CNCF Research User Group

Meetings on the 1st and 3rd Wednesday of every month at 17:00 CET / 8:00 PT / 16:00 UTC / 11:00 ET

- RSVP to the meetings at CNCF Research End User Group
- Fill this form to get invites to the meetings on your calendar
- Slack: #ug-research in http://slack.cncf.io/
- Mailing list: https://lists.cncf.io/g/cncf-research-user-group/topics
- Repo: https://github.com/cncf/research-user-group.
- Bevy Platform: https://community.cncf.io/research-end-user-group/
    - To join a meeting:
        - Log into Bevy
        - Choose the specific date you want to join(make sure the email address of the account and ticket recipient match).
        - Click the blue “JOIN VIRTUAL EVENT HERE!” button on the blue header…

Topic backlog:

- Gateway API
- Cluster API
- NUMA
- Usernetes
- Standards for submitting jobs
    - Annotations
    - Specifying job geometry
    - Handling prereqs including hardware (x86_64 vs arm images)
- Storage API

Lead: Ricardo Rocha (CERN), Jamie Poole (G-Research)

February 19th, 2025

Agenda:

- Managing Kubeflow deployments and updates at CERN

Attendees (please indicate if first time)**:**

- Sylvie Ramos (UCL)
- Andrew Esterson (UCL,UK - first time)
- Hannes Hansen (CERN, first time)
- Amine Lahouel (CERN, first time)

- Raul Chiorescu (CERN, first time)
- Clemens Lange (PSI)
- Ons Hanafi (student, ESPRIT Tunisia)
- Alessandro Vozza (Microsoft)
- Tia Miceli (Fermilab) (first time!)
- Linh Pham (Fermi lab,first time)
- Fan Yang (Bristol, UK, first time)
- Diego Ciangottini - INFN
- Maria Acosta - FNAL
- Darren Boss - University of Victoria
- Tim Kok - SURF
- Ricardo Rocha (CERN)

Notes:

- Deployment configuration to stitch centraldashboard together

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: centraldashboard
spec:
  template:
        spec:
        initContainers:
        - name: combine-js
        image: docker.io/library/python:3.7
        command: ["/bin/sh", "-c"]
        args:
        - cat /parts/part1/* /parts/part2/* > /work-dir/app.bundle.js
        volumeMounts:
        - name: workdir
        mountPath: /work-dir
        - name: bundle-part1
        mountPath: /parts/part1
        - name: bundle-part2
        mountPath: /parts/part2
        containers:
        - name: centraldashboard
        volumeMounts:
        - name: workdir
        mountPath: /usr/src/app/dist/public/app.bundle.js
        subPath: app.bundle.js
```

volumes:
- name: workdir
emptyDir: {}
- name: bundle-part1
configMap:
name: centraldashboard-bundle-part1
items:
- key: app.bundle.js.part-aa
path: app.bundle.js.part-aa
- name: bundle-part2
configMap:
name: centraldashboard-bundle-part2
items:
- key: app.bundle.js.part-ab
path: app.bundle.js.part-ab

September 18th, 2024

Agenda:

- Reference architectures for a cloud native scientific platform

Attendees (please indicate if first time)**:**

- Ricardo Rocha, Brian Major, Jonathan Skone, Piers Harding, Gabe Tocci, Matt Pryor, Alan Mutschelknaus

Notes:

- Collect feedback on template for organizing and intaking information about organization and platform details starting from Bob Killen's description of U Michigan research computing organization and platform details as a reference.
- Brian Major presented CANFAR overview.

September 4th, 2024

Agenda:

How Kubernetes is improving batch api via Job and JobSet, Kevin Hannon

- Update on Job/JobSet

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Kevin Hannon (Red Hat)
- Clemens Lange (PSI)

- Sylvie Ramos (UCL,UK)

August 21st, 2024

Agenda:

- Slinky, Slurm on Kubernetes

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Skyler Malinowski (SchedMD)
- Marlow Warnicke (SchedMD)
- Dennis Marttinen (Aalto University)
- Kevin Fox (PNNL)
- Nick Jones (Nscale)
- Darren Boss (Digital Research Alliance of Canada)
- Clemens Lange (PSI)
- Burt Holzman (Fermilab)
- Raj Ayyampalayam (Genentech)
- Timothy Middelkoop (Internet2)
- Jose E Torres (HPCNow!)
- Drew hudson-Viles (Nscale)
- Shawfeng Dong (Qualcomm)
- Diego Sonaglia (Clustervision)
- Ethan O'Dell (ORNL)
- Paddy O'Neill (Nscale)

Notes:

- Marlow presented the company (SchedMD) and project (Slurm), highlighting after the differences between HPC and Cloud Native worlds
- Skyler then introduced some of the components of Slurm and the Slurm Kubernetes Operator
    - Q: Is the operator joining the two worlds or how does the operator work?
    - A: The operator allows you to deploy Slurm nodes on Kubernetes, with nodes that can live in and out of the Kubernetes world
        - This is the first step to get accounting working on Kubernetes, more advanced features (scheduling, etc) will come later
    - Q: Limitations of running slurmd in Kubernetes?
    - A: cgroups is one of them as both kubelet and slurmd will manage them. One way to prevent this is to restrict workloads in these nodes to Slurm

    - Q: Is the networking hostNetwork or using the k8s stack?
    - A: Currently the k8s stack, but this will change
    - Q: Is there automated discovery of network topologies?
    - A: There is ongoing work to automatically discover the topology via either the NFD or some other way
    - Q: What's the way to add software to the worker nodes?
    - A: Two ways… either add the binaries to the 'worker' image, or rely on OCI containers, or host binaries in a shared filesystem accessible by all nodes
    - Q: Does `procTrackType: "proctrack/cgroup"` work in this model?
    - A: Not currently, submanagement of cgroups by Slurm does not currently work. But there will be work to add this
- Marlow then gave a demo of the operator
    - Q: Is there any particular handling of networking between Slurm nodes and nodes in bursting k8s clusters (for example in the public cloud)?
    - A: No magic, some sort of tunneling is required or ensuring co-location of workloads
- A plan for future work came next
    - Q: The work looks good and it focuses on the backend. Any plan to support a Kubernetes native API as well?
    - A: There is an idea to support Kubernetes Jobs (future planned work)
    - Q: Does this mean a custom resource or the native Job/JobSet/etc?
    - A: Likely the latter
    - Q: Any plans to support additional kubernetes resources, like mounting configmaps, CSI, etc?
    - A: Yes. One particular area that needs improvement in on the networking side of kubernetes. A suggestion to look at what kubevirt is trying to do as well
    - Q: Similar question regarding mapping and matching via labels and selectors?
    - A: NFD should be the source for all of this
    - Q: Any effort to look at alternatives to munge?
    - A: There are efforts in auth/slurm that go in this direction

(CANCELED) August 7th, 2024

Agenda:

- How Kubernetes is improving batch api via Job and JobSet
    - Event has been moved to a later date

Attendees (please indicate if first time)**:**

-

Notes:

July 17th, 2024

Skipped

July 3rd, 2024

Skipped

June 19th, 2024

Agenda:

- Multi cloud, multi cluster orchestration: an update on Karmada

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Eduardo Arango [NVIDIA]
- Nate (SchedMD)
- Skyler Malinowski (SchedMD)
- Darren Boss (Digital Research Alliance of Canada)
- Hongcai Ren (Huawei)
- Jeffrey Sica (@jeefy, CNCF)
- Kevin Wang (Huawei)
- Zach Mayes (ORNL)

Notes:

- Kevin gave a nice overview of the Karmada project
    - Support for native k8s resources (no more CRDs as for kubefed v2), both push (execution controller) and pull (via karmada agent), support to schedule custom CRDs over remote clusters, and much more
    - End users rely on the project for disaster recovery, cloud bursting
- Q: Is Karmada planning to integrate with the cluster inventory API?
    - Started having a look, invested in contributing to it
- Q: Is there anything particular related to network connectivity for workloads across clusters?
    - Some users rely on projects like submariner for the connectivity layer. Higher performance connectivity usually means setting up some sort of VPN to get the different pods in the clusters talking to each other
- Q: Is there usage of the project for batch like workloads as well?
    - Some users like Tencent are integrating Karmada, Volcano and other projects for batch stacks. Ricardo suggests to have a follow up with an end user perspective of integrating these different projects in one stack

June 5th, 2024

Agenda:

- New group lead candidates and selection
- Kubernetes native batch system: update on Kueue and MultiKueue
    - Speaker: Aldo Culquicondor
    - Link to presentation: [PUBLIC] Kueue v0.7.0

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Fengping Hu(Uchicago, first time)
- Lucio Anderlini (INFN Firenze, first time)
- Saurabh Mookherjee (Thoughtworks, first time)
- Clemens Lange (PSI)
- Dennis Marttinen (Aalto University)

Notes:

- Aldo gave an overview of Kueue's job features and basic concepts
- Q: When handling pods directly, are the job steps in the diagram ignored
    - Kueue relies on SchedulingGate for support for Pods. This allows Kueue to handle the assignment, and once passed to the scheduler it behaves like a normal, including cluster autoscaling
- Q: Has there been a request for support of Dask?
    - Not yet, but looking at the description a similar integration as Metaflow should be possible. The easiest way is to pass Jobs or Jobsets eventually, which are directly supported in Kueue. More tight integrations are possible but more complex
- Aldo explained fair sharing in Kueue is based on DRF
    - Fair sharing in Kueue is only applied to the borrowed quota (above the nominal)
- Q: Are things like log retrieval supported by MultiKueue?
    - This can be tricky, but please open an issue with the environment this could be useful on and this can be followed
- Q: When do you expect a 1.0 release?
    - Likely in 3 releases, so 1 or 2 quarters

May 15th, 2024

Agenda:

- Kubernetes Batch with Volcano
    - Speaker: Kevin Wang
    - Link to slides: Cloud Native Batch Scheduling with Volcano-v1.0.pptx

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Timothy Middelkoop (Internet2)
- Piers Harding (SKAO)
- Dagvanoro Lkhagvajav (EVIDEN, UGA/ENSIMAG msc student)
- Zach Mayes (ORNL)
- Nate (SchedMD)
- Kevin Wang (Huawei)
- Angel Beltre
- Yu Bai (Bryan)
- Jonathan Skone (NERSC)
- William Wang (Huawei)

Notes:

- Kevin went through the multiple features supported by the project, including multiple advanced scheduling for co-location, fair share (DRF https://en.wikipedia.org/wiki/Dominant_resource_fairness), preemption, SLO aware scheduling and SLA policies, among many others
- Q&A
    - How is the i/o topology aware scheduling done?
    - Bryan: Can volcano take into account pricing in the scheduling? For example take spot instances if possible instead of on-demand resources?
        - This can be done today with the node type allocation priorities
    - Piers: Are the examples of people moving from SLURM to Volcano?
        - Migrating MPI jobs has been done by different people
        - SLURM migrations have some additional challenges and the project is looking for solutions
    - Piers: One of the plus sides of SLURM often mentioned is the support for singularity. Is this something supported?
        - One of the topics being discussed is two level scheduling. This could mean SLURM is deployed on top of Kubernetes and Volcano could handle some optimizations in this area. This integration is not as native as wanted for now
    - Dagvanorov: what's the deployment model? Volcano on top of k8s or HPC?
        - Right now it's on top of k8s

April 17th, 2024

Agenda:

- Kubernetes HPC + SLURM integration
    - Presentation on Interlink https://github.com/interTwin-eu/interLink
    - e

- Interlink documentation: https://intertwin-eu.github.io/interLink/
- Speaker: Diego Ciangottini

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Dagvanorov Lkhagvajav (EVIDEN & UGA/ENSIMAG msc student, first time)
- Nikos Tsipinakis (CERN)
- Piers Harding(SKAO)
- Dennis Marttinen (Aalto University)
- Nate Rini (SchedMD)
- Skyler Malinowski (SchedMD)
- Alan Mutschelknaus (SchedMD, first time)
- Jonathan Skone (NERSC)
- Diego Bacchin (Do It Systems srl, second time)
- Priyanka Sharma (CNCF)
- Viktoria Spisakova (Masaryk University)
- Camilla Harris (UCL, first time)
- Rohini Joshi (SKACH, first time)
- Eduardo Arango (NVIDIA)
- Kevin Hannon (Red Hat)
- Zachary Mayes (ORNL)
- Jose E Torres (HPCNow, 1st time)
- Pablo Llopis (SKACH/EPFL, first time)
- Timothy Middelkoop (Internet2, nth+1 time)
- Shane Canon (LBNL/NERSC, 1st time)
-
-
-

Notes

- Running plain Kubernetes pods on SLURM with interLink virtual nodes
  - https://docs.google.com/presentation/d/1GKyiH-v0TaTXpNJhXr6oO5nmJHCvu_7OuQFeRG2ebJ8/edit?usp=sharing
  - Diego Ciangottini
  - Interlink documentation: https://intertwin-eu.github.io/interLink/
- Q&A
  - What provider are you relying on from the virtual kubelet?
    - interLink is a specific virtualkubelet implementation, so it's own provider
  - How is data handled since PVs/PVCs are not supported?
    - Scratch/common area is still available for the singularity pods, this would be relying on the usual HPC mechanisms to mount scratch

        - Looking forward supporting PVs would likely have issues due to the limitations regarding intra-cluster networking
    - How is the network setup working in the demo case of a jupyter notebook?
        - There is a ssh tunnel between the hub and the pod, this is something that needs to be setup directly in the image being used in jupyter so not all images would allow this
            - Can be adapted with a simple python init script though
    - How is the trust between the hub and the notebook established?
        - There is a trust between the interlink and the hpc center
    - Is there support for deploying multiple workloads in a single pod, and how is memory and resource sharing done?
        - The current setup with the chosen slurm plugin starts each container as a separate apptainer/singularity container
    - Is there interest in adding this as a provider to the virtualkubelet website?
        - Have not reached the level to want to be published there, but should be done
    - Resource requirements on the interlink cluster? Just etcd?
        - There is a cache layer at the target to reduce the load on the control plane
    - Any specific requirements for network between the entrypoint k8s and the HPC farms?
        - The only requirement is a machine on the “edge” of the HPC with the capability of submitting SLURM jobs and exposing an HTTPs endpoint on a port (where the OAuth2 proxy will be listening for instructions from the VK core part)
    - Any way to set policies on co-location?
        - The annotation on the node is the only way right now. Things like supporting TF distributed has not been tried
    -

April 3rd, 2024

Agenda:

- Looking for a new co-lead to help organize topics and meetings
    - Reach out to Ricardo if you’re interested in helping
- Kubernetes HPC + SLURM integration
    - Presentation on possibilities of integration with KCP
    - Speaker: Carlos Eduardo Arango

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Eduardo Arango Gutierez DE
- Timothy Middelkoop (Internet2)

- Priyanka Sharma (CNCF)
- Dennis Marttinen (Aalto University, first time)
- Giuseppe Gianquitto (SURF, first time)
- Tim Wickberg (SchedMD)
- Skyler Malinowski (SchedMD)
- Bryan Bai (ASML)
- Shawfeng Dong (ASML)
- Piers Harding (SKAO, first time)
- Viktoria Spisakova (Masaryk University)
- Wietze Albers (SURF) First time
- Diego Ciangottini (INFN)
- Alex Scammon (G-Research)
- Nikos Tsipinakis (CERN) (Second time, after a presentation ~ 2 years ago :))
- Davide Pastorino (Do IT Systems s.r.l) First time, here for the talk :)
- Raj Ayyampalayam (Genentech) First time
- Diego Bacchin (Do IT System s.r.l.) First time
- Matteo Sponza (Do IT Systems s.r.l)
- Sanjay Chatterjee (NVIDIA, first time)
- Jonathan Skone (NERSC)
- Matt Anson (StackHPC)
- Diana Gaponcic (CERN)
- Angel Beltre (Sandia National Laboratories)
- Zach Mayes (Oak Ridge National Lab)(first time)
- Kees de Jong (SURF)
- Christopher Desiniotis (NVIDIA, first time)
- Ricardo Aravena (first time)
- Nate (SchedMD)
- Skyler (SchedMD)

Notes

- Eduardo gave an overview of options to do HPC on kubernetes and the relevant groups focusing on this type of workloads
    - Kubernetes
        - https://github.com/kubernetes/community/blob/master/wg-batch/README.md
        - (proposed) wg-serving https://groups.google.com/a/kubernetes.io/g/wg-batch/c/WY10kbuTWLU
        - (proposed) wg-accelerators https://groups.google.com/a/kubernetes.io/g/wg-batch/c/v91ssgsCX-M
        - Kueue: https://kueue.sigs.k8s.io/
    - CNCF
        - Batch System Initiative: https://tag-runtime.cncf.io/wgs/bsi/charter/
- Then followed describing the features of SLURM and a possible integration via KCP

- Goal is to open source the project in the next couple months
    -
- Q&A
    - Integration with Job and JobSet?
    - Giuseppe (SURF): question on how to fragment their supercomputer farms and how Kubernetes can help? What problem is this proposal trying to solve?
        - Eduardo introduced a use case with 100k of nodes where researchers want to use kubernetes, but the cluster is backed by SLURM. The project bridges between the two, from the principle that these SLURM clusters are not going away anytime soon
    - Giuseppe (SURF): for kubernetes people there is always pushback from the HPC/SLURM community, how do people handle this?
        - Eduardo stressed integrating SLURM is the best way forward, making SLURM part of the ecosystem
        - Priyanka followed up that showing good examples of cloud native usage for this type of workloads is often a good way to move forward
        - Dennis gave the example of LUMI, where researchers would like better ways to manage their workloads but the system admins do not want to completely move to k8s. A goal is to bridge the two worlds instead
    - Bryan: Would this tool also help with cluster autoscaling? The use case is for spawning and destroying quickly clusters of 100s of nodes
        - Eduardo replied there might be better ways to spawn this type of setup on public clouds, possibly with direct integrations with the cloud providers
    - Sanjay: How is preemption and reallocation of resources handled? And how are reservations handled (similar to fair share in SLURM)?
        - Eduardo mentioned a few tools are available to provisioning / reprovisioning nodes between the

November 2nd, 2022

Agenda:

- Discuss KubeCon + batch day
- Refresh agenda

Attendees (please indicate if first time)**:**

- Jamie Poole (G-Research)
- Nate Rini (SchedMD)
- Timothy Middelkoop (Internet2)
- Stefan Lasiewski (NERSC/LBNL)
- Christian Kniep (QNIB Solutions)
- Kevin Fox (PNNL)

Notes: Agenda ideas:

- As above, plus:
- Get some actual HPC users to talk to us!
- Storage for HPC, including from K8s/Slurm
-

October 5th, 2022

Agenda:

- Topic: OCI Containers with scrun, Nate Rini, SchedMD

Attendees (please indicate if first time)**:**

- Jamie Poole (G-Research)
- Timothy Middelkoop (Internet2)
- Ricardo Rocha (CERN)
- Clemens Lange (PSI)
- Jonathan Skone (LBNL)
- Stefan Lasiewski (LBNL/NERSC)
- Kevin Hannon (G-Research Open Source)
- Musa Yusuf (Currently unaffiliated)
- Nate R (SchedMD - Presenter)

September 21st, 2022

Agenda:

- Topic: Cilium and eBPF

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Raphaël Pinson (Isovalent)
- Jamie Poole (G-Research)
- Timothy Middelkoop (Internet2)
- Kevin Hannon
- Linda Zhou
- Gregor Krmelj
- Jonathan Skone (LBNL NERSC)

July 20th, 2022

Agenda:

- Open discussion

Attendees:

- Jamie Poole (G-Research)
- Timothy Middelkoop (Internet2)
- Jeffrey Miller (ORNL)
- Nate Rini (SchedMD)
- Dave Gantenbein (G-Research)
- Alex Scammon (G-Research)

July 6th, 2022

Agenda:

- Topic: Cilium and eBPF (POSTPONED, apologies)

Attendees (please indicate if first time)**:**

-

June 15th, 2022

Agenda:

- Topic: Platform Aware Scheduling
    - Alexander: slides that I've used from Batch+HPC day: https://sched.co/10F0n

Attendees (please indicate if first time)**:**

- Timothy Middelkoop (Internet2)
- Jamie Poole (G-Research)
- Jeffrey Miller (ORNL)
- Jonathan Skone (LBNL NERSC)
- Alexander Kanevskiy (Intel)
- Marlow Weston (Intel)
- Ricardo Rocha (CERN)
- Adrian Hoban (Intel)

June 1st, 2022

Agenda:

- Topic: Kubecon roundup

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Jamie Poole (G-Research)
- Luigi Zhou (G-Research)
- Kamil Wilas (G-Research)
- Michel Sumbul (G-Research)
- Florian Goth (University of Würzburg)
- Jonathan Skone (LBNL NERSC)
- Nate Rini (SchedMD)
- Alex Scammon (G-Research)

May 18th, 2022

- Skipped due to Kubecon

May 4th, 2022

Agenda:

- Topic: containerssh.io
    - Janos described the tool as a frontend to containerized work environments accessible via ssh
    - Nikos describing the lxplus service at CERN
    - CERN contributed features for AD/KRB5 integration in containerssh.io
    - Motivations included improved security, improved resource sharing
    - Metadata for user environments includes UID/GIDs, preferred shell, profiles, etc

    - Questions
        - Credential delegation?
        - UID/GID of users in the containers?
        - Did you consider donating the project to the CNCF?

Attendees (please indicate if first time)**:**

- Nikos Tsipinakis (CERN, first time)
- Janos Bonic (Red Hat, first time)
- Jeffrey Miller (ORNL)
- Clemens Lange (PSI)
- Ricardo Rocha (CERN)
- Rémi Voirin (CERN, first time)
- Timothy Middelkoop (Internet2)
- Alex Scammon (G-Research)
- Spyros Trigazis (CERN)
- Doina Cristina Duma (INFN)
- Jonathan Skone (LBNL)

April 20th, 2022

Agenda:

- Topic: Baremetal Deployments
    - Speakers: Jamie Poole, Scott Solkhon from G-Research

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Jamie Poole (G-Research)
- Scott Solkhon (G-Research)
- Stefan Lasiewski (LBNL, NERSC)
- Timothy Middelkoop (Internet2)
- Clemens Lange (PSI)
- Jonathan Skone (LBNL)
- Alex Scammon (G-Research)
- Taylor Dolezal (CNCF)
-

April 6th, 2022

Agenda:

- Topic: Crossplane

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Clemens Lange (PSI)
- Nic Cope (Crossplane)
- Jonathan Skone (LBL)
- Alex Scammon (G-Research)

- Brian Major (NRC / CADC)
- Taylor Dolezal (CNCF)

March 16th, 2022

Agenda:

- Topic: Discussion

March 2nd, 2022

Agenda:

- Topic: Air gapped solutions (TAG App Delivery)
    - https://kccncna2021.sched.com/event/lV9e
    - Q: Is the infrastructure itself air-gapped or also the dev environments?
    - Going all the way left gave the best results. It gave developers confidence that when deploying all the checks had been done with image checks
    - Q: What are the checks being run?
    - Image and code scanning tools (prisma cloud, …) plus functional tests, covering both code and deployment / infrastructure. Also doing full DR tests, including backing etcd and nuking / restoring clusters
    - A lot of gotchas came from the deployments of kubernetes (tanzu, openshift, …) who on day 2 would expect being able to phone home for different purposes
    - Q:

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Alaa Youssef (IBM Research)
- Ian Smith (PNNL)
- Jonathan Skone (LBNL)
- Alex Scammon (G-Research)
- Jordi Mon
- Jeff Saelens
- Nathan Rini (ScheMD)
- Jeffrey Miller (ORNL)

February 16th, 2022

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Eduardo Arango (Red Hat)
- Jeffrey Miller (ORNL)
- Alaa Youssef (IBM Research)
- Aldo Culquicondor (Google)

- Brian Major (NRC/Canadian Astronomy Data Centre)
- Jonathan Skone (LBNL)
- Hanyu Ding (China Mobile) 1st time
- Marco Rorro (EGI Foundation) 1st time
- Clemens Lange (CERN)
- Abdullah Gharaibeh (Google)
- Alex Scammon (G-Research)
- Antoni Bertel
- Gaurav Singh(RedHat)
- Maciek Rozacki (Google)
- Nathan Rini (ScheMD)
- Timothy Middelkoop (Internet2)
- Kevin Fox (PNNL)

Agenda:

- Topic: Kueue Controller and the Kubernetes Batch WG - Abdullah Gharaibeh, Google
    - Kueue
        - https://github.com/kubernetes-sigs/kueue
        - https://docs.google.com/document/d/1jFdQPlGnvjCSOrtAFxzGxEMi9z-OS0VVD1uTfSGHXts/edit?resourcekey=0-BgDvCZcpwFVaCEZj2tlfyw#heading=h.xgjl2srtytjt
    - vs Volcano
        - Attempting to build in the core scheduler
        - Avoid reinventing what's already there regarding resource management, …
    - vs Batch on GKE
        - Similar issues to Volcano
    -
- Kubernetes Batch WG
    - https://github.com/kubernetes/community/tree/master/wg-batch
    - Charter: https://github.com/kubernetes/community/blob/master/wg-batch/charter.md
    - Mailing List: https://groups.google.com/a/kubernetes.io/g/wg-batch
- CNCF Batch System Initiative WG (TAG-Runtime)
    - PR: https://github.com/cncf/tag-runtime/issues/38
    - Charter: https://github.com/cncf/tag-runtime/blob/329a2a29d682b2e084d267d7359f8aa7cebd3904/wg/bsi.md
    - Slack: https://cloud-native.slack.com/archives/C02Q5DFF3MM

February 2nd, 2022

Attendees:

- Jamie Poole (G-Research)
- Stefan Lasiewski (NERSC/Berkeley Lab)
- Alex Scammon (G-Research)
- Bob Killen (Google)
- Timothy Middelkoop (Internet2)
- Sara Jeanes (Internet2)
- Nate Rini (SchedMD)

Agenda:

- Topic: Dealing with Deprecations (SIG Architecture, App Delivery?)
- Kubernetes Deprecation Policy
    - GA API versions may be marked as deprecated, but must not be removed within a major version of Kubernetes
    - **Beta API versions must be supported for 9 months or 3 releases (whichever is longer) after deprecation**
    - Alpha API versions may be removed in any release without prior deprecation notice
- Kubernetes is getting more aggressive with deprecating beta APIs for two primary reasons:
    - Beta APIs became 'perma ga', contributors would work to get them to beta and turned on by default, but not follow up with driving them to GA
    - Users would depend on them as if they were GA and this has led to more disruptions for users when they are removed or incremented.
- Some improvements have been made to better signal deprecations
    - Warning messages in kubectl + event logs: https://kubernetes.io/blog/2020/09/03/warnings/
    - Dedicated blog posts regarding deprecations with each release
    - Working with CNCF and other avenues to communicate deprecations better
- https://relnotes.k8s.io/
    - Can filter by release, area (apiserver vs admission-control), kind (deprecations, bug, feature)
- Now that Helm Charts are decentralized and uncurated, how do they stay on top of deprecations?

January 19th, 2022

Agenda:

- Topic: Cartografos WG

Notes from chat:

From Danielle Cook:

- Prologue: github.com/cncf/cartografos/blob/main/reference/prologue.md

- Supporting spreadsheet docs.google.com/spreadsheets/d/1wn3qxV11wAoOMBFzigeCHEHrZqWHA8wG05zRWhW7RXQ/edit#g...
- Slides from today with links: docs.google.com/presentation/d/1_1tZaT5oRRmU5CGx75r0OeYSw9x2CYoy37aYRMVsNiM/edit#s...
- People: github.com/cncf/cartografos/blob/main/reference/people.md
- Policy: github.com/cncf/cartografos/blob/main/reference/policy.md

Attendees (please indicate if first time)**:**

- Jamie Poole (G-Research)
- Matthew West (Exeter)
- Danielle Cook, Cartografos WG
- Simon Forster, Cartografos WG
- Alex Scammon (G-Research)
- Albin Severinson (G-Research)
- Caterina Rindi (G-Research)
- Timothy Middelkoop (Internet2)
- Michael Crusoe
- CaptainSifff@gmx.de

January 5th, 2022

Agenda:

- Regroup and plan 2022 agenda, #ug-research chat topics

Attendees (please indicate if first time)**:**

- Jamie Poole (G-Research)
- Jonathan Skone (LBNL)
- Timothy Middelkoop (Internet2)
- Ricardo Rocha (CERN)
- Caterina Rindi (G-Research)
- Jeffrey Miller (ORNL)
- Abubakar Siddiq Ango (CNCF Ambassador)
- Michel Sumbul

December 15th, 2021

Agenda:

- Topic: Christmas's Special - Jamie's Jam Session
    - Song?
    - Cartografos WG
- **Attendees** (please indicate if first time)**:**

December 1st, 2021 (Recording)

Agenda:

- Topic: HPC/HTC End User Landscape
    - Why are sites choosing what they're using? Kubernetes, Armada, Slurm, …?
    - How do you build your container images? How do you distribute them?
    - How do you handle data in/out? What filesystems are being used (s3, shared fs, …)?
    - SKA, G-Research, ORNL, CERN, …?
- Notes
    - ACTION: contact / engage with the devs and communities build the tools we talked about (slurm, volcano, htcondor, armada, etc) to make sure they know about our group
    - ACTION: check further on integration of FPGAs with kubernetes
    - Clear motivation on ease of use for devs (try it out locally, scale to hpc)
        - ACTION: ask around how people are using containers in HPC environments, limitations they might have (performance, networking, …) and how their image build workflow looks like (CI/CD, multiple archs, etc)

Attendees (please indicate if first time)**:**

1. Bob Killen (Google)
2. Ricardo Rocha (CERN)
3. Jamie Poole (G-Research)
4. Caterina Rindi (G-Research)

5. Jeffrey Miller (ORNL)
6. Jason Kincl (ORNL)
7. Shri Javadekar (Self)
8. Nate Rini (SchedMD)
9. Alex Scammon (G-Research)
10. Abubakar Siddiq Ango (GitLab, CNCF Ambassador)
11. Michael Crusoe (CWL, ELIXIR-NL)
12. Sara Jeanes

November 17th, 2021(Recording)

Agenda:

- Topic: Armada Update (G-Research)
- Switch to the CNCF Bevy platform
- [Katie Gamanji] - Announcement - I am leaving CNCF, with my last day being 25th of November. Abubakar Ango and Bill Mulligan will cover the End User Community for now. Looking forward to seeing all of you in-person during future KubeCon's. Meanwhile, you can find me on LinkedIn and Twitter :)

Attendees (please indicate if first time)**:**

1. Jeffrey Miller (ORNL)
2. Ricardo Rocha (CERN)
3. Altaz Valani (Security Compass)
4. Michel Sumbul (G-Research)
5. Jamie Poole (G-Research)
6. Caterina Rindi (G-Research)
7. Carlo Camurri (G-Research)
8. Daniel Rastelli (G-Research)
9. Alex Scammon (G-Research)
10. Nate (SchedMD)
11. Abubakar Siddiq Ango (CNCF)
12. Qi Chu (Bloomberg)
13. Timothy Middelkoop (Internet2)
14. Christian Kniep (AWS)
15. Katie Gamanji (CNCF)
16. Shri Javadekar (Self)
17. Florian Goth (University of Würzburg)

November 3rd, 2021 (recording)

Agenda:

- Topic: sssd (ORNL)

Attendees (please indicate if first time)**:**

1. Jeffrey Miller (ORNL)
2. Jamie Poole (G-Research)
3. Ricardo Rocha (CERN)
4. Mauro Sbarigia (G-Research)
5. Rob Smith (G-Research)
6. Caterina Rindi (G-Research)
7. Nate (SchedMD)
8. Timothy Middelkoop (Internet2)
9. Shri Javadekar (Self)
10. Katie Gamanji (CNCF)
11. Christian Kniep (AWS) (first time)

Notes:

- SSSD slide: https://sssd.io/docs/architecture.html

October 20th, 2021 (Recording)

Agenda:

- Qi: Updating time from next month after daylight saving changes
    - Ricardo: I will check and confirm the timing before next meeting.
- Topic: Kubecon Recap
    - Ricardo: Ricardo's Notes from KubeCon
    - Qi: Interesting presentation from Redhat on Kubernetes and Checkpoint restore
    - Jeffrey: Interesting talks:
        - Argo and Tekton: Pushing the boundaries of the possible on Kubernetes
        - The Argo Ecosystem: Tailoring your installation through community add-ons
        - A Multi-Cluster, Multi-Cloud Infrastructure with GitOps at CERN (GitOpsCon)
        - Manage More Clusters with Less Hassle, with ArgoCD Application Sets
    - Stefan: Mostly security talks and a talk on how to explain Kubernetes to others using diagrams for both new and experienced people.
        - Beyond Block Diagrams: Different Ways of Understanding K8s Architecture - Kim Schlesinger, DigitalOcean
            - Summary: Alternatives to regular block-diagrams for explaining Kubernetes workflows.
            - Examples: time-sequence (swimlane) diagrams
            - Slides (Page 13 for a very helpful time-sequence diagram)
        - eBPF Day - Lots of Kernel Internals

- Cilium looks like a very interesting CNI with great observability potential using Hubble
- eBPF is kernel-native and much more performant than IPtables-based CNIs

Attendees (please indicate if first time)**:**

1. Jeffrey Miller (ORNL)
2. Abubakar Siddiq Ango (CNCF)
3. Qi Chu (Bloomberg, first time)
4. Nate (SchedMD)
5. Timothy Middelkoop (Internet2)
6. Michael R. Crusoe (CWL, ELIXIR-NL)
7. Stefan Lasiewski (NERSC, LBNL)
8. Caterina Rindi (G-Research)
9. Ricardo Rocha (CERN)

October 6th, 2021 (recording)

Agenda:

- Topic: Volcano
- Slides

Attendees (please indicate if first time)**:**

1. William-wang (Huawei)
2. Klaus Ma (Huawei)
3. Jeffrey Miller (ORNL)
4. Florian Goth (Uni Würzburg, second time)
5. Michael R. Crusoe (ELIXIR-NL, CWL)
6. Clemens Lange (CERN)
7. Ricardo Rocha (CERN)
8. Nate Rini (SchedMD)
9. Timothy Middelkoop (Internet2)
10. Jamie Poole (G-Research)
11. Gary Conway (G-Research)
12. Katie Gamanji (CNCF)
13. Robert Smith (G-Research)
14. Caterina Rindi (G-Research)
15. Isham Araia (G-Research)
16. James Murkin (G-Research)
17. Carlo Camurri (G-Research)
18. David Yuan (EBI)

19. Alex Scammon (G-Research)
20. Ketan?
21. Mada?

September 15th, 2021

Agenda:

- Topic: Open Discussion
- Upcoming containers WG survey on container runtimes
    - We should try to advertise this when ready, suggestions as well
- Check with Katie for ideas on how to promote the group further

Attendees (please indicate if first time)**:**

22. Ricardo Rocha (CERN)
23. Jeffrey Miller (ORNL)
24. Nate (SchedMD)
25. Jamie Poole (G-Research)
26. Caterina Rindi (G-Research)

September 1st, 2021

Agenda:

- Katie: We want to hear from you! At CNCF, providing a stellar experience for End Users is our top priority. We would like to learn what we have done right and what we can improve on.
    - Share your feedback now 2021 CNCF End User Survey - **deadline 13th of September**
-

August 4th, 2021

Meeting canceled

July 21st (recording)

Attendees (please indicate if first time)**:**

1. Jeffrey Miller (ORNL)
2. Marcos Hermida Mera (ECMWF) (first time)
3. Jamie Poole (G-Research)

4. Oksana Shadura (UNL)
5. Caterina Rindi (G-Research)
6. Cheryl Hung (CNCF)
7. Lukas Heinrich (CERN)
8. Carlo Camurri (G-Research)
9. Ricardo Rocha (CERN)
10. Nate Rini

Agenda

- Topic: Kubernetes API vs using Dask/runners

July 7th (recording)

Attendees (please indicate if first time)**:**

1. Ricardo Rocha (CERN)
2. Cheryl Hung (CNCF)
3. Jamie Poole (G-Research)
4. Rob Smith (G-Research)
5. Sam Clark (G-Research)
6. Caterina Rindi (G-Research)
7. Maxime Mouchet (Sorbonne Université/LIP6, first time)
8. Berat Senel (Sorbonne Université/LIP6)
9. Florian Goth (University of Würzburg)(First Time)
10. Sara Jeanes (Internet2)
11. Manuel Martins (ECMWF) (first time)
12. Doina Cristina Duma (INFN-CNAF)
13. Clemens Lange (CERN)
14. Alessandra Forti (University of Manchester)
15. Doug Benjamin (BNL)
16. Alex Scammon (G-Research)

Agenda

- Kubecon NA Session?
- Topic: EdgeNet - a multi-tenant and multi-provider edge cloud (https://dioptra.io/), Berat Can Senel
    - Paper on the topic presented: https://dl.acm.org/doi/pdf/10.1145/3434770.3459737
    - Do you experience issues with latency between kubernetes components since you deploy all around the world?
        - Not yet, but cluster is rather small

        - etcd can be an issue and replicas should be living in the same region
    - What are the main use cases for this infra?
        - CDN
    - EdgeNet source code
        - https://github.com/EdgeNet-project/edgenet/tree/release-1.0
        - https://github.com/EdgeNet-project/node

June 2nd - Cancelled

Attendees (please indicate if first time)**:**

17. Jeffrey Miller (ORNL)
18. Cheryl Hung (CNCF)
19. Manuel Martins (ECMWF)
20. Caterina Rindi

Agenda

Topic: Securing workloads for researchers

May 19th (recording)

Attendees (please indicate if first time)**:**

1. Jeffrey Miller (ORNL) - first time
2. Moritz Kröger (LLT RWTH Aachen) - first time
3. Ricardo Rocha (CERN)
4. Jamie Poole (G-Research)
5. Gary Conway (G-Research)
6. Cheryl Hung (CNCF)
7. Viktoria Spisakova (ICS MUNI) - first time
8. Carlo Camurri (G-Research)
9. Steve Jahl (Broad Institute) - first time
10. Caterina Rindi (G-Research)
11. Alex Scammon (G-Research)

Agenda

- Topic: Kubecon Recap (Interesting Sessions, Open Discussion)
    - https://docs.google.com/presentation/d/1jPG3LKugYuNiuVZ-2cC_RsaI9gWmSe03NSS2rsMcWhY/edit?usp=sharing
- Idea:

    - Start writing down use cases - start with multi-cluster

May 5th - cancelled due to KubeCon

April 21st (recording)

Attendees (please indicate if first time)**:**

1. Caterina Rindi (G-Research)
2. Ricardo Rocha (CERN)
3. Cheryl Hung (CNCF)
4. Kamil Wilas (G-Research) - first time
5. Jamie Poole (G-Research)
6. Clemens Lange (CERN)
7. Brian Major (NRC/CADC)
8. Owen Thomas (G-Research) - first time
9. Burt Holzman (Fermilab) - first time
10. Nate Rini (SchedMD)

Agenda

- Topic: Multi-Tenancy

April 7th (recording)

Attendees (please indicate if first time)**:**

- Ricardo Rocha (CERN)
- Jamie Poole (G-Research)
- Rob Gardner (UChicago)
- Cheryl Hung (CNCF)
- Lincoln Bryant (UChicago)
- Jeremy Rogers (ORNL)
- Nate Rini (SchedMD)
- Carlo Camurri (G-Research)
- Brian Major (CADC/NRC)
- Diego Davila (UCSD) - First time
- Judith Stephen (UChicago) - First time
- Geert Pingen (TNO)
- Clemens Lange (CERN)
- Adrien Trouillaud (Admiralty)
- John Graham (UCSD)
- Dima Mishin (UCSD) - First time

- David Jordan (Uchicago) - First time
- Cristina Duma (INFN/IT)
- James Murkin (G-Research)

Agenda

- Announcement (Cheryl) - this meeting will be livestreamed to Youtube (TBC)
    - This will be recorded - will figure out Youtube for next time
- Topic: Multi-cluster with Admiralty
    - Adrien Trouillaud (Admiralty)
    - Matthew Madany (UCSD)
- Cheryl - Next meeting will be May 5 which is in the middle of KubeCon - do we want to keep, move or skip the meeting?
    - Skip: +1

March 17th (recording)

Attendees (please indicate if first time)**:**

1. John Long (ORNL/OLCF) - First time
2. Jeremy Rogers (ORNL/OLCF)
3. Enol Fernández (EGI Foundation)
4. Kevin Fox (PNNL)
5. Matthew West (Uni. of Exeter)
6. Jamie Poole (G-Research)
7. Ricardo Rocha (CERN)
8. Brian Major (CADC/NRC)
9. Cheryl Hung (CNCF)
10. Brian Lin (OSG/UW Madison) - First time
11. Chris Hollowell (BNL) - First time
12. Lincoln Bryant (UChicago)
13. Jan Kaspar (G-Research)
14. Carlo Camurri (G-Research)
15. Alexander Götz (LRZ)
16. Rob Gardner (UChicago)
17. Ian Smith (PNNL)
18. Doug Benjamin (BNL)

Agenda

- Topic: POSIX Integrations (how to access hpc storage, slurm jobs etc)
    - Jeremy Rogers ORNL/OLCF
    - Notes (feel free to add notes here)

    - Problem: UID based access to shared filesystems, no root containers allowed
    - Solution: kube-namespace-manager
        - Namespaces keep additional metadata about users (uid, etc)
    - OPA used to mutate workloads
        - Ex: If fs access required, pods get an appropriate securityContext
        - Fs access is expressed as pod annotations
        - Relying on Styra
    - sshexec-operator manages ssh keys to allow cross-submission to other resources
    - Ongoing work to also get sssd and id/login mapping
    - Image csi driver: https://github.com/kubernetes-csi/csi-driver-image-populator
    - Next steps include open sourcing these tools

March 3rd (recording)

Attendees (please indicate if first time)**:**

1. Ricardo Rocha (CERN)
2. Jamie Poole (G-Research)
3. Mauro Sbarigia (G-Research) - first time
4. Krunal Chaudhari (iauro) - first time
5. Cheryl Hung (CNCF)
6. Clemens Lange (CERN)
7. Jeremy Rogers (ORNL/OLCF)
8. Brian Major (CADC)
9. Timothy Middelkoop / Internet2
10. Alexander Götz (LRZ)
11. Sara Jeanes (Internet2)
12. Caterina Rindi (G-Research)
13. Patrick Ross (anynines)
14. Jason Kincl (ORNL)
15. Carlo Camurri (G-Research)
16. Kenny Chen (Brobridge) - first time
17. Matthew West (Uni. of Exeter)
18. Kevin Fox (PNNL)

Agenda

- Topic: OIDC (Lab / Institute SSO) integration with Kubernetes
    - (oob) https://www.keycloak.org/docs/latest/securing_apps/#redirect-uris
- How to follow up on the different activities

    - Triggered by the notebook discussion: UIDs, Storage Access, Dependencies, ...
    - Issues on our github?
        - Yes
- Not always trivial to find speakers on time… should we keep topics for 2 slots in advance?
- Kubecon EU Session preparations

February 17th (recording)

Attendees (please indicate if first time)**:**

1. Jamie Poole (G-Research)
2. Ricardo Rocha (CERN)
3. Alexander Götz (LRZ, first time)
4. Doug Benjamin (BNL, first time)
5. Enol Fernández (EGI Foundation)
6. Matthew West (Uni. of Exeter)
7. Nathan Rini (SchedMD)
8. Rob Gardner (UChicago)
9. Brian Bockelman (Morgridge)
10. Brian Major (Canadian Astronomy Data Centre, first time)
11. Eduardo Arango (Red hat)
12. Sara Jeanes (Internet2)
13. Jan Kaspar (G-Research)
14. Cheryl Hung (CNCF)
15. Jeremy Rogers (ORNL)
16. Matthew Potter (G-Research)
17. Bob Healey (Cuemby)
18. Kevin Fox (PNNL)
19. Andrei Brabete (G-Research, first time)
20. Carlo Camurri (G-Research)
21. Stephan Hachinger (LRZ, first time)
22. Oksana Shadura (UNL, first time)
23. Caterina Rindi (G-Research)
24. Clemens Lange (CERN)
25. Ilija Vukotic (UChicago, first time)
26. Nihir Patel (G-Research, first time)
27. Alessandra Forti (University of Manchester, first time)
28. Lincoln Bryant (UChicago, first time)
29. Alex Scammon (G-Research)
30. Ian Smith (PNNL)

Agenda

- Topic: General Discussion (JupyterHub, Dask, GPU Integration, ...)
    - Enol Fernández: JupyterHub on EGI
    - Brian Bockelman: CMS Analysis Facility (UNL)
    - Ricardo Rocha: Binder, Dask, JupyterHub, Kubeflow on Kubernetes
    - ARDC EcoCommons: https://ardc.edu.au/project/ecocommons-australia/
- Notes
    - EGI OAuthenticator, refresh of credentials: https://github.com/EGI-Foundation/egi-notebooks-hub/blob/master/egi_notebooks_hub/egiauthenticator.py#L129
    - Data Federation powered by onedata: https://onedata.org/ (some code to making this work available at https://github.com/EGI-Foundation/egi-notebooks-hub)
    - Request for post helm customization - follow up in SIG?

February 3rd (recording)

Attendees (please indicate if first time)**:**

- Matthew West (Uni. of Exeter) - first time
- Ricardo Rocha (CERN)
- Jeremy Rogers (ORNL)
- Jason Kincl (ORNL)
- Caterina Rindi (G-Research) - first time
- Cheryl Hung (CNCF)
- Jamie Poole (G-Research)
- Jan Kaspar (G-Research)
- Carlo Camurri (G-Research)
- Nate Rini (SchedMD)
- Alex Scammon (G-Research)

FYI

Research User Group maintainer session at KubeCon EU - need to register this week

1. Cheryl suggestion - present quick history of the group, how Kubernetes and other cloud native computing technologies helps in research, 5-10 unique challenges you face, and what people should expect if they join the group. Goal is to make it seem interesting for people to participate.
2. Deadline to record around mid-April
3. Run survey to gauge people's interests

Topic: General Discussion

- Highlights (2 or 3 key features) that motivate usage of cloud native in research

- Features wanted for the next steps in deployments

Review the topic backlog, assign names to each existing one, add new ones?

Notes
- Suggestion to have a topic on how people are deploying their resources and training their users internally
- Topics of interest: Workflows

(January 20th - CANCELLED)

Attendees (please indicate if first time)**:**
1. Matthew West (UoExeter, first time)
2. Caterina Rindi (G-Research, first time)
3. Alex Scammon (G-Research, first time)

Topic: Event / CI/CD workflows for research users

(January 6th - skipped for holidays)

December 16th (recording)

Attendees (please indicate if first time)**:**
4. Ricardo Rocha (CERN)
5. Jamie Poole (G-Research)
6. Eduardo Arango (Red Hat)
7. Jeremy Rogers (ORNL/OLCF)
8. Clemens Lange (CERN)
9. Nathan Rini (SchedMD)
10. Carlo Camurri (G-Research)
11. Rob Smith (G-Research)
12. Cheryl Hung (CNCF)

Topic: Image distribution / replicated registries / Pull-Through Caching
- Alex Xu: Harbor Open Source Registry
- Upcoming v2.2 features
    - Monitoring metrics with Prometheus
    - Option to install Harbor using an operator (focus on day2 ops)
    - System level robot accounts

- ? To use a proxy cache we need a prefix right? As in registry.mydomain/docker.io/library/ubuntu
    - A mutating webhook might help (github issue focusing on this: https://github.com/goharbor/harbor/issues/11867)
- ? Best way to integrate with gitlab-ci image builds
    - System level robot accounts should help (v2.2)
- ? Possibility to use different storage backends for different projects?
    - This should be possible, no one requested it yet
    - Backend is docker distribution. Project lost a bit of momentum but talk about donating to CNCF and Harbor would be interesting in contributing to the process
    - Opened: https://github.com/goharbor/harbor/issues/13783

December 2nd (recording)

Lead: Ricardo Rocha (CERN)

Attendees (please indicate if first time)**:**

1. Clemens Lange (CERN)
2. Cheryl Hung (CNCF)
3. Ricardo Rocha (CERN)
4. Jeremy Rogers (ORNL)
5. Jan Kaspar (G-Research)
6. Carlo Camurri (G-Research)
7. Sara Jeanes (Internet2)

- Topic: Batch on Kubernetes (Armada)
- Postponed: Image distribution / replicated registries / etc - Harbor? Michael Michael (M2)?

November 18th

- No call - KubeCon

November 4th

- Next call on the 18th is during KubeCon - **Skipping**

Lead: Jamie Poole (G-Research)

Attendees (please indicate if first time)**:**

1. Spyros Trigazis (CERN) - strigazi
2. Eduardo Arango - Red Hat
3. Ricardo Rocha (CERN)

4. Brian Bockelman (Morgridge Institute for Research, first time)
5. Jeremy Rogers (ORNL)
6. Nate Rini (first time - SchedMD)
7. Stefan Lasiewski (LBNL/NERSC)
8. Timothy Middelkoop (Internet2)
9. Carlo Camurri (G-Research)
10. Giuseppe Scrivano (Red Hat, first time)
11. Mat Selmeci (UW-Madison)
12. Artur Jonkisz (G-Research)
13. Jan Kaspar (G-Research)
14. Clemens Lange (CERN)
15. Jamie Poole (G-Research)

Topic: Rootless Kubernetes

- Akihiro presentation Kubecon about rootless containerd: https://kccncna20.sched.com/event/fGWc/containerd-rootless-containers-2020-akihiro-suda-ntt

Topic: Rootless Podman

- Presentation by Giuseppe Scrivano
- Rootless requires user namespaces
- Issue with both mount and user namespaces being cleaned up - podman keeps a pause process to pin them and prevent cleanup
- On the container storage side, it is not yet possible to manage overlayfs rootless (ongoing work for that), a VFS backend would work but no CoW. For CoW fuse-overlayfs was developed and works
- For networking, slipr4netns is used to allow external connectivity
- cgroups v2 is required, both runc and crun have support for it now
- Q: Performance degradation with fuse-overlayfs? Any high iops should use a volume, the degradation is higher the native overlay (~20%). The improvement here will come from rootless overlay in the kernel, fuse-overlayfs should be replaced by kernel overlay - patches already going in
- Q: Performance degradation for networking (slirp4netns)? Much worse than storage, ongoing work to improve this and get close to native performance (requires seccomp notifications). No timeline, Akihiro's work here: https://github.com/rootless-containers/bypass4netns
- Q: Feature parity with docker (for podman)? Docker swarm is one, everything else should be there. Podman is not to be used with kubernetes, cri-o does that
- Q: containerd support? Networking is done in a similar way, overlayfs relies on native overlay support in recent ubuntu kernels, and fallback to VFS if not
- Q: What's the main use case for rootless podman? Immutable images where it's hard to customize / install packages and allows multiple users to have different package versions

- Q: Support for podman in ubuntu / suse? Main difference with ubuntu today is that podman won't make use of native rootless overlay (waiting for the patches to go in the vanilla kernel). Other than that it should work fine

October 21st (recording)

Lead: Ricardo Rocha (CERN)

Attendees (please indicate if first time)**:**

1. Eduardo Arango
2. Cheryl Hung, CNCF
3. Jamie Poole (G-Research)
4. Timothy Middelkoop (Internet2)
5. Ricardo Rocha (CERN)
6. Thomas Hartland (CERN)
7. Jeremy Rogers (ORNL, first time in a while)
8. Stefan Lasiewski (LBNL (NERSC), first time in a while also)
9. Carlo Camurri (G-Research)
10. Spyros Trigazis (CERN) - strigazi
11. Gary Conway (G-Research)
12. Eric Wagner (OSU)
13. James Murkin (G-Research)
14. Jan Kaspar (G-Research)

Topic: Cloud Bursting (Virtual Kubelet and Tensile Kube)

- Presentation by Thomas Hartland (slides)
- Q: Do you have any control on the type of workload that ends up in the virtual kubelet nodes? A: There's a taint that needs to be tolerated and other labels can be used to match workloads
- Q: Which calico version was it? Rather old Calico 3.3, plan to test with something recent. 3.13 now
- Q: Any experience with tackling bursting batch like workloads to external resources? A: There's a talk from the last kubecon about setting up mini VPNs using wireguard to burst to AWS (https://github.com/openshift-psap/wireguard-worker). Every node joins the same VPN that is setup in the master cluster

October 7th (Recording)

Lead: Jamie Poole (G-Research)

Attendees (please indicate if first time)**:**

1. Jamie Poole (G-Research)
2. Ricardo Rocha (CERN)
3. Cheryl Hung (CNCF)
4. Eduardo Arango (Red Hat)
5. Gary Conway (G-Research)
6. Artur Jonkisz (G-Research)
7. James Murkin (G-Research)
8. Carlo Camurri (G-Research)
9. Alessandra Forti (University of Manchester)
10. Michael R. Crusoe (VU Amsterdam; CWL project; ELIXIR-NL)
11. Jan Kaspar(G-Research)
12. Matt Potter (G-Research)
13. Sam Clark (G-Research)
14. Cristina Duma (INFN, first time)
15. Robert Smith (G-Research)
16. Ian Smith (PNNL)
17. Matyas Selmeci (OSG / University of Wisconsin, first time)
18. Geert Pingen (TNO)
19. Clemens Lange (CERN)
20. B

Topic: GPU and GPU sharing +10

- 20min presentation from Eduardo Arango
- Relevant links from the presentation
- Slides link?
- Q&A
    - https://github.com/kubernetes-sigs/node-feature-discovery
    - https://github.com/NVIDIA/gpu-operator
    - https://github.com/NVIDIA/k8s-device-plugin
    - https://github.com/NVIDIA/gpu-feature-discovery
- Q: GPU sharing possible? A: Nothing officially supported yet, work ongoing to use Nvidia MIG in the A100s
- Q: Any recommendations on creating container images in a future proof way? Example would be portable, reproducible images for scientific workloads? A (eduardo): That is hard, we've seen issues with not being able to run older workloads in newer Nvidia architectures
- Q: Experience with other vendors (other than Nvidia)? A: There's a way to deploy AMD GPUs in Kubernetes, but not automated the same way as with an operator
    - https://github.com/RadeonOpenCompute/k8s-device-plugin ?
- Q: Kernels support for this? A: 4.x and 5.x mentioned

- Q: Is there a way to make it use a prebuilt kernel module instead of building on the fly? A (eduardo): Not quite, you can inject a different entrypoint using a configmap

September 16th (recording)

Lead: Ricardo Rocha (CERN)

Attendees:

1. Ricardo Rocha (CERN)
2. Jamie Poole (G-Research)
3. Kohei Tokunaga (NTT)
4. Akihiro Suda (NTT)
5. Clemens Lange (CERN)
6. Jakob Blomer (CERN)
7. Timothy Middelkoop (Internet2)
8. Rob Gardner (UChicago)
9. Cheryl Hung (CNCF)
10. Jan Kaspar(G-Research)
11. Enol Fernández (EGI Foundation)
12. Sara Jeanes (Internet2)
13. Andrea Valenzuela (CERN)
14. Geert Pingen (TNO)
15. Thomas Hartland (CERN)
16. Ryan Bezdicek (Cray/HPE)
17. Harry Staub (SWITCH)
18. Eduardo Arango (Red Hat)
19. Artur Jonkisz (G-Research)
20. Simone Mosciatti (CERN)
21. Matt Potter (G-Research)
22. Murugappan Chetty (Optum)
23. Eric Wagner (Ohio State)
24. Ian Smith (PNNL)
25. Gary Conway(GR)
26. Spyros Trigazis (CERN)

Main Topic: Lazy Loading of Container Images - the containerd remote snapshotter

- Slide: https://www.slideshare.net/KoheiTokunaga/the-overview-of-lazypull-with-containerd-remote-snapshotter-stargz-snapshotter

- Blog Post: https://medium.com/nttlabs/startup-containers-in-lightning-speed-with-lazy-image-distribution-on-containerd-243d94522361
- Kubecon Presentation: https://kccnceu20.sched.com/event/ZepQ/startup-containers-in-lightning-speed-with-lazy-image-distribution-kohei-tokunaga-ntt
- Q: What about security/integrity concerns? A (ktok): The stargz remote snapshotter verifies the digest of every chunk it downloads, for auth/authz are done independently by the stargz snapshotter. It supports ~/.docker/config.json based one and Kubernetes's secret based one.
- Q: How do you populate the registries with these images, any special registry needed? A (ktok): Stargz snapshotter supports standard registies. Other implementations like the cvmfs remote snapshotter can fetch the content from CVMFS repository. You can also implement your remote snapshotter that supports the remote store you want to use.
- Q: Can you achieve similar effects by using small images? A (ktok): Ideally yes, but in some cases it is not possible to achieve that and this covers that use case
    - Even in smaller images this could still help
- Q: Is there a kind of workload where you wouldn't recommend lazy loading? Maybe a long running workload? A (ktok): If a lot of data is needed on startup there is a large prefetch at container startup there will be no improvement (if you use eStargz). But worst case is same as legacy, just with very slightly bigger layers
- Q: How do you come up with the prioritized files? A (ktok): Current strategy is to run a container once and profile the file access during the container startup
- Q: What is the client support to generate stargz/estargz images? A (ktok): Converter available in the stargz remote snapshotter repo, but plans for docker client plugin and buildkit(buildx) - welcoming contributions

Note (Cheryl):
- We will start recording these calls

August 5th (Rotation 1)

Attendees:
- Cheryl Hung, CNCF
- Gary Conway, Jamie Poole (G-Research)
- Bob Killen (Google)

Agenda:
- Cheryl to wind down calls but keep Slack, mailing list and Github repo
- Bob to add notice to Github repo and send notification to Slack and mailing list
- Keep KubeCon slot in case people want to find the Research User Group

July 15th (Rotation 2)

Attendees:

- Cheryl Hung, CNCF
- Jamie Poole (G-Research)
- Bob Killen (University of Michigan)
- Ryan Bezdicek (HPE - Cray)
- Eduardo Arango (Red Hat)

Agenda:

- KubeCon EU presentation (https://sched.co/ZeuD)
    - Jamie: Lots of work happening behind closed doors. G-Research is trying to be more open and share more of what they are doing with K8s. Would like to know what others are doing.
    - Bob: people don't know if they are doing the right things.
    - Ryan: Same in other groups, people not disclosing what they do.
    - Jamie: Can take 5-10 minutes in session to talk about Armada
    - Bob: Discussion via chat with speakers. Can get 5 minutes presentations with others - CERN, Cray
    - Video recording deadline - 31 July
- Book
    - Bob: Ricardo is planning to merge some work in

July 1st (Rotation 1)

Attendees:

- Michael R. Crusoe (ELIXIR-NL, CWL)
- Bob Killen (University of Michigan)
- Clemens Lange (CERN)
- Cheryl Hung (CNCF)

Agenda:

- Progress on modeling CUDA requirements in CWL: https://github.com/common-workflow-language/common-workflow-language/issues/587#issuecomment-648238107
    - Discomfort around the standard practice of injecting the CUDA driver into the containers to match the kernel outside the container

    - Some users currently make containers with the CUDA driver built in; is this widespread or just a moment in time practice? Bob has seen this a little bit as well, but he didn't think it was widespread
    -
- https://github.com/giannisdoukas/ipython2cwl
    - alpha/mvp at the moment
    - Would like to converge on standard tags, feedback via issues on that repo would be appreciated
- AI: Get followup from HPCW

June 17th (Rotation 2)

Attendees:

- Bob Killen
- Ricardo Rocha (CERN)
- Jamie Poole (G-Research)
- Ryan Bezdicek
- Michael R. Crusoe (ELIXIR-NL, CWL)

Agenda:

- Open discussion

June 3rd (Rotation 1)

Attendees:

- Bob Killen
- Jamie Poole (G-Research)
- James Murkin (G-Research)
- Gary Conway (G-Research)
- Ricardo Rocha (CERN)
- Clemens Lange (CERN)
- Cheryl Hung (CNCF)
-

Agenda:

- Kubeflow multi tenancy
    - Notebooks go into different namespaces, but are pipelines in a single namespaces
    - [CERN] Separating out jupyterhub from kubeflow to break them up as you can't integrate another jupyter service with kubeflow
- GPU Usage

        - GPUs cannot be shared within Kubernetes
        - Looking to add “virtual” gpus to pass up the chain
        - GPUs don’t follow a consistent naming/label schema which makes workloads not portable. Is there a standard set of labels out there that are used?
            - Does node-feature-discovery / operator cover this use-case?
    - VM overhead for workloads
        - 3% performance hit for running in VM if you pin-cpu and optimize for virtual workloads
    - For the book, get a DOI
    - ~~AI: Bob Add folders for break out of topics into the repo~~

Next call: May 20th

cancelled

May 6th

Attendees:

- Bob Killen (University of Michigan)
- Michael R. Crusoe (ELIXIR-NL, CWL)
- Jamie Poole (G-Research)
- Jan Kaspar (G-Research)
- James Murkin (G-Research)
- Alex Lovell-Troy (HPE/Cray)
- Clemens Lange (CERN)
- Tadas Bareikis (Vilnius University)
- Cheryl Hung (CNCF)
- Steve Quenette (Monash University)

Agenda:

- Presentation from Vilnius University/CERN on their HTCondor Operator - HTCJob CRD
    - Slides: https://docs.google.com/presentation/d/1v9SbgENUzzws5jNWUMlibeHc39b33eykqo2WI1DNhI8/edit?usp=sharing
    - Repository mirror (CI/CD is running internally at CERN): https://github.com/clelange/htc-operator
    - Why containers at CERN?
        - Ensures code is under version control
        - Data is versioned along with it
    - CERN uses OpenStack to manage their environment
    - Use argo - a kubernetes native workflow engine

    - Kubernetes native
    - Job knowledge between steps
- Problem: We only have 48 cores for our project test cluster at CERN
    - A single analysis might requires hundreds of cores for multiple hours
- Coupled kubernetes with HTCondor,
    - A scheduler that can span multiple clusters and in use in many other places
- HtCondor container jobs
    - Using docker is very limited, they cannot pass arguments to it well
    - They opted to use singularity for easier mounting of filesystems (AFS, CVMFS etc)
- Used Operator SDK from RedHat to write an HTCondor operator
- Converted job submission requirements to HTCJob CRD
    - Added status field with job data such as jobID
    - Spec has a few fields:
        - Image - which container image to execute the job in
        - Command - command to run within the container
        - Script - Large text field that accept script as a part of the HTCJob spec
- Operator has a small sqllite db behind it to manage job status
- Job results are sent with a cloudevents sender
- The goal is to bubble up job status data into Kubernetes so that it can be integrated with an Argo workflow for further analysis
- At CERN storage is tied to the user, they have to kinit within the container as a part of the script.
    - Password for user is stored a secret (could potentially use keytab)
- The Argo workflow monitors the HTCJob status and references that as the success condition and failed condition
- Question: How does Argo+HTCJob CRD compare to REANA? http://www.reana.io/ (also a CERN project?)
    - Reana served as an inspiration for this project.
    - Reana uses the htcondor python api for monitoring job status
    - Operator method is a bit more “kubernetes” / “cloud” native
- Question: When jobs are submitted to condor, are they containerized as well?
    - Yes, the jobs being run in condor are singularity container based jobs.
- Question: Is CERN planning to move more resources to Kubernetes
    - In principle yes, but the driving force is user or researcher based
    - The goal was to build / offer a hybrid solution that integrates with more legacy cluster options.
- Question: Can you say more about the challenges in passing arguments to docker vs singularity? I (M. Crusoe) haven’t had this experience.
    - In CERN HTCondor cluster environment they cannot use docker in a straightforward way

        - Docker runs as a service where they must run it as a process spun up as a specific user
    - Question to the audience: Is anyone using PodMan?
        - [Alex Joseph Lovell-Troy] - yes, they're using it but not for hpc style jobs at this point in time.
    - Question: What is the general usage of Argo at CERN?
        - Right now it's still early / experimental but more groups are interested in using it.
        - Most groups interested in reproducible analysis are using REANA (mainly CWL or yadage)
    - Question: What are your thoughts on the integration of Argo/Flux?
        - Their focus hasn't been on the GitOps style integrations that Argo/Flux provide. They're more interested in the job/workflow side.
    - Question: Have you done anything with Argo / Volcano?
        - Not yet, focus has still been on REANA
- Other business:
    - The CWL+MPI proposal is maturing. Feedback, especially from a MPI in the cloud perspective is very welcome! https://github.com/common-workflow-language/cwltool/pull/1276

April 1st

Attendees:

- Bob Killen, University of Michigan
- Cheryl Hung, CNCF
- Michael R. Crusoe (CWL, ELIXIR-NL)
- Steve Quenette, Monash University
- Susheel Varma, HDR UK

Agenda:

Reverting to 2-rotation meeting times to encourage participation.

- Cheryl: Done
- Rotation 1: 5AM ET / 9AM UTC
- Rotation 2: 11AM ET / 3PM UTC

Book chapters are being managed at https://github.com/cncf/research-user-group/issues especially https://github.com/cncf/research-user-group/issues/16

- Inspired by "The crossroads of Cloud and HPC: OpenStack for Scientific Research"
- Chapter idea: Go into supply-chain vetting for research workloads

- Chapter idea: managing secure workloads, could be part of supply chain or supply chain could be a part of security itself

The Common Workflow Language is an open standard to describe batch data analysis workflows. For a future version of the standard, the CWL community is looking into adding optional support for the Message Passing Interface (MPI) standard. This is important is because there are some scientific applications that require MPI; some can deal with high latency MPI via TCP, others require high performance MPI implementations (very fast interconnect, perhaps infiniband)

- Please contact us via the linked document or PR (or me directly: mrc@commonwl.org) with any insights about combining MPI with traditional dataflow/workflow systems, advice for modeling the configuration with respect to the scientific application, and/or questions

AI: Put together take aways to send both to the ML
AI: Assemble summary take aways for broader community
AI: Send notice to list about switching back to 2 rotation schedule

March 4th - Rotation 2

Time: 1PM ET / 6PM UTC
Zoom: https://zoom.us/my/cncfenduser
Attendees:

- Eduardo Arango, Ryan Bezdicek,

Agenda:

- CFP (ISC/vhpc at http://vhpc.org): there's a session about kubernetes and k8s batch, and CFP is open now :)
- Fyi: (The Australian) National Kubernetes Core Services Workshop [SQ proxied to Bob?]
- Fyi: New CNCF Working Group - Air Gap Installs
    - Meeting time poll: https://doodle.com/poll/vdy3abt5rv5uxedq
    - Meeting cadence poll: https://doodle.com/poll/225ifq7fk5q5qbmr

No meeting - No show

February 19th - Rotation 1

Time: 5AM ET / 10AM UTC
Zoom: https://zoom.us/my/cncfenduser
Attendees:

- Bob Killen
- Eduardo Arango
- Cheryl Hung
- Steve Quenette

Agenda:

- Open discussion around the user group meetings rotation and how to enhance user engagement

February 5th - Rotation 3

Time: 9PM ET / 2AM UTC
Zoom: https://zoom.us/my/cncfenduser
Attendees:

- Noriaki Fukuyasu
- Steve Quenette
- Bob Killen
- Klaus Ma (@k82cn)

Agenda:

- Proposed revised time: 7PM ET / 4PM WC / 11AM ADST / 9AM JPN / 8AM CNA
    - SQ's suggestion: prioritise for pacific rim (i.e both China and US West Coast)
    - Community champions ...
        - US rotation : Bob
        - Pacific Rim rotation : SQ + NF
        - European rotation: ??
- Follow-up from OpenAI: Currently too busy right now, but wishes to be more involved in the future.
- Discuss KubeCon presentation plans
    - Our community falls part of the "Maintainer Track" and a "CNCF SIG" within KubeCon / CloudNativeCon
    - Known community events:
        - KubeCon US (Nov 17+, Boston),
        - KubeCon China (July, Shanghai),
        - Kubernetes Forum Japan (Sept, Tokyo), and
        - KubeCon EU (March 30+)
    - KubeCons : we get 1 presentation session, max 2x speakers
    - Bob: Idea: 1 lead plus a local active member

    - This KubeCon EU - update and progress from book chapters (Eduardo lead). Bob, Klaus [2s sessions: Volcano + Machine Learning working group] and Eduardo will be there.
    - Next KubeCon US : Proposed theme / leads : Bob? + Soheila [data operations chapter of the book]
    - KubeCon China : sessions due February 21st - Klaus to think about - proposed focus: “Introduction to CNCF RUG”
- [Open discussion]: Scheduling Workloads (Eduardo Arango)[optional]
- [Open discussion]: benchmark metrics of batch system (Klaus Ma)[optional]
    - xref:https://docs.google.com/document/d/1fRjeeUUAyuTvk4yA7JPVYJYD5G6aMk4Py-QkeWgvAZA/edit#
    - **ACTION**: Bob to organise USA crowd to provide feedback to Klaus
    - **ACTION**: SQ to organise Aust crowd to provide feedback to Klaus

January 22nd - Rotation 2

Time: 1PM ET / 6PM UTC
Zoom: https://zoom.us/my/cncfenduser
Attendees:

- Bob Killen

Agenda:

- Cancelled (no show)

January 8th - Rotation 1

Time: 5AM ET / 10AM UTC

Zoom: https://zoom.us/my/cncfenduser
Attendees:

- Cheryl Hung, CNCF
- Bob Killen, University of Michigan
- Soheila (DENSO)
- Jamie Poole, James Murkin (G-Research)
- Eduardo Arango (Red Hat / OpenShift)
- Michael R. Crusoe (CWL, ELIXIR-NL)
- Klaus Ma (@k82cn)

Agenda:

- Cloud-Native research book (Carlos)
    - Set action items

        - Best Practices - should include a section on MPI & OCI
        - For simple & scalable jupyter
        - KubeFlow
        - General Workflows (GitOps, workflow engines etc)
            - Research workloads
            - Making Git the source of truth
            - DataOps (Data Lifecycle Management)
- G-Research looking for more feedback on armada:
    - https://github.com/G-Research/armada
- AI (Cheryl): Get this meeting recorded and shared publicly
- Look at / possibly reach out to the folks here:
    - https://github.com/nuest/ten-simple-rules-dockerfiles
    - https://f1000research.com/articles/7-742/v2

December 26th

Cancelled due to holidays.

December 11th

Time: 5AM ET / 9AM UTC
Zoom: https://zoom.us/my/cncfenduser
Attendees:

- Michael R. Crusoe, CWL, ELIXIR-NL; @biocrusoe / @mr-c
- Cheryl Hung, CNCF
- Klaus Ma (@k82cn)
- Bob Killen
- Hippie Hacker (hh@ii.coop)
- Soheila (DENSO)
- James Murkin
- Steve Quenette
- Ricardo Rocha (CERN)

Agenda:

- [Bob Killen] Post KubeCon followup and action items
    - RUG Survey Results
- Proposal / call to action for book chapters
    - OpenStack SIG example: The Crossroads of Cloud and HPC: OpenStack for Scientific Research
        - Best practice recipes / book
        - Across all the components
            - How do I (GPUs etc)
        - A lot of the community is doing things that are similar

        - It's a way for us to have a concrete that writing something to a web page
        - From our convo in San Diego, there were five groups interested in contributed
            - Ricardo had some input
            - Worldwide LHC Computing Grid - k8s topic day, how do I share best practices
                - Chaired by Ian Collier / (<<<who is this group?)
            - Meeting is today
            - Shared experiences
            - To get more folks involved
            - People need recipes
        - Labs
            - Oakridge
            - Pacific Northridge
            - (XX another iuni?)
        - Personas:
            - Running the operations / configuration
            -
        -
- Making operations easier ...
    - Experiences with configuration and life-cycle of clusters
    - Experiences with clusters on top of OpenStack
        - **AI:** @Ricardor Cern contributions
        - **AI:** @bobytables Reaching out to large scale / 2500-5000 nodes in size
    - Experiences with (parallel / HPC) file system access
- Making research cloud native ...
    - For simple & scalable jupyter
        - X Contributors of Zero to JupyterHub
    - For KubeFlow
        - Handles all the lifecycle tasks of ML pipelines
    - General Workflows (GitOps, workflow engines etc)
        - [Topic was CI/CD of Research]
        - Research workloads
        - Making Git the source of truth
        - DataOps (Data Lifecycle Management)
            - **AI:** @Sohelia is interested in helping with
            - @bobytables (west coast US is notable missing, **AI:** email to list forthcoming)
            - Checkpointing
                - Being able to go back to that state in time
            - Very much in demand
            - Needs to integrate current observation / in realtime

            - For ML models, in line with other
            - Snapshots of object stores are harder for those that don't use Ceph etc
            - @Ricardo: we are missing some of these features
    - How do we proceed?
        - Page in github / or freeflow in Google Docs links for there
            - https://github.com/cncf/research-user-group/blob/master/RESOURCES.md
        - **AI**: @steve will reach out to colleges, will reach out to staff who were involved with the Openstack book
            - (**AI:** LINK TO BOOK?)
        - Select some references
    - @Sohelia - https://www.delphix.com/
- NZ - @hh
    - Centralizing the Community College infrastructure over the next five years
    - https://multicore.world
    - Catalystcloud.nz (OpenStack)
    - Any CNCF Uni members, ^ that think this is a good idea ^^
    - **AI:** Start Google doc / PRs against the Repo
    - AI: Connect with Jonas Roland
        - & Josh Berkos / George Casto
- Proposal to move to rotating 3x timezone meeting times
    - Current time works well for US and Western Europe
    - Rotation suggestion:
    - No expectations that everyone makes every meeting
    - Cheryl - Staying every 2 weeks on Wednesday?
        - Yes

	US (EST)	APAC (AEDT)	EU (CET)
Rotation 1 (now)	5 am	9 pm	11 am
Rotation 2	1 pm	5 am, next day	7 pm
Rotation 3	9 pm	1 pm, next day	3 am, next day

-

Notes:

Lack of best practices and documentation
Alot of people with k8s in production

November 27th - **Cancelled**

Time: 11AM ET / 15:00 UTC
Zoom: https://zoom.us/my/cncfenduser - note change of Zoom link
Attendees:

Agenda:

- [Bob Killen] Post KubeCon followup and action items

November 13th

Time: 5AM ET / 9AM UTC
Zoom: https://zoom.us/my/cncfenduser - note change of Zoom link
Attendees:

- Klaus Ma (@k82cn)
- Bob Killen (@mrbobbytables)
- Cheryl Hung (@oicheryl)
- Michael R. Crusoe (@mr-c @biocrusoe)
- Jamie Poole
- Soheila (DENSO)

Agenda:

- [Bob Killen] Pre KubeCon discussion
  - RUG session: Wednesday, November 20 • 11:50am - 12:25pm
    - Presentation link: https://docs.google.com/presentation/d/1ABl3ztaKkxM3m5eZ28zLQoBI8unVypEcFU8ebEOpQcg/edit?usp=sharing
  - You're welcome to join the End User Happy hour, Wednesday 8-10pm
- Volcano being donated to CNCF - https://github.com/cncf/toc/pull/318
- Volcano Example: https://github.com/volcano-sh/volcano/blob/master/example/integrations/mpi/mpi-example.yaml (https://github.com/volcano-sh/volcano/blob/master/example)
- M. Crusoe asks: Any guidance for batch operations when one isn't also the k8s admin?
  - https://github.com/volcano-sh/volcano
  -
- KubeCon Sessions:
  - Multi-tenancy in Kubernetes: Let's Talk - Tasha Drew
  - Present and Future of Hardware Topology Awareness in Kubelet - Connor Doyle
  - Enabling Kubeflow with Enterprise-Grade Auth for On-Prem Deployments - Yannis Zarkadas, Arrikto & Krishna Durai, Cisco
  - Managing Helm Deployments with Gitops at CERN - Ricardo Rocha, CERN

- Introducing KFServing: Serverless Model Serving on Kubernetes - Ellis Bigelow, Google & Dan Sun, Bloomberg
- Managing Apache Flink on Kubernetes - FlinkK8sOperator - Anand Swaminathan, Lyft
- Towards Continuous Computer Vision Model Improvement with Kubeflow - Derek Hao Hu & Yanjia Li, Snap Inc.
- Measuring and Optimizing Kubeflow Clusters at Lyft - Konstantin Gizdarski, Lyft & Richard Liu, Google
- Scaling Kubernetes to Thousands of Nodes Across Multiple Clusters, Calmly - Ben Hughes, Airbnb
- KubeFlow's Serverless Component: 10x Faster, a 1/10 of the Effort - Orit Nissan-Messing, Iguazio
- Advanced Model Inferencing Leveraging KNative, Istio and Kubeflow Serving - Animesh Singh, IBM & Clive Cox, Seldon
- Building and Managing a Centralized Kubeflow Platform at Spotify - Keshi Dai & Ryan Clough, Spotify
- Intro: Scheduling SIG - Wei Huang, IBM & RaviSantosh Gudimetla, Red Hat
- Kubeflow: Multi-Tenant, Self-Serve, Accelerated Platform for Practitioners - Kam Kasravi, Intel & Kunming Qu, Google
- To Infinite Scale and Beyond: Operating Kubernetes Past the Steady State - Austin Lamon, Spotify & Jago Macleod, Google
- Mitigating Noisy Neighbors: Advanced Container Resource Management - Alexander Kanevskiy, Intel
- Batch Capability of Kubernetes Intro - Klaus Ma, Huawei
- Deep Dive: Kubernetes Working Group for Multi-tenancy - Sanjeev Rampal, Cisco
- Improving Performance of Deep Learning Workloads With Volcano - Ti Zhou, Baidu Inc & Da Ma, Huawei
- Networking Optimizations for Multi-Node Deep Learning on Kubernetes - Rajat Chopra, NVIDIA & Erez Cohen, Mellanox
- Tutorial: From Notebook to Kubeflow Pipelines: An End-to-End Data Science Workflow - Michelle Casbon, Google, Stefano Fioravanzo, Fondazione Bruno Kessler, & Ilias Katsakioris, Arrikto
- Building a Medical AI with Kubernetes and Kubeflow - Jeremie Vallee, Babylon Health
- GPU as a Service Over K8s: Drive Productivity and Increase Utilization - Yaron Haviv, Iguazio
- RDMA Enabled Kubernetes for High Performance Computing - Jacob Anders, CSIRO & Feng Pan, Red Hat
- Supercharge Kubeflow Performance on GPU Clusters - Meenakshi Kaushik & Neelima Mukiri, Cisco

October 23rd

Time: 11AM ET / 15:00 UTC

Zoom: https://zoom.us/my/cncfenduser - note change of Zoom link

Attendees:

- Eduardo Arango
- Jamie Poole, James Murkin, Jan Kaspar, Gary Conway (G-Research)
- Cheryl Hung, CNCF
- Ryan Bezdicek, Cray
- Klaus Ma
- Bob Killen

Agenda:

- [Bob Killen] Overview of wg-multitenacy efforts from Ryan Bezdicek
    - Who is wg-multitenancy
        - People who do not have the capability to provision multiple clusters quickly.
        - Generally resource constrained (bare metal clusters)
        - Large amount of end-users
    - How they define a tenant
        - A subset of resources (compute, network, storage, etc.) within a single Kubernetes cluster.
        - “Soft multi tenancy” as a first goal.
            - Different teams within organizations.
            - Reasonable assumption there are no bad actors.
            - Problem: current isolation mechanism is limited to namespace.
        - A tenant is a subset of resources, NOT users.
        - Define standards and conformance tests to validate clusters level of multi-tenancy
    - Personas in a cluster
        - Cluster Admin
            - ALL permissions for the cluster
            - Can create cluster wide resources (bindings, CRDs etc)
            - Create root resources that propagate to all tenants
        - Tenant Admin
            - All permissions within a tenant
            - Create tenant namespaces through the tenant controller
            - Create parent resources to propagate to tenant namespaces
        - Tenant User
            - Can consume resources within a tenant namespace
            - Can deploy resources to a tennt namespace
    - Incubator Projects
        - https://github.com/kubernetes-sigs/multi-tenancy
            - Hierarchical Namespace Controller
            - Virtual Cluster
            - Tenant Controller
    - Demo: HNC

    - Deck: https://docs.google.com/presentation/d/12_FSVs96OQbobg7_BiICtXJIkYUKnseaWhtZlV82Gqg/edit?usp=sharing
    - Demo Script: https://docs.google.com/document/d/1tKQgtMSf0wfT3NOGQx9ExUQ-B8UkkdVZB6m4o3Zqn64/edit#heading=h.rl9sfhfuxyzn

- Guardian project proposal from wg-policy
    - https://github.com/rficcaglia/guardian
    - Mentioned on TOC update
- https://github.com/G-Research/armada (Kubernetes multi-cluster batch scheduler) made public
- CWL on K8s presentation from the recent Common Workflow Language mini-conference https://github.com/Duke-GCB/calrissian

October 9th

Time: 5AM ET / 9AM UTC
Zoom: https://zoom.us/my/cncfenduser
Attendees:

- Bob Killen, University of Michigan
- Ricardo Rocha, CERN
- Jamie Poole, James Murkin, Jan Kaspar (G-Research)
- Cheryl Hung, CNCF
- Soheila Dehghanzadeh (DENSO)
- Michael R. Crusoe (CWL)
- Eduardo Arango
- Brett Milford
- Klaus Ma

Agenda:

- G-Research video demo of Armada
- What is Armada
    - HTC scheduler for multiple kubernetes clusters
    - Fair Share over time
    - Multi-cluster failure aware
    - Designed to be a queueing system for multiple systems, designed similarly to ht-condor
- Demo:
    - 5 clusters, 99 nodes
    - queued 22,000 jobs
        - spread evenly across the clusters
    - lower the priority, higher chance of being scheduled on the cluster

        - submits -just enough- pods to fill a cluster
- Pull based model
- jobs are submitted to armada server and stored in DB
- clusters can be removed without disruption
- jobs are leased by executor
- Why batch on Kubernetes?
    - standard platform
    - large ecosystem with secret management etc
    - better lifecycle management
    - modern data science tooling
- Road Map
    - Finish core functionality (job priority etc)
    - Better AuthN/AuthZ
    - Improved Scheduling
    - Improved user tooling
    - integration with other tooling (kubeflow)
    - Support for heterogeneous hardware
- Questions?
    - Could it support what researchers are used to (condor sub file etc)
    - Will pods be pending in each cluster when they're divided up?
        - Pods are not 'popped' off the queue until the cluster is ready
    - Does not have support for MPI - Armada job would have to be translated to in-cluster MPI job with a scheduler that supports it.
    - Data locality would have be spec'ed out in the job
- Challenges
    - Issues with standard labels
        - slate.ci
        - Node Feature Discovery

Research User Group at KubeCon San Diego

September 25th

Time: 11AM ET / 15:00 UTC
Zoom: https://zoom.us/my/cncfenduser
Attendees:

- Bob Killen
- Jamie Poole, James Murkin, Jan Kaspar (G-Research)
- Klaus Ma
- Eduardo Arango
- Thomas Hartland

Agenda:

- Overview of Volcano (Klaus Ma)
    - Slides
    - http://github.com/volcano-sh/volcano
    - Gaps identified during Kube-batch upstream development
        - Job/Queue Management
            - Queue status/configuration
            - Hierarchical queue
            - Job with multiple pod template
            - Lifecycle management of Job, e.g. restart, suspend/resume
            - Error Handling, e.g. restart job if pod failed (MPI, TFJob)
                - Restart entire pod group if single pod fails
            - Indexed Job
            - Task dependency, e.g. Spark (executor/driver)
            - Delay Pod Creation
        - Scheduler
            - Coscheduling
            - Faire-share
            - Queue
            - Preemption/Reclaim
            - Reserve/Backfill
            - Topology (network, accelerator)
        - Other
            - Throughput
            - Round-trip
            - Data locality (Data Aware Scheduling)
        - Goals of Volcano
            - Provide common services for high performance workloads on top of kubernetes
                - Batch scheduling, fair-share, gang-scheduling
                - Better job management
                - Better accelerator support
                - Kubectl plugins for job/queues
        - Volcano Architecture
        - Adds CRDs to support coscheduled (batch or mpi jobs)
            - Supports ssh and kubectl exec
            - Workers are deleted by job controller, but the pod running mpiexec/mpirun will not.
            - Defines multi-job/template jobs
        - Supports sharing resources by weight
        - Adds cross-namespace queue support based on priority
        - Task based topology + binpack
            - Improved scheduling performance

- Seeing use by greater community
    - Frameworks
        - MPI
        - Horovod
        - Kubeflow/tf-operator
        - Kubeflow/arena
        - Spark-Operator
        - Spark on Kubernetes
        - Cromwell
        - PaddlePaddle
    - Adopters
        - Unisound
        - BIBDR
        - Caicloud
        - Baidu
        - Hauwei
- Upcoming work
    - GPU Share/Topology - high priority
    - Job Management
    - Queue Management
    - Hierarchical Queue
    - Preemption/Reclaim
- Fairshare is based off DRF
- Questions
    - CERN: They have millions of jobs -- how big can a queue be created? How fast can they create jobs? (high level of churn)

September 11th

Time: 5AM ET / 9AM UTC
Zoom: https://zoom.us/j/6600498482
Attendees:

- Bob Killen
- Michael R. Crusoe (CWL)
- Jamie Poole, James Murkin, Jan Kaspar, Alex Scammon, Tim Boudreau, Caterina Rindi (G-Research)
- Brett Milford (Monash eResearch Centre)
- Cheryl Hung, CNCF

Agenda:

- Research User Group meeting at KubeCon San Diego (Cheryl) - https://sched.co/Uaja
- Overview of G-Research's Armada project
    - Looking to move away from Condor
    - Want to achieve the same throughput of Condor on Kubernetes
    - Interested in Federation to support multiple clusters.
    - Looking to maintain queues of jobs (but outside of Kubernetes)
    - Pull based model
    - Agents pop jobs off a work queue
        - If job does not complete within a certain period of time, it goes back into the queue
    - Just support jobs at this time

- Add question about workflow technology/standards usages?
    - Questions have been added
- Federation:
    - Multi-cluster/ Federation Tools / Suggestions
        - Bob: Prometheus for monitoring coupled with Thanos for multicluster aggregation
        - Bob: jwt-to-rbac - possible tool to simplify multi-cluster authn/authz.
    - Other research federation efforts:
        - SLATECI -Nationally distributed multi-cluster effort. Offers a catalog of applications for deployment in joined clusters.
        - Nautilus - Extension of the Pacific Research Platform (PRP) out of the University of San Diego. Globally distributed stretched Kubernetes cluster with rook/ceph for storage.

August 28th

Time: 15:00 UTC
Zoom: https://zoom.us/j/6600498482
Attendees:

- Cheryl Hung, CNCF
- Bob Killen, University of Michigan
- Bill Sparks, Cray
- Alex Lovell-Troy, Cray
- Cody Hammock, TACC
- Jamie Poole, Jan Kaspar, James Murkin, G-Research
- Mike Packard, TACC
- Joe Stubbs, TACC
- Ricardo Rocha, Thomas Hartland, CERN

Agenda:

- Go over survey questions

Survey Questions:

- What is your organization? (name, description, mission)
- How many researchers do you have? or how many are you supporting?
- What areas of research are you primarily supporting using Kubernetes? (e.g. bioinformatics, physics, generic research platform)
- Are your researchers consuming Kubernetes directly (deploy / use kubectl etc)? or do they consume services on top of it?
    - Are your users provisioning their own clusters for use?
    - Being provided a namespace or given a cluster by your support group?
- How many Kubernetes clusters do you run? What size? (node/total core count/total RAM count)
- Do you run other workload managers? (SLURM, HTCondor etc)
    - What size cluster(s) do you have supporting those workloads?
    - Are you considering running them on top of Kubernetes? or pursuing running them natively on Kubernetes?
- Are you using devices (e.g. GPUs) in any of your clusters?
- Where are you running your clusters? On Prem? Cloud Provider? Bare metal? Virtual (if so, what virtualization provider)
- How are you provisioning your cluster(s)?
- What version(s) of Kubernetes are you running?
- What distribution of Kubernetes are you running?
- What other services are you deploying on top of Kubernetes? Argo, Jenkins, JupyterHub? etc
- How are you handling authentication and user management?
- How are you handling access control? RBAC? ABAC?
- Is multi-tenancy important to you? Are you currently supporting a multi-tenant environment?
- Have you explored multi-cluster or federated deployments? If you are actively using it, what route did you go with which tools?
- Have you integrated with more traditional environments? e.g. posix/parallel shared file systems etc?
- What are you doing to provide persistent storage?
- How are you exposing cluster services externally? (service type LoadBalancer etc)
- What types of workloads are you running on Kubernetes? HPC, HTC, service oriented etc?
- What container runtime(s) are you using?
- Do you allow users to deploy their own applications from the internet or other sources?
- Are you using Kubernetes in production? or is it in proof of concept phase?
    - looking into / dev/test / prod
- Are you using Kubernetes to support “new” applications or workloads? or are you migrating traditional workloads.

- What have been the benefits of switching to Kubernetes or what is it enabling you to do?
- What sort of pain points have you experienced with this migration?
- How are you handling training your users to work with Kubernetes?
- What motivated you to explore Kubernetes?
- Were you using containers previously before looking at Kubernetes?
- Do you consider your group or organization a “traditional” HPC site?
- How important is workload portability?
- If you are migrating to Kubernetes, how are you planning to migrate your traditional workloads?

AI: Create doc in root of research user group repo for useful links etc

August 14th

Time: 5AM ET / 9AM UTC
Zoom: https://zoom.us/j/6600498482
Attendees:

- Bob Killen, University of Michigan
- Cheryl Hung, CNCF
- Michael R. Crusoe, mrc@commonwl.org
- Lukas Heinrich, Ricardo Rocha, Tibor Simko, Diego Rodriguez, Rokas Mačiulaitis, Clemens Lange, Thomas Hartland (CERN)
- Brett Milford
- Soheila Dehghanzadeh, DENSO

Agenda:

- Discuss mission statement / charter and decide next steps.

Lead - Bob Killen
Tech Lead - Klaus Ma

-

Mission statement

The CNCF Research User Group’s purpose is to function as a focal point for the discussion and advancement of Research Computing using *“Cloud Native*” technologies. This includes enumerating current practices, identifying gaps, and directing effort to improve the Research Cloud Computing ecosystem.

Potential work items

- Identify common research workflows and tasks that the current-state of kubernetes does not address.
    - MPI workloads
    - Resource matchmaking
    - Data location aware scheduling / Federation / Multi-cluster
    - Lots and lots of jobs
- State of the Art - on application integration platforms and highlight distinction of CNCF projects.
- Document best practices for:
    - Clusters under high churn (job focused)
        - Very little documentation on best practices for running thousands of jobs
        - CERN has some internal documentation
        - OpenAI has some good information on this in 2500-5000 node range
    - Document what doesn't currently work (expectation setting)
    - Quality management systems and information management (security)
- Establish global knowledge of the user base.
    - Create catalog of community member sites, and some public details of site details.
        - AI: Come up with questions to ask for survey
        - AI: Create Github repo (cncf/research-user-group)
            - Done - https://github.com/cncf/research-user-group
            - Have also requested a session at KubeCon San Diego
    - Create catalog (helm charts) of common research applications.
- Push for GA of Kubernetes Device Plugins (currently beta)
    - Support for node topology manager
- Improve Batch Style Computing Capabilities of Kubernetes
    - Ensure system that are being used (e.g. volcano) support multi-tenancy and auth(n,z) for jobs.
        - https://scitokens.org/
    - Work to promote kube-batch to GA (v1.0.0)
    - Work to promote Volcano to GA (v1.0.0)
    - Leverage current tooling (kube-batch / Volcano) with other 3rd party applications (kube-flow, argo, etc) to add or improve batch capabilities.
- Investigate, improve or develop workflow engines.
    - Build an interface that can interact with both traditional Workload Managers (e.g. SLURM) and Kubernetes.
        - interface/translation between HTCondor ClassAd and kubernetes
    - Build, add, or improve support for Common Workflow Language. (M. Crusoe is happy to assist with this)
- Interactive research experience (JupyterLab?, ssh-gateway, …)

- Improve multi-cluster (federation) scheduling and job distribution.
- Work with OCI group / Research Container Groups:
    - Rootless initiative
    - Loading layers from external source (might be containerd)

(For inspiration - see https://github.com/cncf/financial-user-group/issues)

Who else should we collaborate with?

- K8s SIG-Multicluster
- K8s SIG-Scheduling
- K8s SIG-Apps
- WG-Resource-Management
- WG-Machine-Learning
- CERN
- University of Michigan
- TACC
- NYU
- Berkeley Lab (LBL)
- Monash eResearch Centre
- Nvidia
- Intel
- IBM
- Huawei
- Google
- Cray
- Sylabs
- Common Workflow Language project
- AWE (Argonne National Lab Workflow Engine)
- G-Research

Appendix - questions

- What is your organization? (name, description, mission)
- How many researchers do you have? or how many are you supporting?
- What areas of research are you primarily supporting using Kubernetes? (e.g. bioinformatics, physics, generic research platform)
- Are your researchers consuming Kubernetes directly (deploy / use kubectl etc)? or do they consume services on top of it?
- How many Kubernetes clusters do you run? What size? (node/total core count/total RAM count)
- Are you using devices (e.g. GPUs) in any of your clusters?

- What is the baseline level of containers do you expect your researchers to have? (e.g. able to write Dockerfiles)
- What is the baseline level of Kubernetes do you expect your researchers to have?
- What (cloud native) tools do developers interact with and why did you choose them? (e.g. Minikube, Helm, ..)
- What (cloud native) tools did you reject or stopped using and why?
- What do you use for your CI/CD pipeline?
- Who makes decisions on technology? (e.g. does every team choose their stack, is consensus required, is a central team making decisions)
- What regulatory requirements do you have?
- What controls are in place to make production changes?
- What controls are in place to respond to CVEs?
- Who is on-call? (e.g. SRE, ops team, every development team, ..)
- How performance-sensitive are you? (e.g. does 2ms of network latency matter)
- What security precautions do you take? (e.g. network security, do you prevent privileged/root containers, do you use Helm with Tiller?)
- How important is it for you to have cloud-portable/agnostic technologies?
- How much isolation do you require for your workloads? (e.g. to have smaller blast radius, multi-tenancy)
- Do you trust your workloads? (e.g. can someone else run a container on your cluster)
- Do you care about quotas, cost management, granular cost reporting, and/or chargeback?
- How mature does a technology need to be before you start using it? (e.g. service mesh, …)