

# Deployable BIOS pipeline with Ansible & BioConda

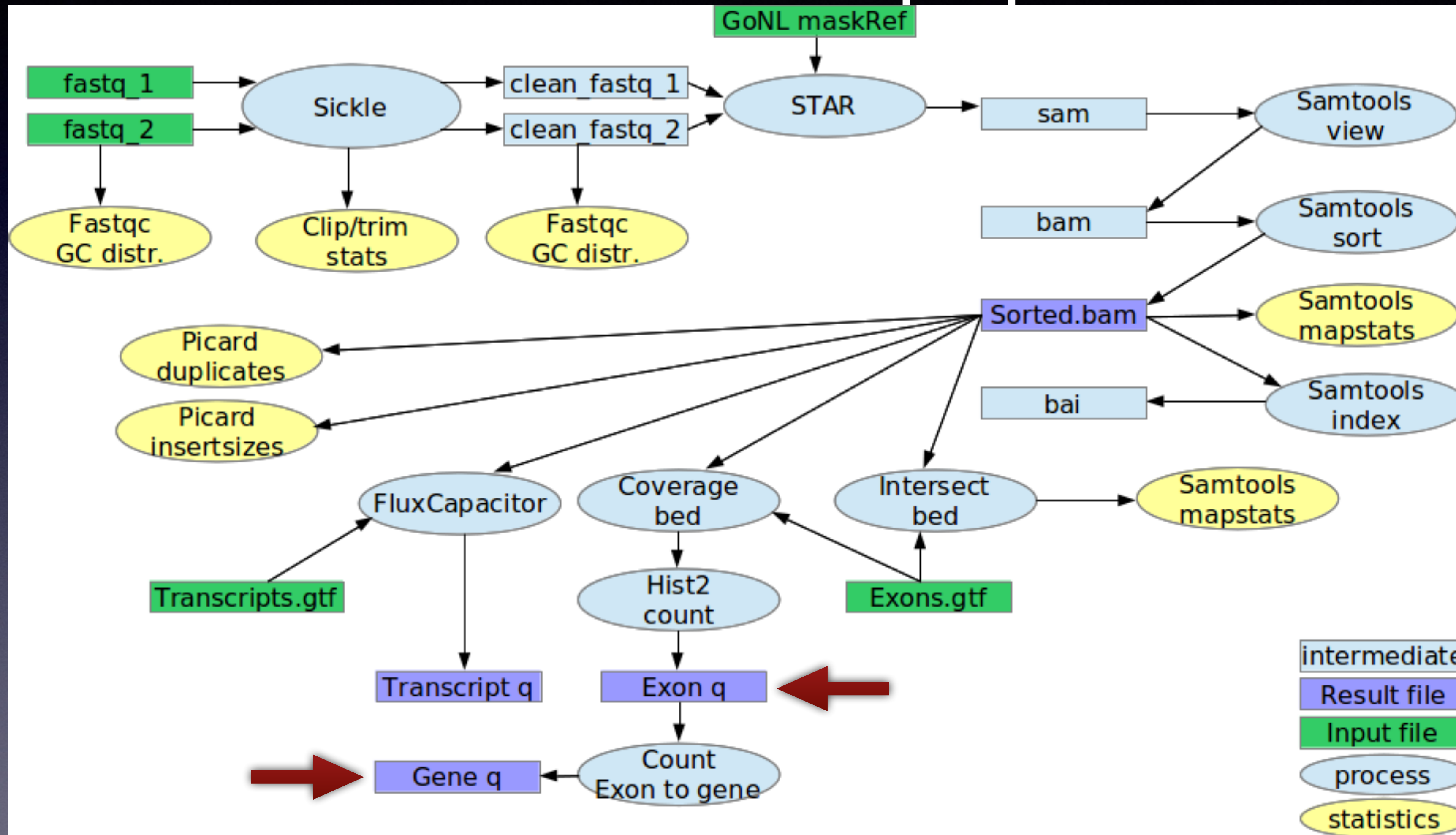
Peter van 't Hof



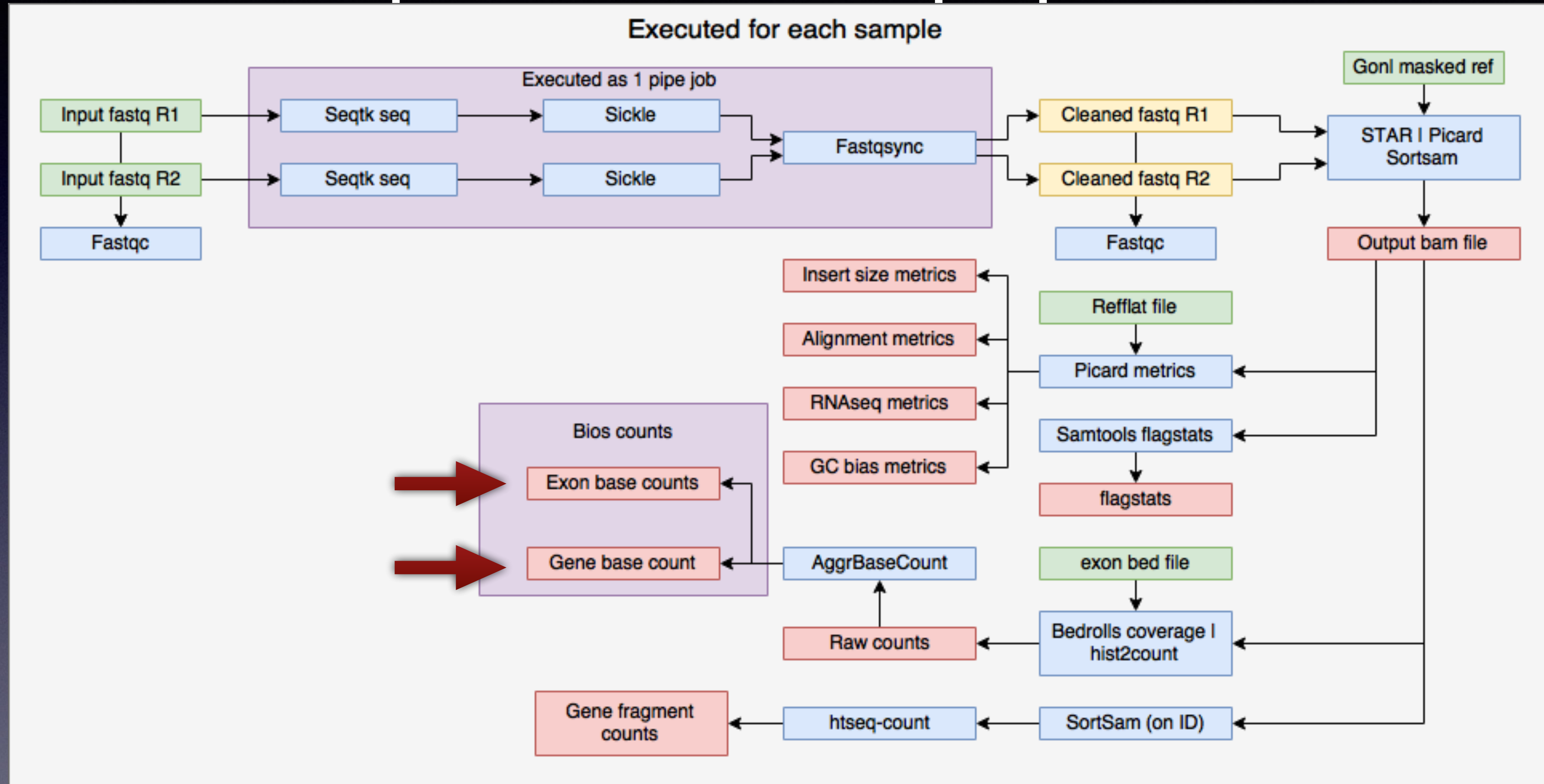
# Introduction

- Goals:
  - Deploy a pipeline on a clean system / VM
  - Install all tools automatically
  - Download reference set automatically
- Use case:
  - Deploy the BIOS pipeline to reproduce the same results

# BBMRI BIOS pipeline



# Biopet BIOS pipeline

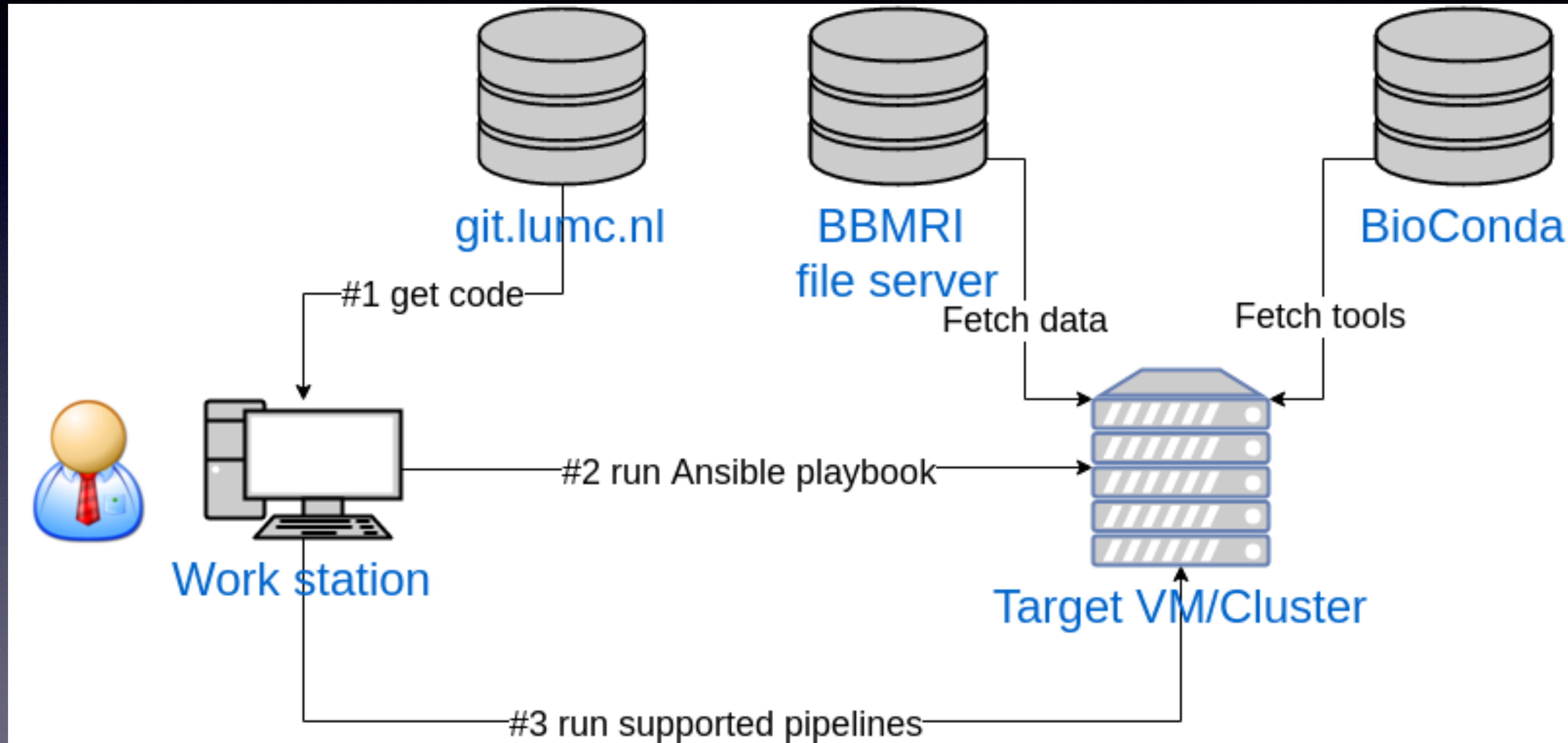


<https://github.com/biopet/biopet/releases/tag/BIOS>

# Ansible & BioConda

- **Ansible:** Industry standard, free software to configure servers
  - Minimal in nature. Management systems should not impose additional dependencies on the environment.
  - Secure. Ansible does not deploy vulnerable agents to nodes. Only OpenSSH is required, which is already tested.
  - Highly reliable & consistent. The idempotent resource model is applied to deployment to prevent side-effects from re-running scripts.
  - Low learning curve. Playbooks use an easy and descriptive language based on YAML.
- **BioConda**
  - A Conda channel to install Bioinformatics software packages
  - Active community (16 contributors with >50commits)
  - Used by Galaxy project

# Architecture/Scenario



# Discussion

- Conda / Bioconda has some system dependencies
- Each tool needs a separated Conda env to prevent shared library collisions
- Not all versions of tools used inside the bios pipeline are available on Bioconda
- Conda activate only sets \$PATH but not the library path

# Demo

- <https://git.lumc.nl/sasc/bios-ansible>