

# KubeVela - Review of Project Telemetry Data Collection and Usage

The following is meant to assist with a review of the project in connection with the project entity's Telemetry Data Collection and Usage Policy. Participants in the project are requested to provide responses to the following questions, regarding telemetry that is collected by the open source project and for use by the open source project community.

## Project: KubeVela

Date: February 18, 2022

KubeVela community serves the developers and aims to improve user experience based on getting a better understanding of them. This attention will be fulfilled via collecting telemetry data. Please review the proposal below and let's get it rightful in compliance with CNCF's Telemetry Data policy.

### 1. Specific data proposed to be collected

- Please fill in the following table with details on the specific data elements to be collected.

<b>Data element</b> e.g., software version; operating system; etc.	<b>Could be personal info?</b> (Yes/No)	<b>Could it be tracking or a unique identifier?</b> (Yes/No)	<b>Could it be end-user / sensitive / business data?</b> (Yes/No)	<b>Notes</b>
version	No	No	No	version of user is using e.g vela-core v1.2.4
platformID	No	Yes	No	an unique ID that represent an individual installation and of course we won't get to know who exactly this is
installTime	No	No	No	date that user install vela-core at the very first time

clusterCount	No	No	No	Index of how many clusters this unique ID possesses. We only know where the counts are located in <10, <50, <100, <500, <2000, <5000, <10000. which helps KubeVela develop and push Multi-Cluster feature to its limit for user's good
appCount	No	No	No	Index of how many KubeVela apps this unique ID possesses. We only know where the counts are located in <10, <50, <100, <500, <2000, <5000, <10000. helping community to improve its performance according the analysis of different scale of each user
enableAddonList	No	No	No	The Addon(KubeVela plugin) lists this user enabled. Then KubeVela community get to know which Addons are frequent used and deserve to be maintaining in a higher

				priority
xDefinitionTopList	No	No	No	The definition list(KubeVela major feature) this user enabled. Then KubeVela community get to know which features are frequent used and deserve to be maintaining in a higher priority
requestTime	No	No	No	The date of the data package is uploaded. Only be sent once a day.
browserInfo	No	No	No	Includes browser type, version, and screen width. This is for VelaUX(UI Dashboard) adaption for better user experience

- If there is public documentation on the project site describing this data, please also provide a URL to that documentation:
  - After this proposal get approved, the KubeVela Community will update this [public reference page](#)

## 2. User notification and opt-in

- Please describe how users are notified (1) that telemetry will be collected; and (2) which specific data elements will be collected:
  - an alert window will be popping up when the user installs VelaUX. **Note that, if they choose not to join the plan, they're still free to use VelaUX. It's not a requirement. Furthermore, the user will have the option to opt-out in VelaUX at any time.**
  - all the elements as above
- If there is public documentation on the project site or in the project source code with the particular notices, please also provide a URL:
  - GitHub repo: [VelaUX](#). **The source code will be under public examination.**

- Is the telemetry only collected and shared if the user voluntarily opts into collection? (As opposed to, collecting data unless the user opts out.)
  - This is true
- Is the user able to select between only sharing certain data elements, but not others?
  - Temporarily they couldn't
- How does notification and opt-in function if the software is installed and runs in a fully-automated installation (e.g., where there is no user who sees the notice and affirmatively clicks the "I consent" button)? Would telemetry data ever be collected in this type of scenario?
  - No. We plan to notice user in each version(e.g. v1.2.x -> v1.3.x) upgrading and get their consent every time

### 3. Storage and use of collected data

- Please describe where the telemetry data is collected and stored (e.g., on which servers / repos; where they are physically located, if known):
  - **Alibaba Cloud account's server. Registered via an gmail account([oamdev2020@gmail.com](mailto:oamdev2020@gmail.com)) that privately shares between the maintainers in [KubeVela Community Contributor Membership List](#)**
- Who administers and has access to the servers where data is stored?
  - the maintainers in [KubeVela Community Contributor Membership List](#)
- Are all participants in the project community permitted to view and use the collected telemetry data? Or only particular participants / community members?
  - all contributors in [KubeVela Community Contributor Membership List](#)

### 4. Security mechanisms

- Is there a documented way that an organization could block the telemetry data from being collected from their systems, even if one of their employees inadvertently approves it?
  - Temporarily not. But **the user has the option to opt-out in VelaUX at any time.**
- Is there a reasonable possibility that including telemetry functionality opens up security vulnerabilities?
  - No. The collecting process will be secured in a way of HTTPS request.
- If so:
  - What steps are taken to mitigate this?
  - If a user does not opt into telemetry data collection, would this risk be fully mitigated?

### 5. Future changes

If the project plans to extend the scope of telemetry collection in the future (e.g. to begin collecting new types of data), or if the answers given above would change, please update this form and notify us so that we can quickly review the updated proposal.

## 2022-03-07 Comments from LF review

Overall, the proposed telemetry data collection and usage generally looks good. There are a few follow-up comments and questions -- please let us know your thoughts in a response below the following.

- 1. Specific data proposed to be collected:
  - **clientIP: Please note that IP addresses in some circumstances and jurisdictions may be subject to data privacy protections, and may reveal personal or business information that should not be incorporated into the telemetry data.**
    - **If the goal is to collect information about e.g. the geographical region where the installer is located, could there be another mechanism to collect that information, perhaps by just asking the user at the time that they opt into participating in the telemetry?**
  - **browserInfo: Just to confirm, is there anything in the metadata collected in this data element that could be used to fingerprint or track users?**
- 2. User notification and opt-in:
  - **Is telemetry only collected if VelaUX is used? If a user were to use Kubevela without VelaUX, is it possible that telemetry data would be collected from them?**
  - **For the question “Is the user able to select between only sharing certain data elements, but not others?”, the response says “Temporarily they couldn’t”. Can you please clarify what this means?**
- 3. Storage and use of collected data:
  - **The service where the collected telemetry data is stored should be a CNCF-owned account. Is that currently the case? It’s fine for the Kubevela maintainers to administer it, but the account should be CNCF-owned and should have access by CNCF/LF IT staff.**
  - **It looks like the data is currently viewable by contributors who are listed on the OWNERS\_ALIASES list. Please note that the telemetry data should be accessible to any participants in the Kubevela contributor and user community. It doesn’t need to be stored in a world-readable fashion, but if it isn’t then there should be a process for interested users to make a request to gain access for appropriate reasons for the benefit of the Kubevela community. Is that something that you can add to the community documentation, for the way that people can make those sorts of requests?**
- 4. Security mechanisms:
  - **We understand that there isn’t currently a way for an organization to block telemetry data across their systems. Is there a particular known address / endpoint to which the telemetry is sent? If so, it would be good to document that, so that organizations could have the option to block that endpoint (using their own firewall setups) if they choose to do so.**

## 2022-03-07 Replies on KubeVela side

Thanks for so much walking through all the details. Let's sort those out one by one.

- 1. Specific data proposed to be collected:
  - clientIP: Please note that IP addresses in some circumstances and jurisdictions may be subject to data privacy protections, and may reveal personal or business information that should not be incorporated into the telemetry data.
    - If the goal is to collect information about e.g. the geographical region where the installer is located, could there be another mechanism to collect that information, perhaps by just asking the user at the time that they opt into participating in the telemetry?
    - **KubeVela: Understand your concern. To ensure personal or business information is secure, we decided to change it from clientIP to region. This means we'll not record the clientIP of users but rather call an IP-to-Region service to locate the region. Let us know if this is more appropriate.**
- 2. User notification and opt-in:
  - Is telemetry only collected if VelaUX is used?
    - **KubeVela: Yes.**
  - If a user were to use Kubevela without VelaUX, is it possible that telemetry data would be collected from them?
    - **KubeVela: No.**
  - For the question "Is the user able to select between only sharing certain data elements, but not others?", the response says "Temporarily they couldn't". Can you please clarify what this means?
    - **KubeVela: Users can only choose to opt-in or not opt-in. But they can't select specified metadatas when choosing to opt-in.**
- 3. Storage and use of collected data:
  - The service where the collected telemetry data is stored should be a CNCF-owned account. Is that currently the case? It's fine for the Kubevela maintainers to administer it, but the account should be CNCF-owned and should have access by CNCF/LF IT staff.
    - **KubeVela: We can do that. Register a new alibaba cloud account, set up the collaboration team of maintainers and transfer the ownership of the account to CNCF/LF staff. The only problem here is that it'll have costs. Would it be possible for CNCF to cover the expense or the community to cover by ourselves? Either way is okay for us.**
  - It looks like the data is currently viewable by contributors who are listed on the OWNERS\_ALIASES list. Please note that the telemetry data should be accessible to any participants in the Kubevela contributor and user

community. It doesn't need to be stored in a world-readable fashion, but if it isn't then there should be a process for interested users to make a request to gain access for appropriate reasons for the benefit of the Kubevela community. Is that something that you can add to the community documentation, for the way that people can make those sorts of requests?

- **KubeVela: Yes, this sure works. We'll add a request process into the KubeVela.io official doc that allows every member of the community to view all datas.**
- 4. Security mechanisms:
  - We understand that there isn't currently a way for an organization to block telemetry data across their systems. Is there a particular known address / endpoint to which the telemetry is sent? If so, it would be good to document that, so that organizations could have the option to block that endpoint (using their own firewall setups) if they choose to do so.
    - **KubeVela: Sure, we're planning to use [telemetry.kubevela.io](https://telemetry.kubevela.io) which is fully under CNCF administration.**

## 2022-03-09 LF responses

Thank you for the replies -- this is extremely helpful. I think this resolves most of the issues we raised, though there are a couple we'd like to follow up on -- see my notes in [blue](#) below:

**Regarding clientIP / region:** *KubeVela: Understand your concern. To ensure personal or business information is secure, we decided to change it from clientIP to region. This means we'll not record the clientIP of users but rather call an IP-to-Region service to locate the region. Let us know if this is more appropriate.*

### **LF response:**

- Thank you -- this is very helpful to switch to region. That said, would the "IP-to-Region service" still involve using a third-party service to determine the region? It would still likely be preferable to avoid sending their IP address data to a third party to determine the region.
- Is there a reason that it would not work to just ask the user their region at the same time when they opt into the telemetry (particularly as this would perhaps be more accurate region data anyway)? If not, then at a minimum you'll want to disclose at the point of opt-in that if they opt into telemetry, then their IP address data will be sent to the third party service.

**Regarding browserInfo:** I didn't see a response to the earlier question about browserInfo -- am I correct in assuming that this won't be detailed enough metadata to enable any form of fingerprinting or tracking?

**Regarding CNCF account ownership:** *KubeVela: We can do that. Register a new alibaba cloud account, set up the collaboration team of maintainers and transfer the ownership of the account to CNCF/LF staff. The only problem here is that it'll have costs. Would it be possible for CNCF to cover the expense or the community to cover by ourselves? Either way is okay for us.*

**LF response:** Thank you -- I've been responding from the LF legal team viewpoint, but I'll raise this to the CNCF folks so that they can coordinate with you on this. Thanks!

## 2022-03-14 Replies on KubeVela side

Thanks and let's wrap up for the last several issues. Here are follow-ups:

### **LF response:**

- Thank you -- this is very helpful to switch to region. That said, would the "IP-to-Region service" still involve using a third-party service to determine the region? It would still likely be preferable to avoid sending their IP address data to a third party to determine the region.
- Is there a reason that it would not work to just ask the user their region at the same time when they opt into the telemetry (particularly as this would perhaps be more accurate region data anyway)? If not, then at a minimum you'll want to disclose at the point of opt-in that if they opt into telemetry, then their IP address data will be sent to the third party service.
  - KubeVela: Agreed. We removed the column *region* from that table. I assume it's a safer and more secure way to do an CNCF-approved publish questionnaire to get the user's region in future.

**Regarding browserInfo:** I didn't see a response to the earlier question about browserInfo -- am I correct in assuming that this won't be detailed enough metadata to enable any form of fingerprinting or tracking?

- KubeVela: My bad, had missed this one. Your assumption is totally accurate and it's impossible to track users via *browserInfo* which only tells about user's browsers not the users themselves, this helps VelaXU(a web application) to polish its user experience.

The rest is all set. Thanks again.

## 2022-03-14 LF final comments and approval

- Thank you for the final set of responses. Per your note above, I'm deleting the "region" row from the analysis table in the first section above. If you can accept that change in the data documentation, then we should be good to go.
- I understand that the CNCF team will likely be working with you to transition over to a storage account for the data that is under CNCF admin and billing. I understand also that there are some updates to the user-facing documentation that you'll be making, as discussed above.
- With those pending changes, LF has approved KubeVela's telemetry collection as described in this document.

## 2022-03-16 KubeVela Community

The KubeVela community shows great appreciation and thanks for all your assistance along the way. Remaining actions will be closed via another service desk ticket. Good day and take care!