SKY COMPUTING WHITEPAPER

New future for cloud computing

Safi

- Previous research
- Services in compatibility set;
- Architecture (how does it work?)
- Example; Cost (is it worth it ?)

- Reasons to use sky;
- Conjectures / Assumptions / Predictions / Rationals about the proposal
- Impact Opportunities / Advantages

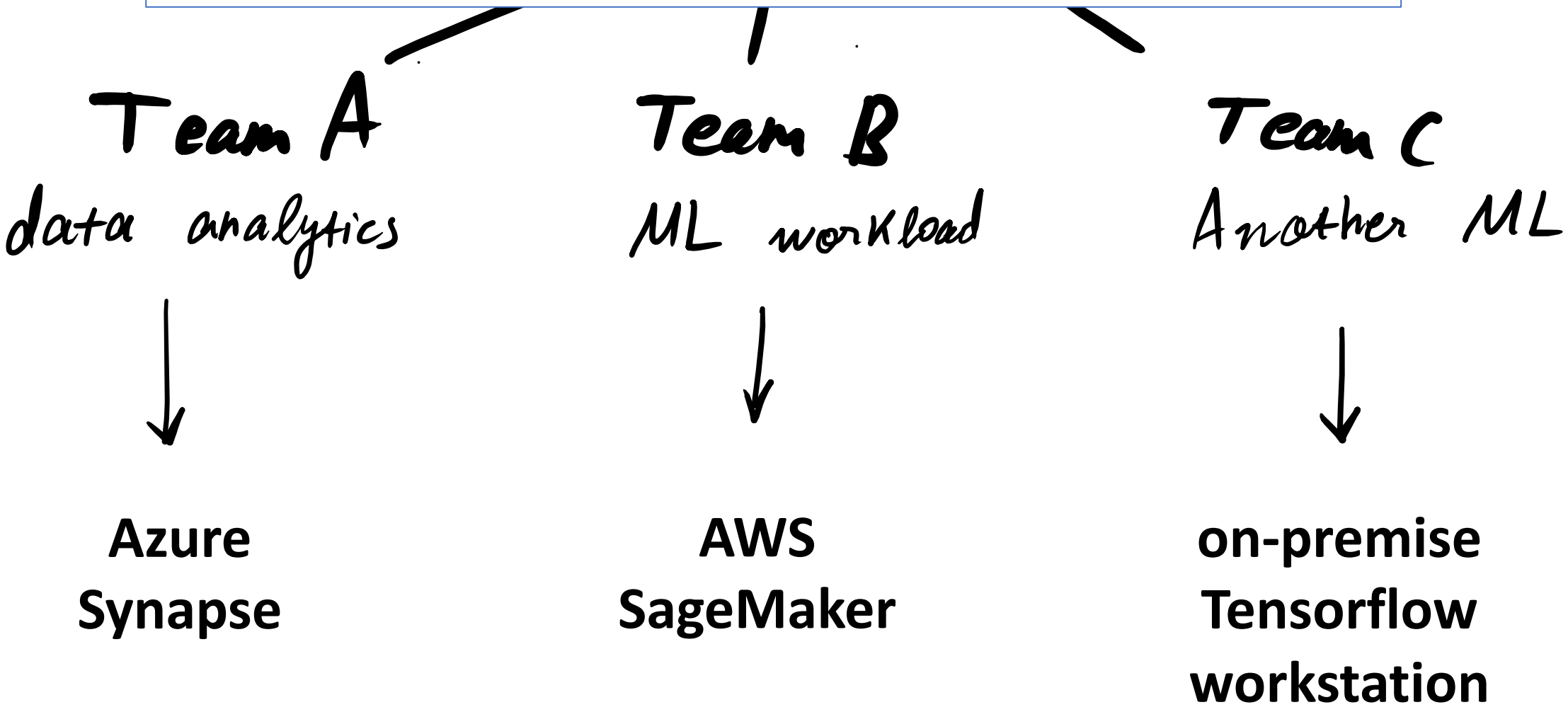
- Objections / Criticism
- Risks

- SkyPilot & SkyPlane, and other projects
- Summary

Multi-cloud or not?

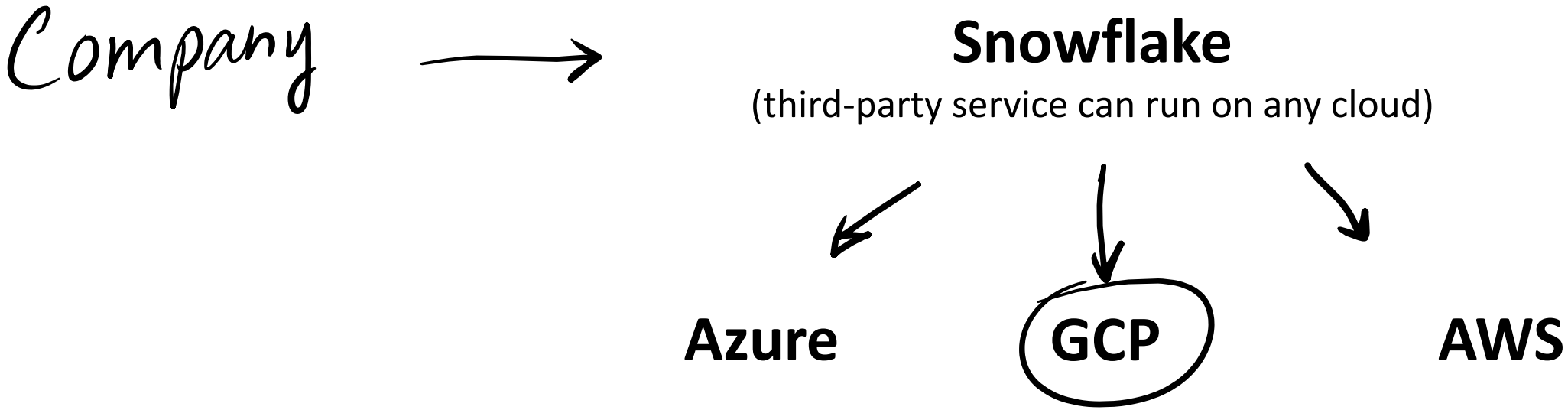
Case 1:

Partitioned multicloud
each workload on a single cloud



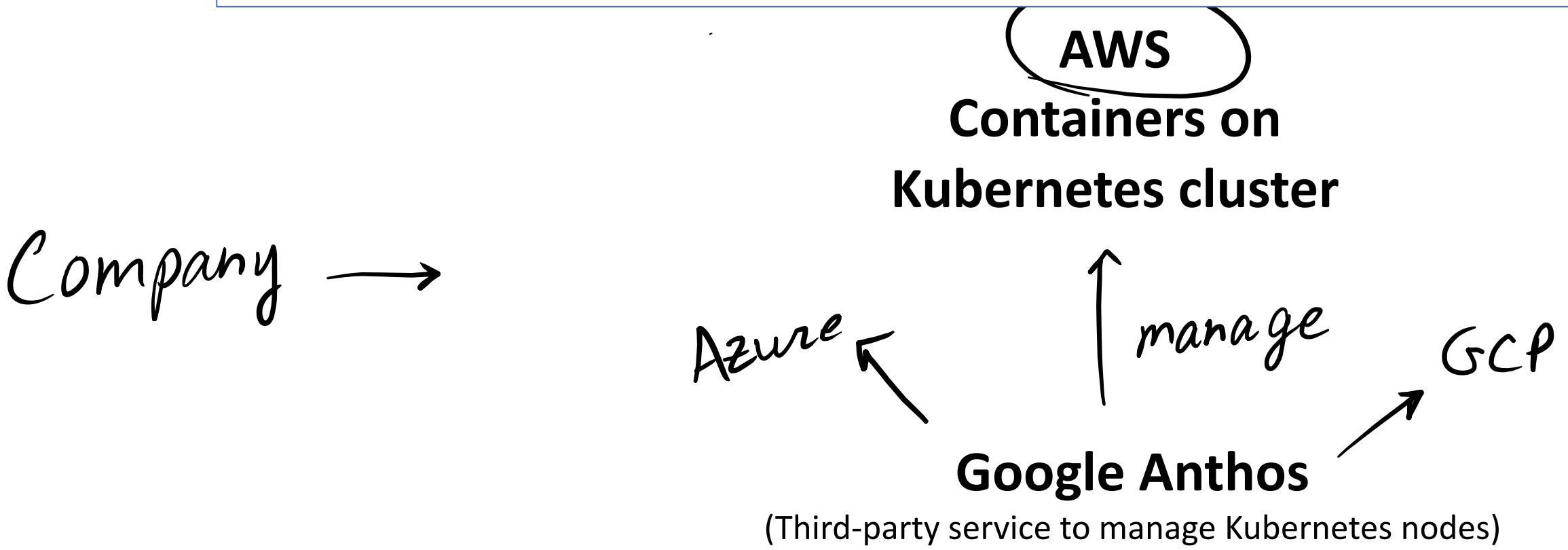
Case 2:

Portable multicloud
Same app can be deployed on any cloud



Case 3:

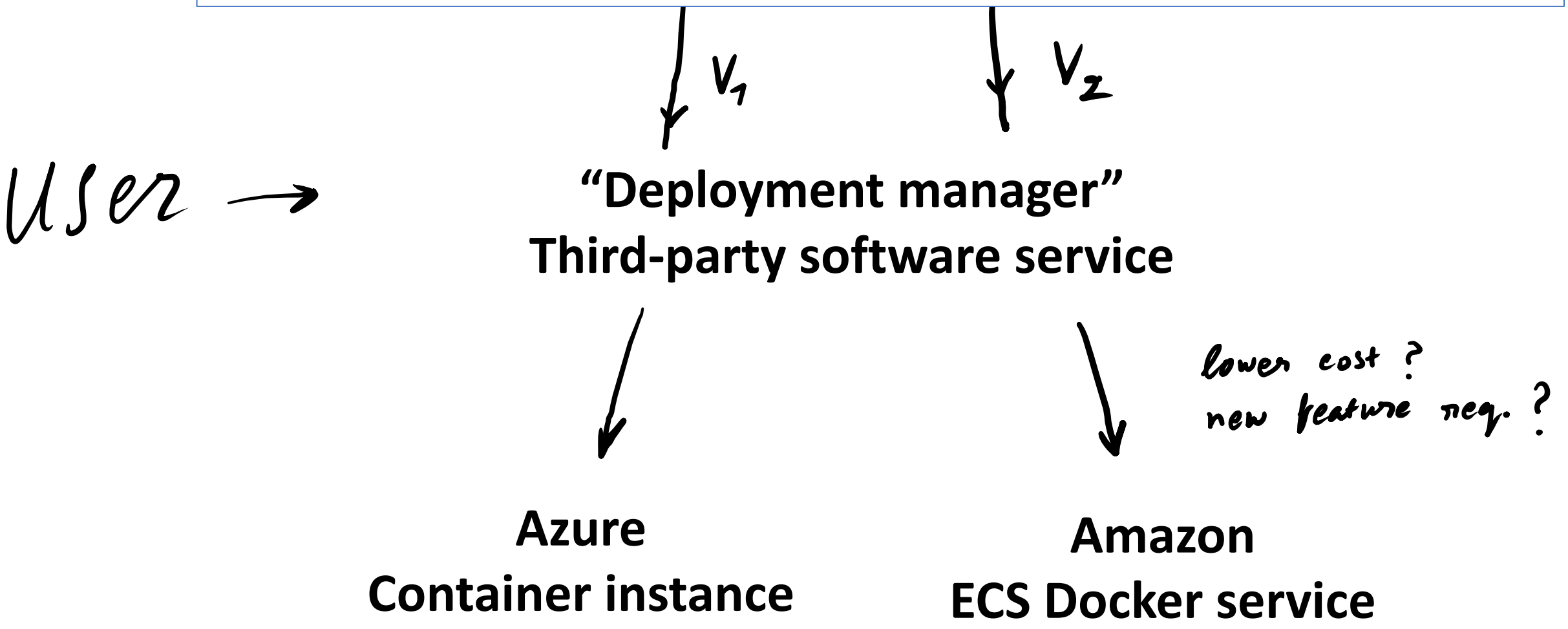
Portable multicloud
Same workload can potentially run on any cloud



Case 4:

Transparent multicloud

Same workload transparently run on any cloud



Types of Multi-clouds

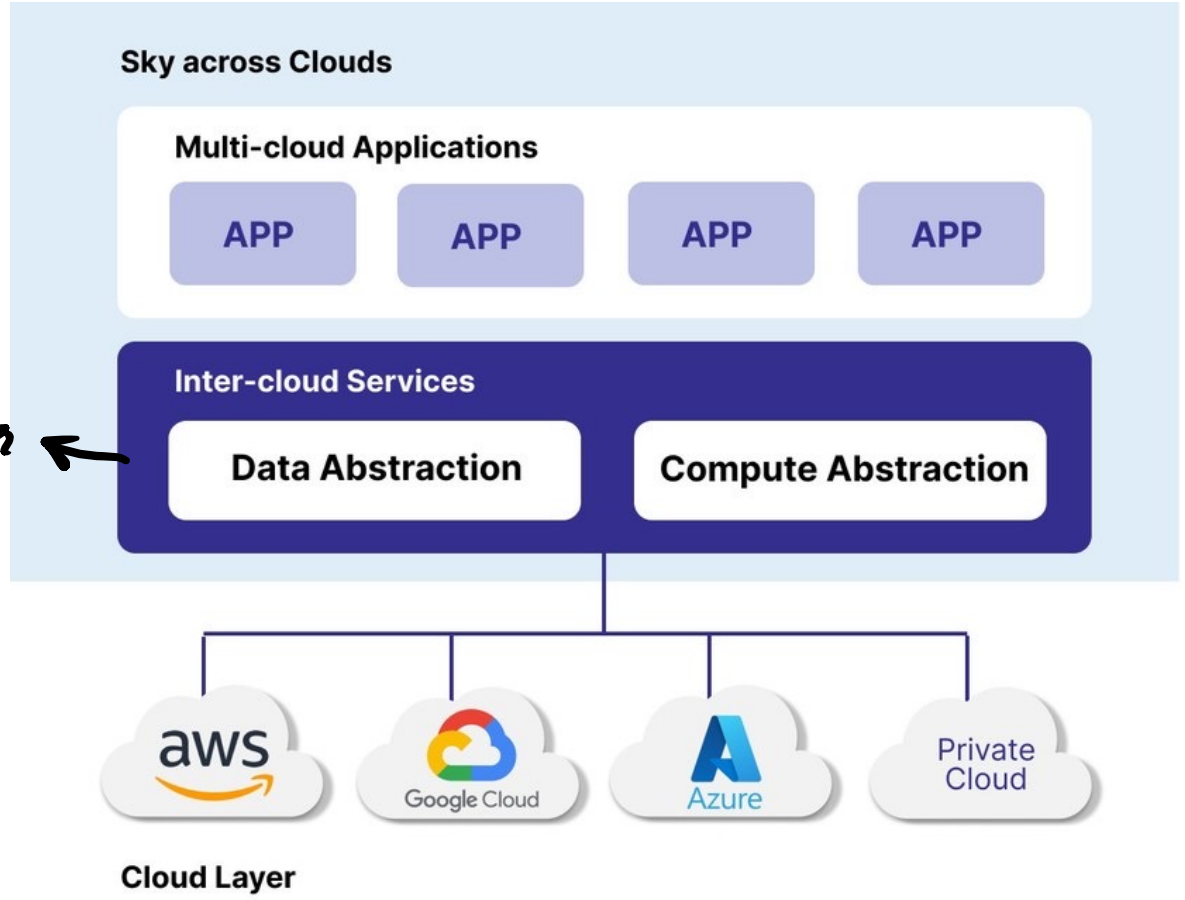
Managed multicloud

	Partitioned multi-cloud	Portable multi-cloud	Sky (transparent multi-cloud)
Same app running on different clouds?	No	Yes	Yes
Cloud transparent?	No	No	Yes
Universal APIs (do all clouds provide the same APIs)?	No	Yes	No
Deep APIs (are APIs at different levels)?	Yes	No	Yes

Table 1: *Comparison between existing types of multi-cloud and Sky.*

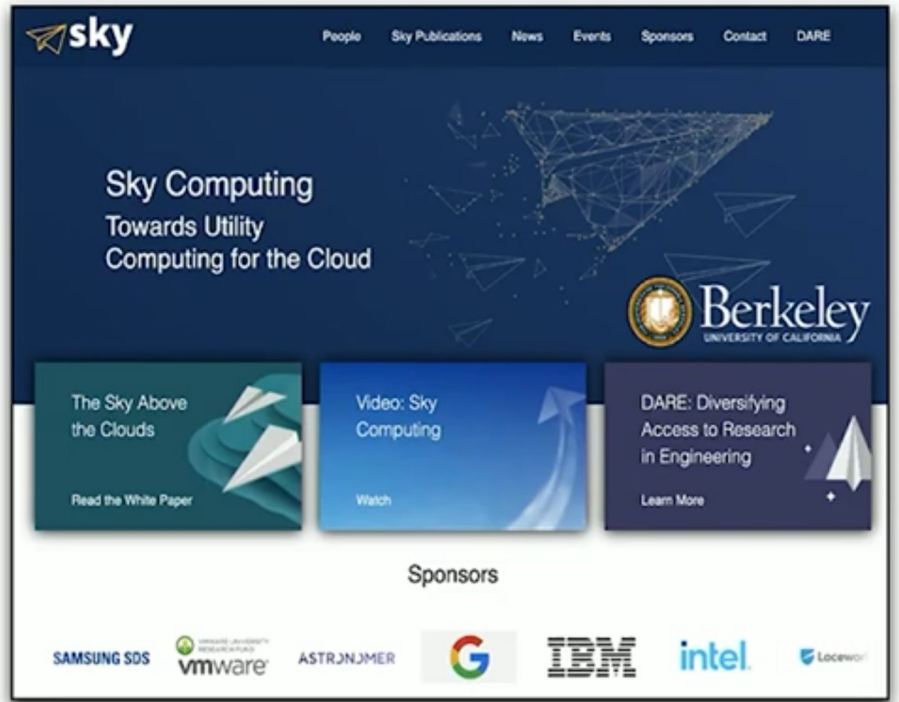
What is SKY-COMPUTING...

A set of software **tools/services** that make it easy to run apps across **multiple clouds** (ideally transparent).



- Berkley lab: sky.cs.berkeley.edu
- Published on: arxiv.org/abs/2205.07147
- Another publication: dl.acm.org/doi/10.1145/3458336.3465301

Sky Computing Lab (2022-)



Successor of RISELab, AMPLab

- 8 faculty
- ~50 students & postdocs
- Sponsored by

IBM intel

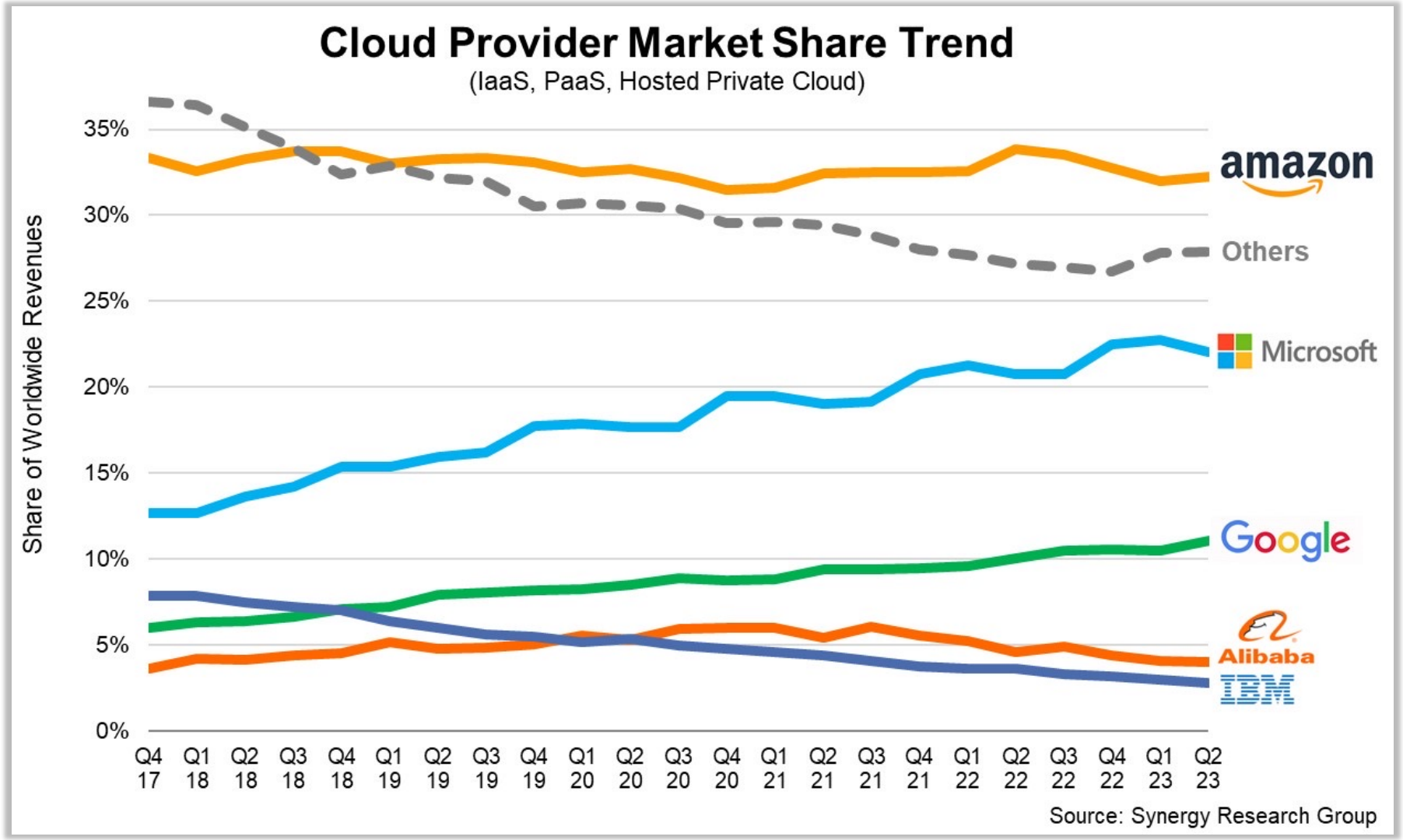
SAMSUNG SDS vmware

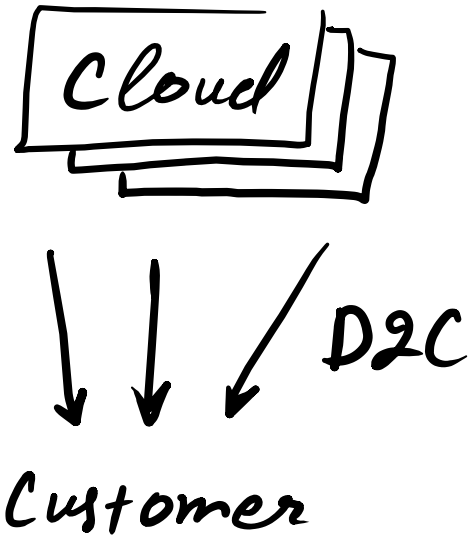
ASTRONOMER Lacework

Prof. Ion Stoica @ Berkley, CA

- Research: Distributed sys., Cloud, Networks
- “Sky computing” lab
- Databricks co-founder
- Video presentation: usenix.org/conference/atc23/presentation/joint-keynote

Current computing **market**: Proprietary-service-oriented business model





Issues with current cloud computing market:

- Non-compatible proprietary interfaces
- Market model discourages competition

  lock-in strategy (egress cost, proprietary API, volume offer/contracts, etc.)
- Difficult to compare services (burden of choice)

  Customers must choose which clouds to use for which workloads
- Complex to setup (configs) & optimize (tunning) the service → high operational cost

  Ever-increasing set of services/configs/parameters; Migration complexity

  e.g. users worry about: resource allocation, query optimization, or excessive configuration and tuning decisions..

Compatibility

Ecosystem actors seek cross-cloud compatibility:

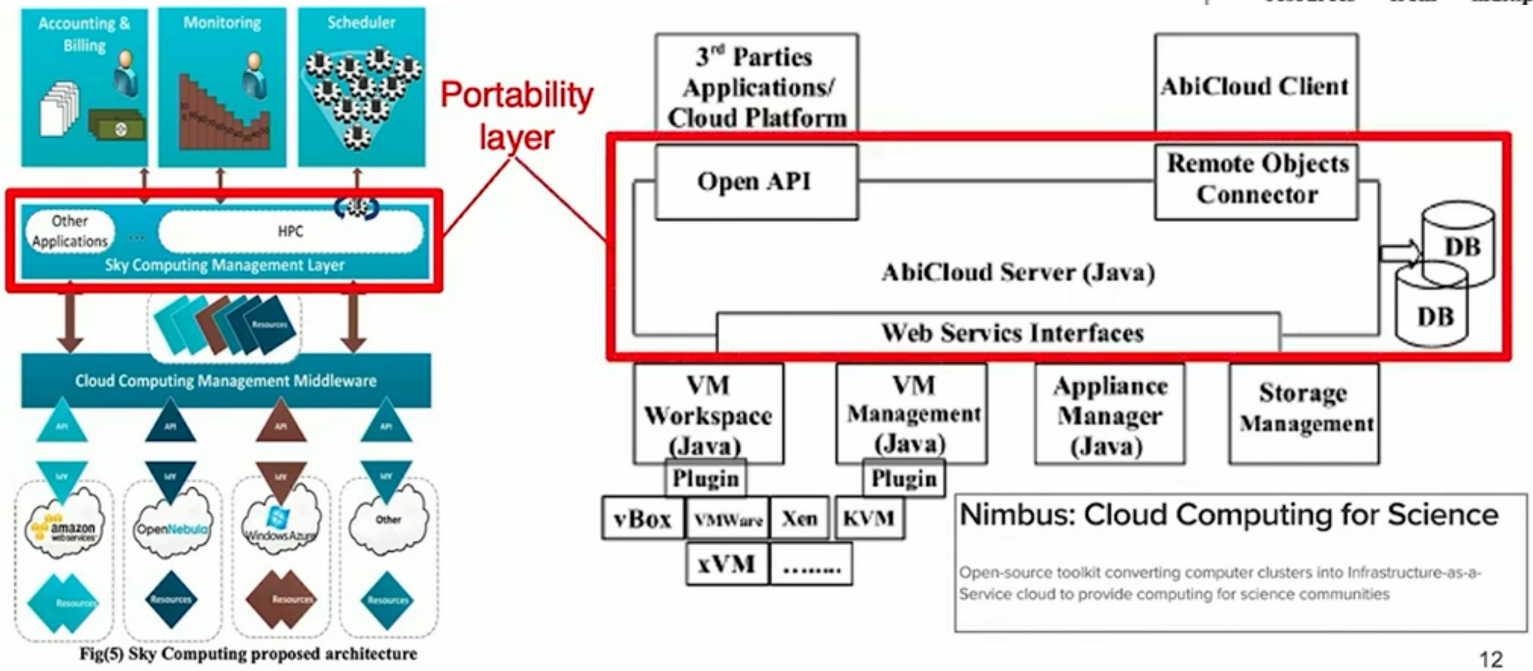
- Corporate users: ability to leverage a combination of services across clouds.
- Third-party software services: need to support multiple clouds as users already make cross-clouds compatibility a requirement.
- Cloud providers: offer service interfaces compatible with other clouds' proprietary interfaces.

OSS drive compatibility

serves as standard (at different layers of the software stack)

Previous proposals: comprehensive compatibility **standard** supported by all clouds

Uniform API



Fig(5) Sky Computing proposed architecture

Sky Computing: The Future of cloud computing

Abha Tewari
Associate Professor
V.E.S Institute of Technology

Pooja Nagdev
Associate Professor
V.E.S Institute of Technology

Aarti Sahitya
PG Scholar, M.E IT
V.E.S Institute of Technology

Abstract-Sometimes, a single cloud isn't enough. Sometimes, you need the whole sky. That's why a number of researchers are developing tools to federate clouds, an architectural concept dubbed "sky computing" .The Sky Computing model allows the creation of large scale infrastructure using resources from multiple cloud providers. These

to run embarrassingly parallel
rmance.
n several cloud infrastructures
oud would be especially useful to
utilize open large-scale compute
ience Grid or the European Grid
h open source cloud computing
order to create their own private
Amazon's EC2.
ut sky computing and differences
sky computing. and various
to sky computing over multiple
s picked up by many projects.

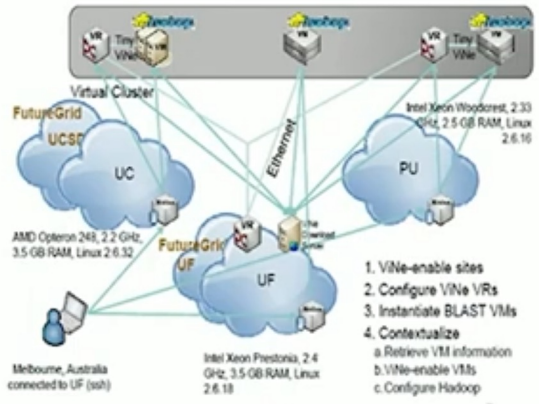


Fig (1) Sky Computing

2015

Nimbus: Cloud Computing for Science

Open-source toolkit converting computer clusters into Infrastructure-as-a-Service cloud to provide computing for science communities

Standardization problem:

- Not feasible (given the amount of interfaces - unlikely to happen)
- Dominant cloud would resist (lessen their competitive advantage)
- Impedes innovation (rigid set of interfaces)

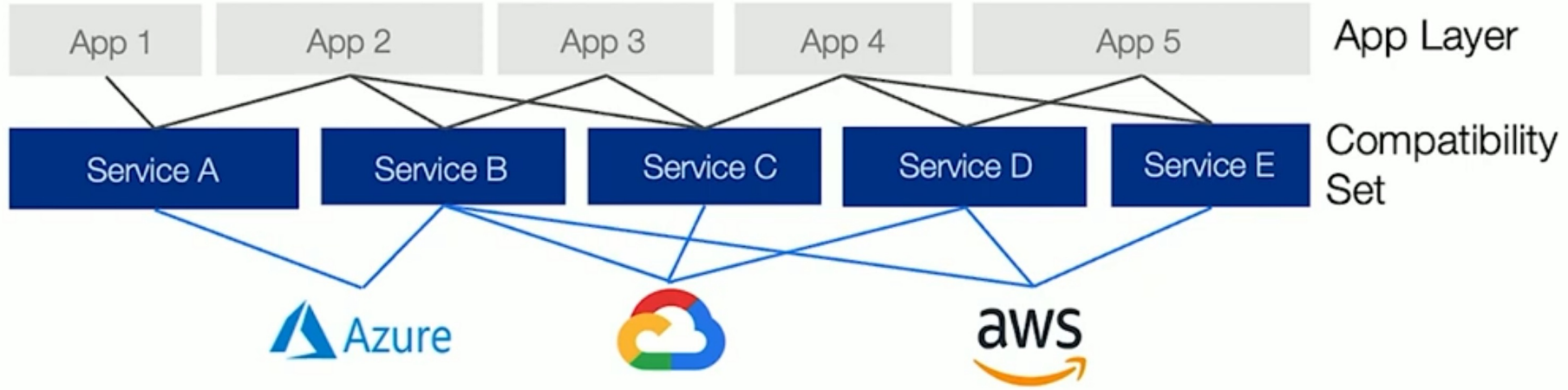
  At what abstraction level? Tradeoff between user-flexibility & operator-innovation.
- Perfect-compatibility of cloud is not necessary (no need for every service to run on all clouds)

New proposal: introduce inter-cloud broker

- Sky is cloud-computing mediated by an **Inter-cloud Broker**.
  = managed mediated multi-cloud
- Inter-cloud broker matches app **demands** & user **preferences** to clouds;

- compatibility set: similar services provided by many clouds (hosted or managed)
  e.g. OSS: Kubernetes, Apache Spark, Apache Kafka
  e.g. cloud-specific: AWS Inferentia, BigQuery

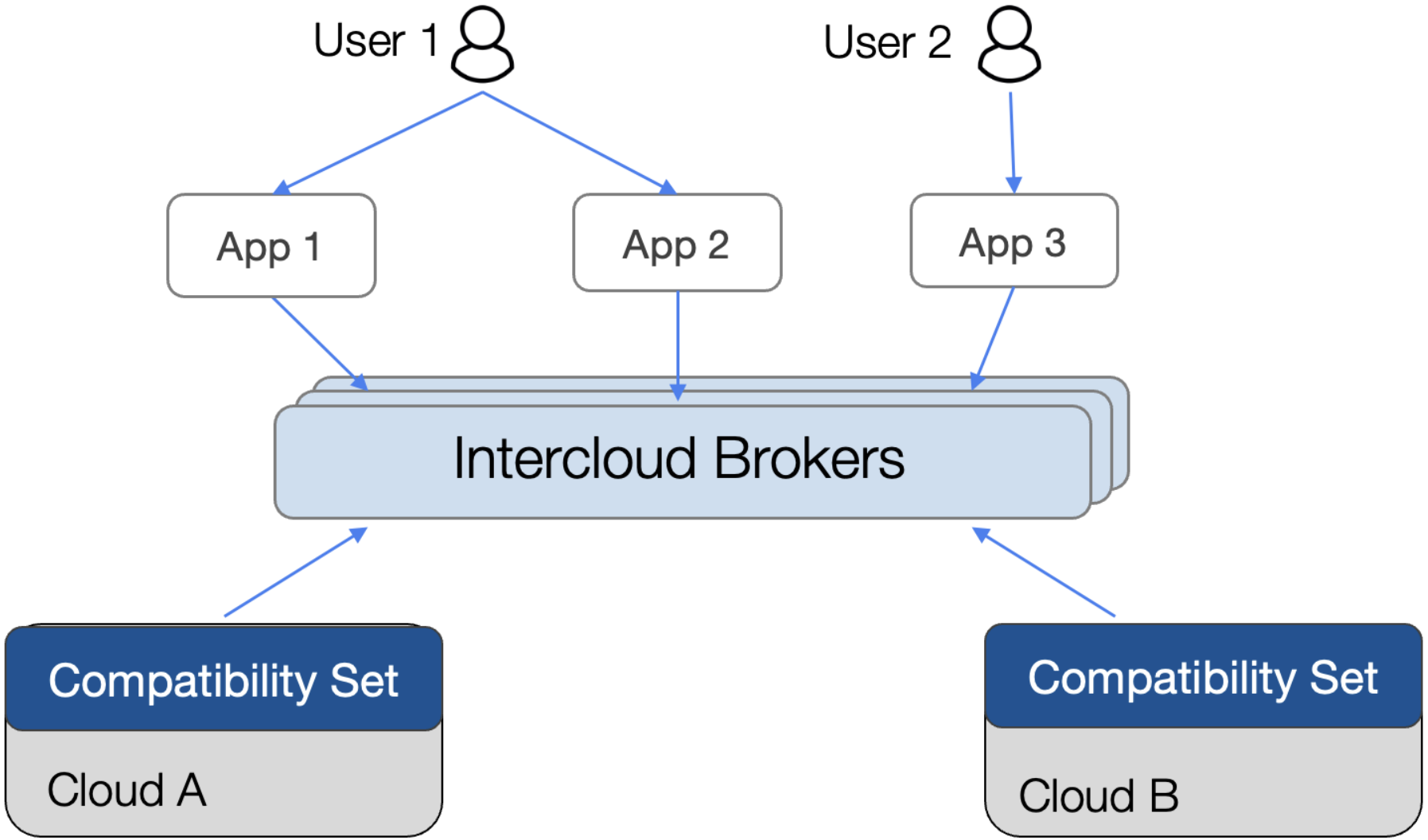
Two-sided Market:



Instead of a portability layer, a two-sided market

- One side: *existing* services running on one or more clouds (aka **compatibility set**)
- Other side: apps using one or more of these services

New proposal: Flexible workload placement



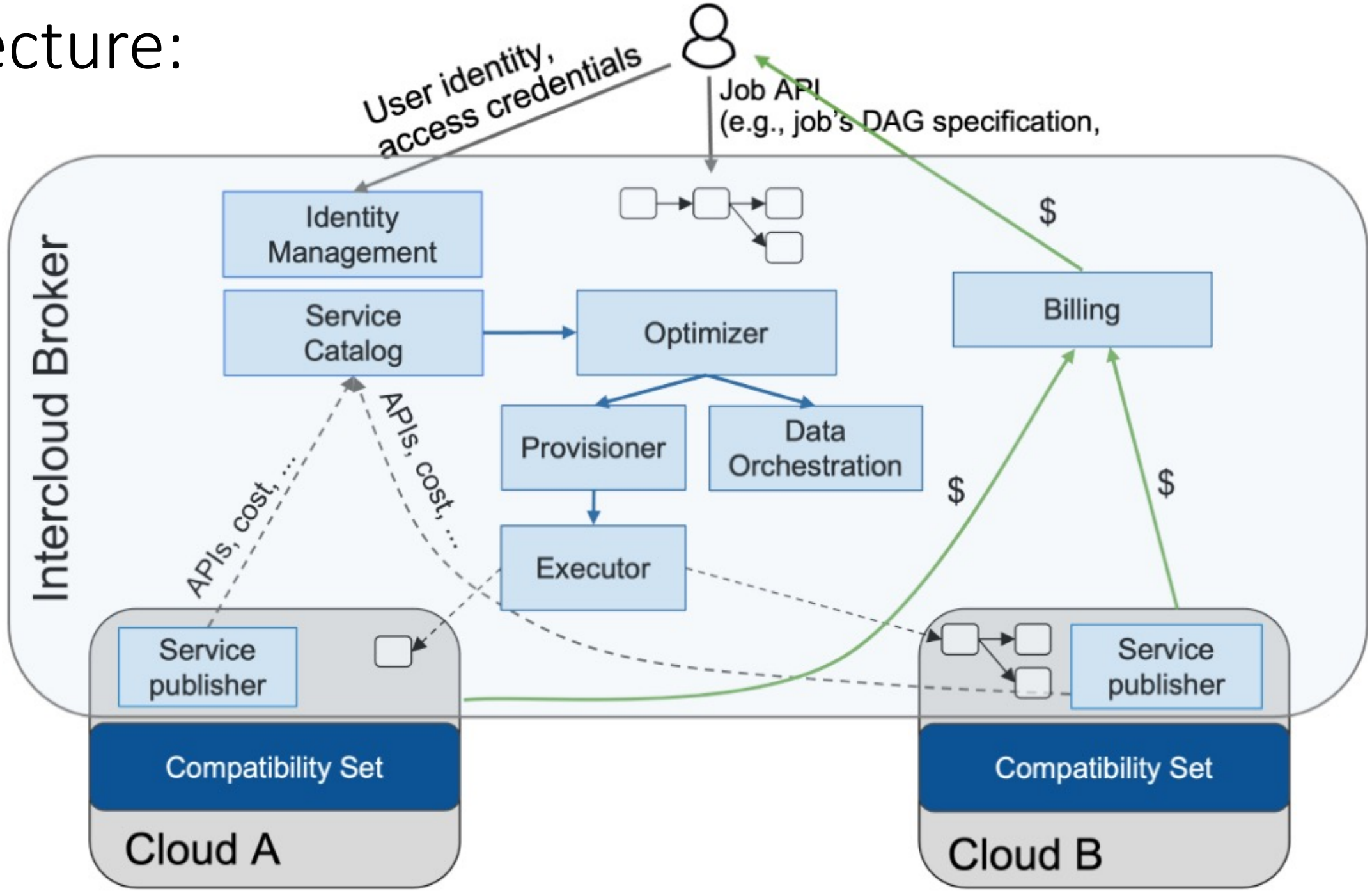
Goals of new proposal:

- Greater/Partial compatibility (encourage expanding of compatibility set)
- Flexible workload placement (through intermediation)

  Allow customers to move/shift workloads between clouds.
- Thriving competitive market (platform serves as marketplace)
- Fully-managed (rely on brokers to optimize desired criteria)

  Self-tuning & self-managing — only need to submit a job description.

Architecture:



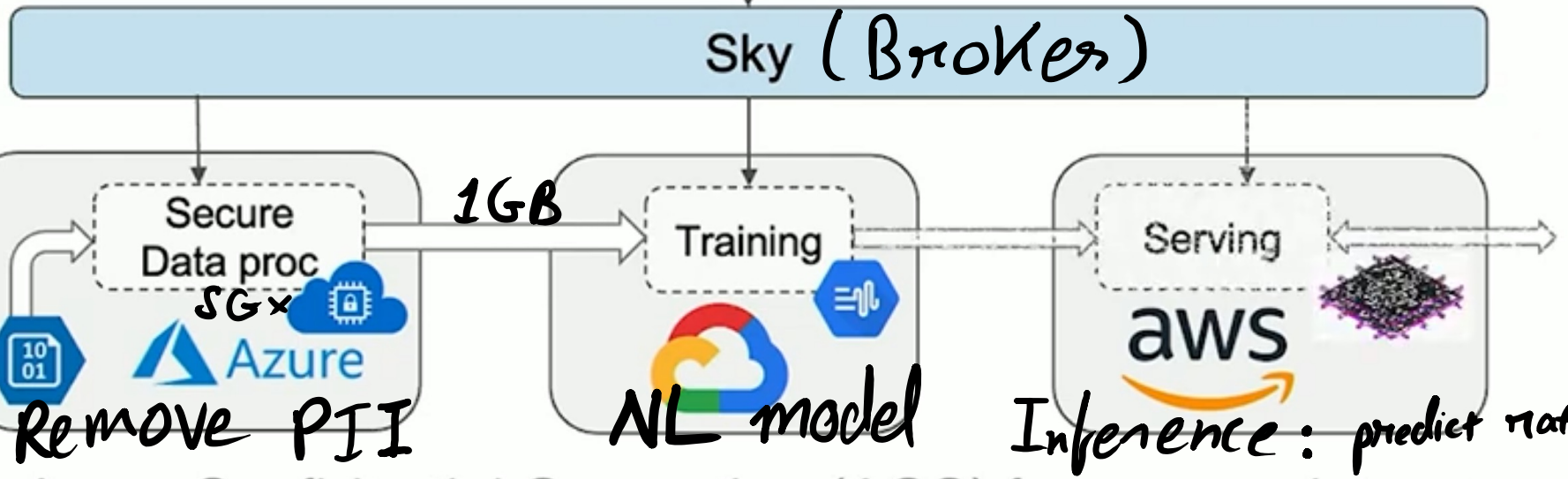
Business model:

- Service fees as intermediator.
- Telemetry data on jobs' execution patterns and providers market share (e.g. frequency of services use).

Sky Example

objective

ML Pipeline + optimize cost

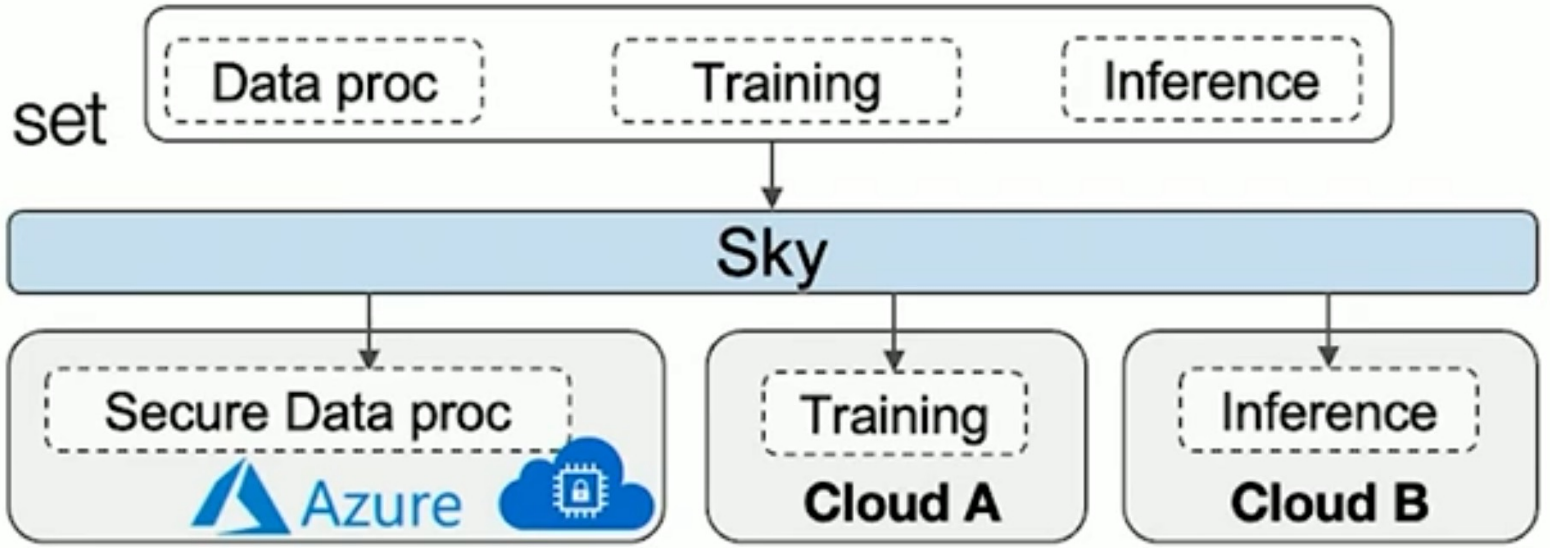


Data set
Azure Blob store

- Use Azure Confidential Computing (ACC) for secure data processing
- Use Google Cloud for training on TPUs (cost-performance)
- Use AWS for serving on Inferentia (low cost)

Does it Work ?

- BERT on user review data set
- 10K queries
- Requirement: process data confidentially
- Preference: reduce cost



		proc	train	infer	egress	Total
Time	**Azure**	0.6	8.4	1.8	–	10.8
(hr.)	**Sky**	0.6	4.0 -47%	1.1 -38%	0.03	5.7 -47%
Cost	**Azure**	0.8	103	1.4	–	105
($)	**Sky**	0.8	39 -62%	0.3 -78%	0.1	40 -61%

Transfer costs

Reduce both latency and cost

Benefits/Opportunities of cross-cloud deployments

- Low barriers to cloud usage → greater cloud adoption
- Easy access to specialized services → Rapid technical innovations
- Integration of various computational options — on-premise, edge, zones, etc.
  - E.g. Massive cost savings of repatriating cloud workloads (to private clouds).
- Enhancing compliance, security, and resilience/reliability.

Conjectures

- Compatibility set: There are enough easy cases to benefit users from Sky computing.
  - Killer apps: DS/ML pipelines (DAG model) & data sovereignty trends;
- No help needed from existing cloud providers.
- Constant evolution: once initiated market forces will create self-reinforcing cycle:
  - More compatibility → Greater supported workloads.
  - More workloads → Larger compatibility set (Clouds offer more services)
- Data transfer cost will drop — offering reciprocally-free data peering

Risks:

- Market devolve in dysfunctional ways: **Collusion** & predatory pricing.
- Inaccurate catalog information.
- Shim layers have a significant drawback: provide the lowest common denominator functionality across services. (remedy: “bolt-on” layers which extend a service’s functionality);
- Sky may remain only a niche market.
- Requires new debugging, monitoring, & profiling tools.

Thoughts?
How large Sky market would become?
will it gain traction?