Sustainable Agricultural Supply Chain Initiative

Trace

Report

INATrace cost estimation

To be fine-tuned with/by the future administrator

Published by:

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH
Registered offices
Bonn and Eschborn, Germany
Friedrich-Ebert-Allee 32+36
53113 Bonn
T +49 228 44600

E info@giz.de
I www.giz.de/en

Programme/project description:
Sector Programme Initiative for Sustainable Supply Chains and Standards (INA)

Authors:
Alessio Vertemati, Gianluca Colombo, Stéphane Henriod

OneOff-Tech UG (haftungsbeschränkt)
Urbanstraße 71 D–10967, Berlin, Germany
Geschäftsführer: Dr. Gianluca Colombo
Handelsregister: Amtsgericht Charlottenburg, HRB 175895 B
VAT number (International) DE305779646
VAT number (German) 37/461/50029

Design/layout.:

OneOff-Tech UG (haftungsbeschränkt)
Urbanstraße 71 D–10967, Berlin, Germany
Geschäftsführer: Dr. Gianluca Colombo
Handelsregister: Amtsgericht Charlottenburg, HRB 175895 B
VAT number (International) DE305779646
VAT number (German) 37/461/50029

On behalf of
German Federal Ministry for Economic Cooperation and Development (BMZ)
Division 122 Agriculture, rural development
Berlin, Bonn, October 2023

On behalf of

Federal Ministry
for Economic Cooperation
and Development

Contents

1 Foreword

This report shows an initial estimate of costs by looking at the INATrace information from the current production and deployment systems, available until March 2023. We assume that no changes will be made to the technology.

The given figures should be seen as an approximation of potential costs, derived from existing information and assumptions. Two scenarios are analysed - one without blockchain and one with blockchain.

Some costs are expressed in terms of expected Working Days (WD) rather than in monetary terms so that the estimate can be updated on a country/region/profile basis.

Cost categories

Considering the analysis in Guideline for setting up a new INATrace instance (v 1.1) and the available database dump at March 2023 INATrace costs factors can be divided in four categories[1]:

1. **Infrastructural costs**. The costs generated by the infrastructure on which the software is running. For example, database hosting costs, compute resources and backups.
2. **Maintenance costs**. The costs of the personnel that ensure service availability
3. **Support costs**. The cost of the agents providing support to users request.
4. **Third party service costs**. The cost of services that are required to run a web-based application, which can be assumed to be fixed up-to a certain threshold. For example, domain name, email server, or third party APIs.

Since this document focuses on the technical aspects of INATrace, the costs for actually using and promoting INATrace are not included. In the current situation, those costs include GIZ's staff salaries (coordination, steering, strategy, onboarding, communication, etc.), external consultants (value chain mapping, interactions with the users, etc.) and other related costs (PR, travels, etc.). The overall success of INATrace however also strongly depends on those "non-technical" activities, which should not be left unaccounted for.

Based on the data from September 2022 to March 2023 collected during the INATrace production instance exploration, the cost factors for INATrace depend on the following dimensions:

- **Number of users**. All users[2] of INATrace, both authenticated and visitors of the publicly available pages (e.g. QR Code landing page).
- **Number of transactions.** The amount of transactions stored and the corresponding payment proof images.
- **Number of cooperatives**: the number of cooperatives that are hosted on a single instance.

The above list is not a ranking, so users are not more important than transactions. As a baseline the currently available analytics from the active INATrace instance are considered.

[1] Some of the categories are adapted from Total Cost of Ownership, Encyclopedia of Production and Manufacturing Management. Springer, Boston, MA . https://doi.org/10.1007/1-4020-0612-8_989 .

[2] The database contained also farmers. Farmers are not considered users of the system as there was no evidence that they were actively interacting with the system. Authenticated users are mostly cooperative personnel and buyers personnel.

Users		Transactions	Cooperatives
112 authenticated	261 Visitors on QR code page (~30 scan/month)	45274 (2854 transaction/month)	19

Table 1. Current available usage statistics as extracted from the database analysis. Period September 2022 - March 2023.

The outlined dimensions are used as the main control variables. It should be noted that there are variables not under the control of INATrace or only foreseeable up-to a certain degree, e.g. price increase, bugs causing excessive hardware resources or service quota usage.

The following assumptions are made throughout the entire document:

- **A single INATrace instance**, hosted in Europe, is considered. If multiple instances are created they all have the same usage patterns and so cost ranges can be multiplied by the number of instances. The cost of migrating an instance to a country can be seen as a multiplier factor of the European costs.
- **INATrace deployment** follows the same approach as the current instance: cloud provider with database backup. Application lifecycle orchestrated on Kubernetes in a single data center inside an availability zone[3].
- **Maintenance** activities includes small bug-fixes and regular maintenance operations based on the estimate received during technical calls with Anteja (i.e. for up-to 2 WDs/month).
- **No additional development** (new features, improvements to the user-interface, etc.) is taken into consideration. As it can be expected that users will gradually ask for new features, additional budget needs to be available.

Limitations:

- Infrastructure resource utilization monitoring was not available prior to the analysis.
- The volume and cost of the support provided to the users were unknown.
- User interactions and monthly usage statistics are known only for a subset of metrics (see table above: number of users, number of scans, and number of transactions).

2 Cost estimation break-down

2.1 Infrastructural cost estimation

Infrastructure costs depend on the number of users and transactions as well as the technology stack (with/without blockchain). The current infrastructure runs on Oracle Cloud managed services. In particular a managed database and a managed Kubernetes cluster with external storage are used.

The cost estimation is based on the following scenarios:

- **Scenario 1**: constant usage with no big increase.
  - 200 total users overall
  - 100 visitors/month across all QR code pages

[3] Data Center vs Availability Zone vs Region – Cloudy For Sure, What are Azure regions and availability zones? | Microsoft Learn.

- overall of 3000 transactions/month[4]
- **Scenario 2**: increase sold products and constant user base
  - 200 total users overall
  - 400 visitors/month across all QR code pages
  - overall of 9000 transactions/month
- **Scenario 3**: increased overall usage
  - 400 total users overall
  - 600 visitors/month across all QR code pages
  - overall of 12000 transactions/month

Each transaction is considered to have a 1 MB picture attached as proof[5].

All the above scenarios can be implemented with various combination of services. To reduce the combinatory explosion, three levels, small, medium and large, are defined for the computational resources. The estimations will indicate what levels are suitable to run the specific scenario.

Item	***small***	***medium***	***large***
Database	1-3 vCPU, 8-16 GB RAM, <150 GB storage	3 vCPU, 24 GB RAM, 150-240 GB storage	3-6 vCPU, 32 GB RAM, > 240 GB storage
Single compute node	2 vCPU, 8 GB RAM	4-8 vCPU, 16 GB RAM	8-16 vCPU, 32 GB RAM
Attached storage	<= 10 GB	10-50 GB	>= 50 GB

The currently used services from Oracle Cloud refer to: medium level for database, small for compute. For the attached storage no data was shared.

Current architecture (no blockchain)

scenario	***database level***	***compute level***	***storage level***	***estimated cost***
scenario 1	small – medium	2 x small	small	100 – 300 €/month
scenario 2	medium	2 x small – 2 x medium	small – medium	200 – 500 €/month
scenario 3	medium – large	2 x small – 2 x medium	medium – large	300 – 700 €/month

If a High Available deployment (SLA up to 99.999% uptime) is envisioned, the above cost estimations should be doubled along all resource levels.

additional services	***estimated use***	***estimated cost***
Kubernetes control plane	24/7	0 – 20 €/month
Backup	sum of database and file storage	10 – 20 €/month

Whenever a Kubernetes cluster is used a control plane must be considered. Some cloud providers offer the service for free and some others don't.

[4] This scenario is comparable to the usage between September 2022 – March 2023 with a peak of 112 users and ~2000 transaction/month.
[5] Based on the current available information this can be considered a worst case scenario and should cover additional file requirement (e.g. product pictures).

With blockchain

Blockchain costs depends on three factors:

1. The number of organizations who take part in the blockchain. A blockchain, in principle, should not be owned by a single organization.
2. The number of peer nodes. Computational nodes that perform the calculation. Each organization must have a minimum of 1 peer node.
3. The number and size of transactions per second. The higher the transactions are the more computational power is required.

For the purposes of this document let's assume that only one organization owns nodes in the blockchain and that there are two peer nodes for that organization. Each peer node runs on a AWS m5.2xlarge comparable machine[6]. This configuration should allow to store 100-200 transactions/second, which should be sufficient to cover the foreseen throughput in the three scenarios. It must be noted that for peak of usage a queue system must be in place.

scenario	*database level*	*compute level*	*storage level*	*estimated additional blockchain cost*[7]	*estimated cost*
scenario 1	small – medium	2 x small	small	500 – 1500 €/month	100 – 300 €/month
scenario 2	small – medium	2 x small – 2 x medium	small – medium	500 – 1500 €/month	200 – 500 €/month
scenario 3	medium – large	2 x small – 2 x medium	medium – large	500 – 1500 €/month	300 – 700 €/month

A managed blockchain solution is considered (e.g. AWS, IBM, Oracle).

If a High Available deployment (SLA up to 99.999% uptime) is envisioned, the above cost estimations should be doubled along all resource levels.

additional Services	*estimated use*	*estimated cost*
Kubernetes control plane	24/7	0 – 20 €/month
Backup	sum of database, file storage and blockchain storage	10 – 50 €/month

Whenever a Kubernetes cluster is used a control plane must be considered. Some cloud providers offer the service for free and some other don't.

2.2 Maintenance cost estimation

Maintenance costs are driven by regular activities. A non-exhaustive list may include the following:

- *Software upgrades*. Keeping all software up-to-date is necessary to avoid security risks and to take advantage of new features that can simplify day-to-day operations.
- *Infrastructure scaling*. Increased traffic and usage of the application may require additional resources, e.g. increase file storage quota to prevent problems when the limit is reached

[6] Selected on purpose to allow comparison with existing benchmarks of Hyperledger Fabric .
[7] For in-depth studies please refer to Total costs of ownership for blockchain solutions (Ernst & Young, LLP, 2019), Where to deploy your Hyperledger Fabric Network and why? | by Ayush Tiwari | Medium, Scalable Blockchain Network – Amazon Managed Blockchain Pricing – AWS .

- *Applying optimizations*. Increased traffic and usage of the application may affect user experience, e.g. as number of transactions increase the processing of database queries can slow down affecting the system performance.
- *Backup monitoring*. Continuously test the disaster recovery procedure
- *Bug-fixes*. Even without addressing new developments, at least the most blocking identified bugs must be resolved, in order to keep the interest and trust of the users.
- *Setup and configuration of new tenants*. Additional configuration might be required as new users and clients are added.

Complexity and cost of maintenance is influenced by number of users, number of transactions, number of cooperatives and the technological stack (database or blockchain).

Source	Infrastructure	Estimation	Estimated person in the devops team
Maintenance	database	2 – 3 WD/month[8]	1 – 2
Maintenance	blockchain	3 – 9 WD/month	2 – 4

Current architecture (no blockchain)

As stated in technical exchanges with Anteja the current maintenance effort is about two WDs/month. If maintenance is performed by a different team other than the one that developed the application, the maintenance cost can be estimated at 2.6 WDs/month[9].

It should be noted that key events leading to increased maintenance time are:

- New tenants added. The complexity of adding and configuring a new tenant (e.g. cooperative) can increase the maintenance costs
- Infrastructure resource limits. When infrastructural limits are approached the maintenance activities effort increases. For example an almost full database storage will require a multi-day operation to migrate the data to ensure no-downtime is caused to the application
- Third party services disruption. Problems to third party services can cause increased maintenance efforts to apply workarounds or mitigation actions

With blockchain

In terms of maintenance, blockchain requires more effort and resources compared to a traditional database. This is because blockchain is a decentralized system. As a result, maintenance tasks are more complex.

The blockchain impacts also the traditional application maintenance and should be summed to the regular activities. Based on available documentation of managed services (.e.g. Oracle Cloud and IBM) the provided estimation is based on the time that a devops team of two person should dedicate monthly.

[8] As provided by Anteja during technical meetings and interviews.
[9] The used scaling factor is 1.3 as expressed in Ivan, Ion, and Mihai Liviu Despa. "Estimating Maintenance Cost for Web Applications." *Informatica Economica* 20.4 (2016).

2.3 Support cost estimation

User Support can be considered independent of the technological implementation behind the scenes, i.e. database vs blockchain. It is expected to be related to the number of active users and cooperatives.

Source	Estimated use	Estimation
2 x Support agent	15 ticket/month, 1 hour average handling	2 – 4 WD/month
Platform to handle Community Support	5 staff members, up to 100,000 page views	0 – 100 €/month
Ticketing platform to handle paid support	5 support agents	50 – 120 €/month

Support agent

Based on the current information all support requests are handled via email or in the field by contact points. Currently, 3 agents (1 in Rwanda, 2 in Central America) provide user support and onboarding, but are also responsible for training supply chain mapping and configuration, etc. Those tasks obviously go well beyond the 2-4 monthly WDs stated in the above table. If the agents are further expected to be the main "face to the customer", for both current and prospective users, and given the fact that users are spread over very different time zones, a realistic order of magnitude would be 2-3 agents for ~500 end users.

It is expected that the number of support tickets will increase as new cooperatives are added and will gradually decrease as users gain knowledge of the application.

It should be noted that the support team is strictly connected to one INATrace instance. Multiple instances will require dedicated support agents.

Community support platform

Given the Open Source nature of INATrace a community based support channel can be considered. Initially this will require one support agent to help the maintainers handle the requests. As the community grows the effort will reduce. This type of support is always on best-effort (whatever the support agent, and other community members, are able to provide) and not suitable for enterprises.

Costs depend on the selected platform. For example GitHub Discussions is free for publicly hosted GitHub repositories, while Discourse, a dedicated forum application, starts at 50 €/month for small communities.

Ticketing platform

It is advisable to use a dedicated ticketing platform to centralize the handling of support requests. This will make the collaboration between agents more efficient and, with a knowledge base, allows for more streamlined problem resolution.

The recommended threshold is 50-100 active users or a team of 5 support agents. These platforms are usually priced based on the number of support agents.

2.4 Third party services cost estimation

Third party services cost can be divided in (1) **fixed**, (2) **usage based** (i.e. pay-per-use). Costs are not influenced by the use of a blockchain.

2.4.1 Fixed

Services that do not depend on any of the variables (user, transaction, cooperative).

Source	Estimation
Domains (currently inatrace.org, inatrace.de and inatrace.com)	3 – 4 €/month[10]
Exchange Rate API (https://exchangeratesapi.io/)	0 – 10 €/month

The pricing of the Exchange Rate API is based on usage, but the use within INATrace - i.e., single calls scheduled periodically - allows to consider it as a fixed cost as it does not depend on the number of users, cooperatives or transactions.

2.4.2 Usage based

Usage based services billing is based on monitoring the assigned quota consumption. Depending on the service the quota could be connected to users or single activities performed by a user.

Source	Estimated use	Estimation
E-Mail service	10 email/month/user => ~ 10000 emails	20 – 40 €/month
Interactive maps via Google Maps	300 QR code scan/day (~ 10 times more than current average)	70 €/month

E-Mail service

Currently managed by GIZ. It is advisable to select a transactional mail provider to ensure email deliverability. It is not recommended to choose providers like Microsoft Exchange or Google Mail when email volume approaches 100 email/day. This is due to service rate limits and terms of use.

Transactional email services are specifically created to send huge amounts of emails and monitor the delivery of such emails. Example of services include: Postmark, Mailgun, Resend or other European based companies.

Estimated usage is based on 100 monthly active users receiving a maximum of 10 emails/month (e.g. password resets, new order notification, etc.).

[10] Domains are usually billed yearly. Estimates are given in months to be comparable with other cost sources.

Interactive maps

Currently provided by Google Maps. Maps are used on the QR Code page and within the authenticated area to show or select location information. For the sake of the estimation all interactions are considered dynamic[11].

The costs of the service depend on the number of users. Current data shows that the 112 users and the 30 QR code scans/day can generate between 30 and 1000 requests to Google Maps per month with an estimated cost of 7 €/month. To represent a significant use case the usage of the QR code scanning functionality is increased to 300 scans/day. The QR Code page contains a map and is therefore the most appropriate candidate for Google Maps costs.

[11] Google Maps divide static maps rendered on a page from maps that can shows point of interests or allow drawing of areas. Specific usage depend on the implementation details and is best to be asked to Anteja.

3 Summary tables

3.1 Evaluation of monthly costs

Item	*small*	*medium*	*large*
Database	1-3 vCPU, 8-16 GB RAM, <150 GB storage	3 vCPU, 24 GB RAM, 150-240 GB storage	3-6 vCPU, 32 GB RAM, > 240 GB storage
Single compute node	2 vCPU, 8 GB RAM	4-8 vCPU, 16 GB RAM	8-16 vCPU, 32 GB RAM
Attached storage	<= 10 GB	10-50 GB	>= 50 GB

Item	**Scenario 1 - small**		**Scenario 2 - medium**		**Scenario 3 - large**	
	Low estimate	**High estimate**	**Low estimate**	**High estimate**	**Low estimate**	**High estimate**
Infrastructure **without** blockchain	100 €	300 €	200 €	500 €	300 €	700 €
Infrastructure **with** blockchain	600 €	1800 €	700 €	2000 €	800 €	2200 €
Kubernetes control plane	0 €	20 €	0 €	20 €	0 €	20 €
Backup	10 €	50 €	10 €	50 €	10 €	50 €
Platform for community support	0 €	100 €	0 €	100 €	0 €	100 €
Ticketing platform	50 €	120 €	50 €	120 €	50 €	120 €
Domains	3 €	4 €	3 €	4 €	3 €	4 €
Exchange rate API	0 €	10 €	0 €	10 €	0 €	10 €
Email service	20 €	40 €	20 €	40 €	20 €	40 €
Interactive maps via Google Maps	70 €	70 €	70 €	70 €	70 €	70 €
Total (without blockchain)	253 €	714 €	353 €	914 €	453 €	1114 €
Total (with blockchain)	753 €	2214 €	853 €	2414 €	953 €	2614 €

3.2 Evaluation of the monthly working effort

Please note that the evaluation below is the absolute minimal effort necessary to ensure that an existing INATrace instance functions properly. It does not include any development or bug fixing, and does not allow for adding new value chains into the system.

Item	Low estimate	High estimate
Database maintenance	2 WDs	3 WDs
Blockchain maintenance	3 WDs	9 WDs
User support (handling tickets)	2 WDs	4 WDs
Total	7 WDs	16 WDs