GitHub Actions

Valeh Valiollah Pour Amiri
CCB Skills Seminar
4/13/2022

source: https://amitlevinson.com/blog/automated-plot-with-github-actions/

Agenda

- Overview of GitHub Actions
- Live demo
- What you can use GitHub Actions for in your research

Overview of GitHub Actions

Overview of GitHub Actions

aka GHA

Why GHA

To streamline DevOps pipelines and automate CI/CD build, test, and deployment steps

Why GHA

To streamline DevOps pipelines and automate CI/CD build, test, and deployment steps

…

Why GHA

To streamline DevOps pipelines and automate CI/CD build, test, and deployment steps

…

Huh?? 😦

Why GHA

- Originally conceived to automate parts of the software development process
    - Used heavily by application/package developers
- But it goes beyond that
- You can think of it as this: something that can do things based on relevant events

Why GHA

- Originally conceived to automate parts of the software development process
  - Used heavily by application/package developers
- But it goes beyond that
- You can think of it as this: something that can do things based on relevant events

The occurrence of a relevant activity within the repo, e.g. the creation of a new issue. It can also be purely chronological, e.g. trigger every week.

We will dive into this in the next slides where I explain *what* GHA is.

The sky is the limit here. It ranges from: adding a label to any issue someone creates in your repo, to: compiling a Jupyter Notebook and publishing it to a site server.

What is GHA

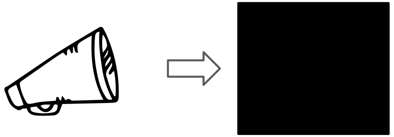
- It is an ecosystem of tools to help you automatically launch workflows based on signals of your choosing
- Here is a step-by-step walkthrough diagram

What is GHA

What is GHA

`event`

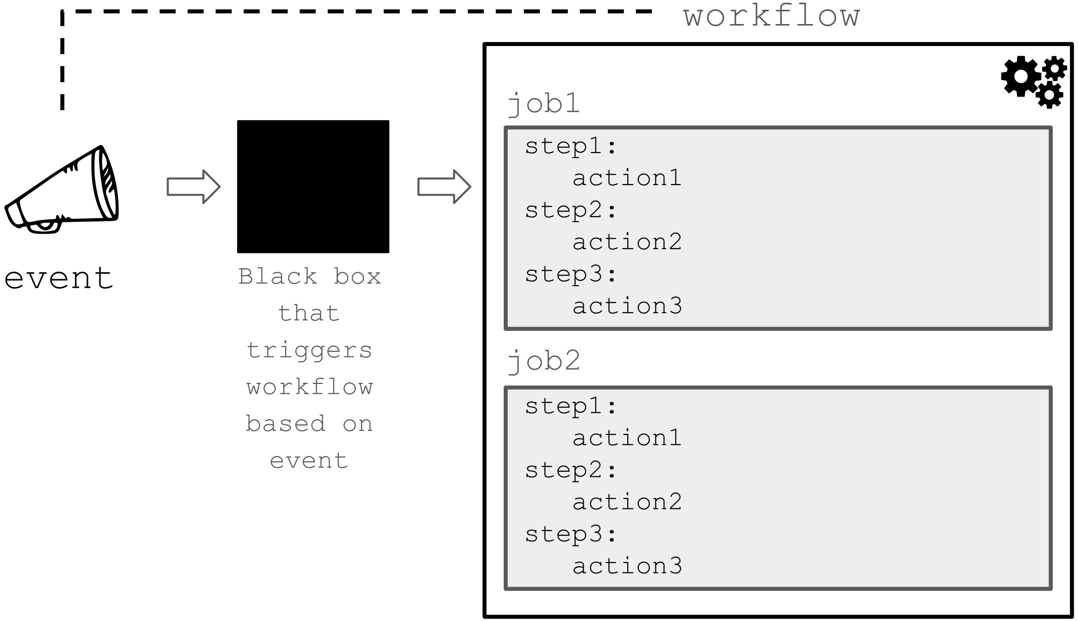
What is GHA



event

Black box that triggers workflow based on event

What is GHA



workflow

event

Black box that triggers workflow based on event

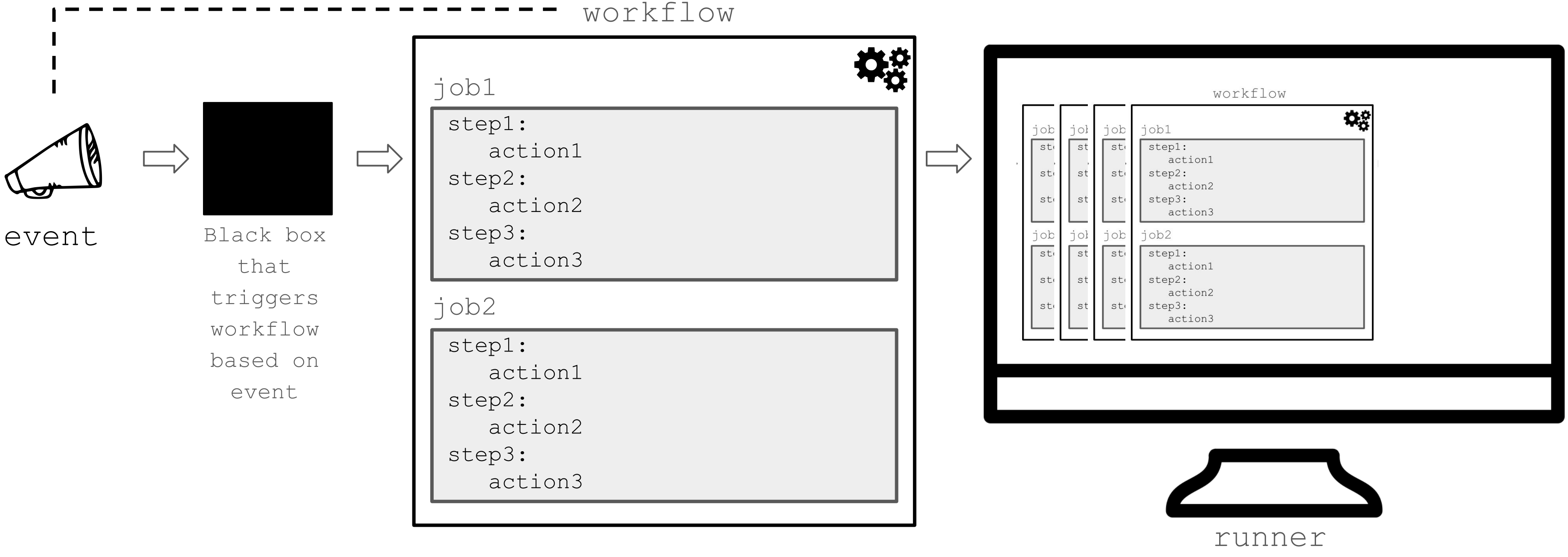
job1

```
step1:
   action1
step2:
   action2
step3:
   action3
```

job2

```
step1:
   action1
step2:
   action2
step3:
   action3
```

What is GHA



workflow

event

Black box that triggers workflow based on event

job1

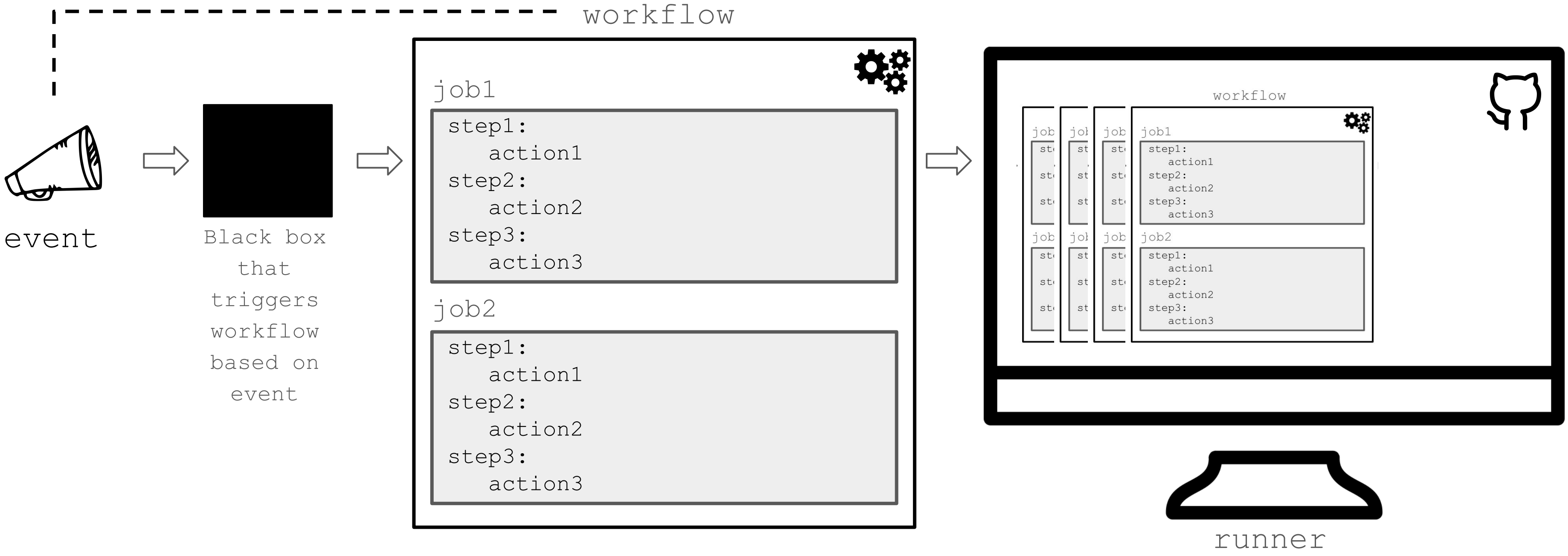
```
step1:
   action1
step2:
   action2
step3:
   action3
```

job2

```
step1:
   action1
step2:
   action2
step3:
   action3
```

runner

What is GHA



workflow

event

Black box that triggers workflow based on event

job1

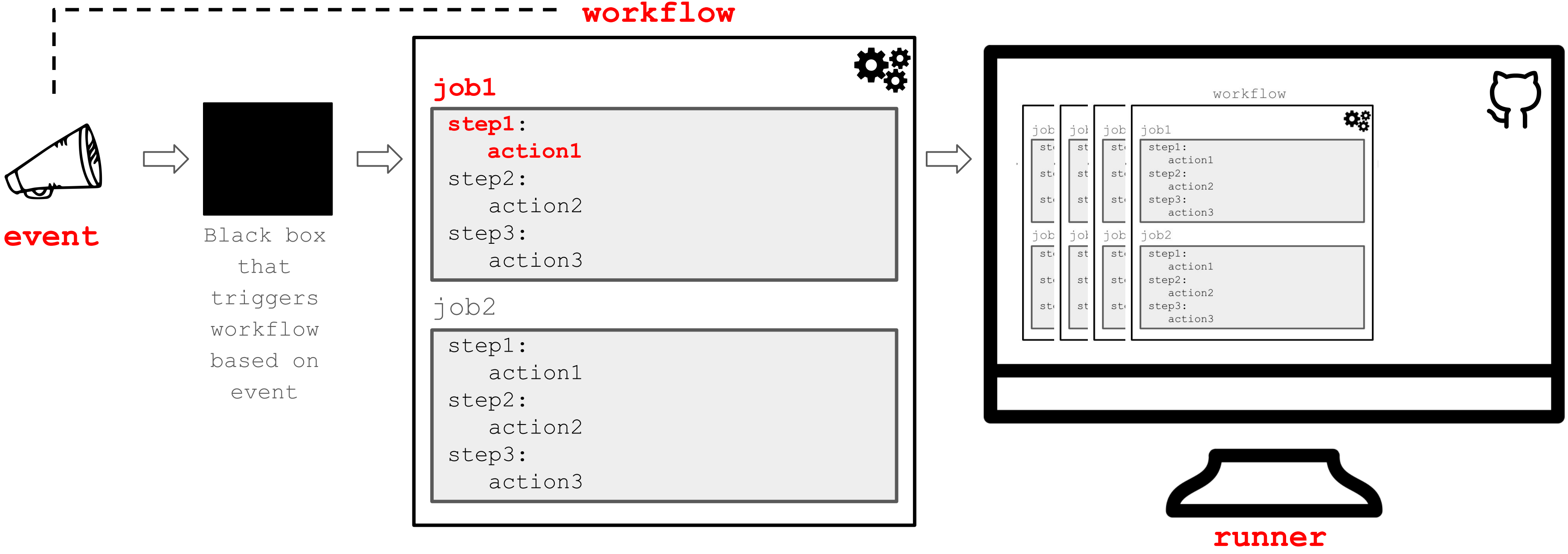
```
step1:
   action1
step2:
   action2
step3:
   action3
```

job2

```
step1:
   action1
step2:
   action2
step3:
   action3
```

runner

What is GHA



event

Black box that triggers workflow based on event

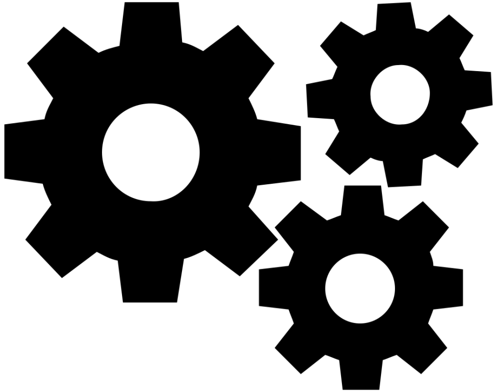
workflow

```
job1
step1:
    action1
step2:
    action2
step3:
    action3
job2
step1:
    action1
step2:
    action2
step3:
    action3
```

runner

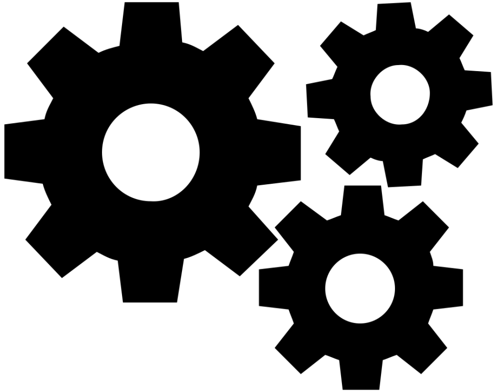
High level properties of GHA elements

High level properties of GHA elements



- Workflows

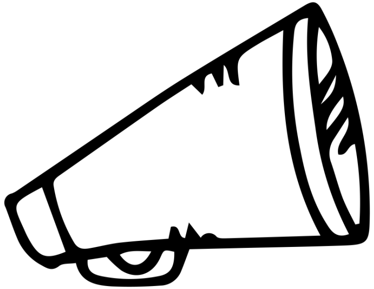
High level properties of GHA elements



- Workflows
  - The meat of GHA
  - Define relevant events that trigger them
  - … and the jobs they run once triggered
  - YAML file under `.github/workflows`
  - Can have more than one workflow
    - e.g. one that tests your code everytime you push to `main`, another that sends “Thanks” to each person that opens an issue in your repo, etc.
  - Workflows can be nested and re-used
  - Concept of starter workflows
    - GitHub-provided: https://github.com/actions/starter-workflows
    - Organization-specific (act like templates)

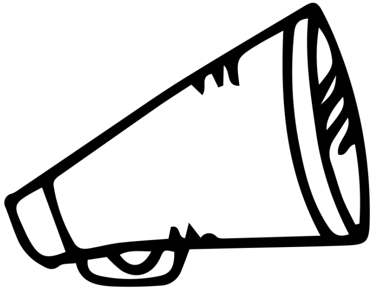
High level properties of GHA elements

- Events

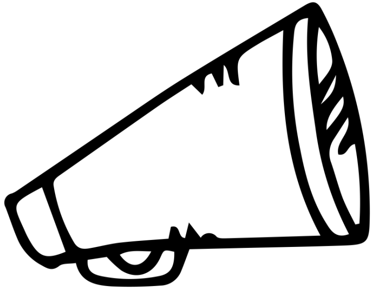


High level properties of GHA elements

- Events
  - Trigger workflows

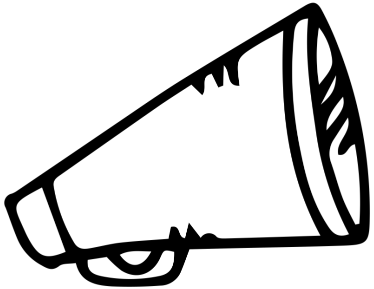


High level properties of GHA elements



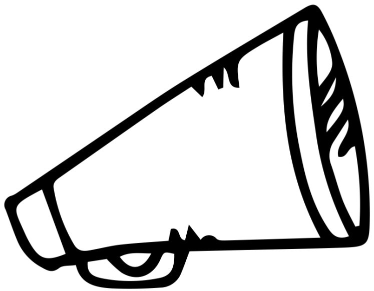
- Events
  - Trigger workflows
  - Based on relevant stimuli: an activity within the repo, or a schedule

High level properties of GHA elements



- Events
  - Trigger workflows
  - Based on relevant stimuli: an activity within the repo, or a schedule
  - Workflows can also be triggered manually via GitHub UI

High level properties of GHA elements



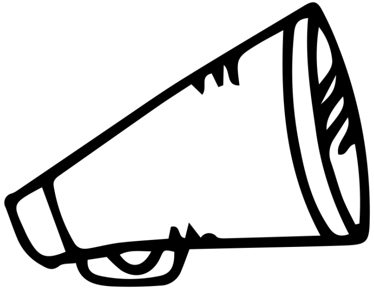
- Events
  - Trigger workflows
  - Based on relevant stimuli: an activity within the repo, or a schedule
  - Workflows can also be triggered manually via GitHub UI
  - Lot of flexibility in defining events, using activity types, filters, etc. Examples:

```
on:
  issue:
    types:
      - opened
      - labeled
```

```
on:
  push:
    branches:
      - main
      - 'releases/**'
```

```
on:
  pull_request:
    branches:
      - 'releases/**'
      - '!releases/**-alpha'
```

High level properties of GHA elements



- Events
  - Trigger workflows
  - Based on relevant stimuli: an activity within the repo, or a schedule
  - Workflows can also be triggered manually via GitHub UI
  - Lot of flexibility in defining events, using activity types, filters, etc. Examples:

```
on:
  issue:
    types:
      - opened
      - labeled
```

```
on:
  push:
    branches:
      - main
      - 'releases/**'
```

```
on:
  pull_request:
    branches:
      - 'releases/**'
      - '!releases/**-alpha'
```

  - Fine tune further using event properties. Example:

```
on:
  issues:
    types:
      - labeled
jobs:
  run_if_label_matches:
    if: github.event.label.name == 'bug'
    steps:
        - run: echo 'The label was bug'
```

High level properties of GHA elements

- Jobs and Steps

High level properties of GHA elements

- Jobs and Steps
  - Workflow = 1 or more jobs

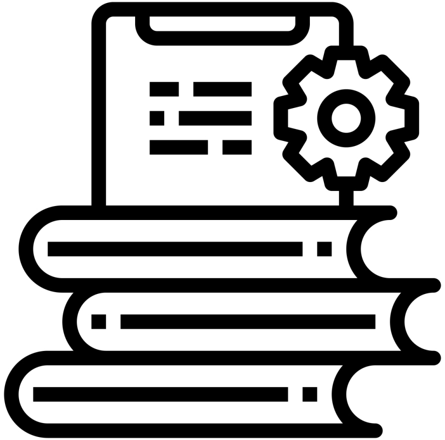
High level properties of GHA elements

- Jobs and Steps
  - Workflow = 1 or more jobs
  - Job = set of steps that run on the same machine (runner)

High level properties of GHA elements

- Jobs and Steps
  - Workflow = 1 or more jobs
  - Job = set of steps that run on the same machine (runner)
  - Step = command/script/action

High level properties of GHA elements



- Jobs and Steps
    - Workflow = 1 or more jobs
    - Job = set of steps that run on the same machine (runner)
    - Step = command/script/action
    - Job concurrency:
        - Default: all jobs run in parallel (possibly on different runners)
        - A job can `needs` another job
        - Concurrency groups => sequential jobs

High level properties of GHA elements

- Jobs and Steps
  - Workflow = 1 or more jobs
  - Job = set of steps that run on the same machine (runner)
  - Step = command/script/action
  - Job concurrency:
    - Default: all jobs run in parallel (possibly on different runners)
    - A job can `needs` another job
    - Concurrency groups => sequential jobs
  - Job matrix: different configs for the same job ->

Python\OS	ubuntu-latest	ubuntu-20.04
3.7	√	√
3.8	√	√
3.9	√	√

```
runs-on: ${{ matrix.os }}
strategy:
  matrix:
    os: [ubuntu-latest, ubuntu-20.04]
    python: [3.7, 3.8, 3.9]
steps:
  - uses: setup-python@v2
    with:
        python-version: ${{ matrix.python }}
```

High level properties of GHA elements

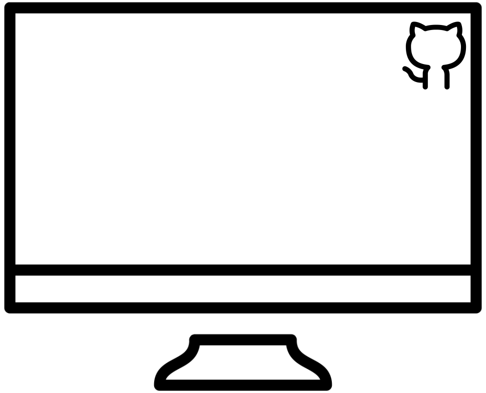
- Jobs and Steps
  - Workflow = 1 or more jobs
  - Job = set of steps that run on the same machine (runner)
  - Step = command/script/action
  - Job concurrency:
    - Default: all jobs run in parallel (possibly on different runners)
    - A job can `needs` another job
    - Concurrency groups => sequential jobs
  - Job matrix: different configs for the same job ->

Python\OS	ubuntu-latest	ubuntu-20.04
3.7	√	√
3.8	√	x
3.9	√	√

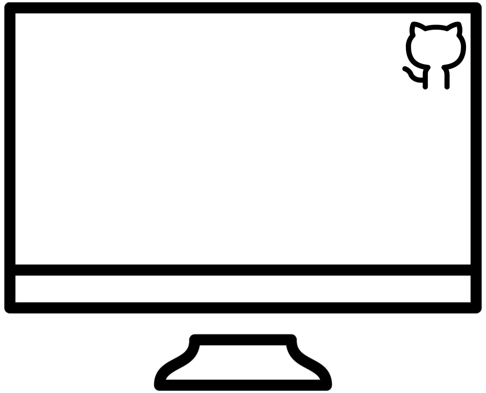
```
runs-on: ${{ matrix.os }}
strategy:
  matrix:
    os: [ubuntu-latest, ubuntu-20.04]
    python: [3.7, 3.8, 3.9]
steps:
  - uses: setup-python@v2
    with:
        python-version: ${{ matrix.python }}

exclude:
  - os: ubuntu-20.04
        python: 3.8
```

High level properties of GHA elements

- Runners

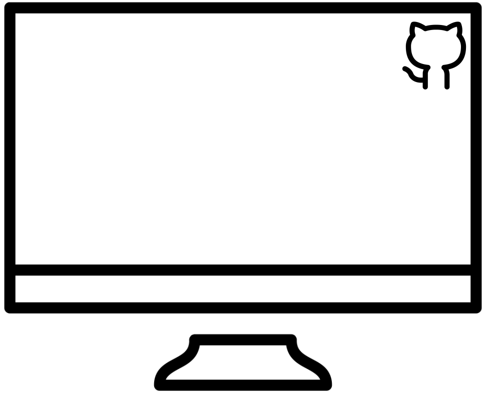


High level properties of GHA elements



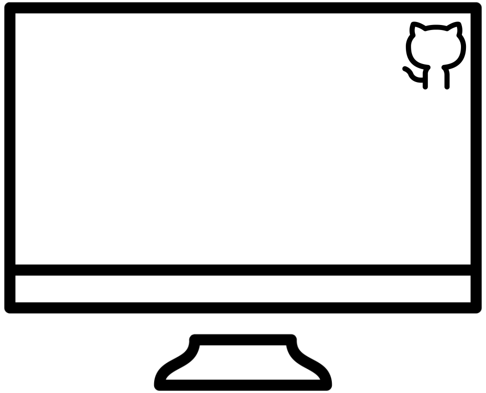
- Runners
  - Machine that your workflow `runs-on`

High level properties of GHA elements



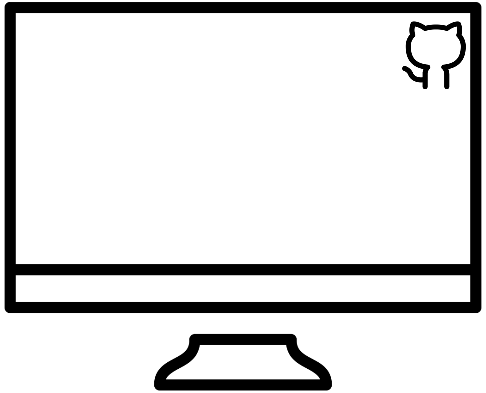
- Runners
  - Machine that your workflow `runs-on`
  - 1 job per machine at a time

High level properties of GHA elements

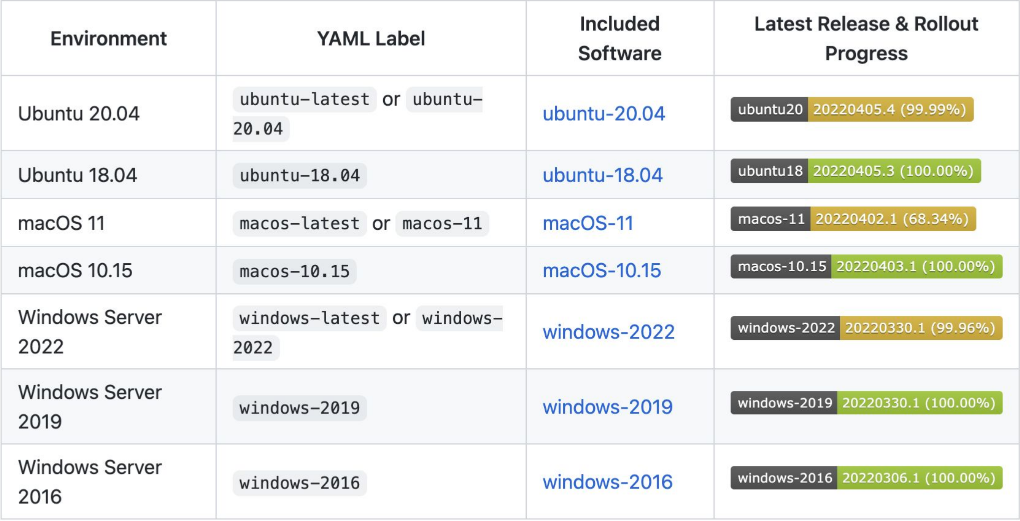


- Runners
  - Machine that your workflow `runs-on`
  - 1 job per machine at a time
  - Fresh virtual environment for each workflow

High level properties of GHA elements

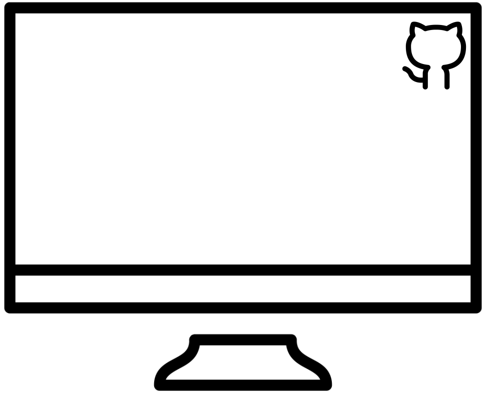


- Runners
  - Machine that your workflow `runs-on`
  - 1 job per machine at a time
  - Fresh virtual environment for each workflow
  - Available OS versions: https://github.com/actions/virtual-environments

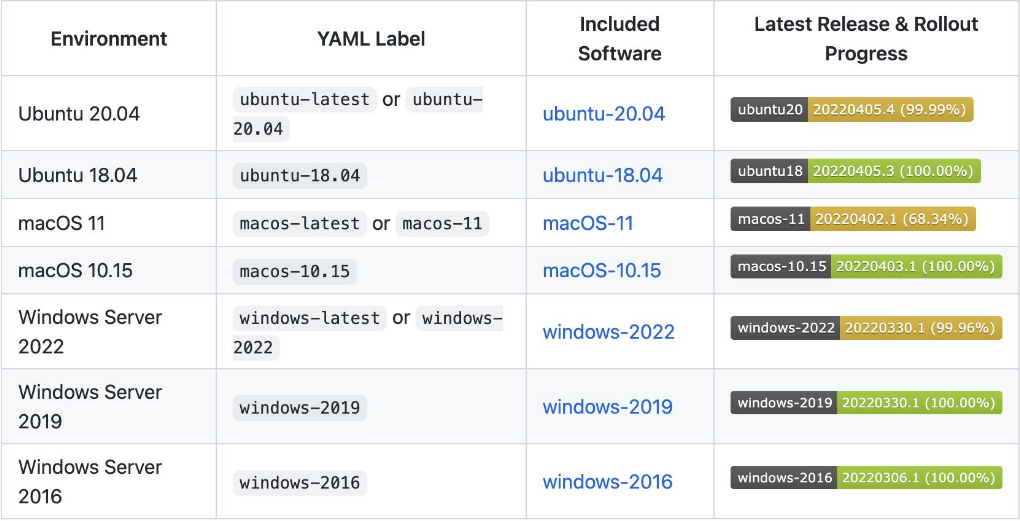


Environment	YAML Label	Included Software	Latest Release & Rollout Progress
Ubuntu 20.04	`ubuntu-latest` or `ubuntu-20.04`	ubuntu-20.04	ubuntu20 20220405.4 (99.99%)
Ubuntu 18.04	`ubuntu-18.04`	ubuntu-18.04	ubuntu18 20220405.3 (100.00%)
macOS 11	`macos-latest` or `macos-11`	macOS-11	macos-11 20220402.1 (68.34%)
macOS 10.15	`macos-10.15`	macOS-10.15	macos-10.15 20220403.1 (100.00%)
Windows Server 2022	`windows-latest` or `windows-2022`	windows-2022	windows-2022 20220330.1 (99.96%)
Windows Server 2019	`windows-2019`	windows-2019	windows-2019 20220330.1 (100.00%)
Windows Server 2016	`windows-2016`	windows-2016	windows-2016 20220306.1 (100.00%)

High level properties of GHA elements



- Runners
  - Machine that your workflow `runs-on`
  - 1 job per machine at a time
  - Fresh virtual environment for each workflow
  - Available OS versions: https://github.com/actions/virtual-environments
  - Can self-host runners
    - Other OS
    - Custom hardware



Environment	YAML Label	Included Software	Latest Release & Rollout Progress
Ubuntu 20.04	`ubuntu-latest` or `ubuntu-20.04`	ubuntu-20.04	ubuntu20 20220405.4 (99.99%)
Ubuntu 18.04	`ubuntu-18.04`	ubuntu-18.04	ubuntu18 20220405.3 (100.00%)
macOS 11	`macos-latest` or `macos-11`	macOS-11	macos-11 20220402.1 (68.34%)
macOS 10.15	`macos-10.15`	macOS-10.15	macos-10.15 20220403.1 (100.00%)
Windows Server 2022	`windows-latest` or `windows-2022`	windows-2022	windows-2022 20220330.1 (99.96%)
Windows Server 2019	`windows-2019`	windows-2019	windows-2019 20220330.1 (100.00%)
Windows Server 2016	`windows-2016`	windows-2016	windows-2016 20220306.1 (100.00%)

High level properties of GHA elements

- Actions

High level properties of GHA elements

- Actions
  - Is the code that steps within a job execute
  - Typically more involved than a script (more like a mini-app). E.g.
    - Set up a Node.js environment
    - Set up authentication to a Cloud provider
    - Upload code coverage reports for your repository to codecov.io
  - Community-driven Actions marketplace: https://github.com/marketplace?type=actions
  - You can write your own Actions and optionally share with the community
    - GitHub provides an Actions toolkit to help you get started: https://github.com/actions/toolkit
  - Piece of cake to add an existing Action to your workflow
    - Live example in a bit

High level properties of GHA elements

- Artifacts

High level properties of GHA elements

- Artifacts
  - Outputs of workflow jobs
  - Used to pass data between jobs in the same workflow
    - Remember steps of a job run on the same runner -> simple IO using the file system
  - … Or upload the results of a run for future access. E.g. Logs from a test run
  - Default storage: 90 days
    - Override using the `retention-days` property of the `upload-artifact` Action
  - Can access artifacts via the UI - we'll see it live in a bit
  - … Or programmatically using Actions:
    - `actions/download-artifact`
    - `actions/upload-artifact`

Live demo

Live demo

https://github.com/watiss/CCB-Skills-GitHubActions

What can I use GHA for in my research?

GHA example use cases

- Typical example is CI (Continuous Integration)
  - Why CI? Keep repo as up-to-date as possible => reduce bugs and merge conflicts
  - Why automated CI? Do you really want to check out your repo and run all your tests on different platforms with different OS/Python version combinations *manually* everytime you push to your repo? … I didn't think so
  - CI workflows do the above and more (linters, code coverage, etc.)
  - GitHub has a plethora of starter CI workflows: https://github.com/actions/starter-workflows/tree/main/ci
  - Let’s look at an example: https://github.com/scverse/scvi-tools/actions

GHA example use cases

- Publish a Jupyter Book to a site (e.g. using GitHub Pages)
- And keep it up to date using GitHub Actions
- Example: https://github.com/vals/single-cell-studies

GHA use cases

- Automatic Rendering of a Plot with GitHub Action: https://amitlevinson.com/blog/automated-plot-with-github-actions/
- Keep the README up to date by having a GitHub Actions re-generate the plot it shows upon every relevant push
- Let’s have a look: https://github.com/AmitLevinson/TidyTuesday

GHA use cases

- Host your up to date Latex files at a permanent URL: https://davidegerosa.com/githubforlatex/
- Idea: GHA workflow compiles your tex files and pushes the artifacts on a branch - the latter is used as the source for the site hosting your paper
- Let’s have a look: https://github.com/dgerosa/writeapaper/tree/main
- … You could even have it be emailed to you once ready! Check out this action on the marketplace: https://github.com/marketplace/actions/send-email

Resources & Acknowledgements

Resources & Acknowledgements

- Official GHA website and guides: https://docs.github.com/en/actions
- Thanks to Adam Gayoso for valuable feedback and suggestions

Thanks for listening!