

squash the flakes!

FOSDEM 2024
[Daniel Hiller](#)

agenda

- about me
- about flakes
- impact of flakes
- flake process
- tools
- the future
- Q&A
- want to help?

about me

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

about flakes

a question:

who can explain what a flake is?

about flakes

a *flake*

...

...

...

about flakes

a *flake*

is a **test** that

without any code change

will either **fail** or **pass** in successive runs

about flakes

another question:

who thinks handling flakes is important?

about flakes

another question:

who has had to deal with flakes?

about flakes

another question:

who has to deal with flakes on a regular basis?

about flakes

flakes are caused

either by production code (a bug)

or flaky test code (also a bug, but handled differently)

about flakes

PR History: kubevirt/kubevirt #9445

9d41878

pull-kubevirt-e2e-k8s-1.25-sig-compute-migrations	1637934812398358528	1636633531918585856	1636403749595385856
pull-kubevirt-e2e-k8s-1.25-sig-compute	1637934815221125120	1636403757321293824	
pull-kubevirt-e2e-k8s-1.25-sig-network	1637934813975416832	1636403756985749504	
pull-kubevirt-e2e-k8s-1.25-sig-operator	1637934815393091584	1636403757992382464	
pull-kubevirt-e2e-k8s-1.25-sig-storage	1637934814088663040	1636633532048609280	1636403756704731136
pull-kubevirt-e2e-k8s-1.26-sig-compute	1637934816471027712	1636404222087925760	
pull-kubevirt-e2e-k8s-1.26-sig-network	16379348160855151744	1636403758015129344	

pull-kubevirt-e2e-k8s-1.25-sig-compute-migrations #1636633531918585856

Job History PR History Artifacts

Test started last Friday at 8:40 AM passed after 1h9m49s. [\(more info\)](#)

JUnit

91/1406 Tests Passed!

1315/1406 Tests Skipped.

Build Log

Show all hidden lines [Raw build log.txt](#)

pull-kubevirt-e2e-k8s-1.25-sig-compute-migrations #1636403749595385856

Job History PR History Artifacts

Test started last Thursday at 7:33 PM failed after 1h18m18s. [\(more info\)](#)

JUnit

1/1406 Tests Failed.

Tests Suite: [rfe_id:393][crit:high][vendor:cnv-qe@redhat.com][level:system][sig:compute] VM Live Migration [Serial]with a dedicated migration network Should migrate over that network 4m18s

90/1406 Tests Passed!

1315/1406 Tests Skipped.

source: <https://prow.ci.kubevirt.io/pr-history/?org=kubevirt&repo=kubevirt&pr=9445>

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”](#)

impact of flakes



impact of flakes

in CI automated testing must give a **reliable signal of stability**

any failed test run signals that the product is **unstable**

test runs failed due to flakes do not give this reliable signal

they only waste time

impact of flakes

Flaky tests waste everyone's time - they cause

- longer feedback cycles for developers
- slowdown of merging pull requests - "retest trap"
- reversal of acceleration effects (i.e. batch testing)

impact of flakes

Flaky tests also cause trust issues - they make people

- lose trust in automated testing
- ignore test results

minimizing the
impact

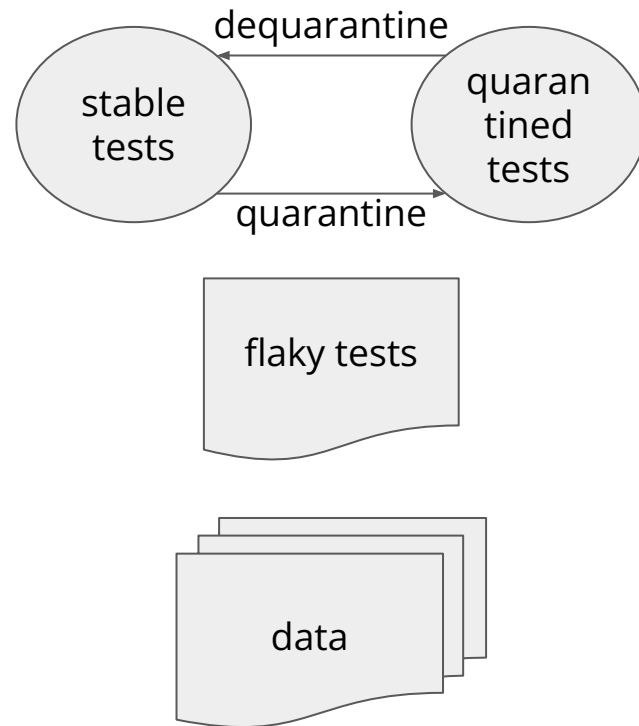
tldr; exclude (aka
quarantine) a flaky test
from test runs **as early
as possible**, but **only as
long as necessary**



minimizing 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



minimizing 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 Status	PR Link
pull-kubevirt-apidocs	1719919846822514688	Success	1717241106646602880
pull-kubevirt-build-arm64	1719919846436638720	Success	171724110521105105912
pull-kubevirt-build	1719919846306615296	Success	1717241105109487616
pull-kubevirt-check-tests-for-flakes	1717241104786526208	Success	
pull-kubevirt-check-unassigned-tests	1719919848957415424	Success	1717241109979074360
pull-kubevirt-client-python	1719919846923177984	Success	1717241107495046592
pull-kubevirt-code-lint	1717241113405820928	Success	
pull-kubevirt-e2e-arm64	1717241112550182912	Success	
pull-kubevirt-e2e-k8s-1.26-sig-compute	1719919853160108032	Success	1717241115981123584
pull-kubevirt-e2e-k8s-1.26-sig-network	1719919851462386432	Success	1717241114173752864
pull-kubevirt-e2e-k8s-1.26-sig-operator	1719919853998968832	Success	171979492280061952
pull-kubevirt-e2e-k8s-1.26-sig-storage	1719919852317052928	Success	1717241115045793792
pull-kubevirt-e2e-k8s-1.27-ipvs-sig-network	1719919849825636352	Success	1717241110838906880
pull-kubevirt-e2e-k8s-1.27-sig-compute	1719919855999651840	Success	1717241119250875008
pull-kubevirt-e2e-k8s-1.27-sig-network	1719919854888161280	Success	1717241117533987584
pull-kubevirt-e2e-k8s-1.27-sig-operator	1719919856494579712	Success	1717241120154450664
pull-kubevirt-e2e-k8s-1.27-sig-storage	1719900030295181312	Success	171976231945555584
pull-kubevirt-e2e-k8s-1.28-sig-compute-migrations	1720001509401300992	Success	1719762314430779392
pull-kubevirt-e2e-k8s-1.29-sig-compute	1719987935857610732	Success	17197623266663800
pull-kubevirt-e2e-k8s-1.28-sig-network	1719919857366994944	Success	171724112092013696
pull-kubevirt-e2e-k8s-1.28-sig-operator	1719919860001017856	Success	1717241123501510656
pull-kubevirt-e2e-k8s-1.28-sig-storage	1719919858210050048	Success	1717241121941229568
pull-kubevirt-e2e-kind-1.27-stdio	1719919845945905132	Success	1717241104606171136
pull-kubevirt-e2e-kind-1.37-vgpu	1719919845832658944	Success	1717241104513896448
pull-kubevirt-e2e-windows2016	17199198445677469696	Success	1717241104309645936
pull-kubevirt-fossa	17199198464633920	Success	171724110576381952

the flake process

regular meeting

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

emergency quarantine

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



tools

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

tools

[flake stats](#) report

the high level overview

([source](#))



tools

[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-groupsv2	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

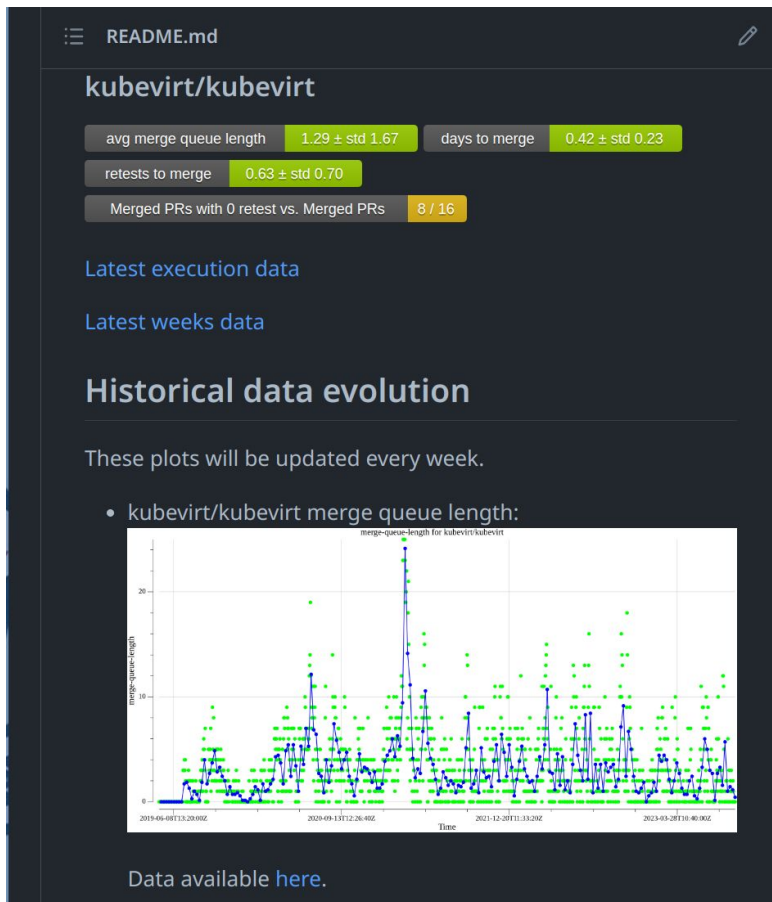
VM Live Migration with a dedicated migration network Should migrate over that network 

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

tools

[ci-health](#)

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



tools

[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

tools

[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

rowjob_name: periodic-kubevirt-e2e-k8s-1.26-sig-storage
 03-22 13:33 CET @1638490818345963520 -- 01-21 19:40 CET @1616686257380175872; 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									

tools

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

tries to catch flakes before they enter the codebase

([source](#))

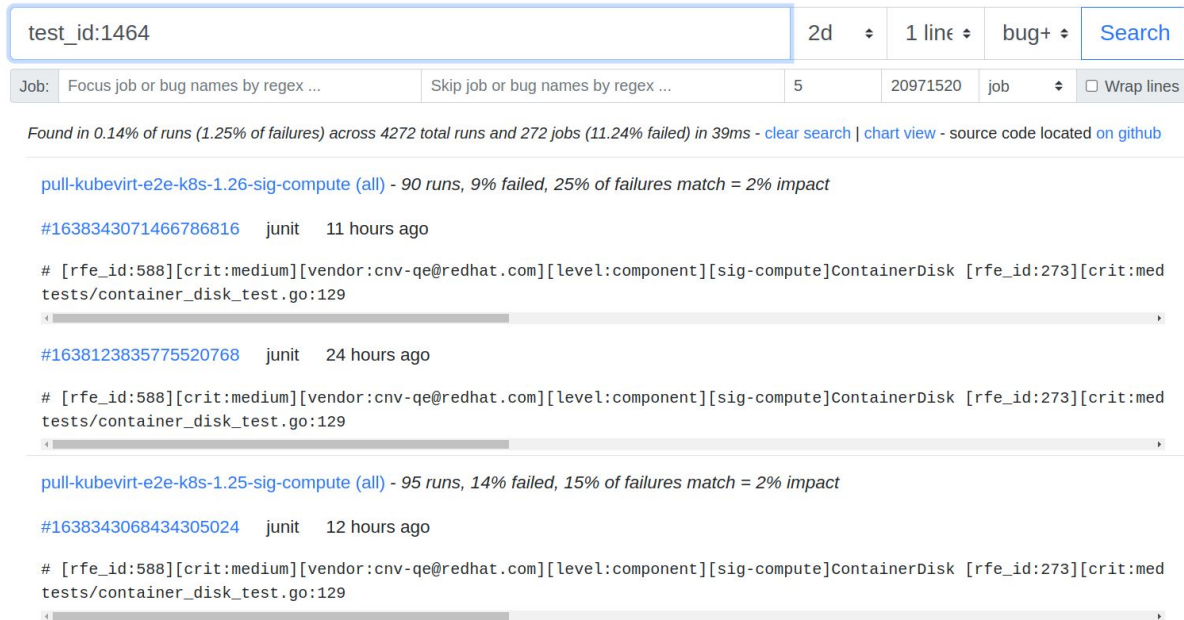
a test lane that

- 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

tools

[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:

- collect more data - run the majority of tests frequently
- close the retest gap - stop retesting after a certain threshold
- get better in detecting new flakes
- long term - automatic quarantine PRs when new flakes have entered the codebase



Q&A

Any questions?

Who else is trying to tackle this problem?

What have you done to solve this problem?

Thank you for your patience!

Feel free to send questions and comments:

mailto: dhiller@redhat.com

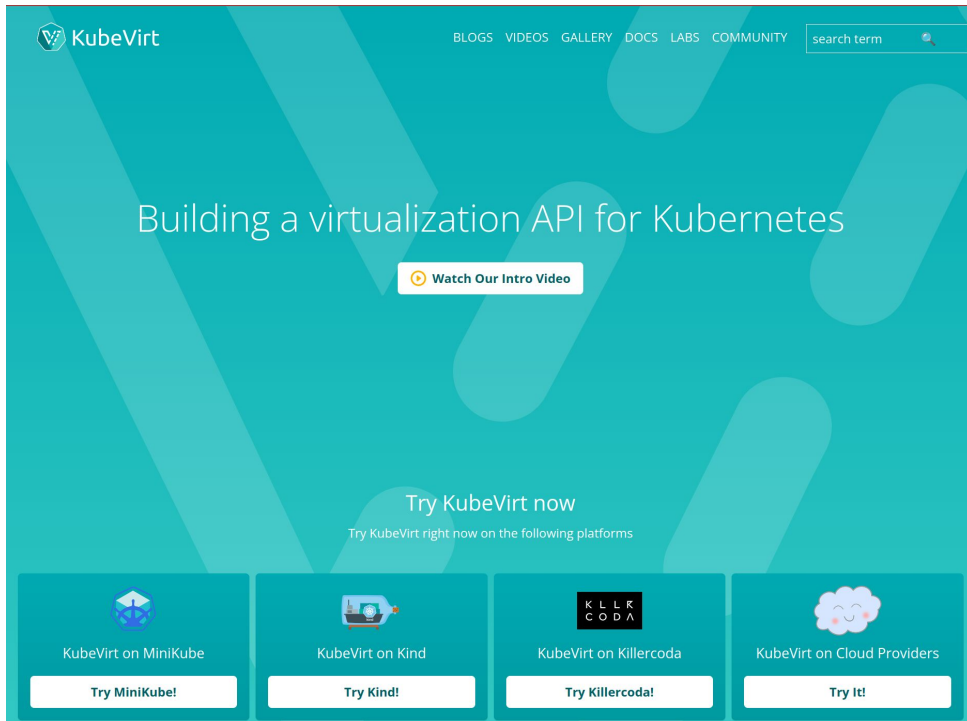
k8s slack: [@dhiller](#)

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 is a "Try KubeVirt now" section with the text "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.