How Pandas use Python

Pamphile T. Roy



Pamphile or @tupui

- https://www.linkedin.com/in/tupui
- https://github.com/tupui

Things I do

- SciPy
- SALib (and a hobbyist researcher)
- Steering Committee (general and SPEC)
 https://scientific-python.org/
- PyData Sphinx Theme
- NumFOCUS SDG and Security
- Stellar Community Fund Pilot

E.g. with SPEC 7

- Deprecate the use of random_state/seed arguments in favor of an rng argument in all functions where users need to control pseudo-random number generation,
- Use numpy.random.default_rng to normalize the rng argument and instantiate a Generator, and
- Deprecate the use of numpy.random.seed to control the random state.

Bitpanda at a glance

Offering











Functionality











Trusted



Footprint



30+ countries

In Europe

Business model





B2B2C offering (Launched in 2022)



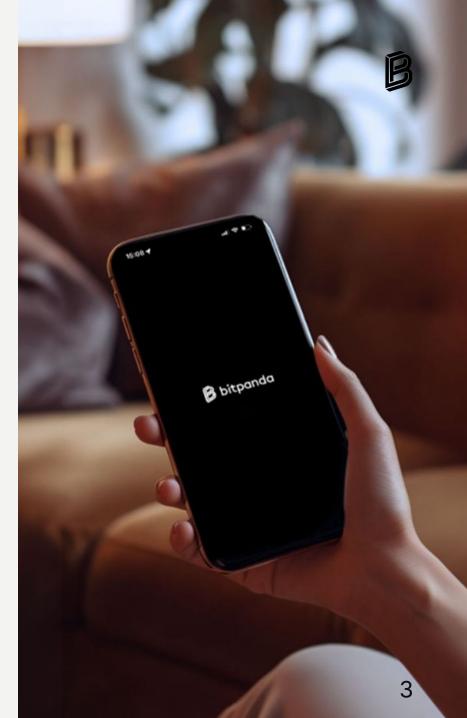
B2C Broker (Launched in 2014)

Our workforce



700+ employees

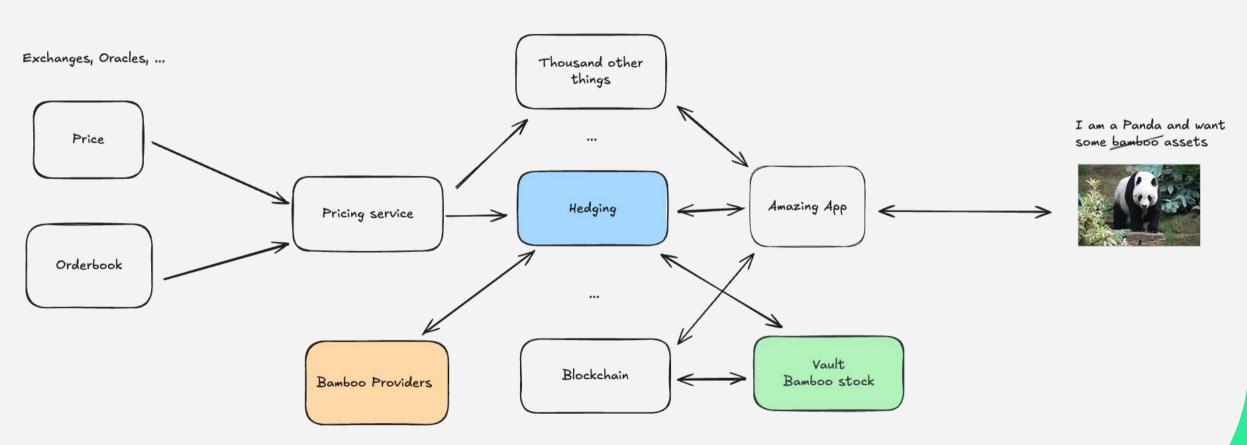
From +50 nationalities

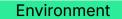






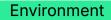
What Pandas do?



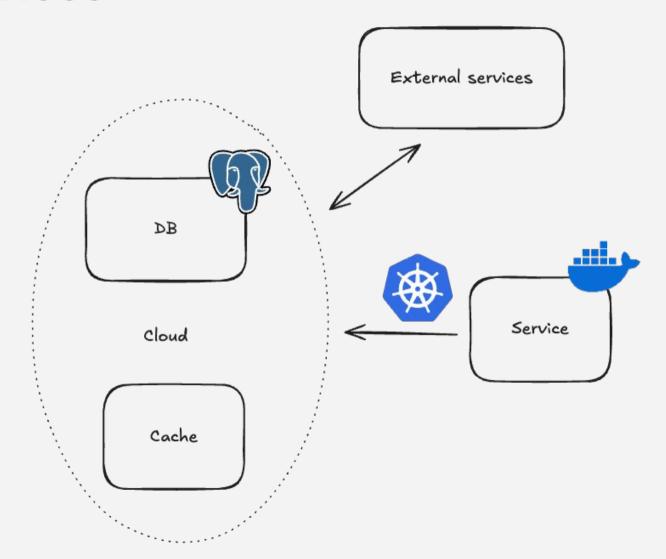


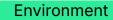




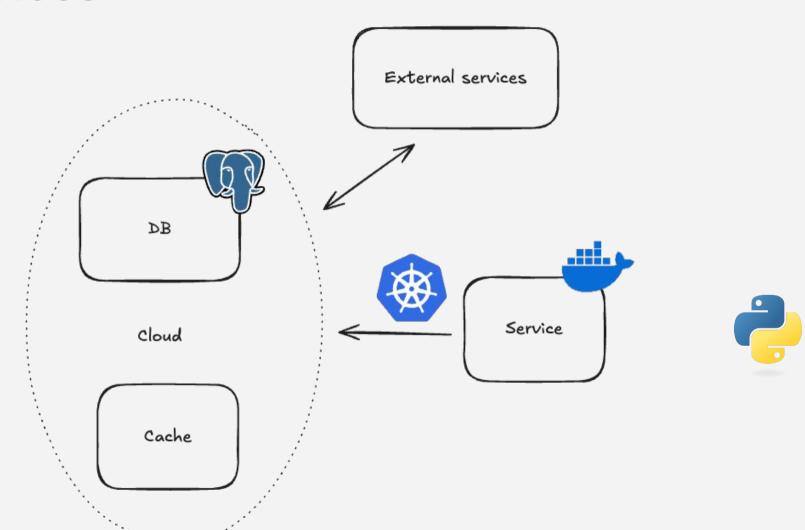


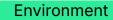




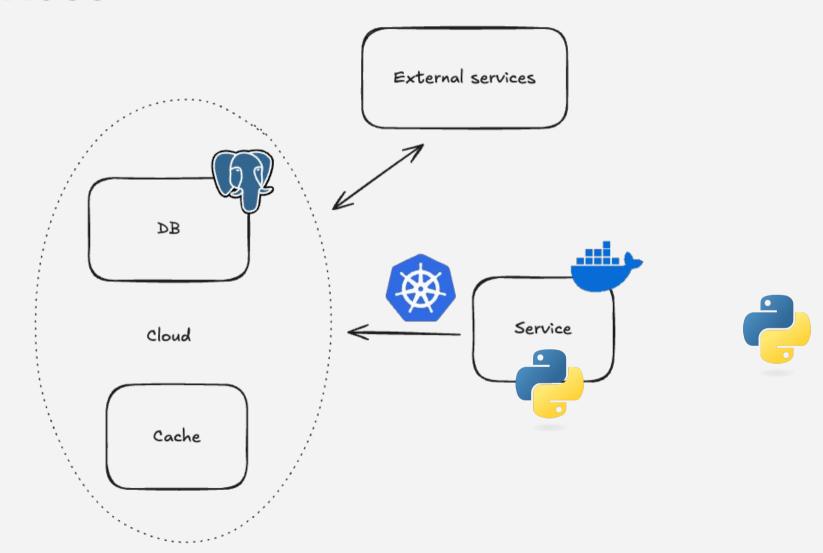


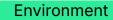




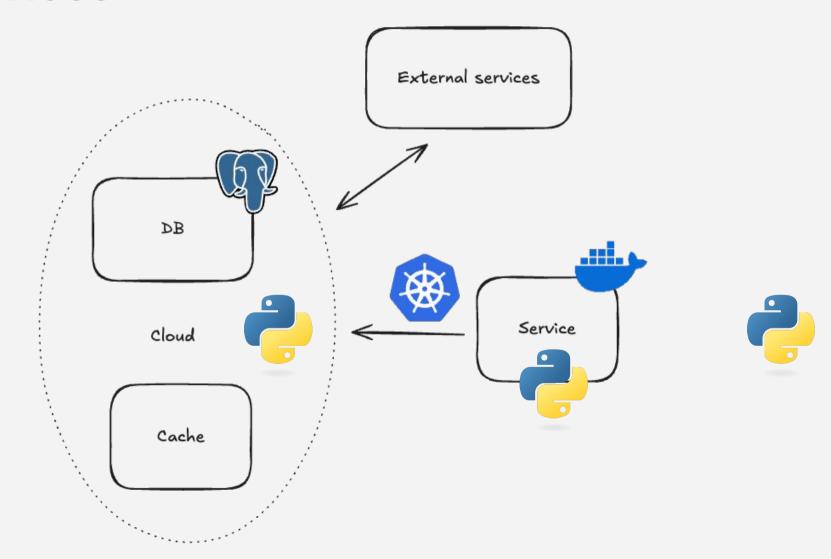


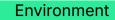




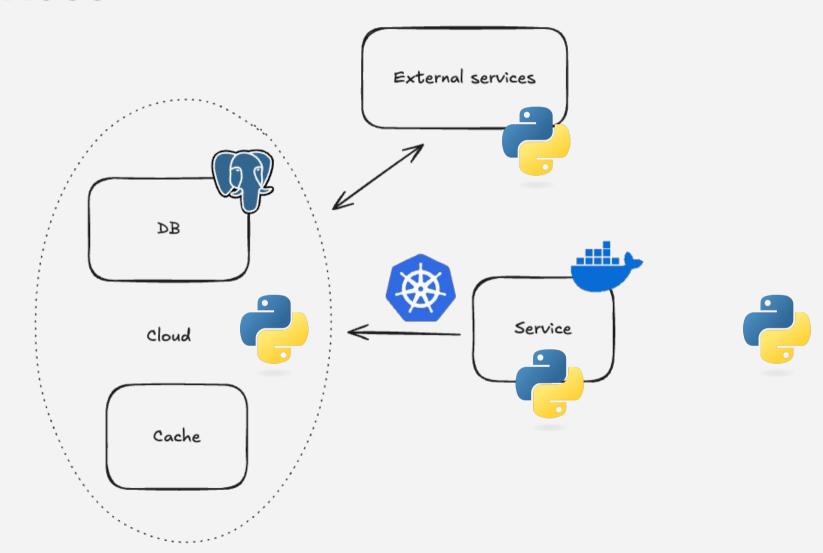


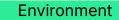




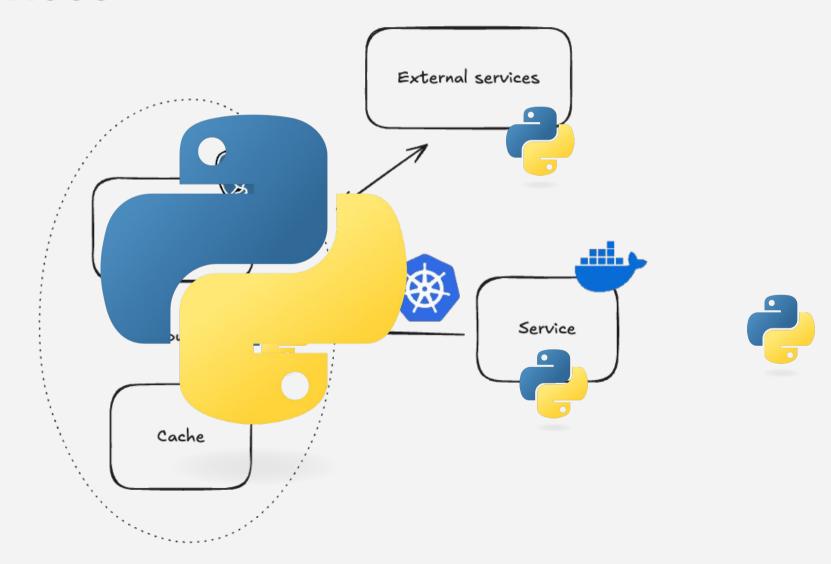












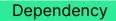


Common stack

- ChatGPT (hard pass)
- Poetry
- Pylint, Black and/or Ruff
- FastAPI
- Pydantic
- Orjson
- Pytest
- Locust
- memray
- Docker and friends in the cloud



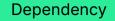






To update or not to update?







Poetry is still there...

- Lockfile
- Strict dependency resolution
- Slow
- Opinionated (not Pythonic, not standart)

 Everything you ever wanted and more, that it.

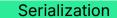






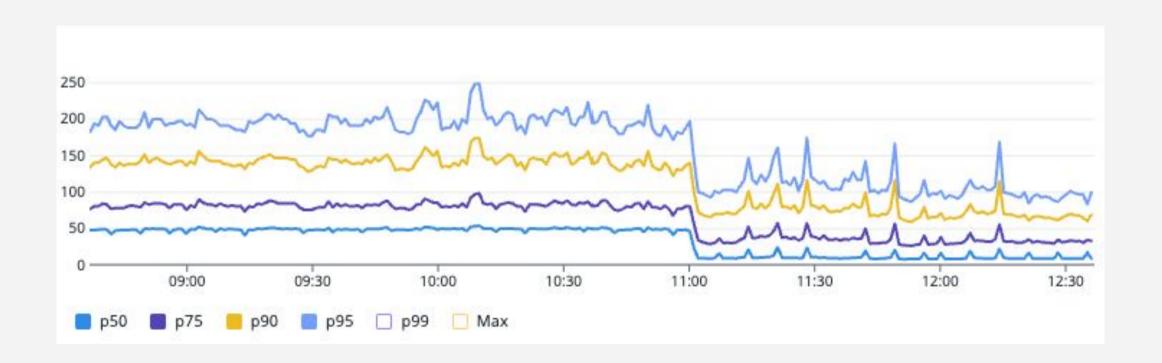
Make it fast without breaking the bank

- Premature optimization is?
- Get to know the concurrency model in Python
- Async all the way, this is hard!
 - Level -1: async def
 - Level 1: Actually async with some awaitable functions
- ORM: use one, learn it well
- Pydantic: use it, learn it too
- Caching: Checkout Cashews! https://github.com/Krukov/cashews
- Benchmarking: Locust and integrate in your CI/CD
- Active monitoring: things do change overtime





Not all JSON are equal



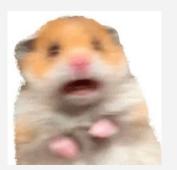


Ship it!

- Multi-stage builds
- Caching is hard
- Secrets handling is hard
- Some security extras (uninstall tools, base image, etc.)

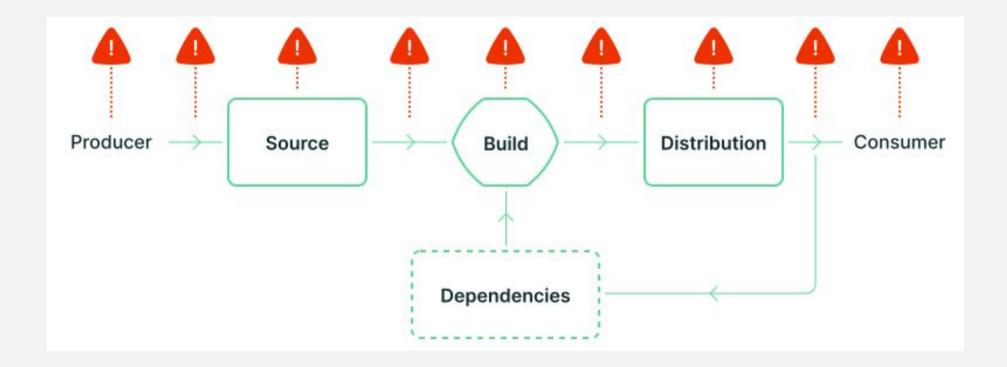


RUN --mount=type=cache,mode=0777,target=/root/.cache/pip pip install scipy



ARG SUPER_SECRET_TOKEN









Going the extra mile

- SLSA
- SBOM
- CVEs
- Active Scans
- Fuzzy testing
- Threat modeling
- Security audits
- etc.



Organization against confusion

