

From Laptop to Lambda: Outsourcing Everyday Jobs to Thousands of Transient Functional Containers

Sadjad Fouladi, Francisco Romero, Dan Iter, Qian Li, Shuvo Chatterjee+,
Christos Kozyrakis, Matei Zaharia, Keith Winstein

Stanford University, +Unaffiliated

<https://snr.stanford.edu/gg>

Stanford

Even today, many applications remain far from interactive

- Software compilation & testing, video processing, simulations, 3D rendering, etc.
- Users who desire the results in seconds need to harness thousands of cores in parallel.



Supercomputing-as-a-service

- Serverless platforms: thousands of functions invoked in just a few seconds.

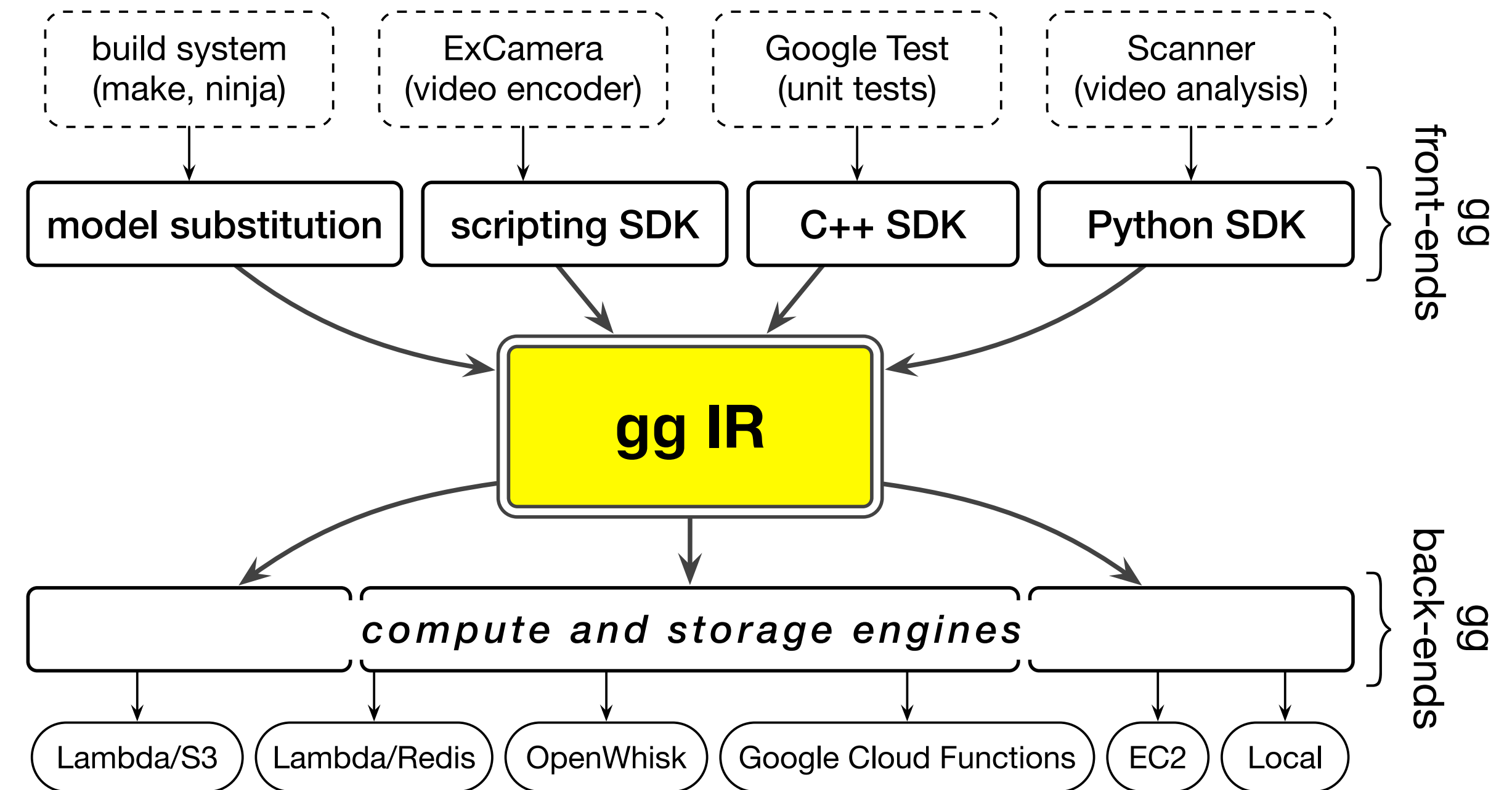
A 10,000-core supercomputer that is billed by the second.

- A number of applications, such as ExCamera, Sprocket and PyWren, exploit this to achieve interactive speeds.
- Building applications on top of these platforms is difficult!

gg is a framework that helps application developers port new and existing applications to serverless platforms and execute them with thousands-way parallelism.

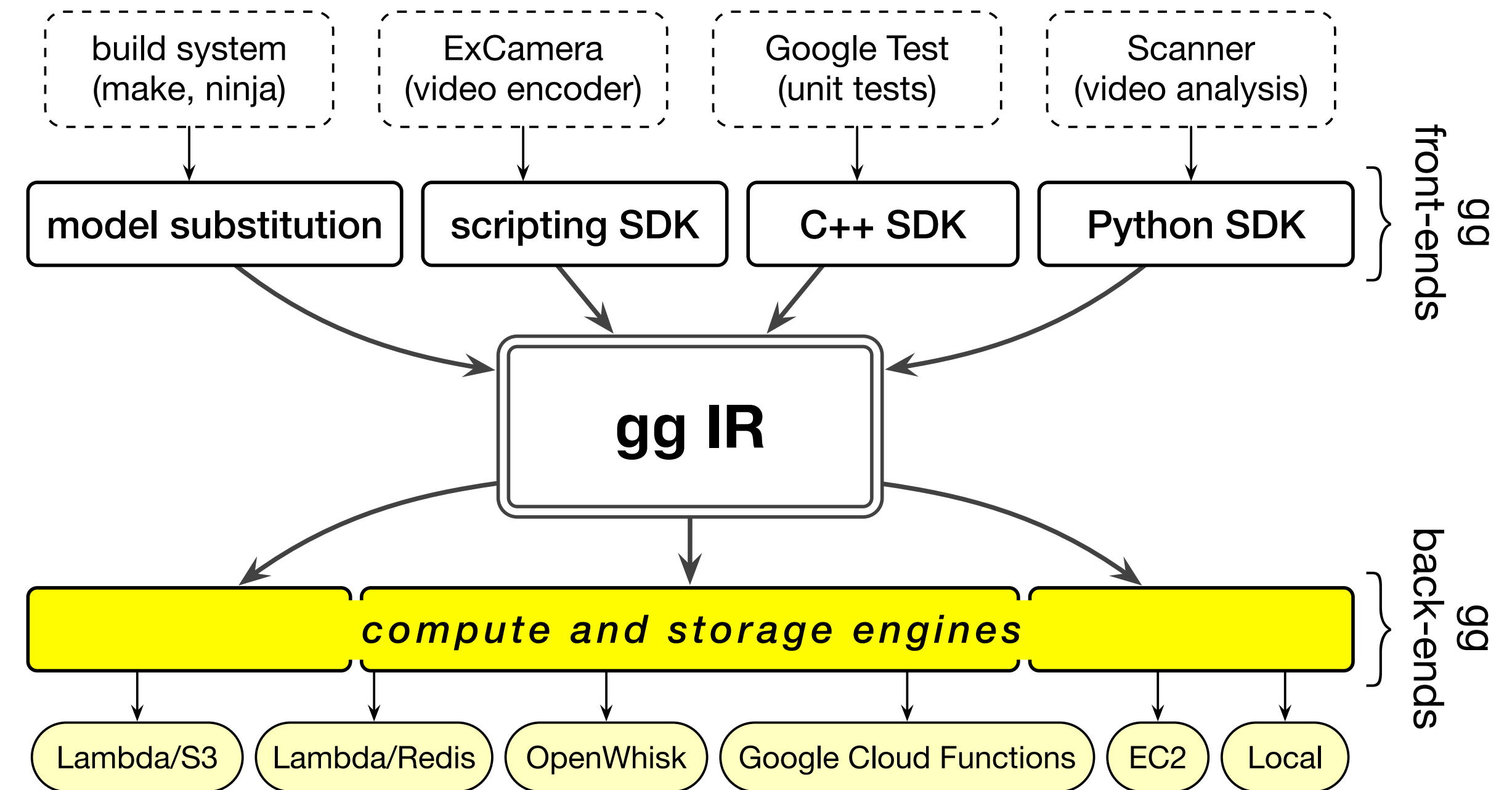
gg decouples application logic from its execution

- Application developers express their jobs in gg's **intermediate representation**, which abstracts the *application logic* from the details of *placement*, *schedule*, and *execution*.



gg decouples application logic from its execution

- gg takes care of the execution and provides runtime features for **dependency management, straggler mitigation, placement, failure recovery and memoization.**



Massively parallel execution can yield significant benefits!

Compiling Inkscape

Tool	Time	Cost
single-core make	32m 34s	—
icecc to a warm 48-core EC2 machine	6m 51s	\$2.30/hr
icecc to a warm 384-core EC2 cluster	6m 57s	\$18.40/hr
gg to AWS Lambda	1m 27s	50¢/run

From Laptop to Lambda:
Outsourcing Everyday Jobs to Thousands of Transient Functional Containers

ATC'19, Track I: Thursday, June 11th at 2:00pm.

<https://snr.stanford.edu/gg>

Stanford