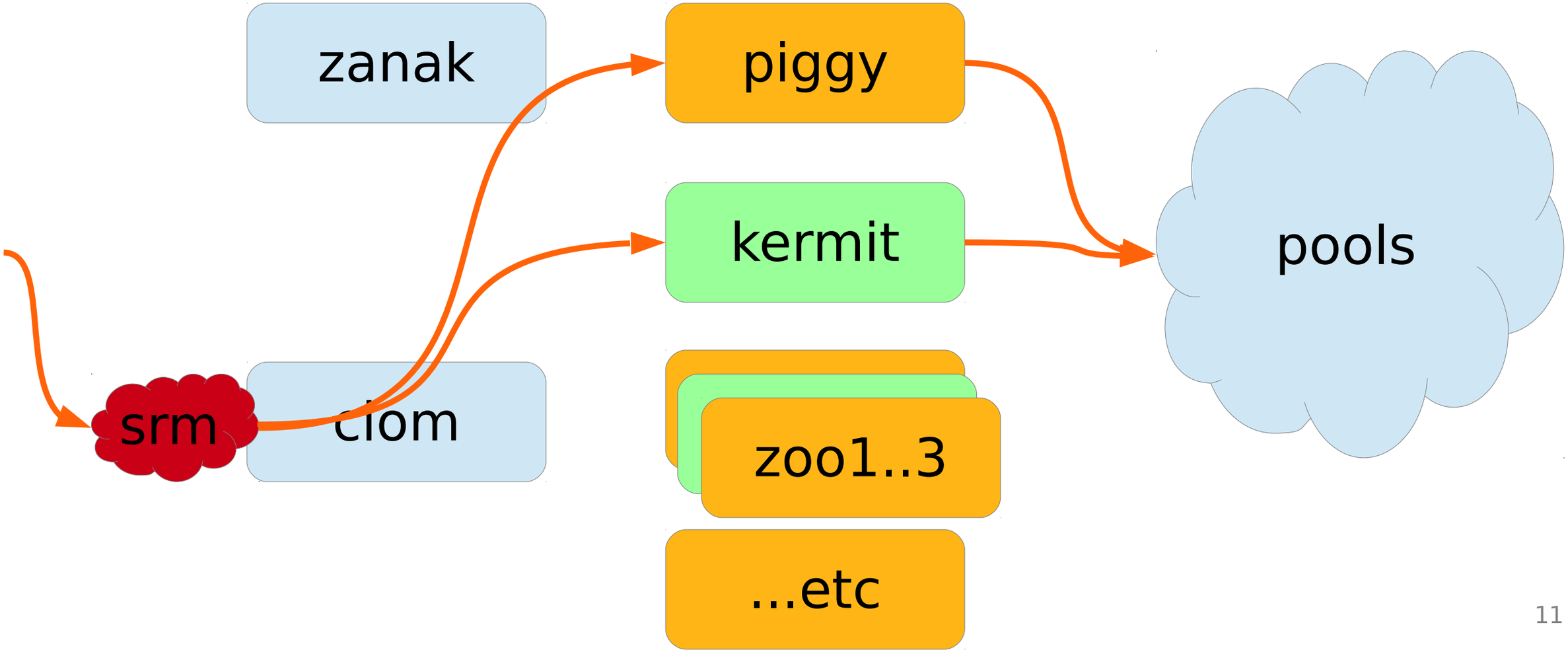
- Reboot or do a dCache upgrade on a headnode (piggy)
- Also without disruption of service
- Preferably not done on the same day as hardnodes



Upgrading headnodes

- Disable the frontends from getting haproxy requests:

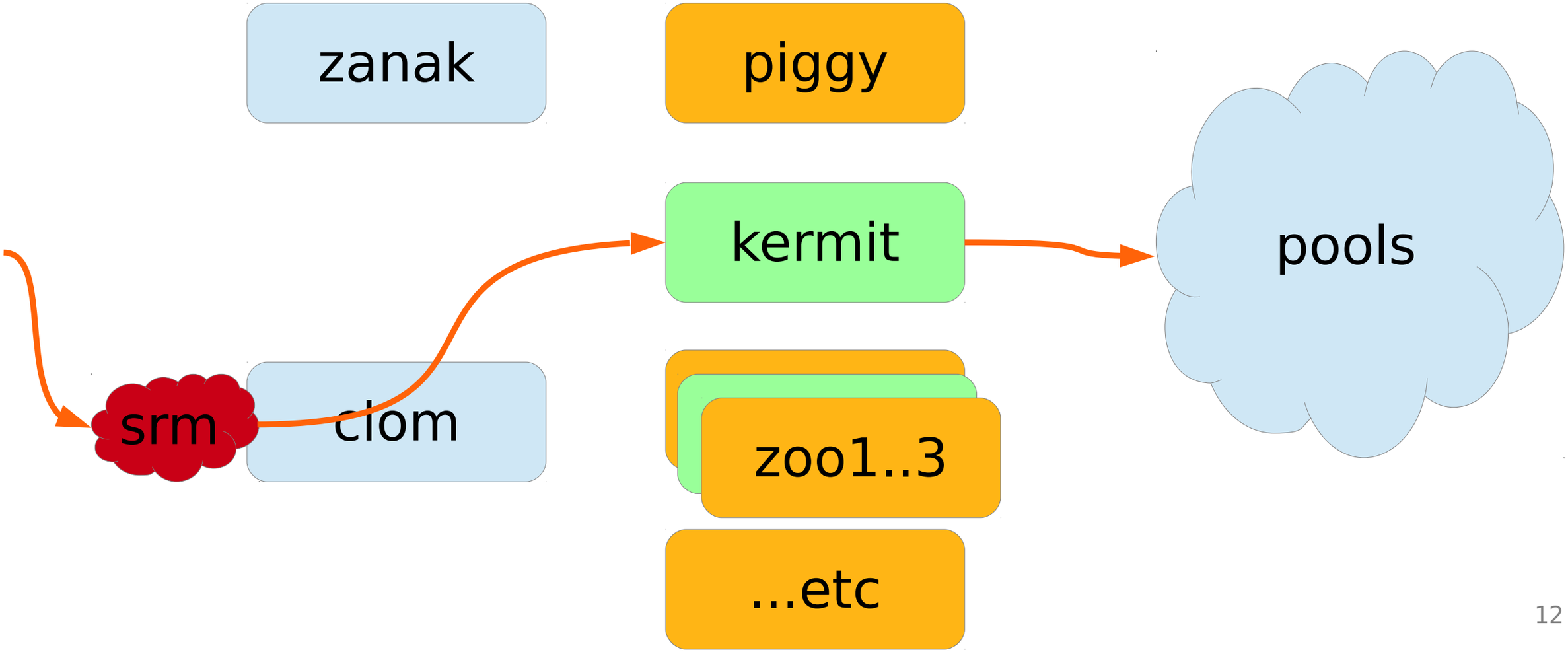
```
echo 'disable server xrootd-alice/piggy' | sudo socat /run/haproxy/admin.sock -
echo 'disable server srm/piggy' | sudo socat /run/haproxy/admin.sock -
echo 'disable server gsiftp/piggy' | sudo socat /run/haproxy/admin.sock -
```



Upgrading headnodes

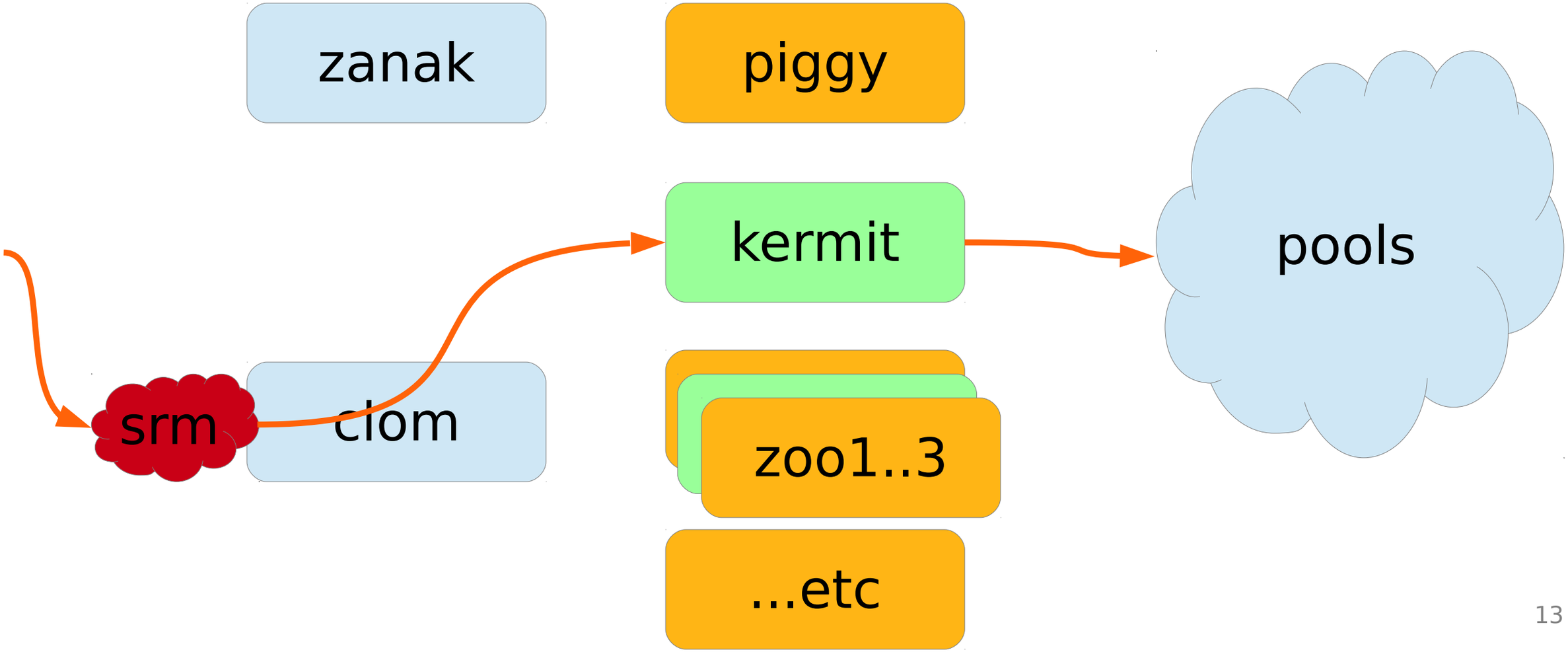
- Disable piggy internally in dCache for SRM turls

```
\s FTP-*-piggy lb disable
\s WebDAV*piggy lb disable
\s Xrootd*piggy  lb disable
```



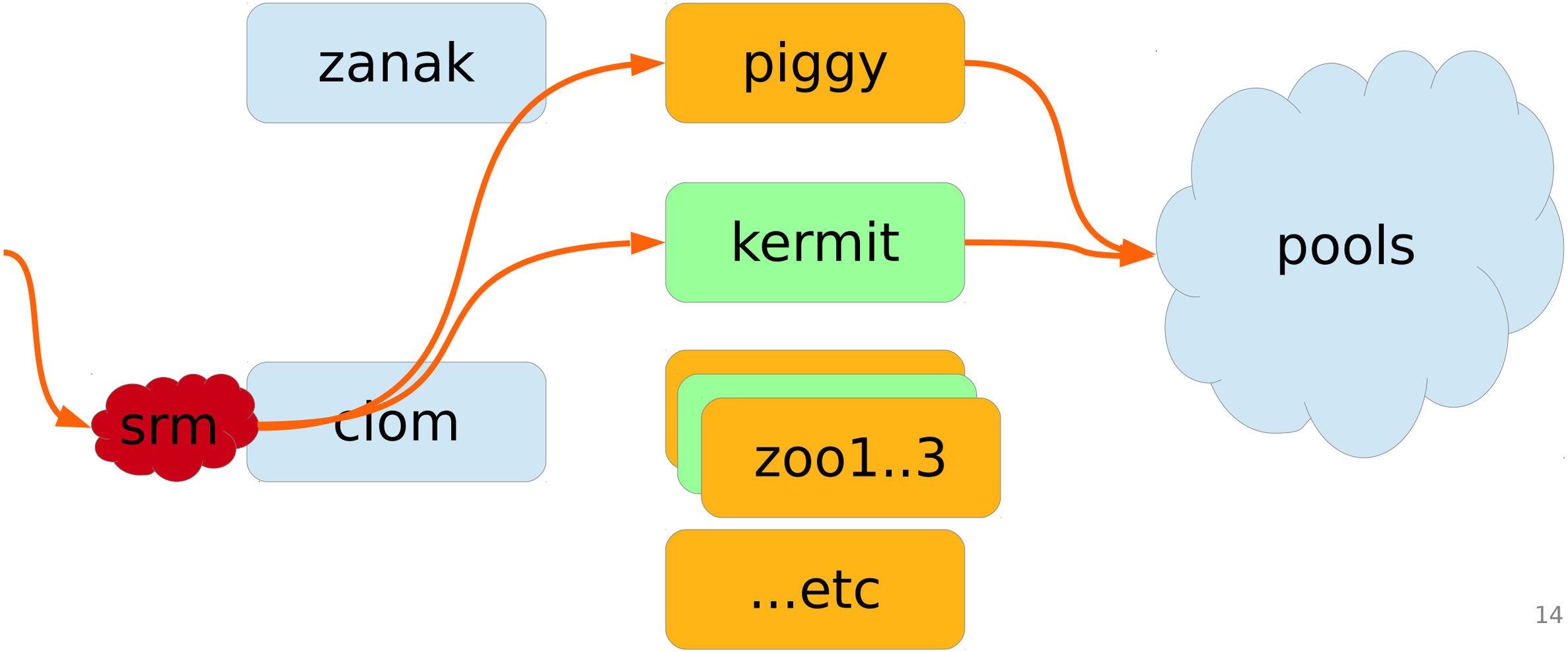
Upgrading headnodes

- Repoint DNS for gsi-xrootd door (old clients still aronud)
- Wait approx 24h (default SRM turl lifetime)
- Check in dCache and haproxy for 0 requests to piggy



Upgrading headnodes

- Upgrade dCache or OS, including possibly reboot
- Enable piggy in haproxy (lb resets with dCache restart)
- Repeat for kermit



Upgrading headnodes

- This is how we have done upgrades of both dCache and OS (including kernel updates with reboots) for the last 6 months
- And we’re running on Ubuntu that releases kernel patches early and often
  - A couple of reboots per month, on average
- No user inconvenience
- No need for planning or scheduling downtime

HPC2N: Resource news

- Bought more dell 730xd dCache pool servers
  - 8 servers, same specs as last report
- Upgrading the Abisko cluster from Ubuntu 12.04 to 16.04
  - Very much in progress this week
  - Will include singularity to provide ATLAS-compatible OS image

HPC2N: Recruiting news

- As advertised at last HEPiX
- In case some of you have more staff than funding
  - Or someone wants to move for other reasons, better to keep them in the community than losing them to industry...
- One full-time HPC sysadmin
  - https://www.hpc2n.umu.se/node/627
  - Exact focus depends on the candidate, looking to enhance the team, not fill a specific gap.

HPC2N: Performance regression on HPE raid controller

- When updating the dCache tape pools to a newer Ubuntu version, we discovered a severe performance regression
  - 1.8GB/s -> 0.6GB/s sequential IO
  - https://bugs.launchpad.net/ubuntu/+source/linux/+bug/1668557
- This applies to anything with a change that came with mainline kernel 3.18.22
  - Linux now respects what the driver/HW says for max_sectors_kb
  - But the hpsa cards we have access to say 4096, but perform way worse with max_sectors_kb > ~1k
  - Workaround udev script that sets it to 512 linked in above bug
  - For those of you with RHEL-derivatives, 3.10.0-327.36.3.el7.x86_64 is reported to be the last fast version in CentOS7

IJS: New stuff

- Additional ~P for dCache pools (currently at 3.7 PB, 10 servers)
  - 7 for ATLAS (in srm.ndgf.org), 3 for Belle (in dcache.ijs.si)
- CEPH used for cluster cache at SiGNET (0.7 PB, 9 servers), very good perfomance for cephfs as ARC cache, plans to implement on other sites in Slovenia
- Network upgrades for LHCone to 20 and 30 Gb/s for the two main clusters - important because of dCache pool usage

IJS: Singularity

- Pioneering and ongoing work on Singularity use for ATLAS and as generic container for our clusters
- Switched to using Singularity environment for executing jobs (from previous schroot based execution)
- Used singularity to enable the common cluster at Jozef Stefan institute for ATLAS jobs - IPv6 only! - and potentially include more resources in Slovenia

IJS

- Continue moving computing resources to the new centre at JSI Institute (~200 m2)
- Waiting for AMD to release new processors to acquire some more cores
- Progress with the National Supercomputing Network / NGI in Slovenia and hopes for more regular funding and more people - things are looking up!

NSC: Funding problems

- Re-organization of Swedish funding agency SNIC.
- => Even more chaos than before...
- Funding for replacement of current academic cluster delayed.
- Funding for running current academic cluster reduced.
  - Cluster shrunk from 1594 to 1017 worker nodes.
- WLCG part still at 52 nodes, though.
- Funding for replacement cluster finally in place!
  To be installed in 2018. Procurement work started.
-

NSC: Burning PDUs

- Total of four PDUs in current weather forecast system burning and losing power to rack.
- Not very well designed internally.
- Supplier has now replaced all PDUs with revised (better) design

NSC: Met.no and SMHI

- Agreement with Norwegian meteorology office (met.no).
- Second cluster for Swedish & Norwegian met-offices at NSC. Already had one for SMHI. Second cluster previously in Norway.
- Built from leftovers of academic cluster
  Plus new Lenovo GSS, and some new Infiniband switches.
- In a separate computer room from existing one, with separate power feed.
- To be in production in May.

NSC: More disk for WLCG (NDGF-T1)

- 3 × HP DL360 with HP D6020 external SAS box and dual RAID controllers. 40 Gbit/s ethernet.
- 70 × 8 Tbyte disk in each.
- Ordered two weeks ago; not yet delivered.

NSC: New SUNET (Swedish NREN) network

- Dual 100 Gbit/s connections to university (previously dual 2×10 Gbit/s)
- NSC now has logical systems in SUNET CPE routers for LHC-OPN.
- Finally gotten around to have proper redundant OPN connection with BGP. (Previously had single wavelength to Stockholm, running layer 2, with default router for NSC net in Stockholm.)

NSC: Power outage

- Monday March 6th, 07:30
- Excavator cut a cable, large parts of city lost power. Including both NSC computer rooms, despite separate power feeds.
- Power restored at 08:05.
- Storage systems (including dCache pools) on priority UPS power, so survived, but all worker nodes in all clusters went down after 4 minutes.
- Rack with system servers in met-office cluster incorrectly cabled after replacing bad PDUs. Worker nodes got priority UPS power, while system servers lost power after 4 minutes...

NSC:Another power outage, part 1

- Tuesday March 28th, 02:30
- Detector for electric arcs in power substation triggered
- Power lost for one of our computer rooms; the one with our large academic cluster, and current weather forecast system.
- Power company could not find reason for detector triggering, and restored power 03:55.
- One sysadmin still awake when power was lost, went onsite and brought systems online when power was restored.

NSC:Another power outage, part 2

- Wednesday April 5th, 16:45
- Detector for electric arcs triggered
- Same detector as last time!
- Power restored and systems brough online again after a couple of hours.
- Still no sign of actual arcing. Suspected bad detector. Power company disabled detector.

NSC: Another power outage, part 3a

- Thursday April 6th, 06:50
- Detector for electric arcs triggerd. This time a different detector. Also saw overcurrents in upstream power station.
- After several hours of measuring and thinking, electrician crawled into space around power rail.
- Found nut and bolt supposed to hold the power rail, below it. Also saw sooting.
- This space was not designed for servicability (difficult to access, no access holes for fiber cameras).

NSC: Another power outage, part 3b

- Power company/landlord ordered diesel generator for us (and some others at the university).
- Took long time to arrive (normal truck driver on vacation...)
- Arrived ca 18 o'clock.
- Two 450 kVA (~360 kW) diesel generators. Should be enough for current load in our computer room, barely.
- Power central in our computer room not really designed to attach external power source.
- Had to pull cables through open door, and find some nuts and bolts for power rail to attach to.
- Could only attach one of the generators.
- Power restored ca 20:00. Met-offices finally able to run their weather forecasts...
- Climate research and academic cluster run at <25% capacity.
- Some UPS batteries died, including giving off smoke, when they started getting charged again.

NSC: Another power outage, part 3c

- Monday April 10th: Power company managed to repair substation.
  Could turn off diesel generator and get real power again.
- Wednesday April 12th: Scheduled power outage to be able to disconnect diesel generator. Academic and climate research worker nodes drained, but weather forecast nodes running on our (battery) UPS.

NSC: Another power outage, part 3d

- The lose nut and bolt was on the customer side in the power substation. Power company wants to inspect the incoming side as well to see if the same fault exists there. Need to schedule a power outage for this, sometime in May.

NeIC 2017 Conference Umeå 2017-05-29 – 2017-06-01

- 2 days of workshops
  - dCache, sensitive data, data management, KNL programming, ...
- 2 days of conference
  - On the e-Infrastructure needed to enable science
    - Sessions on: Society, Science, Technology, Industry, Online computing
- Excellent location :)
- More info on: http://neic2017.nordforsk.org

Questions?