Facebook Data Caching Akamai CDN - Edge and Facebook Scalable Testing

Dr. Animesh Chaturvedi

Assistant Professor: IIIT Dharwad

Post Doctorate: King's College London & The Alan Turing Institute

PhD: IIT Indore MTech: IIITDM Jabalpur

Facebook Data Caching Akamai CDN - Edge

A. Content Delivery Network, or
     Content Distribution Network (CDN)
B. Hybrid Akamai and Facebook

Content Delivery Network, or Content Distribution Network (CDN)
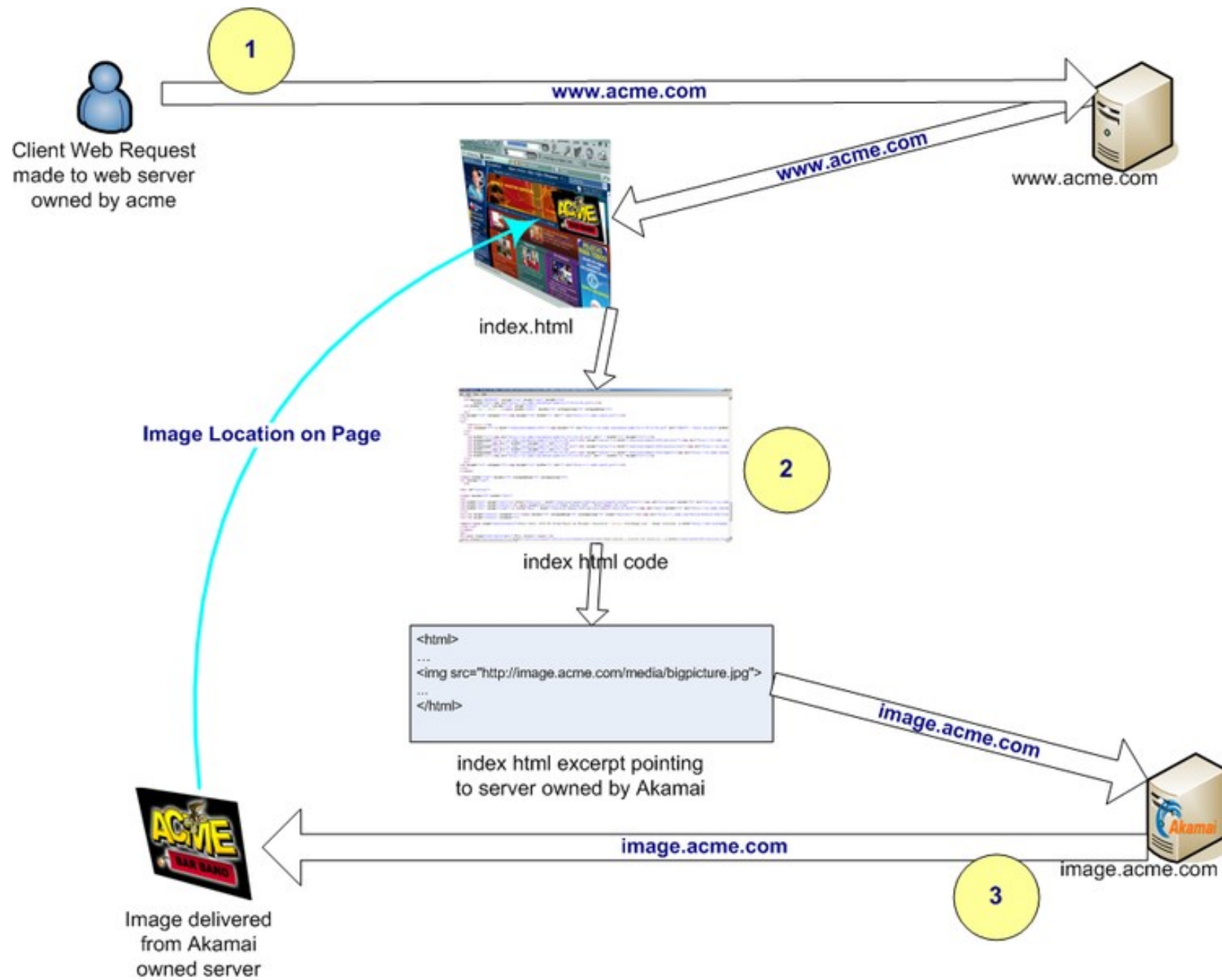
- CDN is a geographically distributed network of proxy servers and their data centers
  - to provide high availability and performance by distributing the service spatially relative to end users.
  - came in late 1990s to improve the performance bottlenecks of the Internet
- Peer-to-peer CDNs: Consumers provide resources as well as use them. Different from client-server
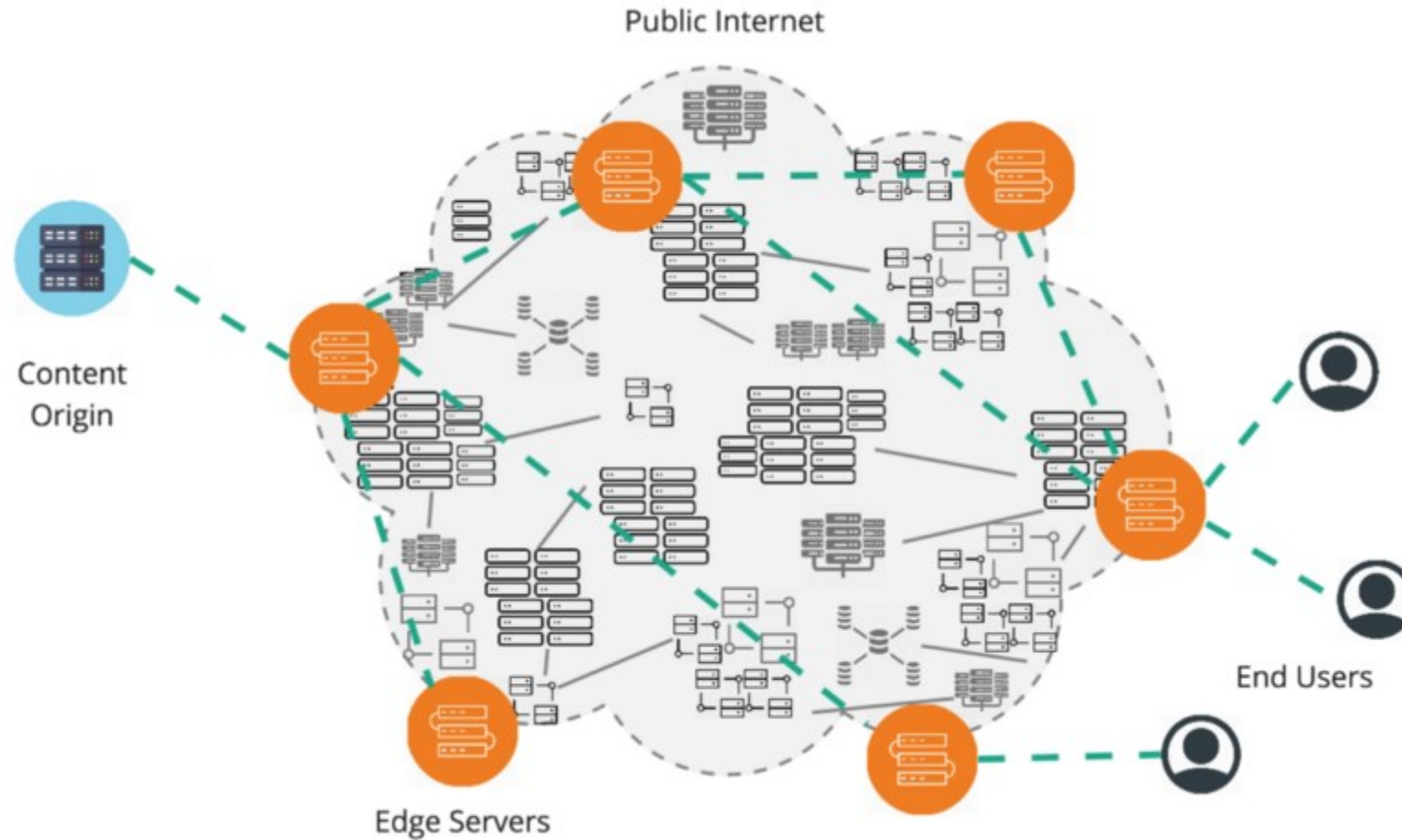- Private CDNs: Content owners create their own CDN.
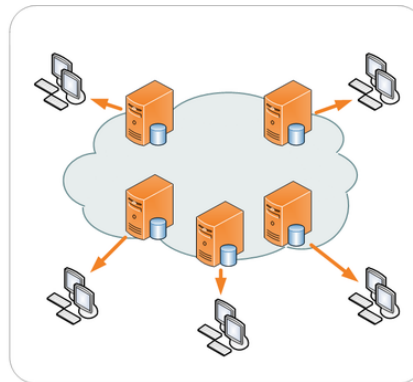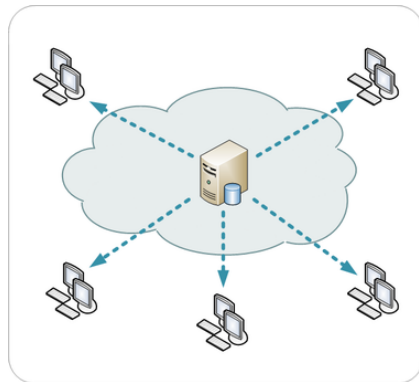
(Left) Single server distribution
(Right) CDN scheme of distribution

https://en.wikipedia.org/wiki/Content_delivery_network

Akamai CDN

- Global CDN, Cybersecurity, and Cloud service provider.
- Akamai CDN is one of the largest distributed computing platform
  - Serving between 15% and 30% of all web traffic
  - Operates and rents worldwide network of servers
  - Consumer websites work faster by distributing content from locations near the end-user.
  - End-user navigates to the URL of an Akamai's consumer; their browser is redirected to one of Akamai's copies of the website.
- Facebook as Consumer of Akamai CDN uses many edge servers closer to FB consumer (i.e. FB users).

https://en.wikipedia.org/wiki/Akamai_Technologies

https://en.wikipedia.org/wiki/Akamai_Technologies

Akamai Content Delivery Network

Public Internet

Content Origin

Edge Servers

End Users

Hybrid Akamai and Facebook

- Facebook Photo Serving Stack: components are linked to show the photo retrieval work-flow.

- Desktop and Mobile clients initiate request traffic, which routes either directly to the Facebook Edge or via Akamai CDN depending on the fetch path.

- The Origin Cache collects traffic from both paths, serving images from its cache and resizing them if needed.

- The Haystack backend holds the actual image content.

Huang, Qi, et al. "An analysis of Facebook photo caching." *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles*. 2013.

Facebook's Photo-Serving Stack

- **Browser:** The first cache layer is in the client's browser

- **Edge:** The Facebook Edge is comprised of a set of Edge Caches that each run inside points of presence (PoPs) close to end users.

- **Haystack:** The backend layer accessed when there is a miss in the Origin cache (co-located with storage servers), the image can often be retrieved from a local Haystack server.
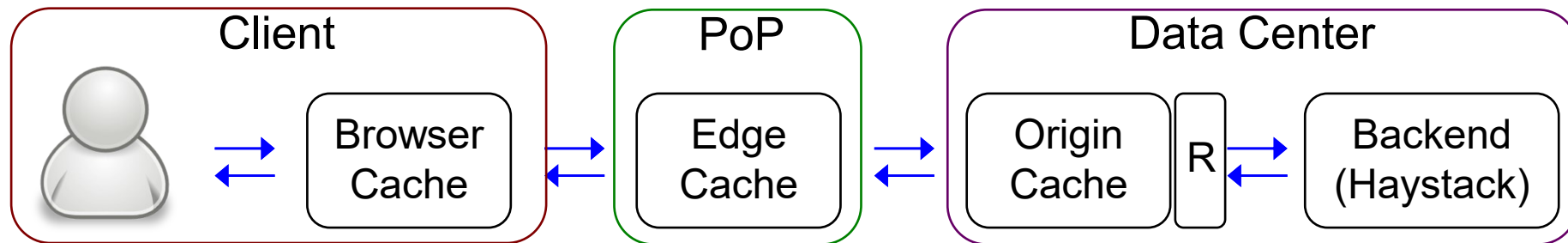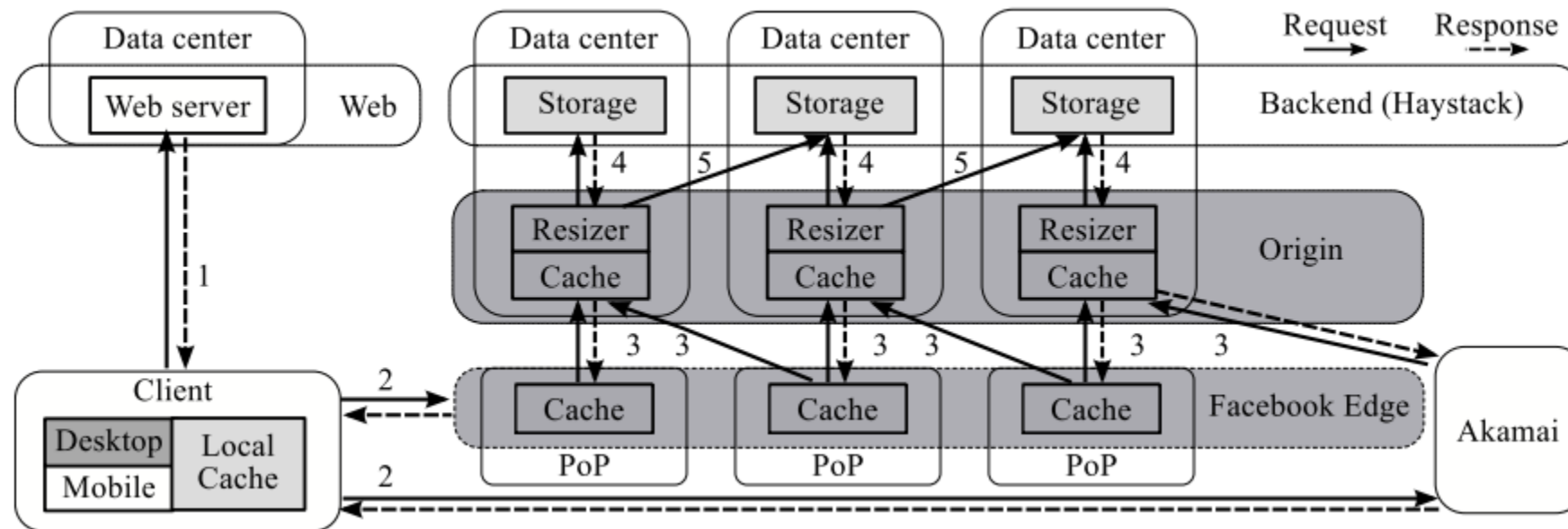

Huang, Qi, et al. "An analysis of Facebook photo caching." *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles*. 2013.

Hybrid Akamai and Facebook

Architecture and full life cycle of a photo request

- Photo Transformations: Resizers transform photos that are requested by the Akamai CDN,

- Caching Stack

Huang, Qi, et al. "An analysis of Facebook photo caching." *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles*. 2013.

Facebook Scalable Testing

A. Infern and
B. Sapienz
C. FBLearner
D. Facebook's OneWorld platform

Scalable Facebook Testing

- Scale means different things to different techniques to scale (or fail to scale) in different dimensions
- Scalability needed to assess the degree of deployability
- Challenges and Opportunities when deploying
  - Static analysis and Dynamic analysis at scale
- Facebook deploy two technologies
  - the Infer for static analysis and
  - the Sapienz dynamic analysis
  - Both are research-led start-up that was subsequently deployed at scale, impacting billions of people worldwide.

Harman, Mark, and Peter O'Hearn. "From start-ups to scale-ups: Opportunities and open problems for static and dynamic program analysis." *2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. IEEE, 2018.
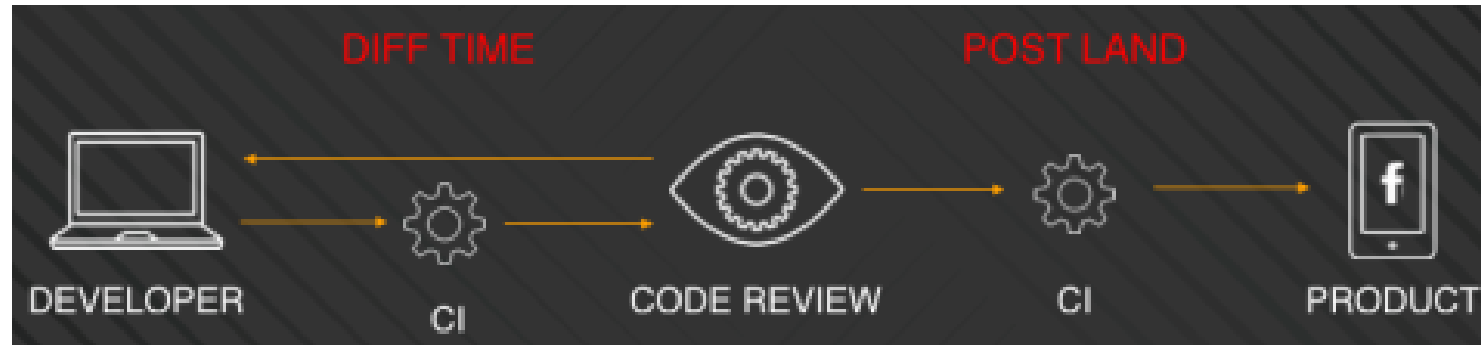
Infer: Static Analysis at Scale

- At Facebook, Infer is a 'deep' static analysis tool applied to
  - Java, Objective C, and C++ code bases
- Infer at Facebook with acquisition of startup Monoidics in 2013
- Infer is open source and used at several companies
  - AWS, Mozilla, JD.com and Spotify
- Based on Separation Logic to scale algorithms for reasoning "states"
  - of a *store* and a *heap* about memory,
  - local (or stack-allocated) variables and dynamically-allocated,
  - implements a compositional program analysis: inter-procedural analysis,
  - follows pointer chains, safety of programs with embedded pointers,
  - scales to large code bases
- Result: 10s of 1000s bugs fix by Facebook developers before production.

Harman, Mark, and Peter O'Hearn. "From start-ups to scale-ups: Opportunities and open problems for static and dynamic program analysis." *2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. IEEE, 2018.

Infer: Static Analysis at Scale

- Facebook practices Continuous Development and Deployment i.e. Continuous Integration (CI) of "Diffs" (a code change) for code-review.

- Developers shares a single code base to alter and commit. Then, prepares a diff and submits it for the code review.

- Infer runs as a bot during diff time, writing comments for the developer and other human reviewers. Post-land incorporate the master build of the system.

Harman, Mark, and Peter O'Hearn. "From start-ups to scale-ups: Opportunities and open problems for static and dynamic program analysis." *2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. IEEE, 2018.

Sapienz: Dynamic Analysis at Scale

- Sapienz is deployed to run continuously testing the most recent internal builds of the Facebook apps.

- Sapienz is a multi-objective automated test case design system, that seeks to maximize fault revelation while minimizing the debug effort to fix.

- It uses
  - FBLearner (Facebook's Machine Learning infrastructure)
  - Facebook's OneWorld platform

Harman, Mark, and Peter O'Hearn. "From start-ups to scale-ups: Opportunities and open problems for static and dynamic program analysis." *2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. IEEE, 2018.

Sapienz: Dynamic Analysis at Scale

- FBLearner (Facebook's Machine Learning infrastructure)
  - Machine Learning (ML) perform at a scale, e.g., ML inference phase executions run into the tens of trillions per day.
  - FBLearner is a ML platform designed to support at a global scale
  - Sapienz is 1 of 100s services deployed on FBLearner framework
  - FBLearner Flow component in Sapienz
    - deploys detection of crashing behaviour directly into the work flow,
    - integrates with Phabricator for reporting and actioning fixes
    - corrects the failures detected

- Facebook's OneWorld platform
  - support scalable deployment on many of emulators, at a time
  - Sapienz deploys on 1000s of emulators to test the Android app alone

Harman, Mark, and Peter O'Hearn. "From start-ups to scale-ups: Opportunities and open problems for static and dynamic program analysis." *2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. IEEE, 2018.

FBLearner

ML-as-a-Service Inside Facebook: FBLearner Feature Store, FBLearner Flow, FBLearner Predictor.
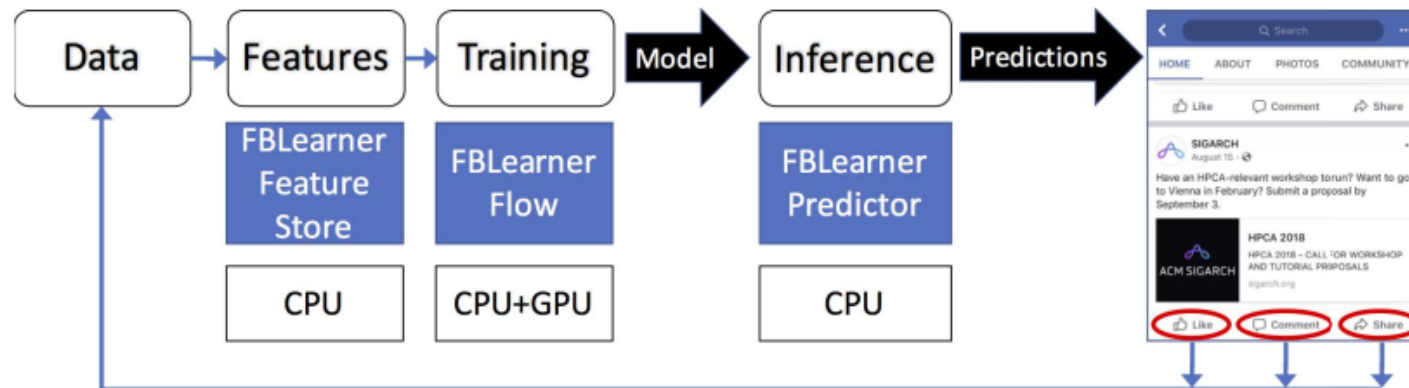
MACHINE LEARNING ALGORITHMS LEVERAGED BY PRODUCT/SERVICE

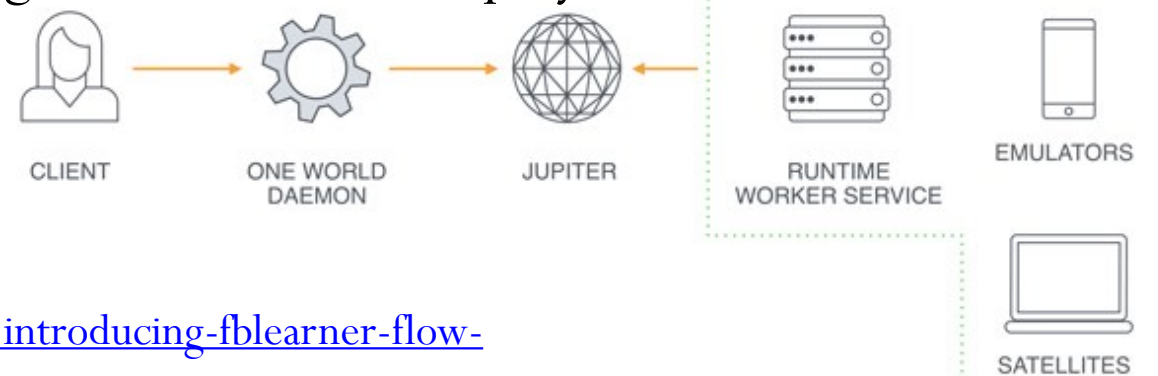| Models	Services
Support Vector Machines (SVM)	Facer (User Matching)
Gradient Boosted Decision Trees (GBDT)	Sigma
Multi-Layer Perceptron (MLP)	Ads, News Feed, Search, Sigma
Convolutional Neural Networks (CNN)	Lumos, Facer (Feature Extraction)
Recurrent Neural Networks (RNN)	Text Understanding, Translation, Speech Recognition

Hazelwood, Kim, et al. "Applied machine learning at facebook: A datacenter infrastructure perspective."
2018 IEEE International Symposium on High Performance Computer Architecture (HPCA). IEEE, 2018.

Facebook's OneWorld platform

Support standard communication mechanisms like ADB (Android Debug Bridge) and Provide the illusion that remote devices are connected locally
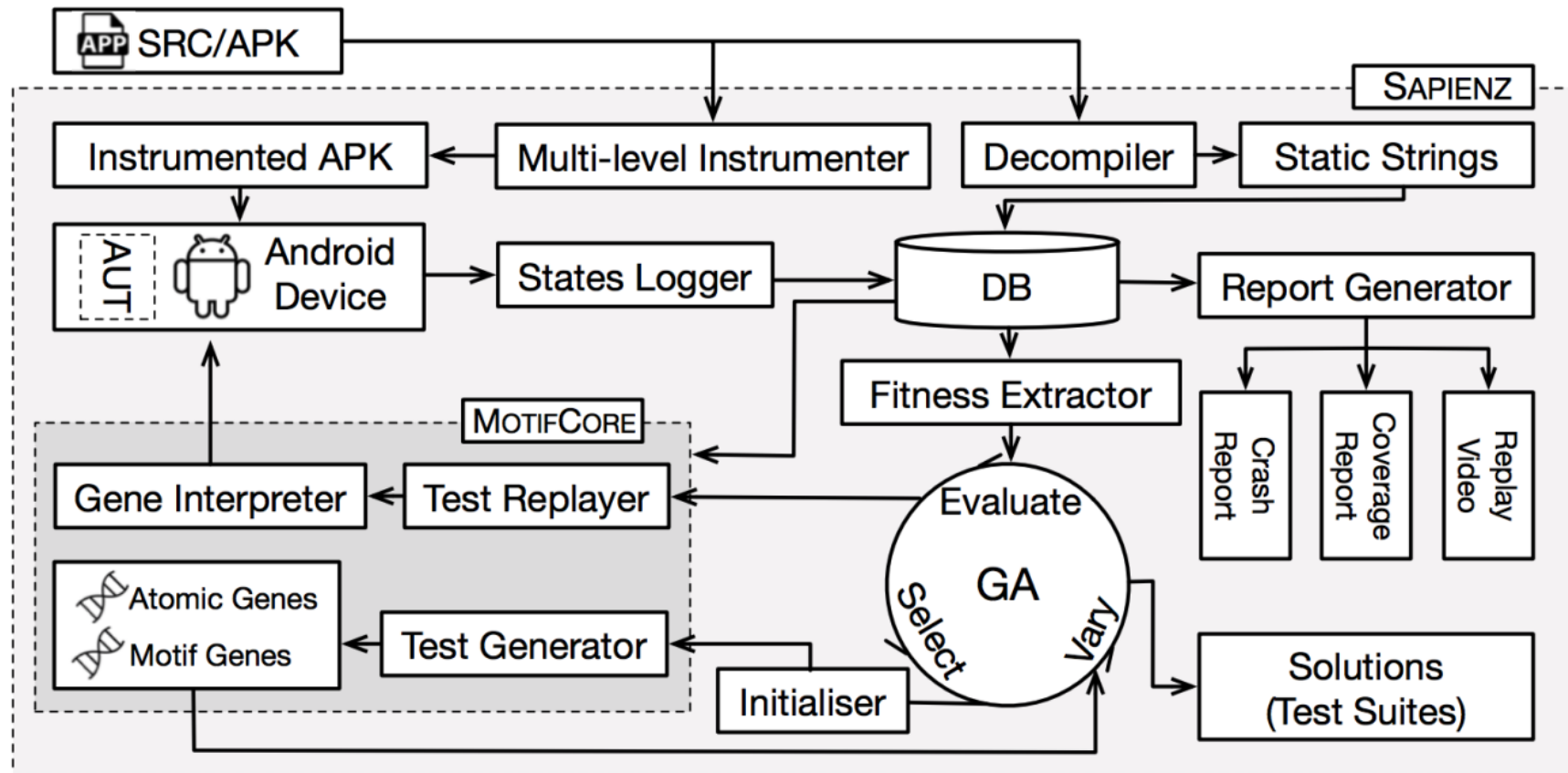
- **Runtime worker service:** it manages the life cycle of the resource and responds to requests from clients to use its resources.
- **One World daemon:** it connect to remote resources with protocol to communicate with workers and sets up the environment.
- **Scheduler:** Uses Jupiter, a job-scheduling service at Facebook to match clients with workers' resources as per specified requirements.
- **Satellite:** it connect local resources to the global One World deployment.

CLIENT → ONE WORLD DAEMON → JUPITER ← RUNTIME WORKER SERVICE

DEVICE LAB

EMULATORS

SATELLITES

https://engineering.fb.com/2016/05/09/core-data/introducing-fblearner-flow-facebook-s-ai-backbone/

Sapienz workflow

Sapienz Automated test design workflow

Harman, Mark, and Peter O'Hearn. "From start-ups to scale-ups: Opportunities and open problems for static and dynamic program analysis." *2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. IEEE, 2018.

References

- Huang, Qi, et al. "An analysis of Facebook photo caching." *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles*. 2013.

- Harman, Mark, and Peter O'Hearn. "From start-ups to scale-ups: Opportunities and open problems for static and dynamic program analysis." *2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. IEEE, 2018.

- Hazelwood, Kim, et al. "Applied machine learning at facebook: A datacenter infrastructure perspective." *2018 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2018.

- https://en.wikipedia.org/wiki/Content_delivery_network

- https://en.wikipedia.org/wiki/Akamai_Technologies

- https://engineering.fb.com/2016/05/09/core-data/introducing-fblearner-flow-facebook-s-ai-backbone/

ขอบคุณ
Thai

Grazie
Italian

תודה רבה
Hebrew

ధన్యవాదగళు
Kannada

धन्यवादः
Sanskrit

Ευχαριστώ
Greek

Thank You
English

Gracias
Spanish

Спасибо
Russian

Obrigado
Portuguese

شكراً
Arabic

Merci
French

多謝
Traditional
Chinese

धन्यवाद
Hindi

Danke
German

多谢
Simplified
Chinese

நன்றி
Tamil
Tamil

ありがとうございました
Japanese

감사합니다
Korean