
vos.ai
compute

We thank you for your consideration and welcome the opportunity to embark on this journey together for a better future.

C O N T E N T S

1. Problem Statement

This section describes the opportunity and context for the problem being solved. It will clearly articulate how the problem will be solved.

1.1 Problem

The following enumerates the problems faced with building a system such as VOSAI AGI. Preceding this list will be in depth explanations of each item in the list.

- Cloud providers lack the infrastructure at a cost effective rate for
    - Machine learning
    - Many low power CPU cores
- An infrastructure in a data center to support an AGI is costly due to
    - GPU dependence
    - Hybrid data store dependence (e.g. graph, document, key/value)
- Infrastructure changes between evolutions are highly disruptive
- Privacy and security
- Performance and throughput
- Existing enterprise application support lacking in blockchain
- Mining the blockchain

1.1.1 Cloud Infrastructure is Lacking

VOSAI AGI requires extensive compute power. The compute power required is costly, custom and not completely available in modern day cloud providers. Cloud providers of today focus on delivering an infrastructure that works very well for standard websites, e-commerce and enterprise systems. However, where it falls short is in the ability to compute massive amounts of data using GPUs as well the ability to scale to thousands of CPU cores at a cost effective price. The reasoning behind this is really they have designed these cloud systems to work for the most commonly used systems in the world today. These system include websites, web services, databases, caching, messaging and general application servers (e.g. WWW, FTP, SSH, SMTP, SNMP).

"How many GPUs are in the world computer?"

1.1.2 Custom Infrastructure is Costly

Machine learning systems are expensive to build. They often require thousands of GPU based servers processing petabytes of information as fast as possible. Moreover, the network and storage requirements alone for this infrastructure is astronomical. It has become such a problem that it has led to new developments in server architecture that deviate from the CPU / GPU paradigm. The advent of TPUs

and DPUs are promising but has yet to prove its worth in the market. Many of these solutions are not yet readily available to the masses for deployment and use.

"WHY BUILD AN INFRASTRUCTURE WHEN IT CAN BE DECENTRALIZED DRIVING COSTS TO A MINIMUM?"

1.1.3 Infrastructure Evolutions / Changes / Dependence

The past few decades we have seen evolutions of infrastructures take place. From main frames, client server, having a server room, having a cage in a data center, make your own data center or leveraging a 3rd party data center to cloud providers like AWS, GCP and Azure. These advancements have been great and have saved organizations millions in the process while costing them millions at the same time. Organizations have been forced to continually upgrade, change or migrate systems between these environments time and time again. Moreover, if your enterprise systems are developed using too many of the tools provided by these infrastructures they become married forever and difficult to migrate thereafter. For example, AWS offers many great tools for infrastructure but once you build applications around them you are stuck with that forever (e.g. Lambda functions in AWS).

"WHAT IF SOFTWARE INSTANTLY UPGRADED WITH NO CHANGES TO CODE?"

1.1.4 Privacy and Security

Many cloud providers and data center provide privacy and security mechanisms to make nearly everyone happy enough. However, when it comes to your privacy and the data related to you nothing is ever really safe. Governments across the world storm in and demand access to all data. We have seen this countless times in the USA and other countries where email providers provide data to the government without your permission. Furthermore, your data is inter-sold between companies for monetary gain. What do you think Alexa from Amazon is actually doing? How is Amazon really using this information and who gets to use it?

"IS NO COUNTRY SAFE?"

1.1.5 Performance and throughput

Billions of requests if not more would be made to VOSAI AGI. The sheer performance and throughput required by this system would be one of the largest

hurdles to overcome aside from the actual creation of the AGI itself. Moreover, the AGI itself is forever learning and growing on its own as well. That means it's almost double the amount of work going on to achieve its objectives.

"How many data centers does it take to screw in a lightbulb?"

1.1.6 Lack of enterprise application support

Blockchain is sure to disrupt much more than we are aware of. Yet, it is still in its infancy. There is limited support for enterprise applications running in a decentralized manor. For example, should an enterprise system require a data server such as MySQL it's not readily available on a decentralized system.

"What if enterprise applications were available on a decentralized system?"

1.1.7 Mining the blockchain

Recent advancements in mining have proven to differ between tokens. In some cases a miner is simply performing a hashing function while in other cases it's completely different. The problem today is that many miners with GPU rigs are not able to secure the network as they once were. ASICs has far overcome mining and rightfully so.

The secondary problem we see is conceptual, should GPUs really just be sitting there running a hashing function all day long? Yes this calculation is used to secure the network however what if the processing actually being done was useful outside of just securing the network. We currently have "proof of work" (POW) and "proof of stake" (PoS) – what we see is possibly an alternate evolution called "proof of product".

"How can a compute layer be developed future proofing the software residing within?"

1.2 Opportunity

We have an opportunity to create a generic compute layer (VOSAI Compute) which completely abstracts the underlying infrastructure at the application level. VOSAI Compute would be agnostic to infrastructure. This means that it could be backed by AWS, GCP,

Azure, Bluemix, Golem, World Computer, Super Computers, Hybrid Data Centers and/or all of the above.

Moreover, VOSAI Compute would serve as the primary infrastructure of any AI based system. The first system being that of VOSAI AGI. VOSAI Compute allows VOSAI AGI to scale infinitely where resources are automatically added at a far lower cost. It allows VOSAI AGI to flourish with a far lower cost in infrastructure that ever seen before.

VOSAI Compute starts with the basics of decentralizing compute power at the GPU level such that VOSAI AGI can harness the power of the world computers GPUs rather than its own in some locked up data center or cloud provider. VOSAI Compute will grow over time to include enterprise applications such as data stores, caching mechanisms, message queues and much more.

"If we worked on the assumption that what is accepted as true really is true, then there would be little hope for advance."

– Orville Wright

1.3 Why is it timely and important?

There are many technological revolutions occurring today. Some that are far away from the scope of what we will achieve. Those revolutions specific to VOSAI revolve around machine learning, computer vision, robotics, drones, infrastructure, and blockchain. All of which are at turning points or already turning.

For example, it was only a few years ago that drones was only reserved for a military application. Yet today, drones are a common buzzword among many organizations that are using quadcopters for surveying buildings, farms and infrastructure.

None of us can go one day without seeing something in the news about any of these topics which indicates to us that there is a lot of momentum and innovation occurring in the space. In the next decade, we will experience more automation and technological progress than we have in all the existing of human kind.

The reason why it is timely and important today is because all of these technologies have reached a maturing point where they can modularized, plugged in, assembled and integrated into cohesive units that make it possible. Similar to the evolution of blockchain, AI also had the same trends that span decades of research and developments.

"If man realizes technology is within reach, he achieves it. Like its damn near instinctive." - Motoko Kusanagi, Ghost in the Shell

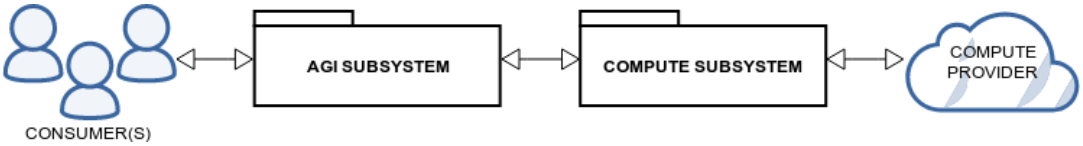
2. Proposed Solution

This section explains the approach used to solve the problem defined in Section 1 in detail.

2.1 Overall approach

VOSAI Compute is a layer between the AGI and the actual back end compute executing AGI code. The reasoning for creating this layer to enable the AGI with the ability to be agnostic to compute back end while rapidly adapting to the latest technologies. This ability allows for delivering better performance and quality of output from the AGI to its consumers with minimal effort and costs. The effort and cost would typically reflect at the VOSAI organization as well as cost pass thru to the end consumer.

The following diagram illustrates a high level approach to the VOSAI Compute layer:



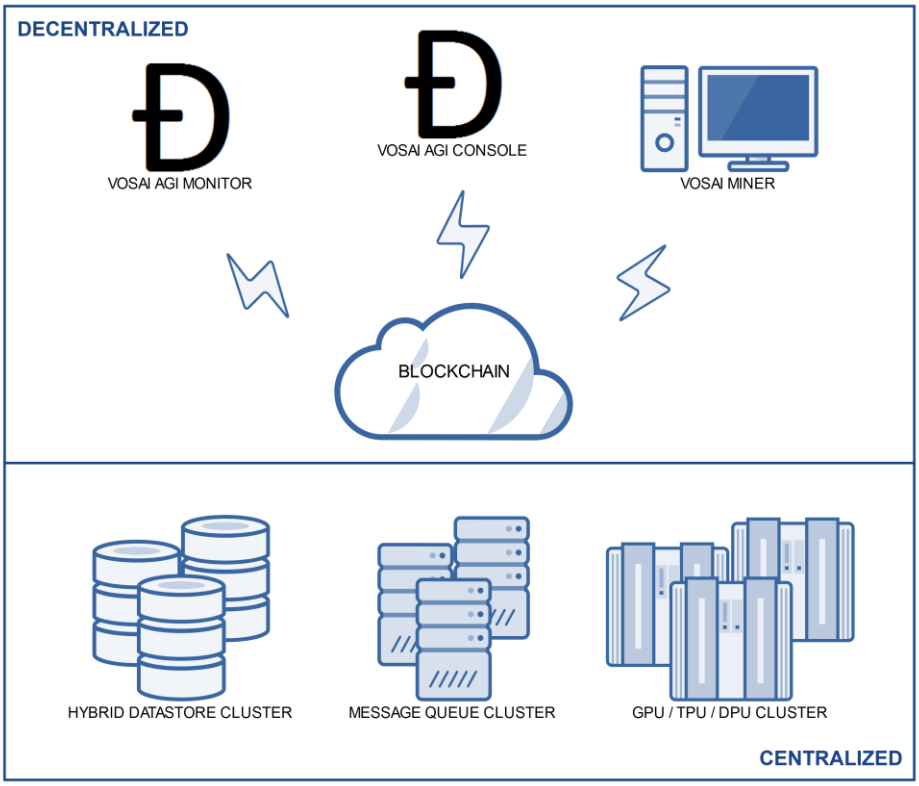
- A consumer is a system that leverages the AGI to perform some task. It may be using it for conversations between its end users and the purpose of their business.
- The AGI Subsystem is the module(s) responsible for handling inbound requests from the outside world and relaying them accordingly to the AGI internal systems.
- The Compute Subsystem is where the AGI actually resides. The AGI design and components are outside the scope of this section. They are further explained in the AGI repository.

To ensure success while following a gradual iterative approach, the compute layer is implemented in phases. These phases do not cover mining (e.g. performing machine learning on a miner's PC). The reader should assume the miner is part of this infrastructure from the start.

2.1.1 Phase 0

Phase 0 sets the foundation to build upon for all subsequent phases. Our intent is to leverage the world computer for processing of machine learning functions for the AGI. At least initially that is our intent. Over time, we want to see this grow to the point where we can have the entire AGI system including message queues, caching layers and data stores running on the world computer.

The following diagram illustrates our intent with Phase 0.



During this phase we have segmented out central and decentral silos. The current technology in place for a purely decentralized system does not exists. Therefore, we must have a starting point. Our intent, is to move to a purely decentralized infrastructure over time as the technology matures. VOSAI Compute has this vision in mind – therefore – if the technology available does not provide this then VOSAI Compute will evolve to be this infrastructure.

The following list briefly explains each component in the diagram above.

- **VOSAI AGI Monitor** is a decentralized application (DAPP) used to monitor the current state of the AGI
- **VOSAI AGI Console** is a DAPP allowing users to interface directly with the system in a simple chat like interface
- **VOSAI Miner** is a standard application running on a mining rig which takes requests from the network to process machine learning functions which may be CPU, GPU, DPU or TPU intensive. Technologies like ASICs would not apply here at all.
- **Hybrid Datastore Cluster** is a specialized set of data stores which combine relational, graph, document and key/value store databases into one unified system. This hybrid data store is required for VOSAI AGI and will amass data infinitely. Eventually this will scale to the world computer once the infrastructure can support it.

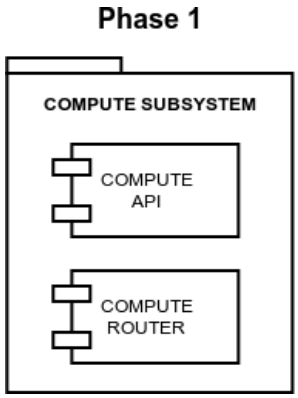
- **Message Queue Cluster** is a supporting cluster to the VOSAI AGI system. VOSAI AGI is designed as a purely message based and distributed set of systems. Given the nature of a messaging system, queues are a requirement.
- **GPU / TPU / DPU Cluster** is a supporting cluster which performs machine learning functions much like that of the **VOSAI Miner**. We do not foresee doing centralized mining on our own nor is it a primary function of what we intend to do. However, we do need to take part in mining just like all remaining miners for R&D purposes. In the future, this cluster will evolve as technology evolves to support R&D of machine learning algorithms and specific hardware required for these.

Conclusion

There is an initial infrastructure that going to be created for both hardware and software levels to support the creation of VOSAI. Some may appear to be transient (temporary) but in fact will forever evolve and grow with the system. We do ultimately foresee moving the entire VOSAI system(s) onto a decentralized infrastructure once the proper technology is made either through our efforts or the community's efforts.

2.1.2 Phase 1

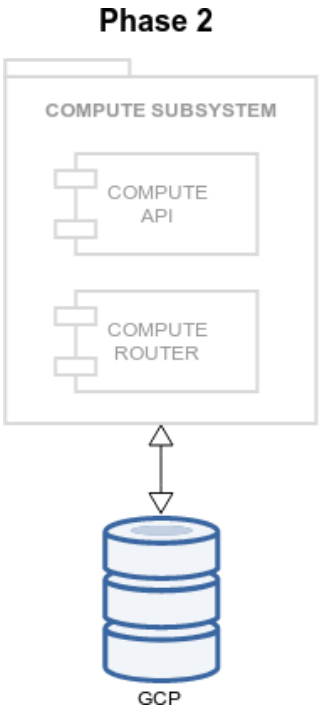
This diagram illustrates Phase 1 of VOSAI Compute. This is a preliminary design and will evolve over time.



- Compute API - a restful API which handles request from the AGI subsystem.
- Compute Router - a daemon responsible for routing the request to the appropriate compute back end

2.1.3 Phase 2

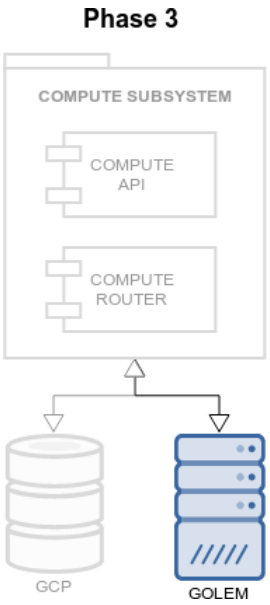
This diagram illustrates Phase 2 of the VOSAI Compute. The purpose of this phase is to integrate with one cloud provider as the compute back end. For this back end, we have chosen to integrate with GCP - Google Cloud Platform.



- GCP - the first of many cloud providers to integrate as a compute back end
- Additional components are developed in this phase to allow plugging in new compute providers easily

2.1.4 Phase 3

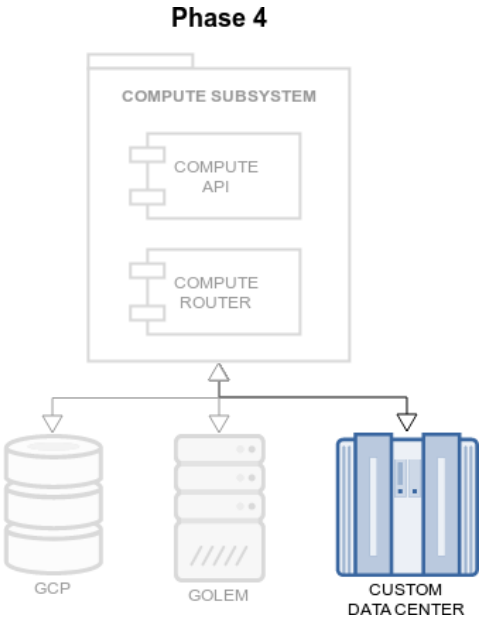
This diagram illustrates Phase 3 of the VOSAI Compute. The intent of this phase is to integrate with Golem or similar compute back end which heavily leverages blockchain as well as the crowd sourced approach to computing. This integration is subject to change based on the progress of Golem. We have yet to see substantial progress with Golem that would completely warrant us leveraging their infrastructure. Upon arriving this Phase we will determine if we leverage Golem and expand upon it or create our own solution specialized for VOSAI Compute.



- Similar integrate effort here as in Phase 2
- Port existing applications and frameworks to work appropriately over to Golem (e.g. Tensorflow)

2.1.5 Phase 4

This diagram illustrates Phase 4 of the VOSAI Compute. The intent of this phase is to create a data center which is highly specialized for both machine learning and mining applications. Today's current cloud offerings do not fully provide what is required for these two tasks.



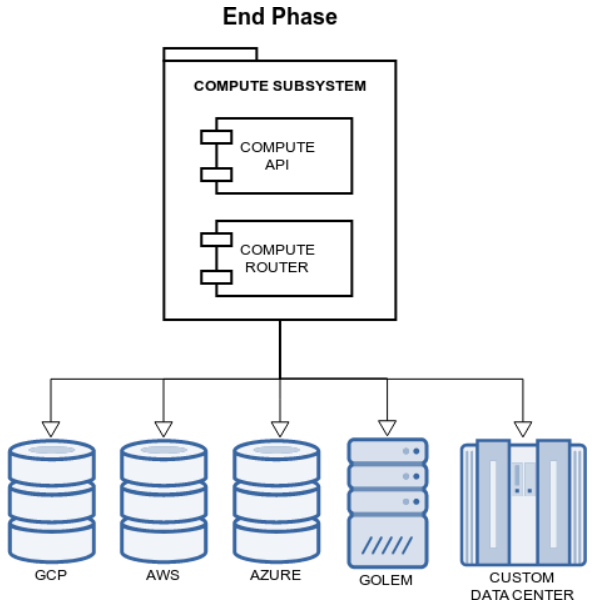
- ARM / Intel CPUs
- GPU Agnostic (AMD / NVIDIA)

- DPU Backed
- FPGA Possibilities
- Alternate means of power and cooling

In this phase, the custom data center used to create the requirement compute layer modules will be developed against the infrastructure created in Phase 0.

2.1.6 End Phase

This diagram illustrates Phase 5 - the end Phase of the VOSAI Compute. The intent of this phase is to add additional cloud providers or remove them if they are deemed useless.



2.1.7 Beyond

There are a few unknowns when it comes to the underlying hardware for compute. That of which what new technology will be available and which direction to take when designing the systems that run on them.

Super Computers

There are various organization spanning the globe that provide access to super computers. This is early in discussions but we are considering Oak Ridge National Laboratory.

Quantum Computers

There are only a few worldwide that claim to have a Quantum Computer. At the time of this writing - Quantum Computers are still in their infancy and not quite ready for the mainstream.

2.1.8 End Notes

There are many topics we have covered in this section. The key take away we want to hammer home is our intent to use the world computer in addition to cloud providers and data centers. Over time, as the world computer grows become ever more prevalent we foresee VOSAI moving completely to the world computer.

3. Annual Progress Plan

This section describes and defines our progress plan. Our progress plan indicates our expected annual progress towards achieving our goals related to the problem stated in Section 1.

3.1 AGI Progress Plan

Year 1

PHASE 0-2

1. Final Infrastructure design, Software Architecture Design
2. Development commence on VOSAI Learner v1, Data store, Messaging
3. Deploy Infrastructure to Test DC
4. Deploy Infrastructure to Final DC
5. **Release Candidate 1 – VOSAI Learner, Monitor, Console**
6. Community Review RC1, adjustments as required
7. **Release Candidate 2 – VOSAI Learner, Monitor, Console**
8. Community Review RC2, adjustments as required
9. **Release 1 – VOSAI Learner, Monitor, Console**
10. System monitoring, maintenance and automation
11. Development commence of VOSAI Compute v1
12. Final Infrastructure design for Cloud Providers
13. Infrastructure deployment in GCP or chosen Cloud Provider
14. Infrastructure testing in chosen Cloud Provider
15. **Release Candidate 1 – VOSAI Compute**
16. Community Review RC1, adjustments as required
17. Performance tuning and monitoring
18. **Release Candidate 2 – VOSAI Compute**
19. Community Review RC2, adjustments as required
20. **Release 1 – VOSAI Compute**
21. **Release 2 – VOSAI Learner, Monitor, Console**

Year 2

PHASE 3

1. Finalize infrastructure design and software architecture
2. Determine existing blockchain tech for VOSAI Compute (e.g. Golem)
3. Commence development of VOSAI Compute for existing/new tech
4. Deployment to existing/new blockchain tech
5. **Release Candidate 1 – VOSAI Compute**
6. Community Review RC1, adjustments as required
7. **Release Candidate 2 – VOSAI Compute**
8. Community Review RC2, adjustments as required
9. **Release 2 – VOSAI Compute**
10. **Release 3 – VOSAI Learner, Monitor, Console**

Key notes to take away from Year 2 are as follows. Golem has been pre-selected but it has yet to be seen that it can support an infrastructure for a modern day enterprise system let alone VOSAI. We will determine during this year if Golem has progressed enough to support our efforts. Should it be found that it cannot work for our findings then we will selected alternative solutions with the intent of choosing a decentralized solution.

Should no solution exist then we will extend VOSAI Compute to handle for this infrastructure.

Year 3

PHASE 4

1. Redesign / re-architect Phase 0 infrastructure for scale
2. Final design of scaled infrastructure
3. **Release Candidate 1 – Scaled Infrastructure Design**
4. Community Review RC1 – adjustments as required
5. **Release Candidate 2 – Scaled Infrastructure Design**
6. Deploy scaled infrastructure to data center(s)
7. **Release 1 – VOSAI Infrastructure**
8. **Release 3 – VOSAI Compute**
9. **Release 4 – VOSAI Learner, Monitor, Console**
10. Performance tuning, monitoring, automation

The intent of building a custom infrastructure for VOSAI Compute surrounds the details of privacy and security. The purpose of VOSAI Compute is to handle the computational aspects of VOSAI AGI. With that comes a tremendous amount of data flowing in and out of the system(s). The actual data could be images, audio, video, or transcribed conversations between individuals. Therefore, security and privacy is of utmost priority to us and we will ensure at all costs it is kept safe and secure.

With that in mind, if a solution is found to decentralize this information (which we prefer) then such a solution will be devised if at all possible.

For each year, it is expected that the plan may be changed slightly. We do not foresee year 1 changing dramatically. However, year 2 and 3 could be altered based on immediate findings.

We foresee additional years of ongoing development efforts with the VOSAI Compute. As with any software system, it must be maintained, expanded, improved and changed to fit the demands of the future. The previously illustrated roadmap simply guides the reader and sets expectations for what is to come.