Pine64: Synapse on a SBC Cluster

Matthew Petry
@fire219:matrix.org
(or @fire219:pine64.org)

Agenda

- The Two Sides of Pine64
- The RockPro64 Cluster
- Synapse on the Cluster?

The Two Sides of Pine64

(how to make FOSS-friendly devices)

Pine64

- So what is it?
- According to Wikipedia:

Pine64

From Wikipedia, the free encyclopedia

Pine64 is an organization which designs, manufactures and sells single-board computers, notebook computers and smartphones.

- I guess that's not really wrong, but there's a bit more to it...

Pine64 Hardware

- Do one thing and do it well: Make FOSS-friendly ARM hardware at lowest possible cost

Pine A64

Pinebook Pro

PinePhone

Pine64, a bit more in depth

- Rather unusual: zero employees!
- Everyone within the organization is either a contractor, or volunteer.

Two distinct entities:

- Business: Pine Store Ltd. *(Manufacturing, Logistics, Sales, Official Support)*
- Community: Pine64 *(Main website, forum, wiki, chat rooms, social media)*

It takes two to tango

- Hardware is nice and all, but it's not much good without the software to run on it!
- Most people watching this probably agree: Hardware businesses and FOSS usually don't mix well
  - Official firmware often full of blobs, undecipherable without access to NDA-gated documentation

Separation of Responsibilities

- The solution is rather simple: **have the community handle software development!**
- In short:
  - Business → Pine Store Ltd. → Hardware
  - Community → Pine64 → Software
- The two sides run independently, but work together to make a complete experience.

Here's a couple more benefits...

Community Software is Diverse Software

- As an example, the PinePhone has firmware images available spanning several Linux distributions *(PostmarketOS, UBports, Mobian, etc.)*

- Shown Right: PinePhones running three different desktop environments – *Phosh, Lomiri (Unity8), and Plasma Mobile*

Community Software is Up-To-Date Software

- Many embedded platforms end up relying on (semi-)proprietary software
  - BSPs (Board Support Packages) built around a specific Linux kernel version, with blobbed modules
  - Very quickly end up obsolete: missing out on performance, feature, and security upgrades

- Businesses update software when it'll make them money, enthusiast communities update it just because they want the latest and greatest.

It’s not a one way street…

- At this point, you might think this is all just a ploy to get free labor
  - *Software R&D departments are expensive, after all*

- But here’s the other side of the deal: **The community (especially developers) get to shape the future of Pine64 devices.**

- This can be as simple as polling preference between spring-loaded or friction-fit microSD card slots, or influencing an entire hardware development cycle…

The PinePhone development cycle

- Phase 1: “Project Anakin”
  - Basically just a bunch of off-the-shelf Pine64 products put together with an included USB cell modem
  - Given to developers late 2018

The PinePhone development cycle

- Phase 2: Project “Don’t Be Evil”
  - Hardware refined using feedback from Phase 1 developers, now into something actually smartphone shaped
  - Given to developers in early 2019

The PinePhone development cycle

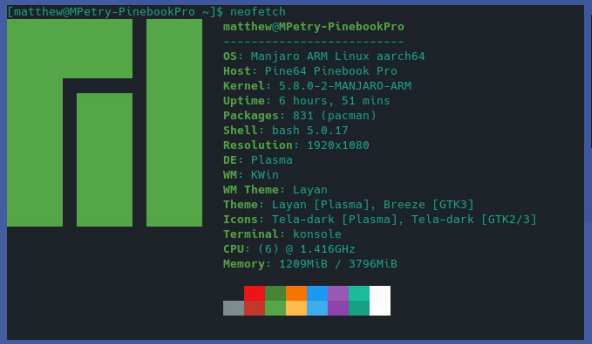
- Phase 3: Production-ready PinePhone
  - Mainboard revision 1.0: Dev
    - Late 2019
  - Mainboard revision 1.1: “Braveheart”
    - January 2020
    - First version ready for consumers
  - Mainboard revision 1.2(a): Community Editions
    - June 2020
    - Fix all known remaining bugs
    - Ready for daily use

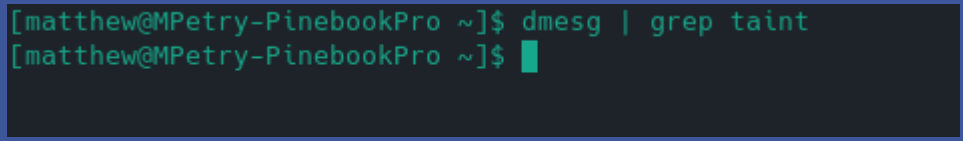
Practice what you preach…

This presentation is being given from a Pinebook Pro…

…Running community-developed software (Manjaro ARM)…

…with a fully-FOSS (untainted) build of Linux kernel 5.8





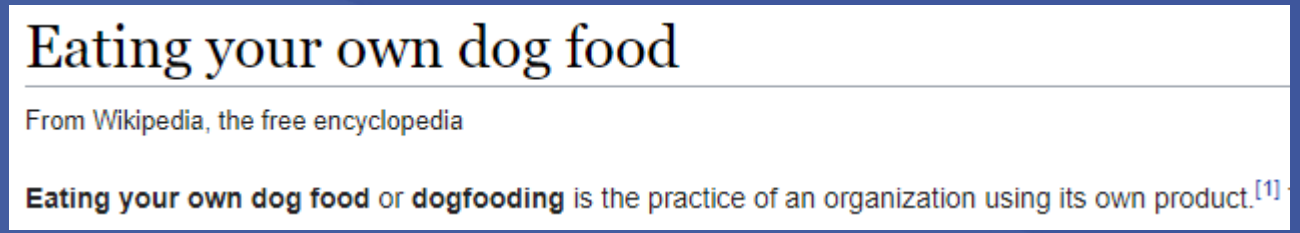
:)

The RockPro64 Cluster

(how to dog-food your own products)

The RockPro64 Cluster

(how to dog-food your own products)



Eating your own dog food

From Wikipedia, the free encyclopedia

Eating your own dog food or **dogfooding** is the practice of an organization using its own product.[1]

People have been asking for a long while...

People have been asking for a long while...

We've been promising a blog write-up on our cluster set-up for a long while.

Still hasn't happened: perfectionism getting in the way of "good enough".

This isn't a substitute for that blog post.

But it is the first time the configuration has been discussed in public.

The Cluster: Hardware

- 24x RockPro64 (4GB) SBCs
- 2x 1TB Samsung 970 EVO SSDs
- Custom designed 2U rackmount chassis
- Built in Gigabit Ethernet switching
- Backplane for Remote Reboot capability
  - *(not implemented yet… this will be an issue later)*
- *Colocated at BBXNET*

The Cluster: Software *(in people's dreams)*

Whenever the cluster has been mentioned in our social media, people's expectations go wild.

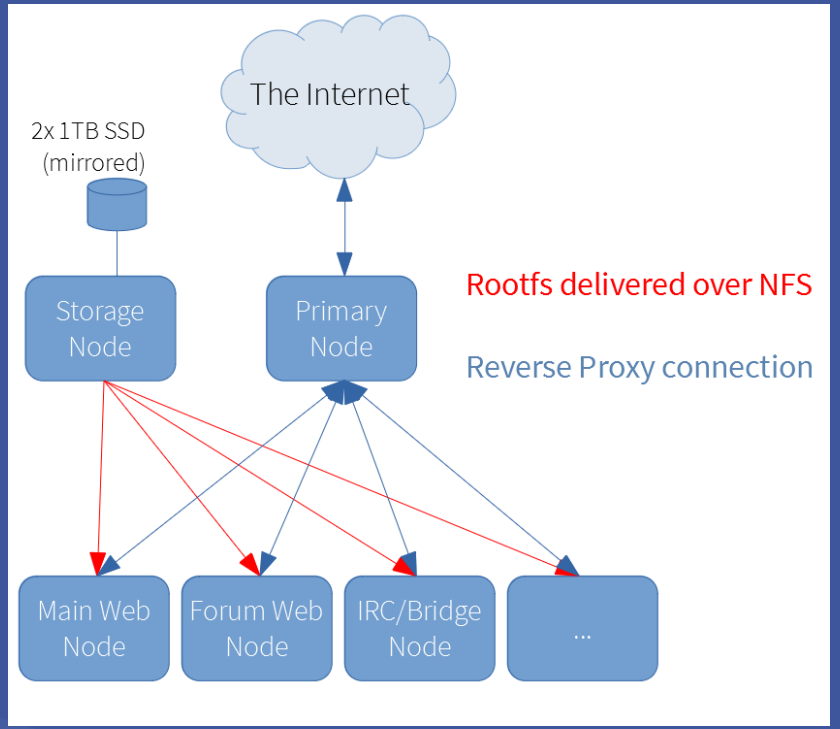
They expect we've made some incredible distributed-load system.

Maybe using cool stuff like Kubernetes?

Sorry to let you down. We ~~got lazy~~ ran out of time.

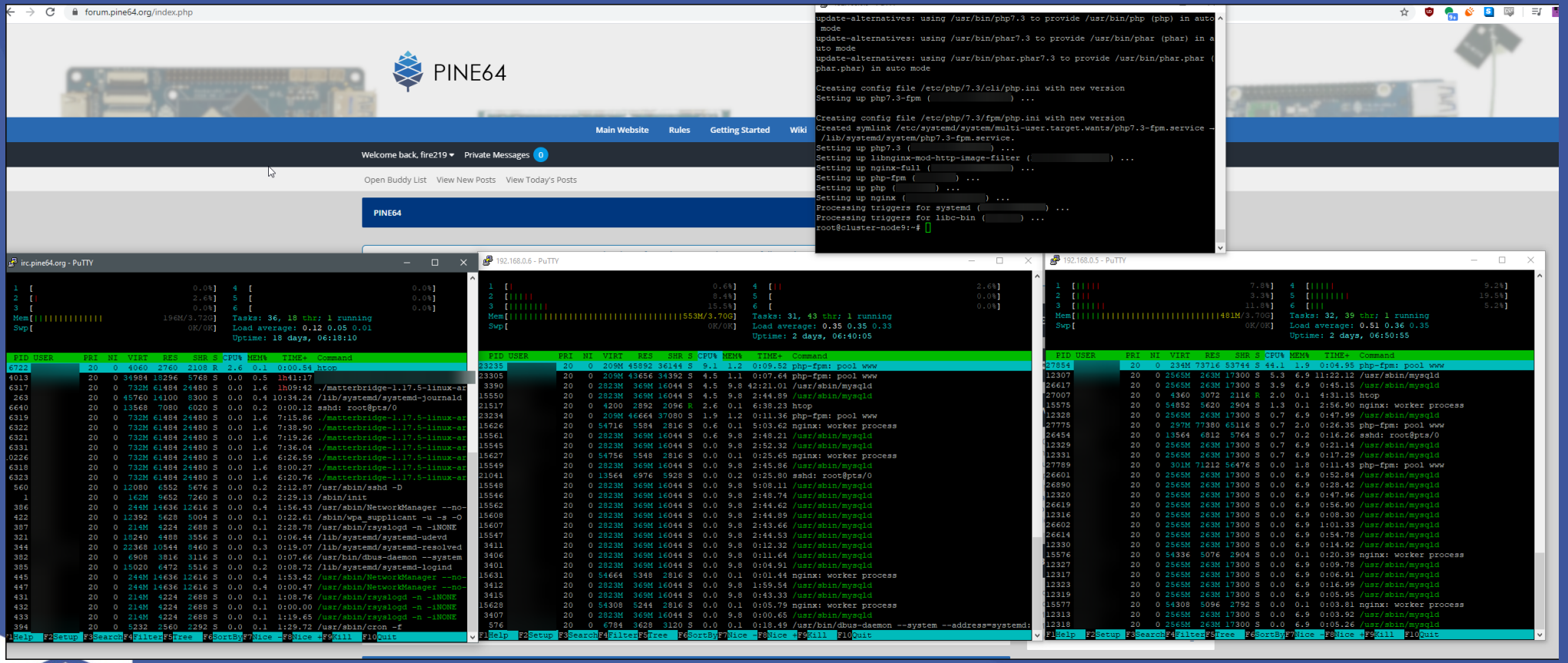
The Cluster: Software *(in reality)*

- All nodes run/compute independently
- Custom Debian 10 images with kernel 5.7
- Storage node handles all root filesystems, delivers to other nodes via NFS
  - *(will change to iSCSI soon… for good reason)*
- All external (internet) access to services is NGINX reverse proxied through primary node



The Cluster: Software *(in reality)*

- This relatively simple setup was much faster to bring up than a “proper” distributed cluster.
- Want to bring up another service? Just bring up another node, set up the software, and route reverse proxy accordingly.
- The downside? Well...



It can get a bit overwhelming when monitoring resource utilization. :)

Synapse on the Cluster

(The trials and tribulations of 4GB RAM)

Long Time Coming...

For quite a while now, we've had our various chatrooms bridged on several platforms

- *Matrix, Discord, IRC, Telegram*

For even longer than that, people have been asking when we'll have an official Pine64 Matrix homeserver.

Until recently, we just didn't have the resources to try.

But now we have *twenty-four* RockPro64 boards sitting in a datacenter. After initial deployment, only six were in use.

So, **why not?**

Doing it by the book.

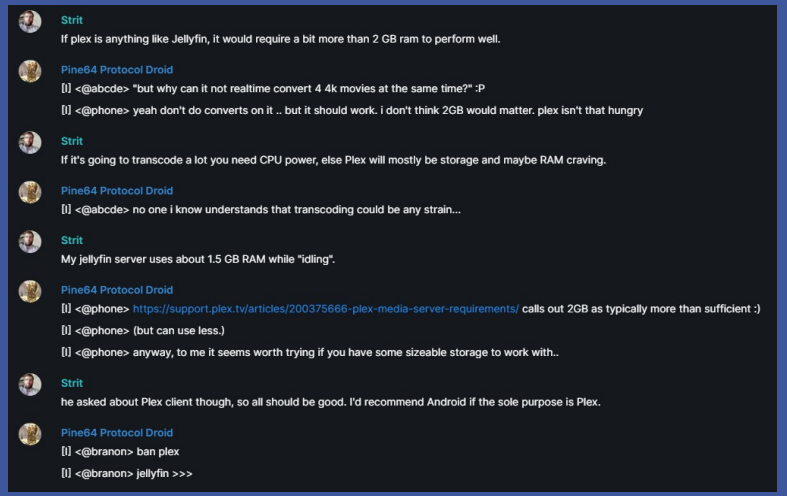
Soon after our initial cluster deployment, I get to work setting it up by the ~~book~~ INSTALL.md

- Fight with dependencies
- Install Synapse
- Make a config
- Start Synapse
- Make another config after remembering .well_known exists
- Restart Synapse
- Realize that sqlite is awful
- Install postgreSQL
- Restart Synapse

```
mkdir -p ~/synapse
virtualenv -p python3 ~/synapse/env
source ~/synapse/env/bin/activate
pip install --upgrade pip
pip install --upgrade setuptools
pip install matrix-synapse
apt install zlib-dev
apt search zlib
apt install zliblg-dev
pip install matrix-synapse
ls
cd ~/synapse
python -m synapse.app.homeserver --server-name matrix.pine64.org
omeserver.yaml --generate-config --report-stats=no
ls
nano homeserver.yaml
```

It Works!

- After that sequence of events, the pine64.org homeserver is running!
- Happily ever after, right?
- *Not so fast...*



The Pine64 Community is quite large. *</humblebrag>*

If this were a small personal homeserver, I'm sure the story could end here.

But it's not.

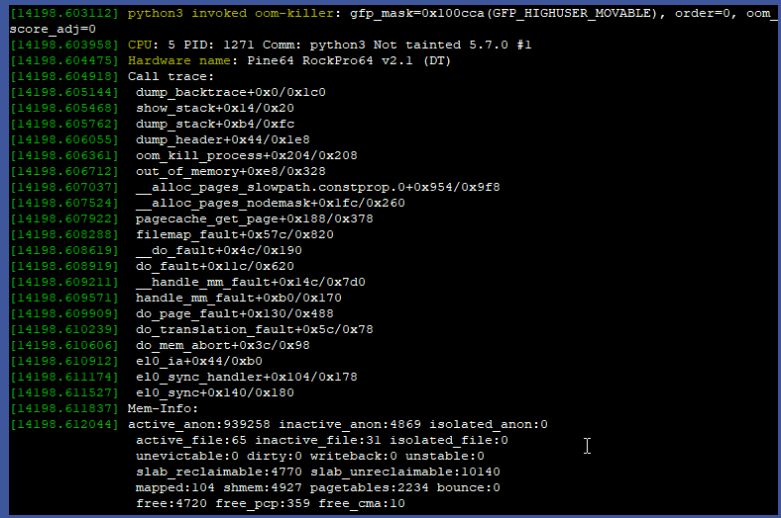
Pine64 Matrix room sizes:

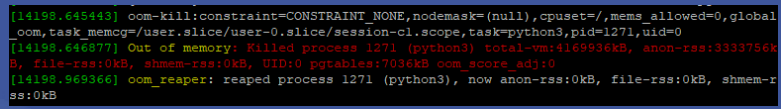
- Pinebook: **~400 users**
- Pine64 (General): **~430 users**
- PinePhone: **~1340 users**

Many different homeservers among those.

It Works… for about 2 hours.

- All seemed well. Until all of a sudden…
- oom-killer'd.
- Long story short, Synapse apparently **really** likes to eat RAM when told to deal with big rooms.
- Why not just implement swap space?
  - On an NFS share? Nope.
  - Give node local storage? Maybe, but rather not.



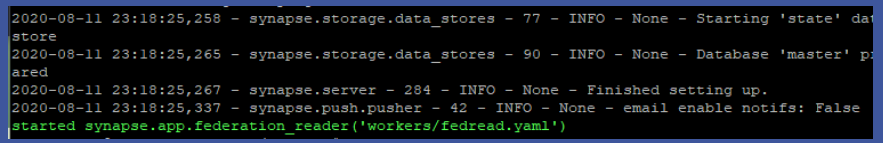


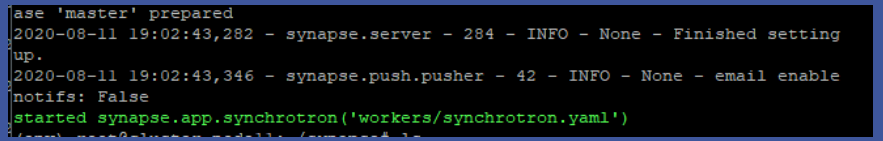
Take Trouble, and Make it Double

Well, Synapse does support splitting into workers. Lets try that.

Chain of events:

- Read workers.md
- Spin up another RP64 node
- Write worker scripts for both nodes
- Get lazy for a month
- Get a Matrix DM about an event called OTWSU
- Scramble to try to get Synapse up and running in time
- Crash node 2 hard enough to require physical intervention
- Finally get Synapse running again! … mostly.





Right when I felt good about myself...

Synapse is running, but federation seems a bit broken.

It's talking properly to most other homeservers as far as I can tell.

But not matrix.org for some reason?

This results in many missing messages.

(I guess that's why we need more homeservers. :))

When I start trying to troubleshoot that…

Right when I felt good about myself...

```
[791829.784962] [  22751]   106 22751    53513    35634   421888        0             0 postgres
[791829.785714] [  22755]   106 22755    53419    31866   421888        0             0 postgres
[791829.786464] [  22757]   106 22757    53521    33549   421888        0             0 postgres
[791829.787213] [  22759]   106 22759    53505    32910   421888        0             0 postgres
[791829.787963] [  22773]   106 22773    53909    36434   425984        0             0 postgres
[791829.788713] [  22776]   106 22776    53781    36101   421888        0             0 postgres
[791829.789464] [  22778]   106 22778    53895    36320   425984        0             0 postgres
[791829.790214] [  22780]   106 22780    53787    36104   421888        0             0 postgres
[791829.790963] [  22782]   106 22782    53763    36151   421888        0             0 postgres
[791829.791728] [  22785]     0 22785     1083      242    53248        0             0 htop
[791829.792454] [  23184]   106 23184    53438    16688   421888        0             0 postgres
[791829.793206] [  23257]   106 23257    53015     1380   167936        0             0 postgres
[791829.793957] oom-kill:constraint=CONSTRAINT_NONE,nodemask=(null),cpuset=/,mems_allowed=0,global
_oom,task_memcg=/user.slice/user-0.slice/session-c8.scope,task=python3,pid=22656,uid=0
[791829.795469] Out of memory: Killed process 22656 (python3) total-vm:2281768kB, anon-rss:1600780
kB, file-rss:0kB, shmem-rss:0kB, UID:0 pgtables:3332kB oom_score_adj:0
[791829.980302] oom_reaper: reaped process 22656 (python3), now anon-rss:0kB, file-rss:0kB, shmem-
rss:0kB
```

¯_(ツ)_/¯

That's All, Doc!

That's the status as of the night before OTWSU 5.

I wanted to have this all working in time, but that's how life goes.

If anyone has any ideas as to what's going on with the federation issue, I'm listening.

Totally not a cry for help. :)

If you still want to give it a try, Synapse is running (albeit with the questionable federation.)

matrix.pine64.org

Registration is open.

Questions?

(Sorry for the likely excessively long presentation.)