Insights on the use of WRF in Cloud Systems

Diego A. Pérez Montes < kabute@uvigo.es >
Juan A. Añel Cabanelas < j.anhel@uvigo.es >
Javier Rodeiro < jrodeiro@uvigo.es >

Prior Work

- Running large ensembles for ClimatePrediction.net (CPDN) on different public cloud providers [1] [2]

Montes et al. - Enabling BOINC in infrastructure as a service cloud system. Geoscientific Model Development, 2017

Añel et al. - Evaluation and intercomparison of Cloud Computing solutions for climate modelling. 2018. Submitted to PLoS ONE.

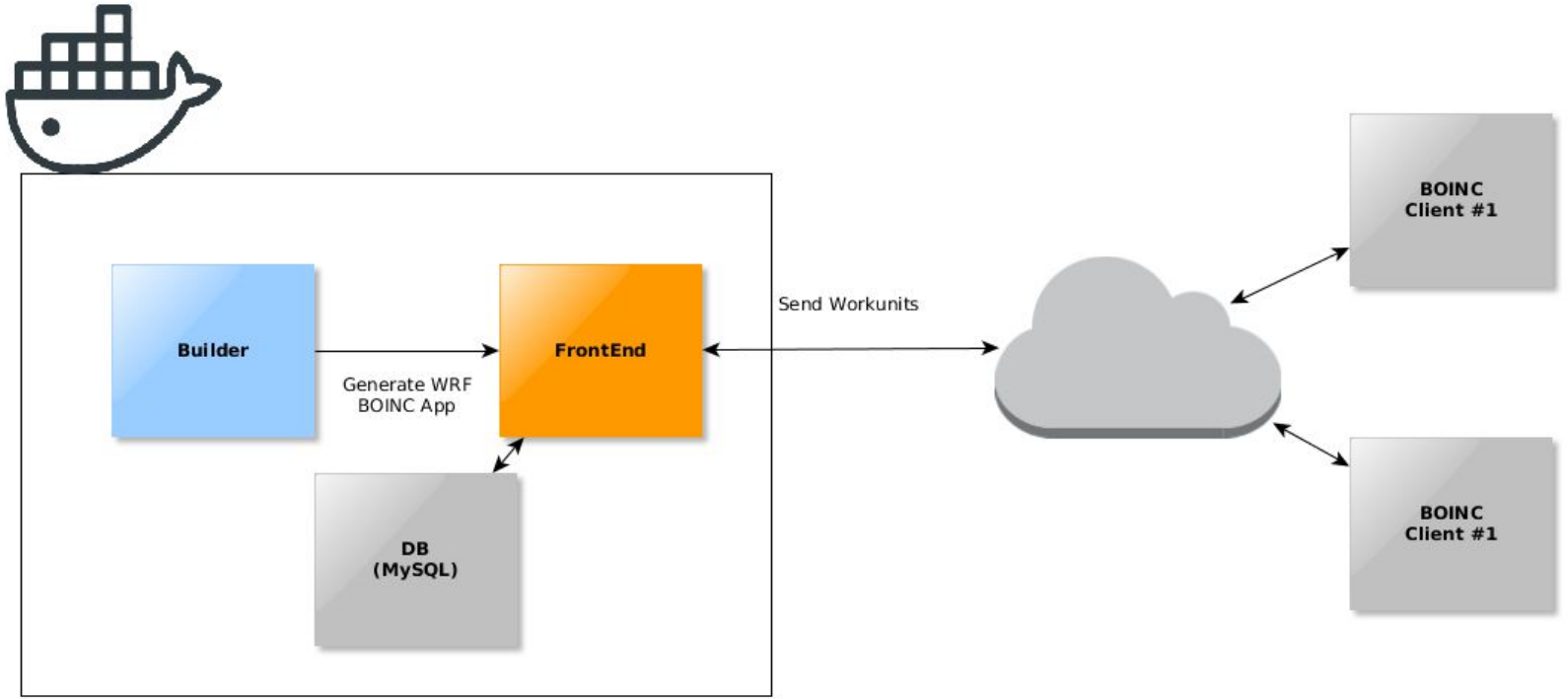
What? (Current Goals)

- Run WRFv3 on heterogeneous and dynamic (and ephemeral!) environments (e.g. Cloud).
- Run over Berkeley Open Infrastructure for Network Computing (BOINC).
- Fully Automated pipeline (Build → Computation).

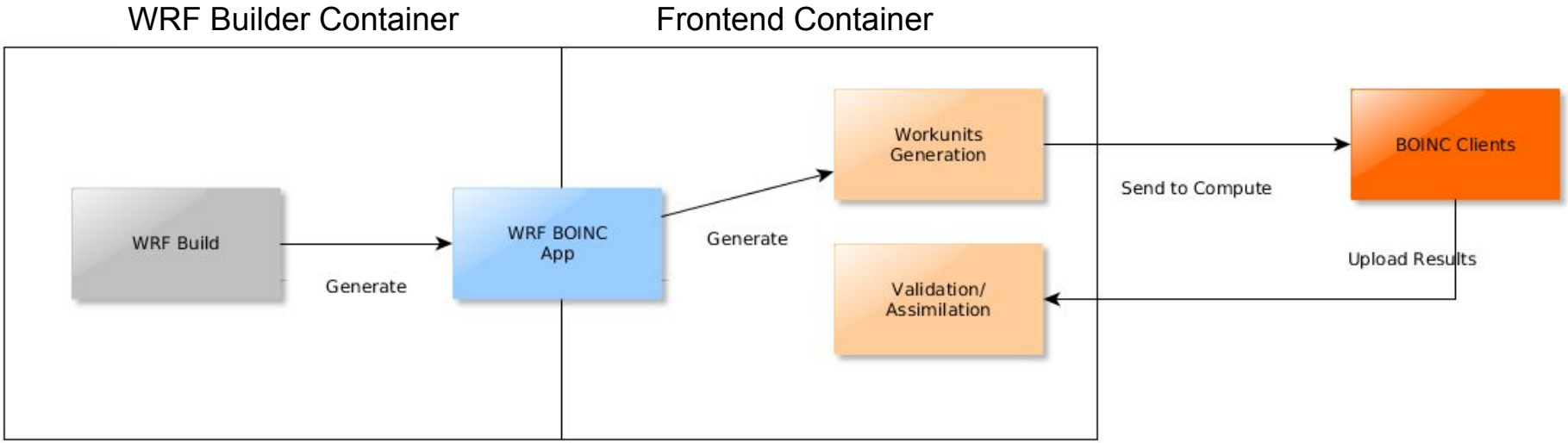
Why?

- WRF popularity and it is Free Software.
- Add more possibilities on how to run WRF.
- Add the ability to run on the Cloud where no “local” HPC available.

Proposed Architecture



Pipeline


WRF Builder Container
Frontend Container
WRF Build
Generate
WRF BOINC App
Generate
Workunits Generation
Validation/ Assimilation
Send to Compute
BOINC Clients
Upload Results

Some Implementation Details...

WRF:

- MPICH3, NetCDF4 and GCC/GFortran 4.9 (x86_64 Linux)
- Modified configurators and Makefiles.
- Static compilation: big binaries > 40MB

Some Implementation Details...

Infrastructure:

- Based on (very modified) images by Marius Millea [3].
- docker-compose and kubernetes
- Ran over instances but could be done over PaaS like GKE.

Work In Progress: Automation/UI

- Tweaks on pipeline automation.
- Improving BOINC Validation and Assimilation stages.
- UI Integration for workunit generation (e.g. variable perturbations).

Conclusions

- Completed a functional pipeline for WRF over BOINC.
- BOINC seems to be an adequate middleware for WRF distributed computing.
- Volunteer computing addition is going to present interesting challenges.

Thank You!

(Further discussion on Screen 11)

Appendix

Some Implementation Details...

BOINC:

- Big results and workunits.
- WRF needs/expects all input files on the slot (*<copy_file/>*)
- Running Inside the Wrapper (App).
- Target: x86_64-pc-linux-gnu

Running

```
======== Applications ========
1) -----------
   name: wrf_full
   Project: boinc_wrf

======== Application versions ========
1) -----------
   application: wrf_full
   version: 1.00
   project: boinc_wrf

======== Workunits ========
1) -----------
   name: wrf_full_567_1523217695.675265
   FP estimate: 3.600000e+12
   FP bound: 8.640000e+13
   memory bound: 476.84 MB
   disk bound: 953.67 MB
2) -----------
   name: wrf_full_515_1523217549.609134
   FP estimate: 3.600000e+12
   FP bound: 8.640000e+13
   memory bound: 476.84 MB
   disk bound: 953.67 MB
3) -----------
   name: wrf_full_507_1523217382.011911
   FP estimate: 3.600000e+12
   FP bound: 8.640000e+13
   memory bound: 476.84 MB
   disk bound: 953.67 MB
4) -----------
   name: wrf_full_575_1523217696.006498
   FP estimate: 3.600000e+12
   FP bound: 8.640000e+13
   memory bound: 476.84 MB
   disk bound: 953.67 MB
```

References

[1] *Enabling BOINC in infrastructure as a service cloud system.* D Perez Montes, et al. - Geoscientific Model Development, 2017

[2] *Use of several Cloud Computing approaches for climate modelling: performance, costs and opportunities.* D Perez Montes, et al. - EGU General Assembly Conference Abstracts, 2017

[3] Boinc Server Docker, Millea M. , https://github.com/marius311/boinc-server-docker/

[*] Front image by: Victor Hermida Prada