No More Vendor Lock-In? The Rise of Sky Computing

BYTEBYTEGO

OCT 5, 2023 · PAID

Cloud computing has unleashed a wave of innovation, powering industry giants like Netflix, Airbnb, and countless others. Yet despite its promise, cloud's full potential remains constrained, locked within vendor-specific silos. What if you could break free from these limitations? Imagine an open sky above the clouds, where your applications can freely soar and shift between clouds at will. This vision is now within reach. The next evolutionary leap beyond cloud computing is showing great promise - welcome to the era of Sky Computing.

In Sky Computing, applications are not bound to any single cloud provider. You can develop cloud-agnostic applications and optimize for performance, cost, latency - on your terms. Initial implementations have already shown the immense benefits, from cutting costs in half to being able to run high-throughput batch workloads across clouds. With open frameworks replacing vendor lock-in, the possibilities are endless. Join us as we explore how Sky Computing works, the promising benefits it unlocks, and how you can start soaring beyond the limits of cloud today.

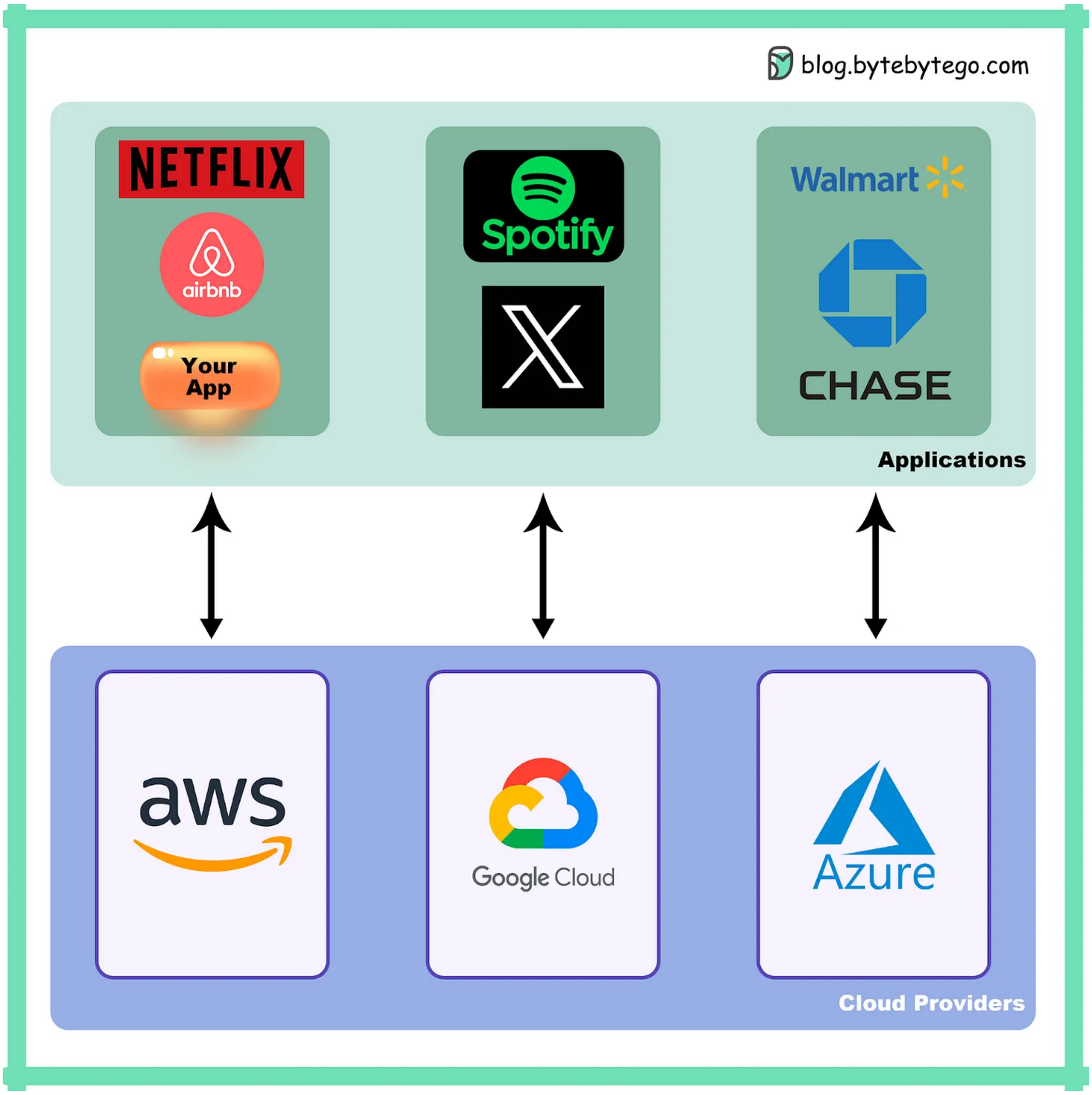
Sky Computing relies on key technical abstractions that intelligently distribute workloads across diverse cloud environments. Initial open source implementations like SkyPilot and SkyPlane demonstrate the viability of these concepts. They provide a seamless multi-cloud experience, while optimizing for cost, performance and latency based on application needs.

In this issue, we will:

- Analyze the incentives for cloud providers and users to participate in this next wave of innovation
- Examine how early pioneering innovations in Sky Computing are already demonstrating promising benefits
- Outline an action plan for you to start experimenting with these architectures - unlocking the promise of an open sky above the clouds

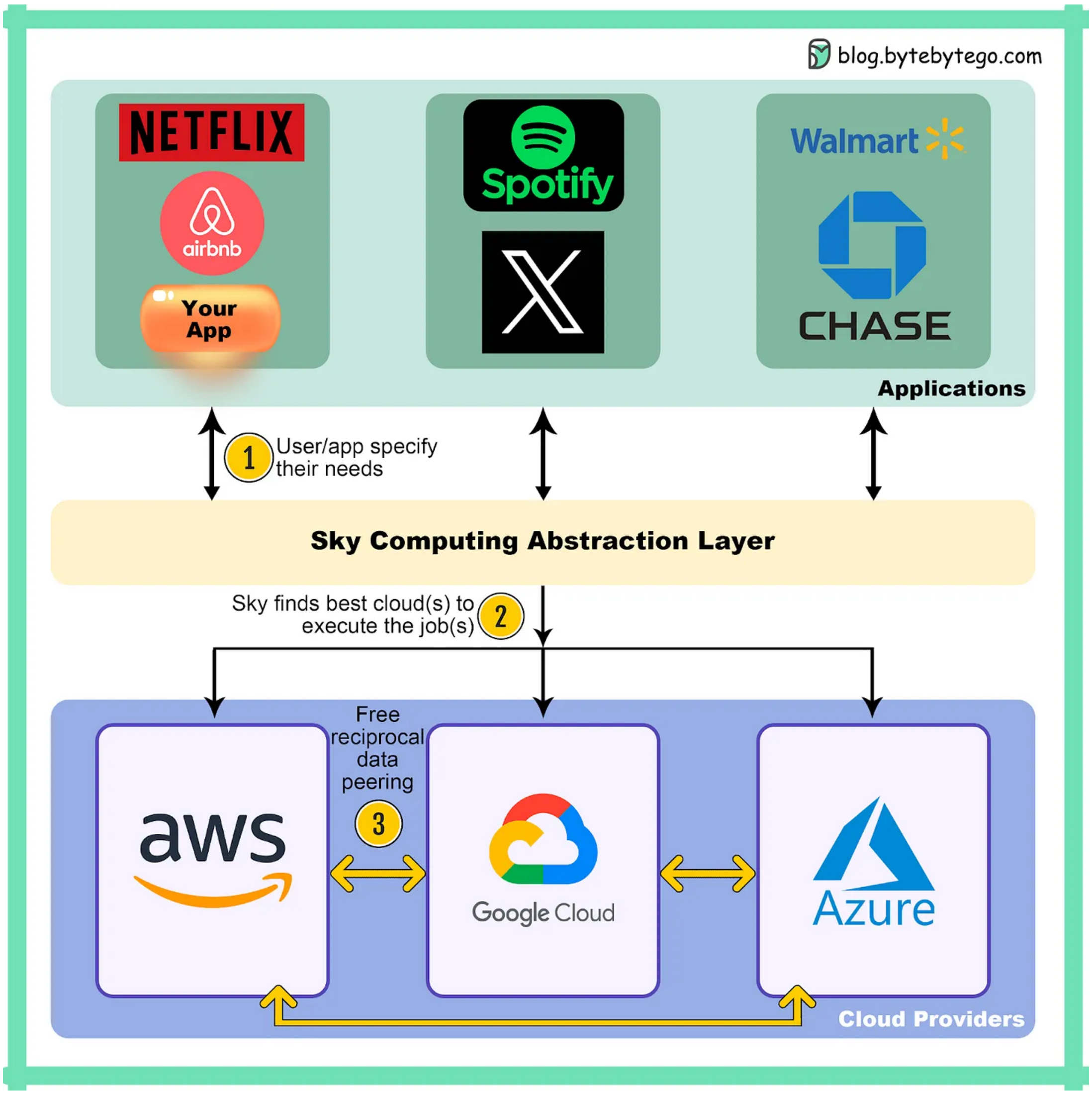
Cloud Computing Landscape Today

The following illustration shows how most applications use cloud computing today.



Most applications today use just one cloud

The following illustration shows what sky computing promises us and its three main enablers—(1) job specification by apps, (2) inter-cloud abstraction layer, and (3) network peering between the clouds.


blog.bytebytego.com
NETFLIX
airbnb
Your App
Spotify
X
Walmart
CHASE
Applications
1 User/app specify their needs
Sky Computing Abstraction Layer
Sky finds best cloud(s) to execute the job(s) 2
Free reciprocal data peering
3
aws
Google Cloud
Azure
Cloud Providers

Sky computing makes applications cloud provider agnostic by using three primary enablers (numbered in the above diagram)

Let's first understand the reasons why many applications depend on one cloud provider.

Why do most applications rely on just one cloud today?

Most cloud applications today utilize one of the major cloud providers. Applications built in this manner are tightly coupled with the underlying cloud provider's API. In a sense, such applications are silos—they cannot easily transition from one cloud to another. Organizations choose cloud providers—often for multiple years—that serve their current and future needs at a favorable cost. The cloud providers frequently leverage such opportunities to lock in customers they believe will be profitable. Here are some common reasons why many applications opt for a single cloud provider:

1. **Simplicity and ease of management:** Using a single cloud provider can streamline the management and operation of an application. Organizations only have to deal with

one set of tools, one billing platform, and one set of resources to manage, which can reduce complexity and overhead. For many organizations such as start-ups, keep technical complexity and associated risks in check is imperative for survival.

2. **Cost considerations:** Adhering to a single cloud provider can make it easier to predict and control costs. Each cloud provider has its pricing model, and managing multiple providers can be more challenging for optimizing cost efficiency.

   Predicting pricing in a cloud provider's spot market can be challenging due to high price and availability fluctuations. Some models even treat such spot markets as stock exchanges and use those models for prediction. According to one study, Apple spent more than $30 million a month in 2019 at AWS for iCloud storage, based on a multi-year contract between the two companies. Large customers like Apple have substantial leverage to negotiate prices.
   Often cloud providers can offer subsidized rates for organizations if they commit to staying with the provider for many years. Such deals are seen as a win-win - the customers get their computing done at a competitive price, and the cloud provider secures important long-term customers. However, once agreed-upon, the customer is locked into that specific provider on the agreed-upon terms.

   For an organization, architecting an application to run on multiple cloud providers can be *n* times as costly (where n is the number of clouds supported) in terms of required technical expertise, training, business deal negotiations, and challenges debugging and fixing issues.

3. **Integration and compatibility:** Cloud providers offer a wide range of services and tools. By using one provider's ecosystem, it's often easier to ensure seamless integration between services, which can improve application performance and reliability.
4. **Service differentiation:** Some cloud providers excel in specific areas or industries. If an application has particular requirements that align well with a provider's strengths (e.g., AWS for ARM-based virtual machines for lower cost, Azure for enterprise integration with Microsoft-specific solutions, GCP for globally distributed, transactional Spanner database, etc.), it may make sense to use that provider.
5. **Data gravity:** Applications that deal with massive data volumes often find it more efficient to keep their data within one cloud provider's ecosystem due to data transfer costs, latency, and compliance considerations.

   Even if a customer did not get a special volume discount, most cloud providers make it harder to move data out (while data ingress rates are often much lower compared to egress). Additionally, artificial throttling limits might be in place to slow egress throughput. For example, a recent study showed that AWS throttled egress data at 5 Gbps.

6. **Technical expertise:** Teams may have specialized expertise with a specific cloud provider's services and tools, making it more efficient to continue using that provider for new projects.

Some customers operate on very slim profit margins and are highly cost-sensitive. So much so that at times, some applications choose the cheapest data center among many while ignoring redundancy. For example, AWS's North Virginia's datacenter is the most affordable for some services, so they only deploy there. However, there is an inherent risk of failure when a data center-wide outage occurs, as has happened in the past. Still, many customers seem to ignore such risks because cloud providers have become very reliable over the years. Some organizations are willing to take this gamble.

> An aside: Artificial fault injection to keep the clients honest to the SLA
>
> As an interesting aside, Google reports artificially inducing errors to its Chubby locking service so internal customers don't start believing Chubby has 100% availability. Such practices are not usually employed for public facing services.

7. **Vendor Lock-In:** Inertia exists in computing as well—as an organization's footprint grows within a cloud provider, it becomes increasingly difficult to switch providers.

   While vendor lock-in can be a concern with a single cloud provider, some organizations prioritize short-term efficiency over long-term flexibility. They may opt to address potential lock-in issues later, once their application matures. However, that moment may never arrive as applications and workloads can become further intertwined with proprietary APIs over time.

> An aside: Multi-cloud and Sky computing
>
> Some organizations use more than one cloud, either for distinct, unrelated applications or different business units. This strategy is also called multi-cloud. For our discussion, we will consider an application's

```
ability to freely transition across any cloud as either multi-cloud or
sky computing.
```

Ultimately, the decision to use one or multiple cloud providers should be based on the application’s specific needs and goals, as well as the organization's tolerance for complexity and risk.

Why would anyone want Sky Computing?

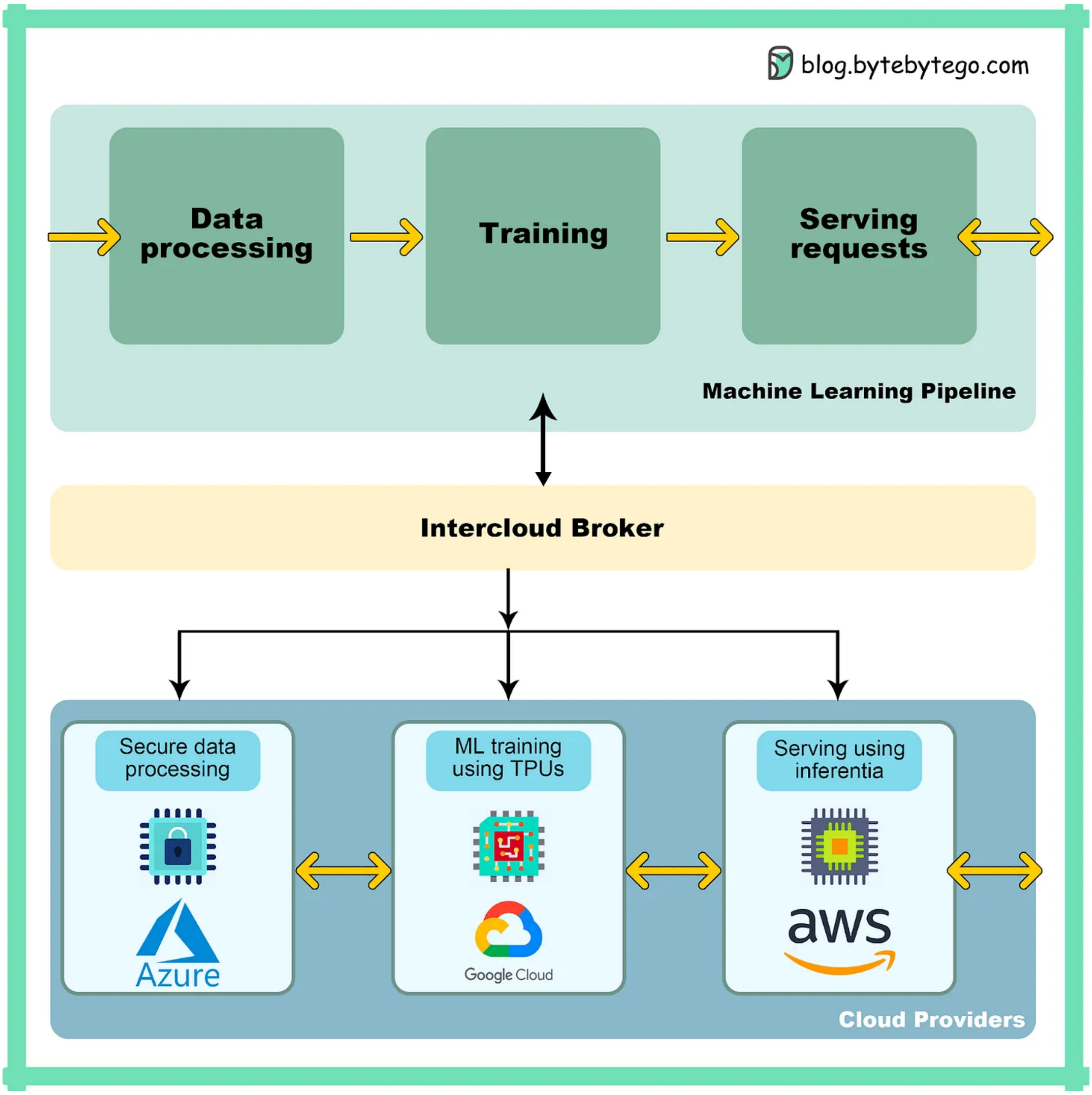
The obvious question is: why would anyone want Sky Computing to happen? What incentives exist for stakeholders?

Multi-cloud architectures, where applications utilize multiple cloud providers simultaneously, offer several benefits depending on an organization’s specific needs and goals. Here are some of the key incentives and benefits of multi-cloud apps:

1. **Redundancy and high availability:** Spreading workloads across multiple cloud providers and regions enhances the application redundancy and high availability. If one cloud provider experiences an outage, the application can seamlessly failover to another, minimizing downtime and ensuring business continuity.
Recent events demonstrate that cloud provider outages happen. Applications that dynamically move workloads between providers can mask an outage.
2. **Risk mitigation:** Multi-cloud strategies mitigate vendor lock-in risks. Organizations avoid over-reliance on a single provider, reducing vulnerability to changes in pricing, service quality, or strategic shifts by a single provider.
3. **Optimization for specific services:** Different cloud providers excel in different areas. By using multiple providers, organizations can choose the best services for each sub-task within their application. For example, one provider may offer superior machine learning solutions while another excels at IoT or analytics services.

There will always be differentiated services among cloud providers. Later we will examine an example ML pipeline where each stage runs on a different cloud.

The following illustration shows a workload with three major phases—data processing, ML training, and client serving. We might figure that Azure’s secure processing service is the best choice for the first phase (data processing), GCP’s TPU processors can give us superior performance per dollar for the second phase (ML training), while we can serve economically using AWS Inferentia service for the third phase (client serving). The output of one phase provides the input to the next by utilizing egress routes. Prototyping confirmed that this approach reduced costs up to 80% and latency up to 60% compared to using a single cloud.



Data processing, ML training, and data serving using three cloud providers.

4. **Cost optimization:** Multi-cloud deployments allow organizations to optimize costs by selecting the most cost-effective cloud provider for each workload. This can lead to potential cost savings, as different providers may offer better pricing for specific services or regions.

```
An aside: Sky computing / Multi-cloud provides economic benefits and
accelerates innovation.
```

> Initial studies (see Skypilot and Skyplane) suggest substantial cost savings using multiple cloud providers, even after taking egress fees into account.

Not every organization gets volume discounts—those are reserved for a select few. For others, it often makes sense to move services between cloud providers. Overall costs can still be lower, even after factoring in data egress fees. We will examine those results later in this newsletter.

Spot market price and availability fluctuations also differ among cloud providers—at different times spot markets of different cloud providers might offer better cost performance.

5. **Compliance and data sovereignty:** Compliance requirements vary by region and industry. Multi-cloud strategies enable organizations to place data and workloads in specific geographic regions to meet compliance and data sovereignty requirements. For example, some countries in the European Union might demand that data placement and processing of its citizens must happen inside the physical boundaries of the country. A single cloud provider might not have its data center in a specific country. Using another cloud provider for that country can resolve the issue.
6. **Performance optimization:** Geographical distribution across cloud providers can improve application performance for users in different parts of the world. Organizations can strategically place resources closer to their end-users to reduce latency and enhance user experience.
7. **Disaster recovery:** Multi-cloud architectures can simplify disaster recovery planning. If a disaster affects one provider, failover to another ensures availability.
8. **Innovation and best-of-breed solutions:** Organizations can leverage innovations and best-of-breed solutions from multiple providers, staying competitive and capitalizing on emerging technologies.
9. **Flexibility and scalability:** Multi-cloud environments offer greater flexibility and scalability. Organizations can scale resources as needed, avoiding the constraints of a single provider's resource limits.
By leveraging sky computing, a customer has the ability to pool resources from multiple cloud providers to get their work done. For example, with GPU shortages, for many ML training sessions, we could combine GPUs from multiple cloud providers to get our work done instead of waiting for the resources to be available.
10. **Negotiating leverage:** Using multiple cloud providers may provide negotiating leverage with pricing and service agreements. Competition between providers can

lead to better terms for customers.

The decision to adopt a multi-cloud approach should align with an organization's specific goals, requirements, and resources. It's not necessarily the right choice for every application, but it can provide significant advantages in the right circumstances.

Incentives for cloud providers

Cloud providers have incentives to embrace the multi-cloud paradigm, as it can increase customer engagement and revenue opportunities. Here are some of the key incentives and benefits for cloud providers:

1. **Increased market share:** When clients adopt a multi-cloud strategy, they are more likely to use services from multiple cloud providers. Cloud providers can capture a larger share of the market by offering services that cater to different aspects of the clients' multi-cloud architecture.
2. **Revenue diversification:** By providing services that support multi-cloud deployments, cloud providers can diversify their revenue streams. This reduces their reliance on any one client or market segment. This makes them more resilient to market fluctuations.
3. **Cross-selling and up-selling:** Cloud providers can cross-sell and up-sell additional services to clients who embrace a multi-cloud strategy. For example, they can offer tools and services for managing multi-cloud environments, security solutions, and data integration services to clients. For example, GCP's Anthos and Azure's Arc projects are a step in that direction.
4. **Partnerships and ecosystem expansion:** Cloud providers can establish partnerships with other cloud providers or technology vendors to create a more extensive ecosystem. These partnerships can lead to joint marketing and revenue-sharing opportunities. For example, GCP's Anthos collaborates with VMware and HP to enable Anthos across providers.
5. **Customization and flexibility:** Offering customizable solutions that cater to a client's multi-cloud needs allows cloud providers to differentiate themselves in the market. Clients often seek providers who can adapt to their unique requirements.
6. **Resource optimization:** Cloud providers can optimize their infrastructure and resource allocation based on client demand for multi-cloud solutions. This optimization can lead to cost savings and better resource utilization.
7. **Innovation and competitive advantage:** Cloud providers that invest in multi-cloud capabilities and technologies can gain a competitive advantage by staying at the

forefront of innovation in this evolving space. This can attract more clients looking for cutting-edge solutions.

It's important to note that while there are benefits for cloud providers, embracing the multi-cloud paradigm also presents challenges, including increased competition and the need to ensure interoperability with other cloud providers. To be successful in this space, cloud providers must continually adapt their offerings to meet the evolving needs of multi-cloud clients and maintain a strong focus on customer satisfaction.

In the ever-evolving cloud market, no single player is too currently big (AWS's share 32%, Azure's 22%, and GCP's 11%.), and smaller providers will be more willing to embrace sky computing than the large players. Once again projects like Anthos and Arc from GCP and Azure are an example for such a phenomenon.

An aside: John McCarthy's vision of computing becoming a public utility.

"Computing may someday be organized as a public utility just as the telephone system is a public utility, … Each subscriber needs to pay only for the capacity he actually uses, but he has access to all programming languages characteristic of a very large system … Certain subscribers might offer service to other subscribers … The computer utility could become the basis of a new and important industry."

A quote by Professor John McCarthy at MIT's centennial celebration in 1961

John McCarthy (the Turing award winner of 1971 for his contributions to AI) had the vision of computing becoming a public utility where customers could use as much of it as needed and pay only for the time they used the resources. The invention of public cloud circa 2006 popularized part of McCarth's vision where clouds provide huge resources where customer pay-per-use. However, computing becoming a public utility has yet to be realized. Just like we can plug our devices to wall sockets without worrying which electric company produced the power, we need applications that could use the infra and services without worrying which cloud provider is offering it, as long as that offering meets customer's needs.

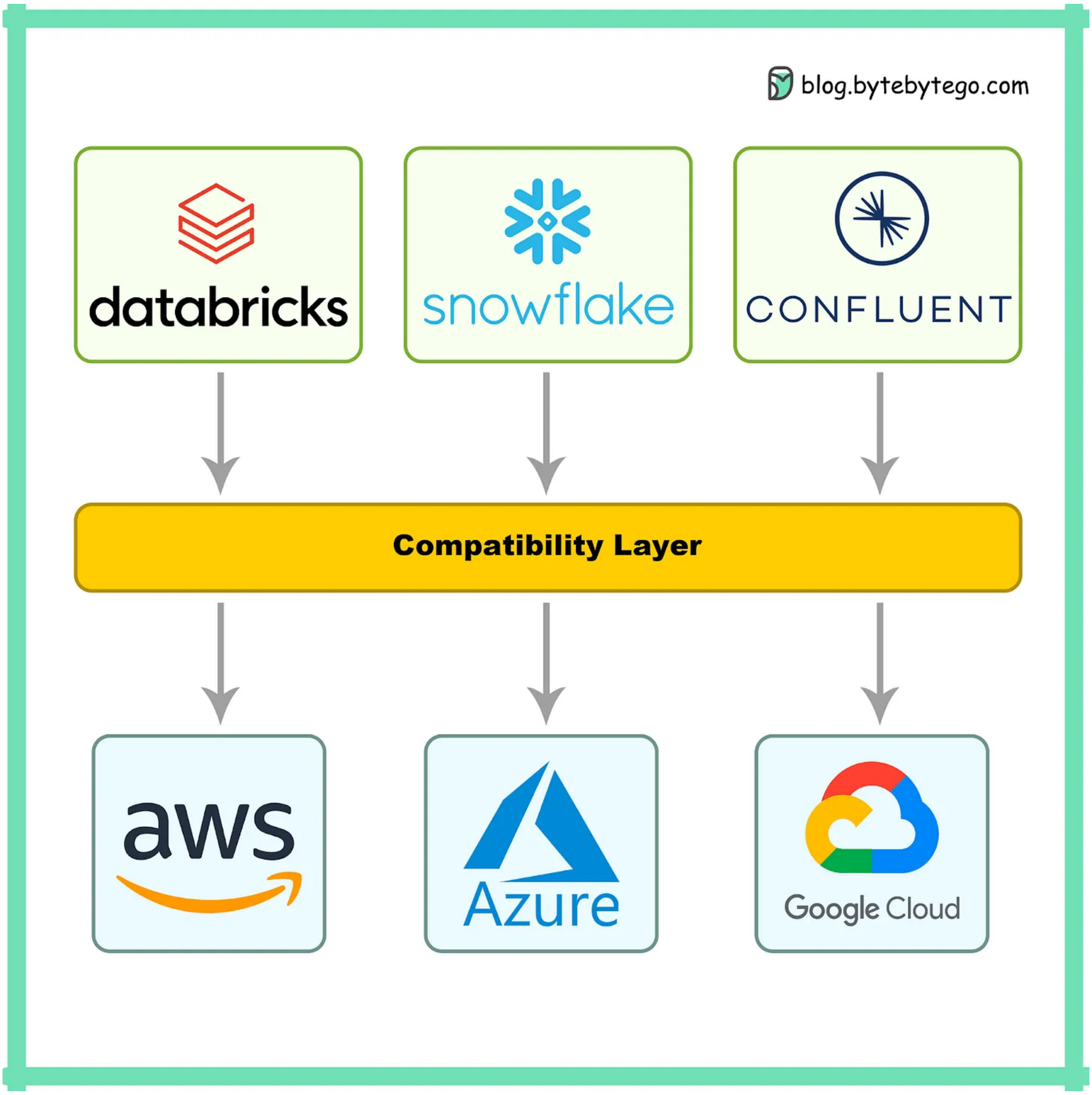
Cloud providers might be averse to the idea of them becoming easily replaceable by anyone else. Each large cloud tries hard to differentiate itself by offering something unique—GCP has TPUs (Tensor Processing Unit) that have better cost-performance for ML training, AWS has low cost, ARM-based virtual machines and AWS Inferentia for economical deep learning inference, and Azure has services like secure enclaves for processing sensitive data.

Let’s now see how organizations can architect sky computing.

How we will realize Sky Computing

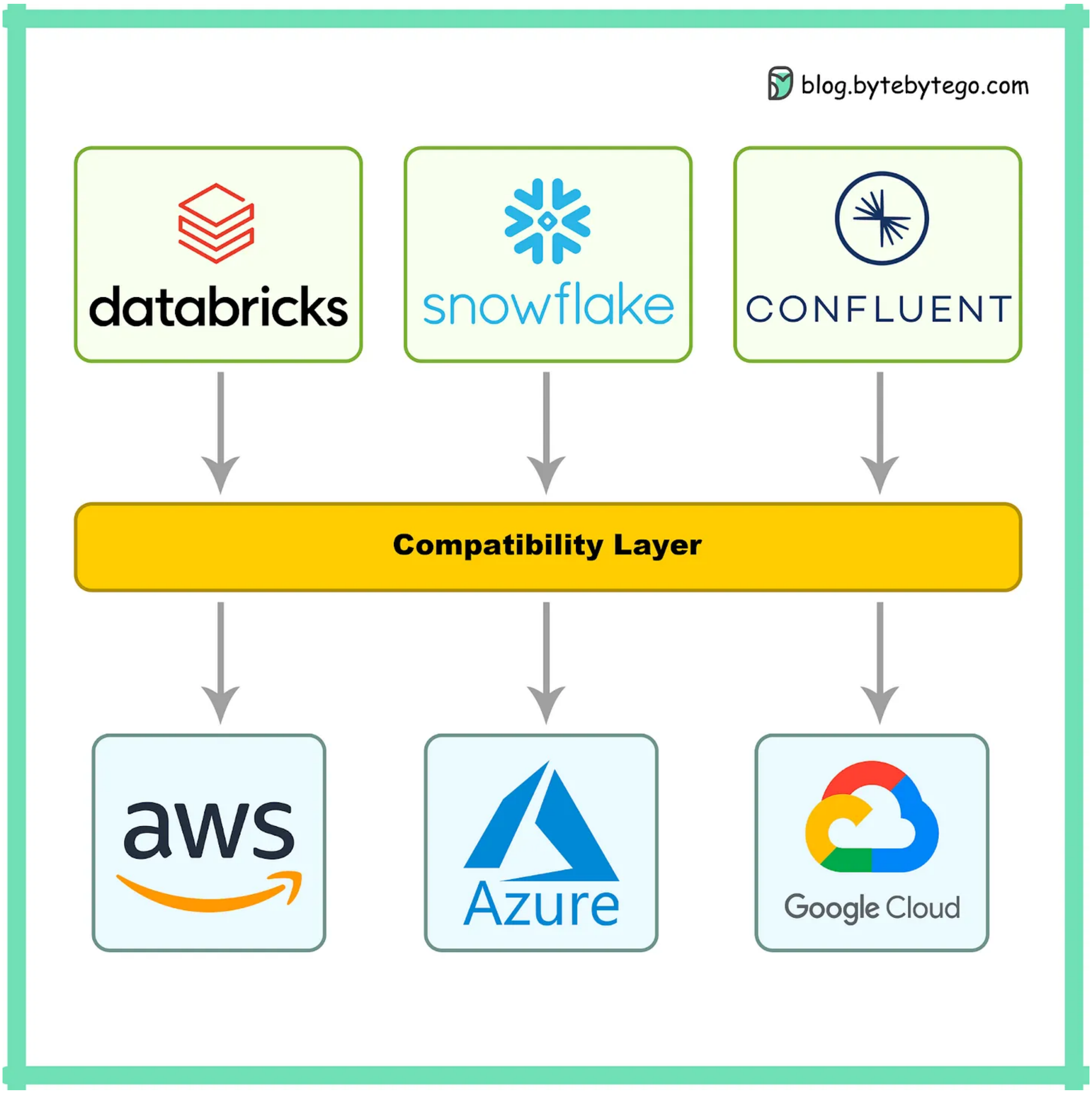
Now let's put our designer hats on. How would we architect Sky Computing to meet our stated goals? There are a few options to meet our stated goals, though only one good choice emerges.

1. Porting each application to every cloud provider is an impractical m x n solution. With m providers and n applications, each application must conform to every cloud's unique API. For example, Databricks reportedly required many person-years of effort just to port their application to Azure. Only a few large organizations could afford this for a handful of major cloud providers.



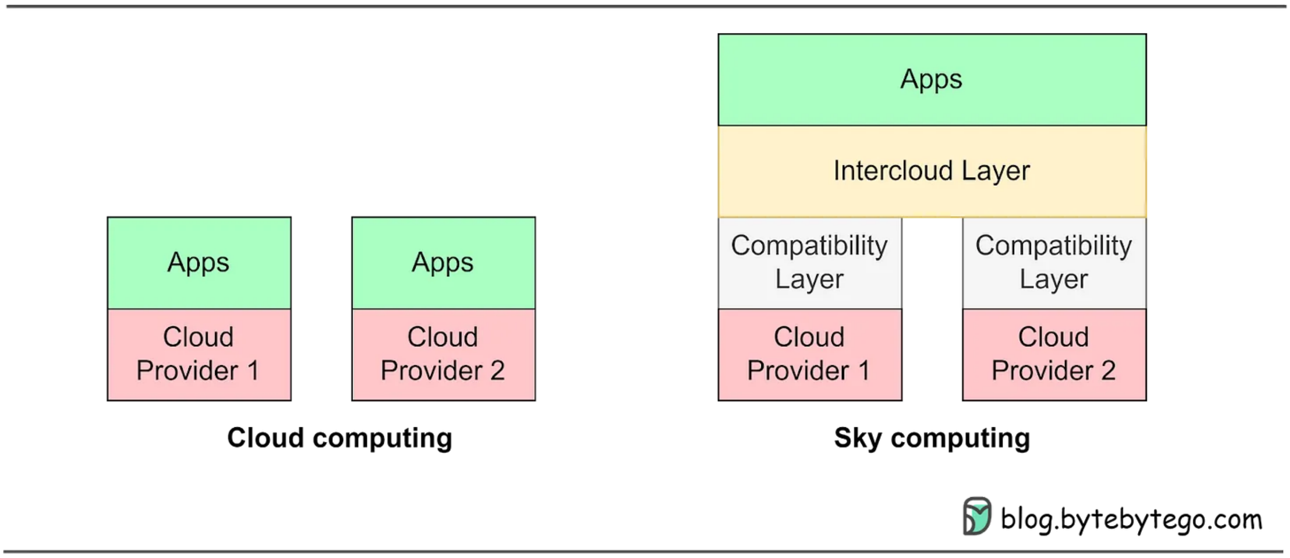
Porting each app to every cloud provider. Databricks took tens of person-years to port it to Azure.

2. Standardizing a common API implemented by all providers also has challenges. Cloud services are too diverse to find consensus on a single API. Nimbus provided a unified API for High Performance Computing applications but doesn't generalize beyond this niche use case. Unlike networking protocols which converged on packet transport, cloud providers don't share a simple common denominator.
3. A better approach is an open source compatibility layer abstraction. This understands a user's workload needs, plans job execution across clouds to meet goals, and executes the workload. Multiple compatibility layers can exist, each optimized for use cases like machine learning or big data processing. This creates an m+n problem space: each application leverages the compatibility layer's API, which in turn utilizes the necessary underlying cloud APIs. Kubernetes exemplifies this strategy - portable across clouds while simplifying application portability.



A better strategy: an abstraction to break tight coupling between the applications and clouds

We can go a step further. In the above scheme an application still needs to explicitly use a compatibility layer for the required cloud provider and needs to manage for any cross-cloud data movement. To generalize this idea, three key components are required—an intercloud layer to take apps' needs on one side and translating them to compatibility layers for specific clouds, compatibility layers across providers, and ability to move data around.



Generalized abstraction layer consisting of intercloud and compatibility layers

Let's discuss the three enablers of sky computing in more detail.

Compatibility layer

This layer utilizes a software stack that is often available on all major clouds. Such software makes it viable to actually execute customers' jobs. Fortunately there are substantial open source products across all the layers of the stack. For example:

1. Linux for the operating system
2. Kubernetes for cluster orchestration
3. Docker for application packaging
4. Apache Spark for data processing
5. MySQL, Postgres for databases

Anyone can run and manage these systems in all the major clouds, and on premise data centers if need be.

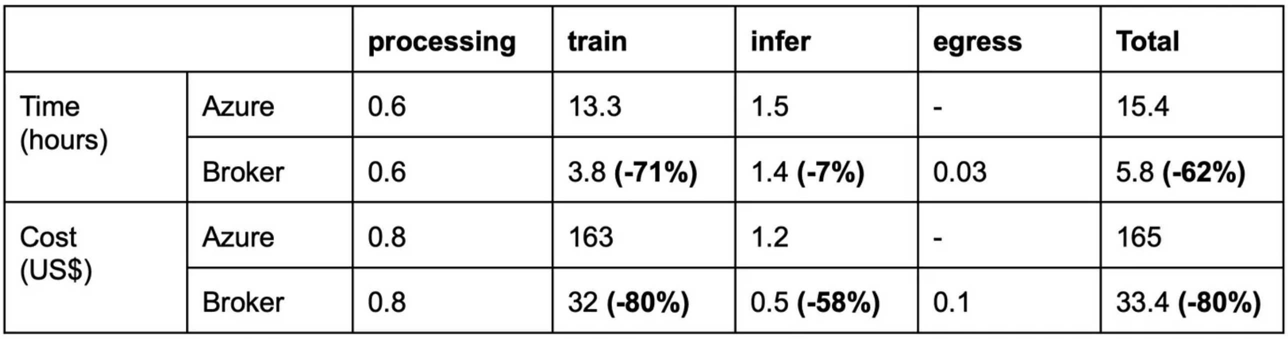
Google's Anthos and Azure's Arc can be considered proprietary forms of such a compatibility layer.

Intercloud broker

The purpose of an intercloud broker is to collect applications' requirements in a standardized way. It has been shown that many computations can be described as a workflow or a DAG (directed acyclic graph) of steps. The intercloud broker takes the goals (e.g. lowest cost or latency) and solves an optimization problem to meet them by utilizing different services by different clouds. This plan is similar to what a database engine generates for a SQL query. The broker then executes the workflow using the compatibility layers, possibly moving data between clouds, and returns the result.

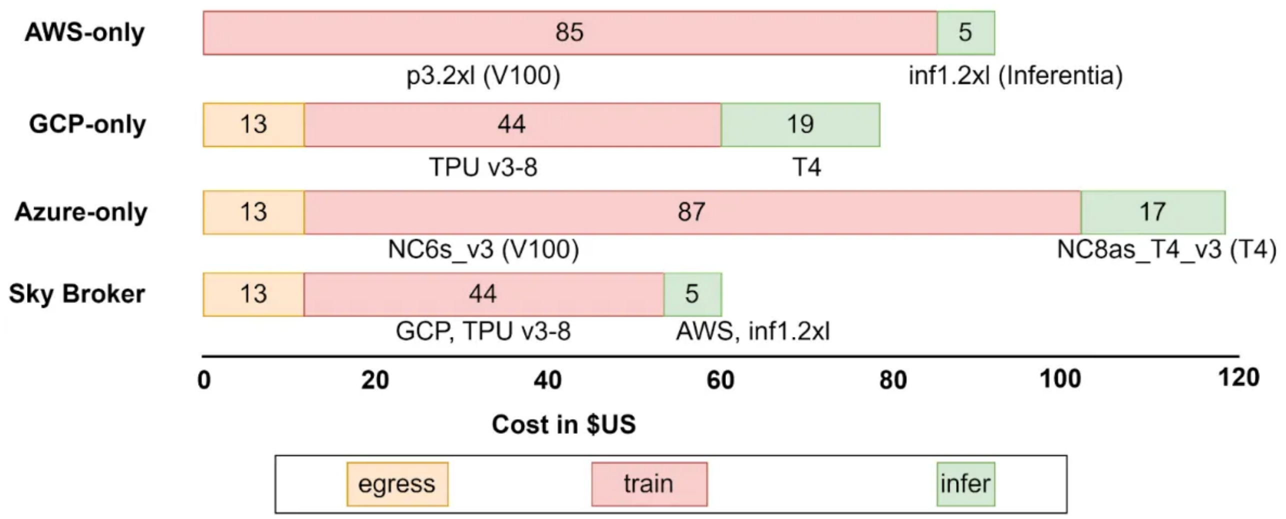
SkyPilot is an inter-cloud broker and compatibility layer implementation. As of October 2023, SkyPilot can run workloads like LLMs, AI, and batch jobs on major clouds including AWS, GCP, Azure, IBM Cloud, Oracle Cloud, Cloudflare, Lambda cloud, Samsung, and Kubernetes clusters. SkyPilot maximizes cost savings, GPU availability, and managed execution. SkyPilot supports existing GPU, TPU, and CPU workloads without code changes. A quick start guide is available to get started with SkyPilot. An example is available where an LLM model is trained on different clouds.

As we saw earlier a workflow (data processing, ML training, and inference and data serving) where different phases were run on different clouds were actually produced by the SkyPilot. This saved 80% on cost and completed 62% faster versus using just Azure. Such projects have a multi-million dollar budget and a cost saving of 80% is a huge deal. Saving time also translates to improved customer satisfaction and faster innovation speed. This demonstrates the promise of sky computing for customers.



		processing	**train**	**infer**	**egress**	**Total**
Time (hours)	Azure	0.6	13.3	1.5	-	15.4
	Broker	0.6	3.8 **(-71%)**	1.4 **(-7%)**	0.03	5.8 **(-62%)**
Cost (US$)	Azure	0.8	163	1.2	-	165
	Broker	0.8	32 **(-80%)**	0.5 **(-58%)**	0.1	33.4 **(-80%)**

The initial study of SkyPilot also compared costs across cloud providers for a vision pipeline. The following illustration shows the results. It is clear that the intercloud broker beats all the cases where we run the workload on just one cloud.



Source: NSDI 2023 SkyPilot. AWS does not have an egress cost because data was placed there originally.

An important point to note is that SkyPilot worked without cooperation from cloud providers by using their publicly available services.

Peering between cloud providers

The above prototype of the broker incurred egress costs when data was moved from one cloud provider to the next. Each cloud provider has different egress characteristics, so they need to be thoroughly studied. While we said earlier that free, reciprocal data peering between cloud providers will super-charge sky computing, we are far from it. Meanwhile we need another solution to reduce the egress costs.

An initial investigation for fast and economical cross-cloud egress communication was recently conducted by SkyPlane. They used overlay paths for data transfer instead of direct paths. As we argued that free data exchange between the clouds with reciprocal peering arrangements will expedite sky computing. However, we are not there yet. The question is this: *Can we exchange data economically across clouds?* SkyPlane answers that by providing a plan to copy data from source to destination via intermediate cloud zones, meeting overall cost and latency goals. SkyPlane also provides a tool to carry out the actual copying as per its plan.

SkyPlane is a bulk data copier across clouds, or across regions of the same cloud. Its customers can specify needs, for example: *Copy x terabytes of data from AWS west1 to GCP's asia-south1 consuming no more than $y*. Or, *copy x terabytes as fast as possible from some source to some destination with a cost ceiling of $z*. As of October 2023, SkyPlane supports copying between AWS, Azure, GCP, and IBM Cloud.

SkyPlane can:

- Copy between object stores within a cloud provider (e.g. AWS us-east-1 to AWS us-west-2)
- Copy between object stores across multiple cloud providers (e.g. AWS us-east-1 to GCP us-central1)
- Copy between local storage and cloud object stores (experimental support as of October 2023)

A quickstart guide is available to try SkyPlane.

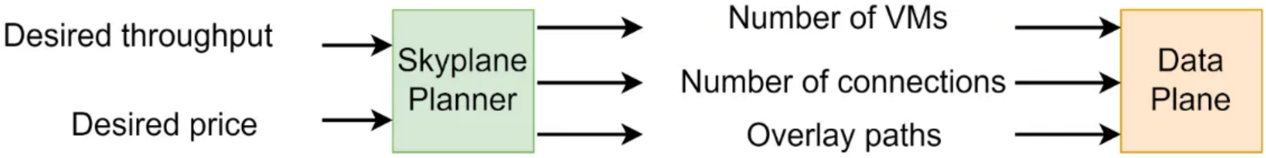
```
An aside: GCP’s cross-cloud interconnect

What we want is free, bilateral peering agreements for data movement
between clouds. Google in Next 2023 announced across-cloud networking.
Currently this service is not free and needs some manual steps to start.
However, we believe cost will reduce over time, especially due to
SkyPlane’s successful feasibility of bulk data transfer using overlay
paths.
```

All major clouds provide bulk data transfer tools. For example, AWS DataSync, GCP Storage Transfer Service, and Azure AzCopy. Google’s cross-cloud internetconnect allows dedicated, high-speed connections between major cloud providers. As of September 2023, it costs $18.05 per hour per 100-Gbps circuit.

So the main question of interest is: How can we optimize network cost and throughput for cloud bulk transfers? SkyPlane’s main findings are:

1. Most cloud providers throttle the egress traffic.
2. Overlay paths, possibly using indirect routes via intermediate regions, can provide much better throughput and cost.

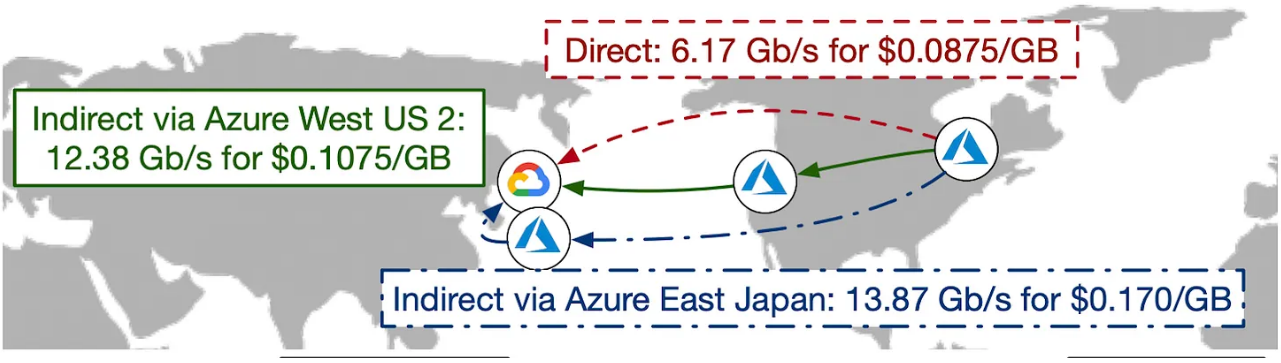
Customers provide desired throughput and price goals to the Skyplane planner. It utilizes gathered information about different cloud provider’s data centers, data rates between regions, and data movement cost. Using an algorithmic technique called linear programming, Skyplanner solves for the given constraints and provides an overlay path, number of connections, and number of virtual machines to use to achieve the desired goal.



Finding an economical, low latency route for bulk data copy across cloud providers.

Throughput on a direct path from Azure's central Canada region to GCP's asia-northeast 1 region was recorded at 6.2 Gbps. Skyplane found an overlay route via an intermediate hop at Azure's US west2 with a throughput of 12.4 Gbps (about 2X speed up!). There is a tradeoff between price and throughput. Cloud providers charge egress per overlay hop. One of the jobs of Skyplane is to achieve the customer's stated cost ceiling for best throughput. While in a traditional network, the bandwidth between two hops does not fluctuate much, for cloud-overlay hops, we can often get more throughput by employing more virtual machines to send data in parallel.

Skyplane prototype outperforms AWS DataSync by up to 4.6X and GCP Storage Transfer by up to 5.0X.


Direct: 6.17 Gb/s for $0.0875/GB
Indirect via Azure West US 2: 12.38 Gb/s for $0.1075/GB
Indirect via Azure East Japan: 13.87 Gb/s for $0.170/GB

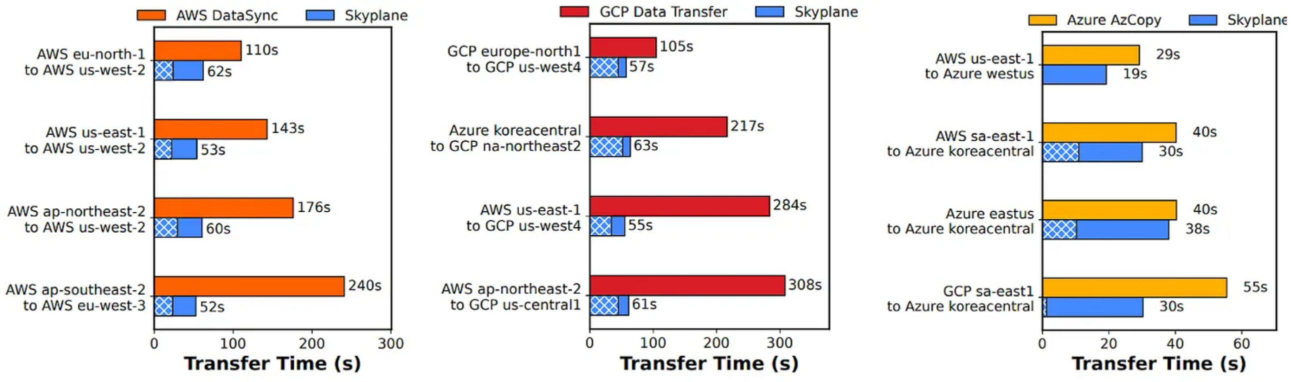
Source: Skyplane study from NSDI 2023

The Skyplane study also found that inter-cloud links are consistently slower than intra-cloud links for network routes from Azure and GCP. GCP throttles inter-cloud egress to 7 Gbps, while AWS throttles all egress traffic to 5 Gbps (even between their own data centers). Azure showed no set throttling limit for inter-cloud traffic (the only constraint was NIC card capacity).

Skyplane needs to find an optimal path from a source to a destination such that an application's throughput, latency, and cost constraints are satisfied. However, the input to that process is a matrix stating maximum achievable throughput, latency, and date movement costs across all pairs of cloud regions. While data movement cost can be relatively easily determined, throughput and latency (especially across cloud providers) need to be measured periodically. This periodic measurement is necessary because available throughput and latency dynamically change due to factors such as shifting traffic conditions.

Skyplane constructs a costs, throughput, and latency matrix for all region pairs across all clouds. The current Skyplane implementation expends $4000 to compute this matrix. Data from major cloud providers shows that throughput and latency remain relatively stable over a few days. Skyplane uses synthetic data to derive throughput and latency estimates. As an optimization, they plan to incorporate actual application data readings from routes, hence reducing the need for synthetic data to update the matrix.

For inter-cloud and across-cloud workloads, Skyplane overlays also have better latency, more so when regions are physically farther apart. Since Azure showed no hard service limit on egress throughput, its difference is less than the other two (AWS and GCP). The illustration below shows transfer latencies of each cloud provider's own bulk data tool versus Skyplane. Skyplane beats AWS and GCP by large margins and can easily bypass their egress throttling



(a) AWS DataSync comparison (b) GCP Storage Transfer comparison (c) Azure AzCopy comparison

Data source: NSDI 2023 paper on Skyplane

Skyplane study shows that economical, high-speed egress data movement is possible across different clouds. Additionally overcoming artificial throttle limits by some of the cloud providers is also possible.

Innovation strategy by embracing sky computing

We are encouraged that sky computing is promising economically and can supercharge technical innovation. Here is a roadmap that we can follow going forward.

1. Explore open source tools like Skypilot and Skyplane for firsthand sky computing experience.

2. For batch workloads, use skypilot cloud broker and skyplane data movement software. Their source code is freely available. We can extend the software as needed.
3. Initiate new open source projects for new intercloud brokers in other domains you need and collaboratively build them.

Conclusion

Sky computing is the next wave of evolution for cloud computing where customers can write provider-agnostic applications easily, economically, and without delving deep into the proprietary APIs of each cloud provider. Sky computing can supercharge the speed of innovation along with cost benefits. The sky computing movement is still in its infancy and lots needs to be done. There are numerous opportunities for new business models and architectural designs.

The initial prototypes of Skypilot and Skyplane show great promise. They demonstrated cost benefits of up to 80% and time reductions up to 62%. They also showed that cooperation from the cloud providers is not a prerequisite for sky computing to happen.

102 Likes · 2 Restacks

1 Comment

Write a comment...

1 more comment...