

squash the flakes!

SFSCon 2023
[Daniel Hiller](#)

agenda

- about me
- introduction to flakes
- how do we minimize the impact
- how do we exercise the flake process
- what tools do we have
- the future
- want to help?
- Q&A

about me

- Software Engineer @ [Red Hat OpenShift Virtualization](#) team
- [KubeVirt](#) CI, automation in general

what is a flake

a flake is a test
that without any code change
will either fail or pass in successive runs

what is the impact of flakes

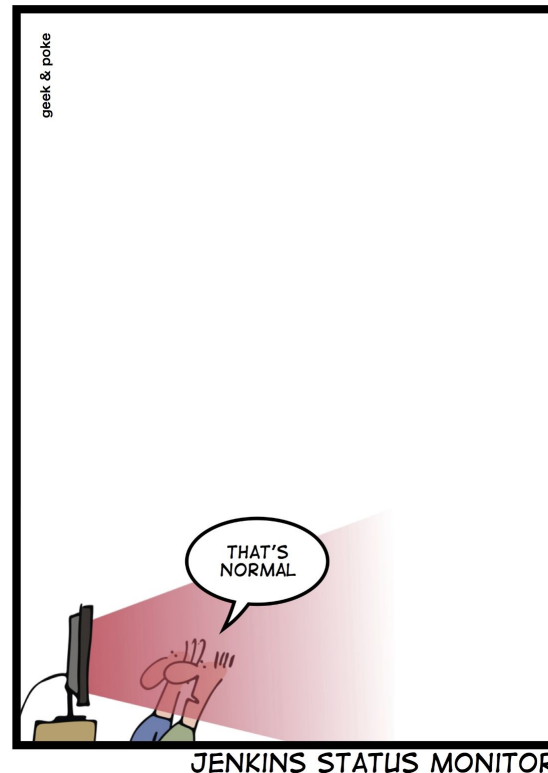
from “a survey of flaky tests”:

- 97% of flakes were false alarms, and
- more than 50% of flakes could not be reproduced in isolation

this leads to the conclusion:
“ignoring flaky tests is ok”

source: [“A survey of flaky tests”](#)

SIMPLY EXPLAINED



what is the impact of flakes

failed test runs signal that the code is not stable

a flake can invalidate the entire test run

thus failed test runs with flaky tests have no value in CI

what is the impact of flakes

Flaky tests lead to:

- longer feedback cycles for developers
- slowdown of merging pull requests - “retest trap”
- reversal of acceleration which batch testing provides
- **lost trust in automated testing**
- **ignoring test results**

how we minimize
the impact

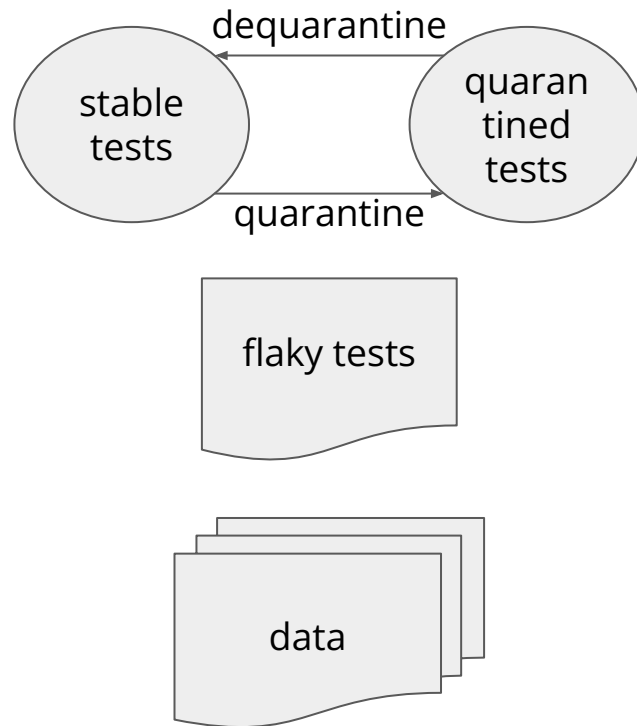
tldr; exclude (aka
quarantine) a flaky test
from test runs **as early**
as possible, but **only as**
long as necessary



how we minimize the impact

what do we need?

- ability to move a test from **set of stable** tests into **set of quarantined** tests and back
- a report over possible flaky tests
- enough runtime data to triage flakes
 - devs decide whether we quarantine right away or they can fix them in time



how we minimize the impact

how can we find flaky tests?

any merged PR had all tests succeeding in the end,

thus any test run with test failures from that PR *might* contain execution of flaky tests

PR History: kubevirt/kubevirt #10634

PR Title	PR Number	PR Number	PR Number
pull-kubevirt-apidocs	1719919846822514688	1719762310676877312	1717241106644602880
pull-kubevirt-build-arm64	1719919846436638720	1719762310169366528	1717241105210150912
pull-kubevirt-build	1719919846306615296	1719762310098063360	1717241105109487616
pull-kubevirt-check-tests-for-flakes	1717241104786526208		
pull-kubevirt-check-unassigned-tests	1719919848957415424	1719762312748863488	1717241109979074360
pull-kubevirt-client-python	1719919846923177984	1719762310819483648	1717241107496046592
pull-kubevirt-code-lint	1717241113405820928		
pull-kubevirt-e2e-arm64	1717241112550182912		
pull-kubevirt-e2e-k8s-1.26-sig-compute	1719919853160108032	1719762316993499136	1717241115981123584
pull-kubevirt-e2e-k8s-1.26-sig-network	1719919851462386432	1719762315433218048	1717241114177572864
pull-kubevirt-e2e-k8s-1.26-sig-operator	1719919853998968832	171979492280061952	171996748920101888
pull-kubevirt-e2e-k8s-1.26-sig-storage	1719919852317052928	1719762316116889600	1717241115045793792
pull-kubevirt-e2e-k8s-1.27-ipvs-network	1719919849825636352	1719762313587724288	1717241110838906880
pull-kubevirt-e2e-k8s-1.27-sig-compute	1719919855999651840	1719762320290222080	1717241119250875008
pull-kubevirt-e2e-k8s-1.27-sig-network	1719919854888161280	1719762318616694784	1717241117553987584
pull-kubevirt-e2e-k8s-1.27-sig-operator	1719919856494579712	1719762321137471488	1717241120154450664
pull-kubevirt-e2e-k8s-1.27-sig-storage	1719900030295181312	1719815842004209664	1719762319455555584
pull-kubevirt-e2e-k8s-1.28-sig-compute-migrations	1720001509401300992	1719976014571900928	1719919850681274368
pull-kubevirt-e2e-k8s-1.29-sig-compute	1719987935857610732	1719919859040322240	171976232666636800
pull-kubevirt-e2e-k8s-1.28-sig-network	1719919857366994944	171976232195955072	171724112092013696
pull-kubevirt-e2e-k8s-1.28-sig-operator	1719919860001017856	1719762324488720384	1717241123501510656
pull-kubevirt-e2e-k8s-1.28-sig-storage	1719919858210050048	171976232284038912	1717241121941229568
pull-kubevirt-e2e-kind-1.27-stov	1719919845945905152	1719762309670244352	1717241104606171136
pull-kubevirt-e2e-kind-1.27-vgpu	1719919845832658944	1719778795453419520	1717241104513896448
pull-kubevirt-e2e-windows2016	17199198445677469696	1719772088580968448	1717241104309645936
pull-kubevirt-fossa	17199198464633920	1719762310567825408	171724110576381952

how does the flake process work

regular meeting

- look at flakes
- decide: fix or quarantine?
- hand to dev
- bring back in

emergency quarantine

source: [QUARANTINE.md](https://quarantine.md)



what tools do we have

ci honoring QUARANTINE label

- presubmits **skip** quarantined tests
- periodics **execute** quarantined tests to check their stability

```
# If KUBEVIRT_QUARANTINE is not set, do not run quarantined tests. When it is
# set the whole suite (quarantined and stable) will be run.
if [ -z "$KUBEVIRT_QUARANTINE" ]; then
  if [ -n "$KUBEVIRT_E2E_SKIP" ]; then
    export KUBEVIRT_E2E_SKIP="${KUBEVIRT_E2E_SKIP}|QUARANTINE"
  else
    export KUBEVIRT_E2E_SKIP="QUARANTINE"
  fi
fi
```

```
176
177 It("[QUARANTINE]should successfully upgrade virt-handler", func() {
178     var expectedEventsLock sync.Mutex
179     expectedEvents := []string{
180         "maxUnavailable=1",
181         "maxUnavailable=10%",
182         "virt-handler=ready",
183         "maxUnavailable=1",
184     }
185
186     ds, err := virtCli.AppsV1().DaemonSets(flags.KubeVirtInstall
```

sources:

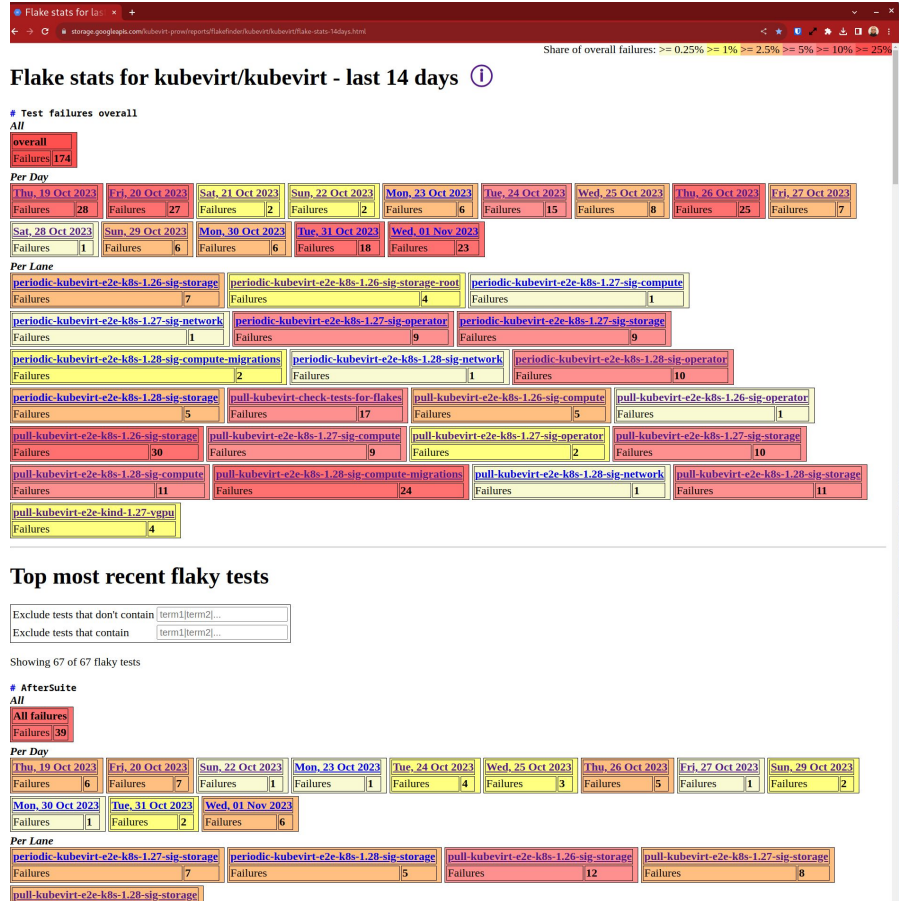
- <https://github.com/kubevirt/kubevirt/blob/38c01c34acecfafc89078b1bbaba8d9cf3cf0d4d/automation/test.sh#L452>
- <https://github.com/kubevirt/kubevirt/blob/38c01c34acecfafc89078b1bbaba8d9cf3cf0d4d/hack/functests.sh#L69>
- https://github.com/kubevirt/kubevirt/blob/38c01c34acecfafc89078b1bbaba8d9cf3cf0d4d/tests/canary_upgrade_test.go#L177

what tools do we have

[flake stats](#) report

the high level overview

([source](#))



what tools do we have

[flakefinder](#) report

the detail overview

gives an overview of the current flaky tests

flakefinder report for kubevirt/kubevirt

Data range from 2023-03-21T00:00:00Z till 2023-03-21T23:59:59Z
 Source PRs: #9472, #9455, #9392, #9420, #9462, #9445.


	periodic-kubevirt-e2e-k8s-1.24-sig-storage-groups2	periodic-kubevirt-e2e-k8s-1.25-sig-compute-migrations	periodic-kubevirt-e2e-k8s-1.26-sig-compute	periodic-kubevirt-e2e-k8s-1.26-sig-compute-root	periodic-kubevirt-e2e-k8s-1.26-sig-storage	periodic-kubevirt-e2e-k8s-1.26-sig-storage-root	pull-kubevirt-e2e-k8s-1.24-sig-storage	pull-kubevirt-e2e-k8s-1.25-sig-compute	pull-kubevirt-e2e-k8s-1.25-sig-compute-migrations	pull-kubevirt-e2e-k8s-1.25-sig-compute-migrations	pull-kubevirt-e2e-k8s-1.25-sig-compute-migrations
0	0/0/1	1/0/0	0/0/1	0/0/1	0/0/2	0/0/1	0/0/1	0/0/1	0/0/1	3/0/1	
1	0/0/1	0/0/1	0/1/0	1/0/0	0/0/2	0/0/1	0/0/1	1/0/0	0/0/4		
2	0/1/0	0/0/1	0/0/1	0/0/1	0/2/0	1/0/0	0/1/0	0/0/1	0/0/4	1/2/0	2/5/0
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0/0/2	1/1/0	2/4/0

Serial **sig-compute**

rfe_id : 393 **crit : high**

vendor : cnv-qe@redhat.com

level : system

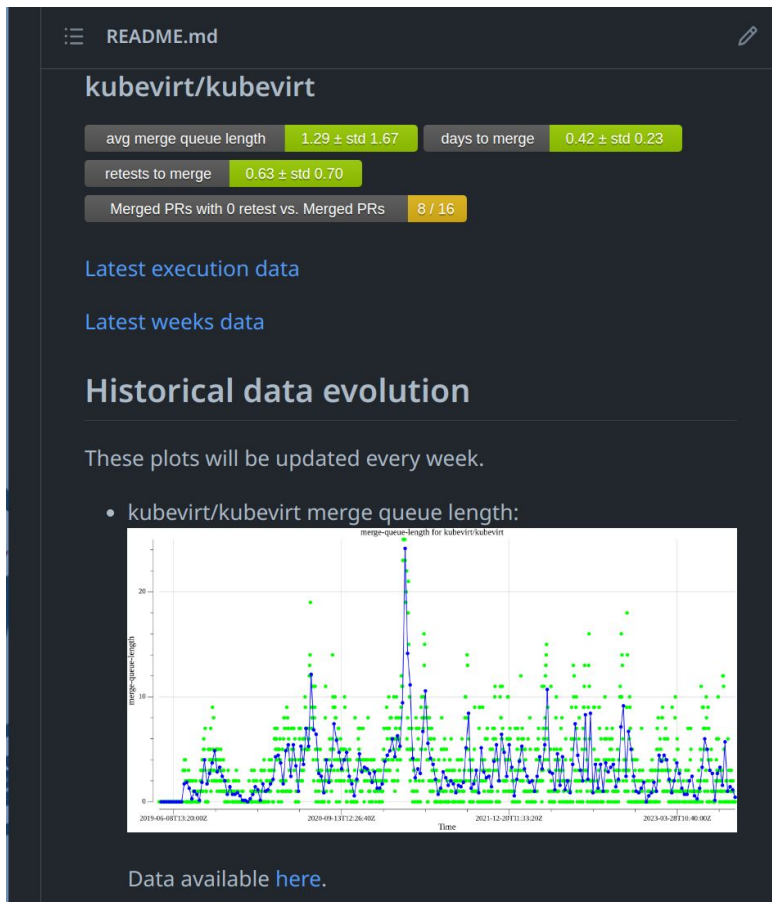
VM Live Migration with a dedicated migration network Should migrate over that network 

1/2/0	2/5/0
1/1/0	2/4/0

what tools do we have

[ci-health](#)

record metrics over merge-queue-length, time-to-merge, retests-to-merge and merges-per-day



what tools do we
have

[quarantine overview](#)

([source](#))

Overview of Quarantine tests

Total: 1 tests

<https://github.com/kubevirt/kubevirt/tree/c35be138d85864b17946ae1fe07f99a75445d501/tests/operator/operator.go>

```
var _ = Describe("[Serial][sig-operator]Operator", Serial, decorators.SigOperator, func() {  
    Describe("[rfe_id:2897][crit:medium][vendor:cnv-qe@redhat.com][level:component]Dynamic feature  
detection", func() {
```

2 months **It("[test_id:3153][QUARANTINE] Ensure infra can handle dynamically detecting
DataVolume Support", func() {**

Last updated: 2023-09-08 10:18:02.458076498 +0000 UTC m=+3.056989100

what tools do we have

[testgrid](#)

drill down on all jobs for kubevirt/kubevirt that are running inside KubeVirt Prow


[kubevirt](#)
[kubevirt-containerized-data-importer-presubmits](#)
kubevirt-presubmits
[kubevirt-periodics](#)

✓	periodic-kubevirt-e2e-k8s-1.24-operator: PASSING Tab stats: 10 of 10 (100.0%) recent columns passed (490 of 490 or 100.0% cells)	Last update: 03-22 14:44 CET Tests last ran: 03-22 12:55 CET Last green run: 38c01c34a
⊖	periodic-kubevirt-e2e-k8s-1.24-sig-compute: FLAKY Tab stats: 9 of 10 (90.0%) recent columns passed (6725 of 6727 or 100.0% cells)	Last update: 03-22 14:43 CET Tests last ran: 03-22 11:38 CET Last green run: 211553d1b
✓	periodic-kubevirt-e2e-k8s-1.24-sig-compute-cgroupsv2: PASSING Tab stats: 10 of 10 (100.0%) recent columns passed (6709 of 6709 or 100.0% cells)	Last update: 03-22 14:21 CET Tests last ran: 03-21 22:31 CET Last green run: 3679b6014
✓	periodic-kubevirt-e2e-k8s-1.24-sig-network: PASSING Tab stats: 10 of 10 (100.0%) recent columns passed (2014 of 2014 or 100.0% cells)	Last update: 03-22 14:43 CET Tests last ran: 03-22 08:10 CET Last green run: 211553d1b
⊖	periodic-kubevirt-e2e-k8s-1.24-sig-storage: FLAKY Tab stats: 9 of 10 (90.0%) recent columns passed (2548 of 2550 or 99.9% cells)	Last update: 03-22 14:43 CET Tests last ran: 03-22 09:38 CET Last green run: 211553d1b

jrowjob_name: periodic-kubevirt-e2e-k8s-1.26-sig-storage
 03-22 13:33 CET @1638490818345963520 -- 01-21 19:40 CET @1616868257380175872; Serv

About ▾ Size ▾ Options ▾ Graph ▾ Local Time: ON

Display Clustered Failures List

03-21 03-20 03-19
 3bee... c453

[Show 7 stale tests](#) (no results in last 10+ runs)

Merged Test Suite.[sig-storage] DataVolume Integration Fedora VMI te	F								
Merged Test Suite.[sig-storage] Export Should recreate the exporter pr	F								
periodic-kubevirt-e2e-k8s-1.26-sig-storage.Overall	R								
Merged Test Suite.[sig-storage] VirtualMachineRestore Tests [storage-				F					
Merged Test Suite.[sig-storage] DataVolume Integration Fedora VMI te				F					
Merged Test Suite.[sig-storage] DataVolume Integration Fedora VMI te						F	F		
Merged Test Suite.[sig-storage] Storage Starting a VirtualMachineInsta									
Merged Test Suite.[sig-storage] Storage Starting a VirtualMachineInsta									
Merged Test Suite.[sig-storage] [Serial]ImageUpload [storage-req] Upl									

what tools do we have

[check-tests-for-flakes test lane](#)

tries to catch flakes before they enter the codebase

([source](#))

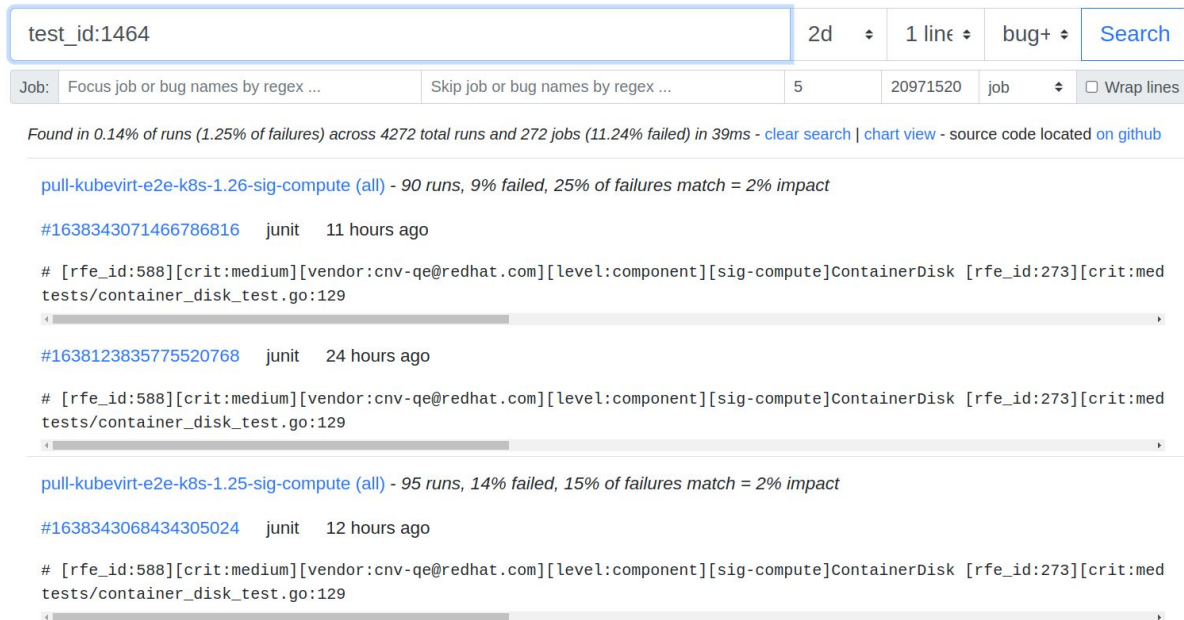
a test lane that

- selects the changed e2e test files from the commit set
- runs changed e2e tests five times
- runs in random execution order to catch order-dependent test

what tools do we have

[ci-search](#)

search for terms in prow job logs (see [openshift ci-search](#))



test_id:1464 2d 1 line bug+ Search

Job: Focus job or bug names by regex ... Skip job or bug names by regex ... 5 20971520 job Wrap lines

Found in 0.14% of runs (1.25% of failures) across 4272 total runs and 272 jobs (11.24% failed) in 39ms - [clear search](#) | [chart view](#) - source code located [on github](#)

[pull-kubevirt-e2e-k8s-1.26-sig-compute \(all\)](#) - 90 runs, 9% failed, 25% of failures match = 2% impact

#1638343071466786816 junit 11 hours ago

```
# [rfe_id:588][crit:medium][vendor:cnv-qe@redhat.com][level:component][sig-compute]ContainerDisk [rfe_id:273][crit:med
tests/container_disk_test.go:129
```

#1638123835775520768 junit 24 hours ago

```
# [rfe_id:588][crit:medium][vendor:cnv-qe@redhat.com][level:component][sig-compute]ContainerDisk [rfe_id:273][crit:med
tests/container_disk_test.go:129
```

[pull-kubevirt-e2e-k8s-1.25-sig-compute \(all\)](#) - 95 runs, 14% failed, 15% of failures match = 2% impact

#1638343068434305024 junit 12 hours ago

```
# [rfe_id:588][crit:medium][vendor:cnv-qe@redhat.com][level:component][sig-compute]ContainerDisk [rfe_id:273][crit:med
tests/container_disk_test.go:129
```

in a nutshell

In regular intervals:

- follow up on previous action items
- derive action items from data available
- hand action items over to dev teams
- revisit and dequarantine quarantined tests

the future - more data, more tooling

gaps we want to close:

- a quick dashboard overview of how exactly we are progressing (beyond ci-health)
- overview of current number of flakes in quarantine per sig
- automatic quarantine PRs when new flakes have entered the codebase



Q&A

Feel free to send questions and comments:

mailto: dhiller@redhat.com

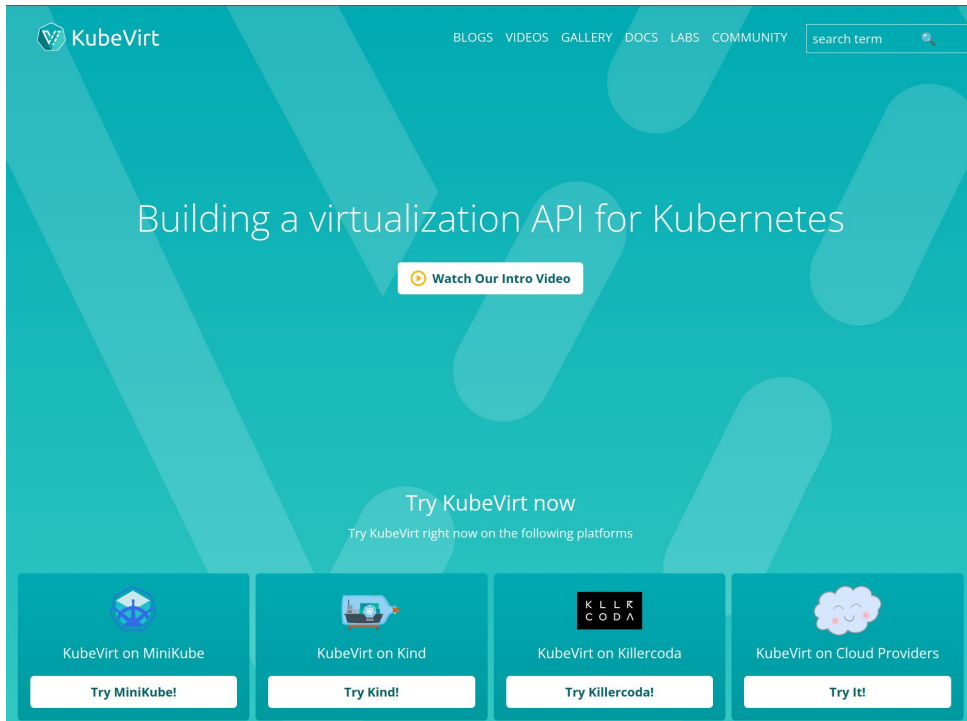
k8s slack: [@dhiller](#)

mastodon: [@dhiller@fosstodon.org](https://fosstodon.org/@dhiller)

web: www.dhiller.de

interested in kubevirt.io?
want to help?

- join [#kubevirt-dev](#) Slack channel
- join [kubevirt-dev](#) Google group
- fix flakes on [kubevirt/kubevirt](#)



The screenshot shows the KubeVirt website homepage. The header includes the KubeVirt logo, navigation links for BLOGS, VIDEOS, GALLERY, DOCS, LABS, and COMMUNITY, and a search bar. The main heading is "Building a virtualization API for Kubernetes" with a "Watch Our Intro Video" button. Below this, it says "Try KubeVirt now" and "Try KubeVirt right now on the following platforms". At the bottom, there are four cards for different platforms: KubeVirt on MiniKube, KubeVirt on Kind, KubeVirt on Killercoda, and KubeVirt on Cloud Providers, each with a "Try" button.