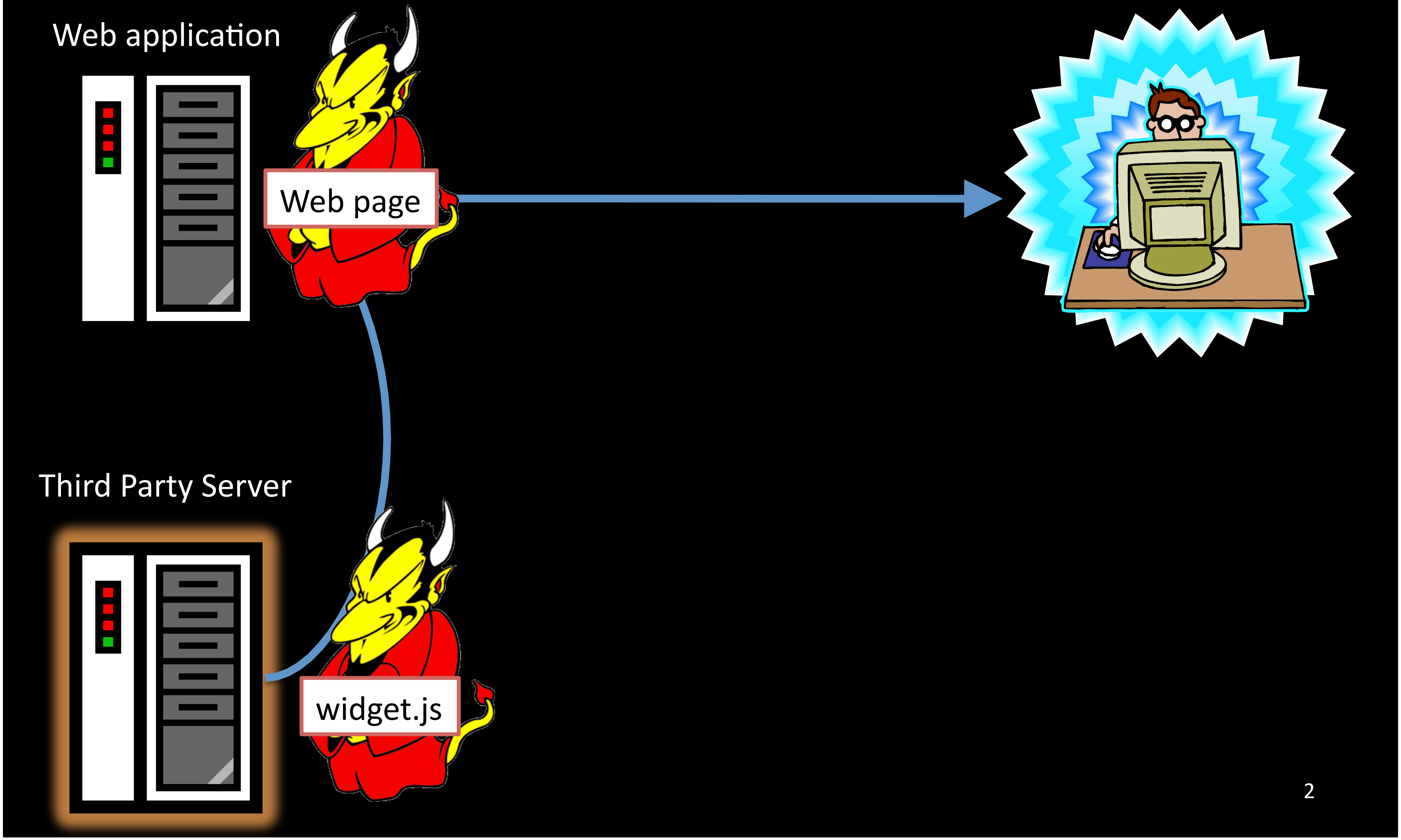
Gulfstream

Staged Static Analysis for Streaming JavaScript Applications

Salvatore Guarnieri
University of Washington

Ben Livshits
Microsoft Research



Web application

Web page

Third Party Server

widget.js

Safe Code Inclusion In JavaScript

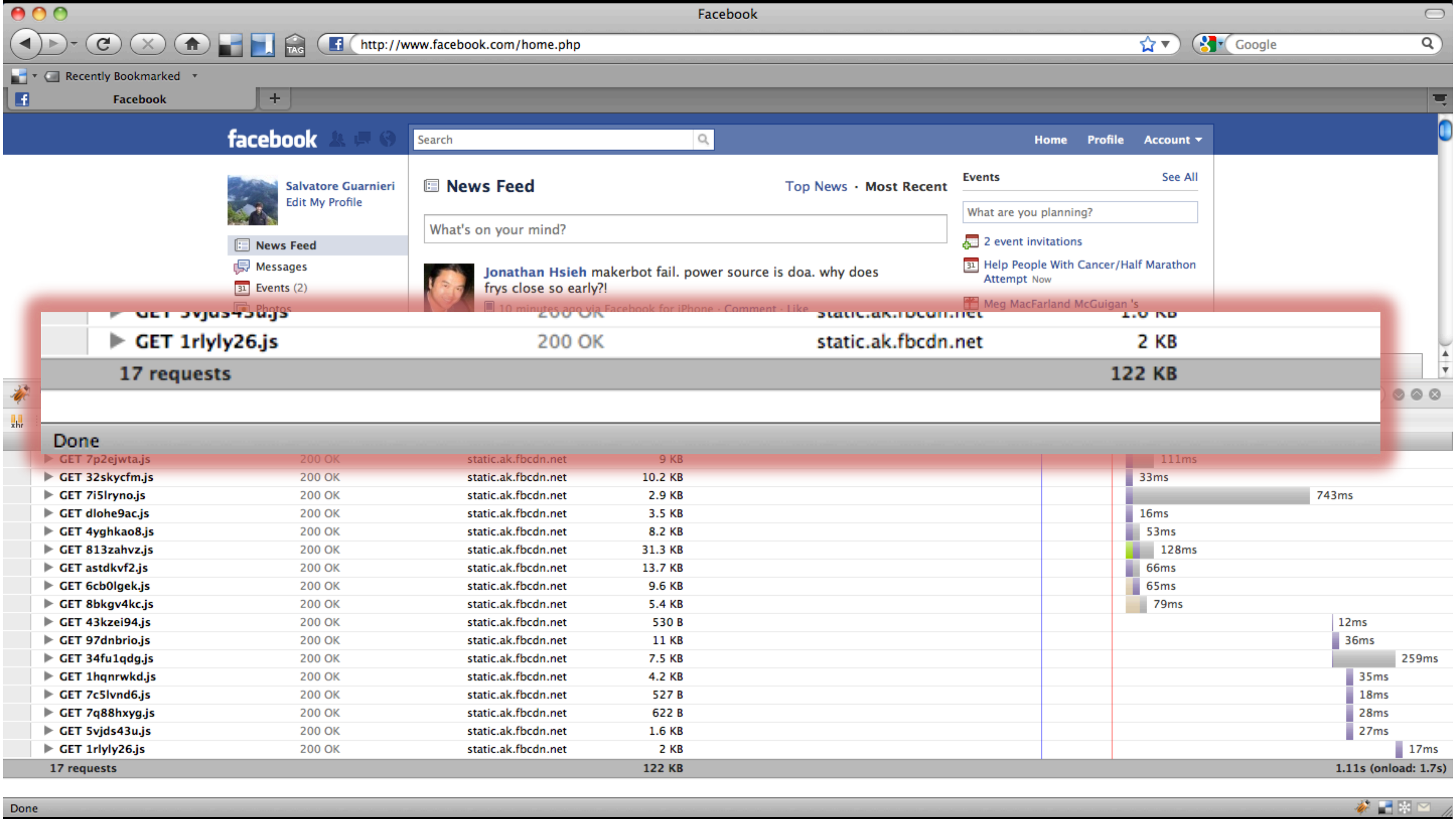
Runtime Enforcement

- Conscript [Oakland 10]
- BrowserShield [OSDI 06]
- Caja

Static Analysis

- Gatekeeper [USENIX Sec 09]
- Staged Information flow for JavaScript [PLDI 09]

Whole program analysis approaches require the entire program



Jonathan Hsieh makerbot fail. power source is doa. why does frys close so early?!

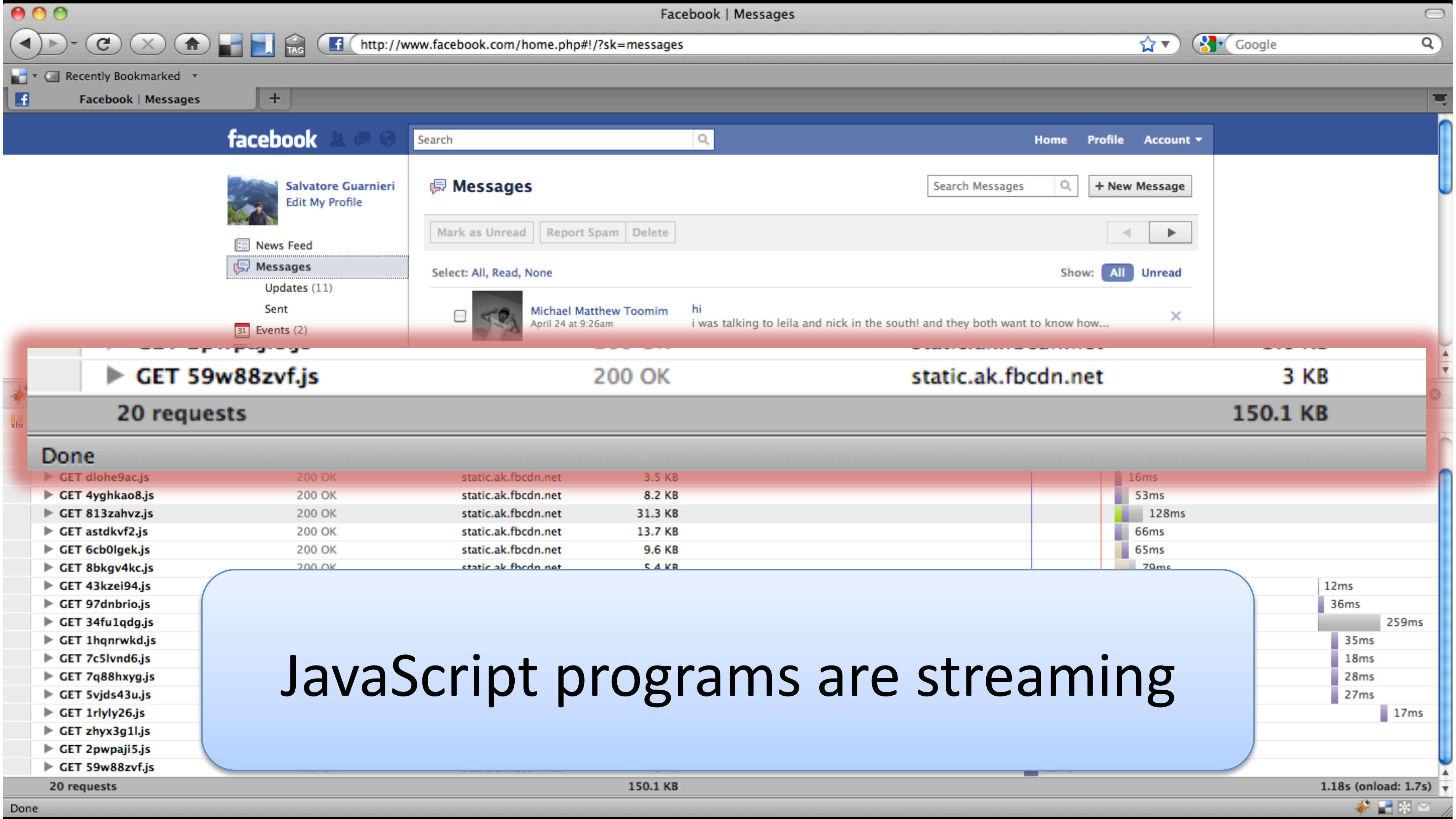
10 minutes ago via Facebook for iPhone · Comment · Like

```
▶ GET 1rlyly26.js        200 OK    static.ak.fbcdn.net    2 KB
  17 requests                                             122 KB

Done
```

```
▶ GET 7p2ejwta.js    200 OK    static.ak.fbcdn.net    9 KB      111ms
▶ GET 32skycfm.js    200 OK    static.ak.fbcdn.net    10.2 KB   33ms
▶ GET 7i5lryno.js    200 OK    static.ak.fbcdn.net    2.9 KB    743ms
▶ GET dlohe9ac.js    200 OK    static.ak.fbcdn.net    3.5 KB    16ms
▶ GET 4yghkao8.js    200 OK    static.ak.fbcdn.net    8.2 KB    53ms
▶ GET 813zahvz.js    200 OK    static.ak.fbcdn.net    31.3 KB   128ms
▶ GET astdkvf2.js    200 OK    static.ak.fbcdn.net    13.7 KB   66ms
▶ GET 6cb0lgek.js    200 OK    static.ak.fbcdn.net    9.6 KB    65ms
▶ GET 8bkgv4kc.js    200 OK    static.ak.fbcdn.net    5.4 KB    79ms
▶ GET 43kzei94.js    200 OK    static.ak.fbcdn.net    530 B     12ms
▶ GET 97dnbrio.js    200 OK    static.ak.fbcdn.net    11 KB     36ms
▶ GET 34fu1qdg.js    200 OK    static.ak.fbcdn.net    7.5 KB    259ms
▶ GET 1hqnrwkd.js    200 OK    static.ak.fbcdn.net    4.2 KB    35ms
▶ GET 7c5lvnd6.js    200 OK    static.ak.fbcdn.net    527 B     18ms
▶ GET 7q88hxyg.js    200 OK    static.ak.fbcdn.net    622 B     28ms
▶ GET 5vjds43u.js    200 OK    static.ak.fbcdn.net    1.6 KB    27ms
▶ GET 1rlyly26.js    200 OK    static.ak.fbcdn.net    2 KB      17ms
  17 requests                                          122 KB    1.11s (onload: 1.7s)

Done
```



Facebook | Messages

http://www.facebook.com/home.php#!/?sk=messages

Recently Bookmarked

facebook

Salvatore Guarnieri
Edit My Profile

News Feed
Messages
Updates (11)
Sent
Events (2)

Messages

Michael Matthew Toomim
April 24 at 9:26am

hi
i was talking to leila and nick in the south! and they both want to know how...

GET 59w88zvf.js	200 OK	static.ak.fbcdn.net	3 KB
20 requests			150.1 KB

Done

GET dlohe9ac.js	200 OK	static.ak.fbcdn.net	3.5 KB	16ms
GET 4yghkao8.js	200 OK	static.ak.fbcdn.net	8.2 KB	53ms
GET 813zahvz.js	200 OK	static.ak.fbcdn.net	31.3 KB	128ms
GET astdkvf2.js	200 OK	static.ak.fbcdn.net	13.7 KB	66ms
GET 6cb0lgek.js	200 OK	static.ak.fbcdn.net	9.6 KB	65ms
GET 8bkgv4kc.js	200 OK	static.ak.fbcdn.net	5.4 KB	79ms
GET 43kzei94.js				12ms
GET 97dnbrio.js				36ms
GET 34fu1qdg.js				259ms
GET 1hqnrwkd.js				35ms
GET 7c5lvnd6.js				18ms
GET 7q88hxyg.js				28ms
GET 5vjds43u.js				27ms
GET 1rlyly26.js				17ms
GET zhyx3g1l.js				
GET 2pwpaji5.js				
GET 59w88zvf.js				
20 requests			150.1 KB	1.18s (onload: 1.7s)

Done

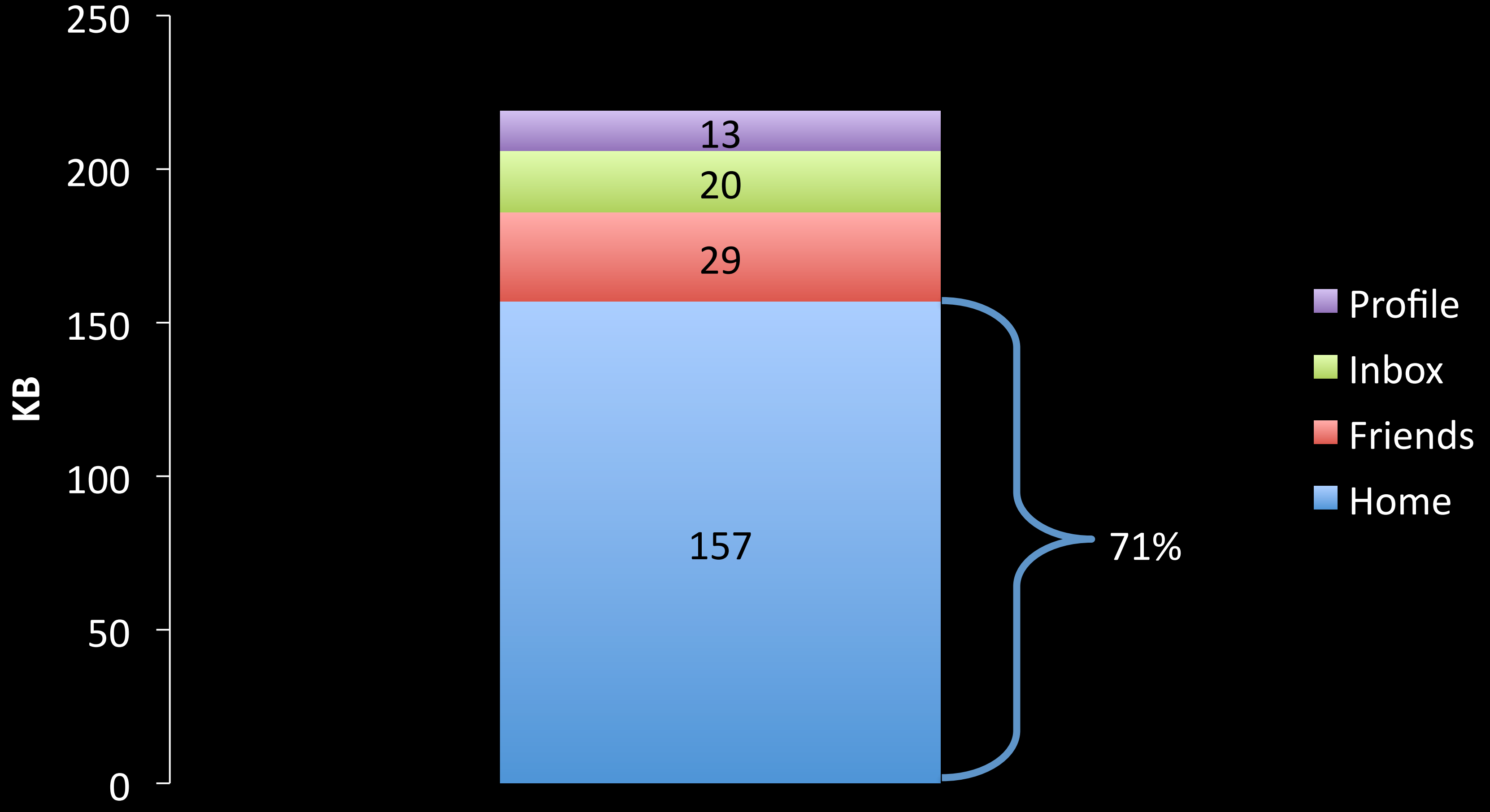
JavaScript programs are streaming

Script Creation

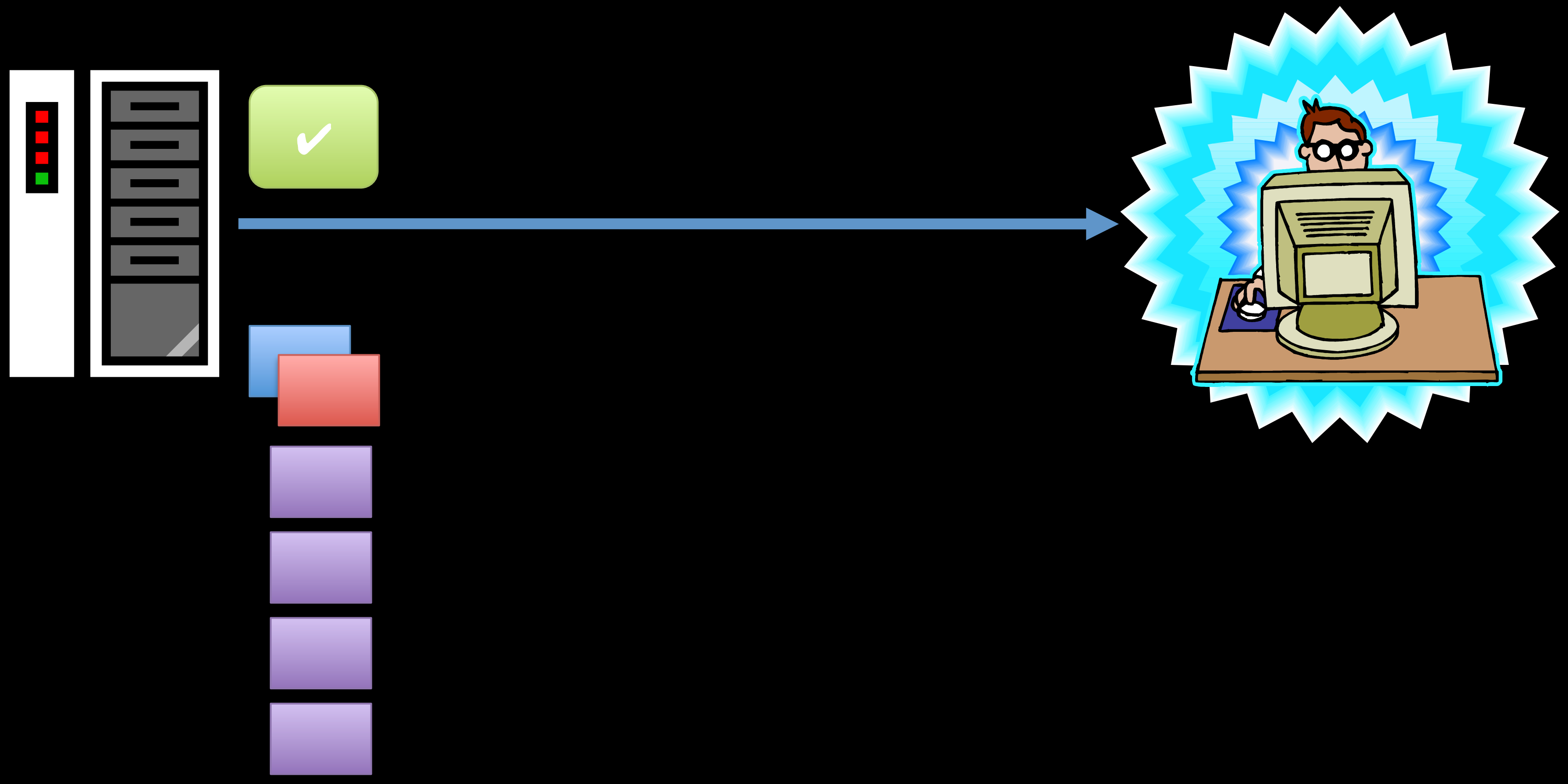
```
<HTML>
  <HEAD>
    <SCRIPT>
      function foo(){...}
      var f = foo;
    </SCRIPT>
    <SCRIPT>
      function bar(){...}
      if (...) f = bar;
    </SCRIPT>
   </HEAD>
  <BODY onclick="f();"> ...</BODY>
</HTML>
```

What does f refer to?

Incremental Loading in Facebook



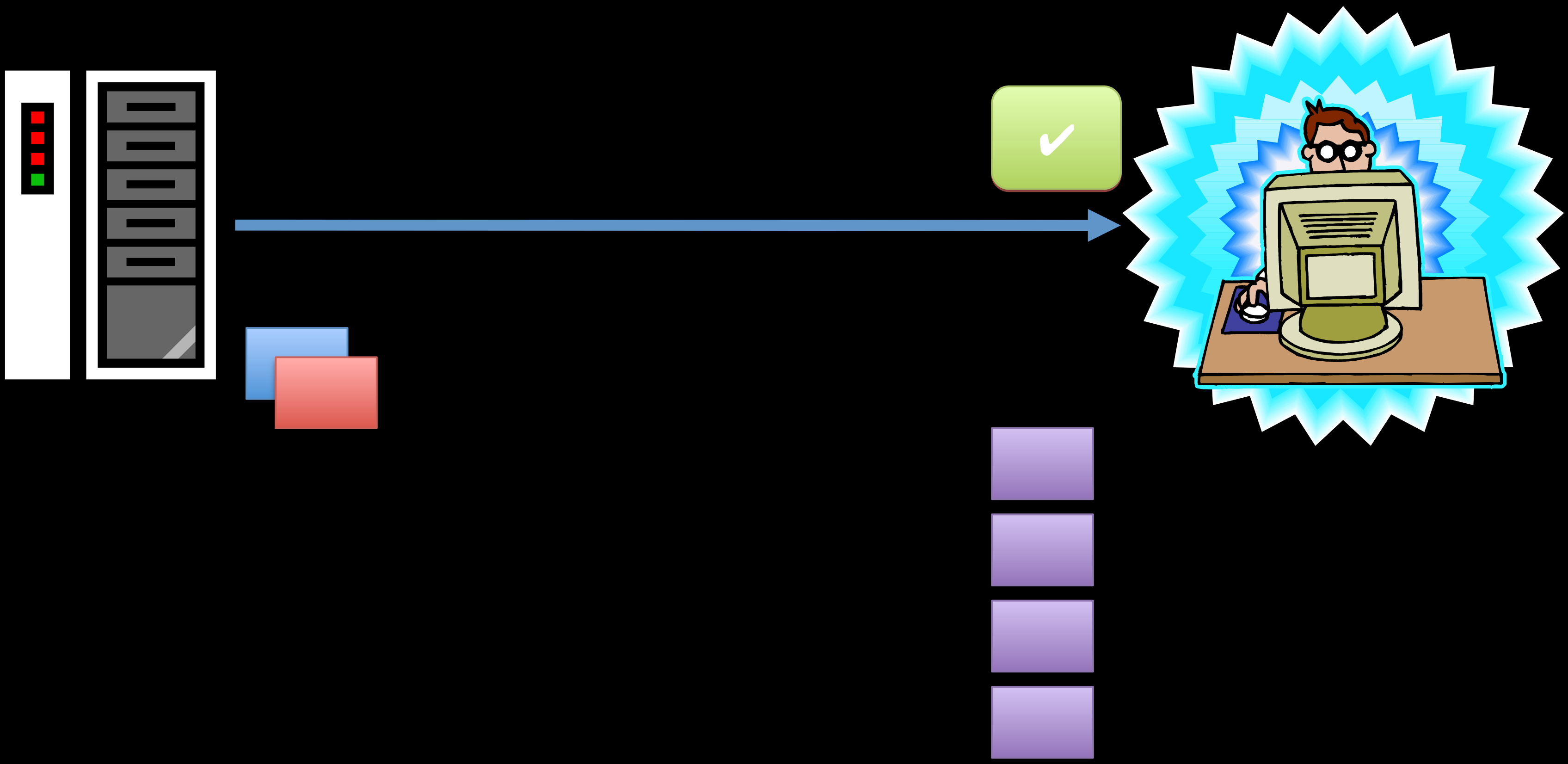
Gulfstream In Action



Offline

Online

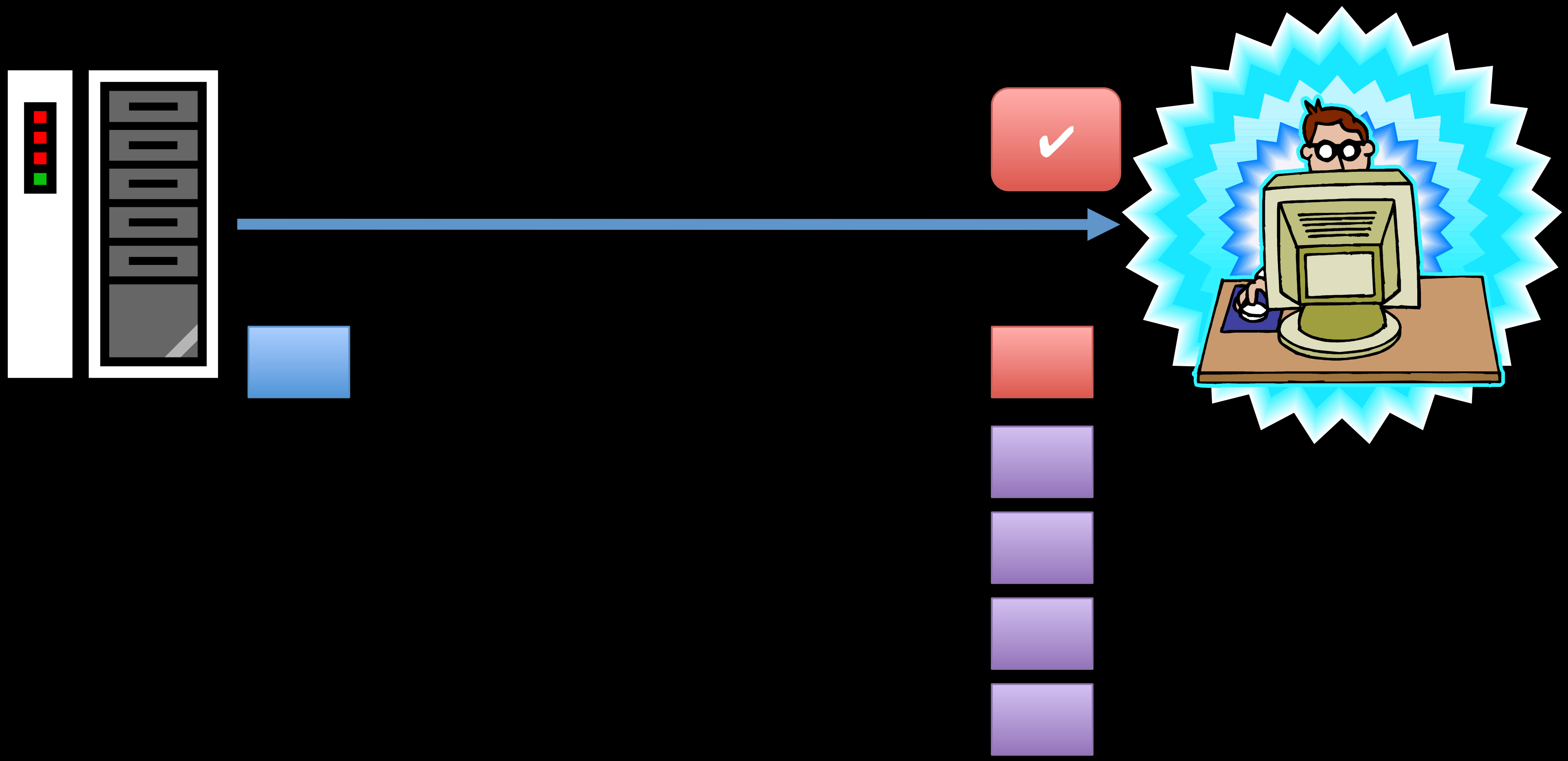
Gulfstream In Action



Offline

Online

Gulfstream In Action



Offline

Online

Outline

- ~~Motivation~~
- Implementation
- Evaluation
- Conclusions

Queries

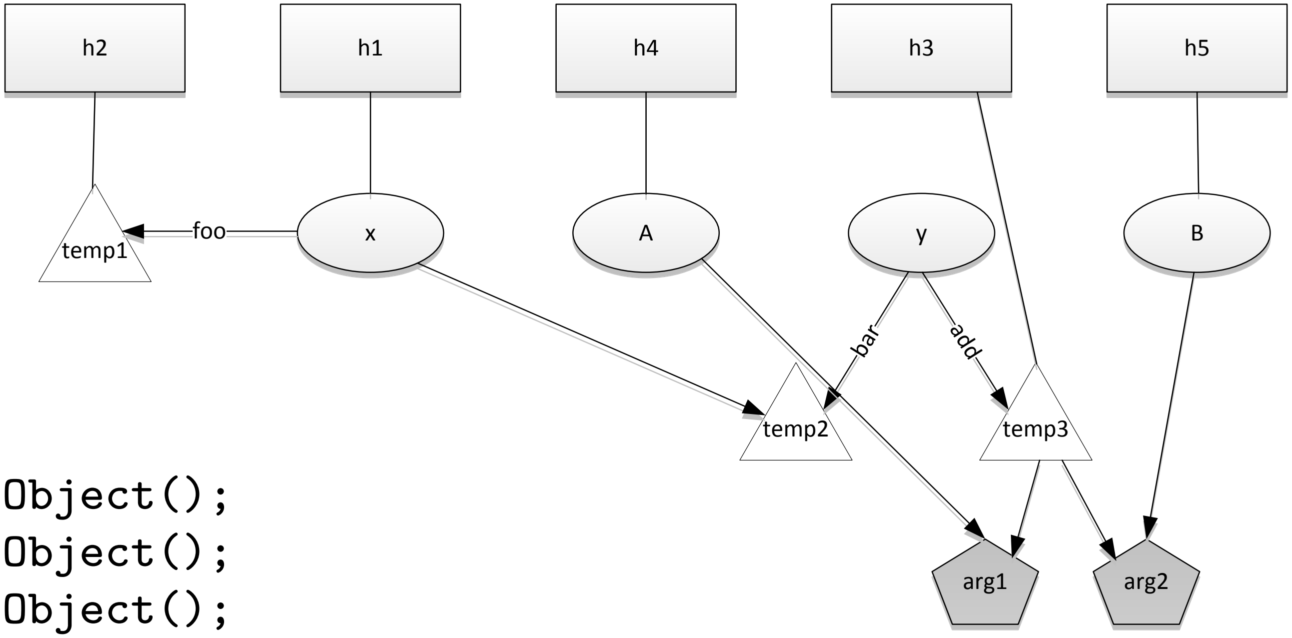
- We want to determine something about the program

- Example
  - What does f() refer to
  - Detect alert calls
  - Does this program use setTimeout

Points-To Analysis

- Provides deep program understanding
- Can be used to construct call graphs
- Is the foundation of further analyses
- Answers a simple question: What heap locations does variable *x* point to

Points-To Example



```
1.  var A = new Object();
2.  var B = new Object();
3.  x     = new Object();
4.  x.foo = new Object();
5.  y     = new Object();
6.  y.bar = x;
7.  y.add = function(a, b) {}
8.  y.add(A, B)
```

Implementation Strategies

Datalog with bddbddb

+ Fast for large programs

+ Highly tuned

- Large startup cost

- Difficult to implement in the browser

- Used in Gatekeeper [USENIX Sec 09]

Graph-based flow analysis

+ Very small startup cost

+ Customized to work with Gulfstream

- Does not scale well

Implementation

- Normalize JavaScript
  - Turn program into a series of simple statements
  - Introduce temporaries as necessary
- Create flow graph – Use normalized program to generate flow constraints
- Serialize flow graph – Encode the flow-graph so online analysis can use it to update results

Implementation Continued

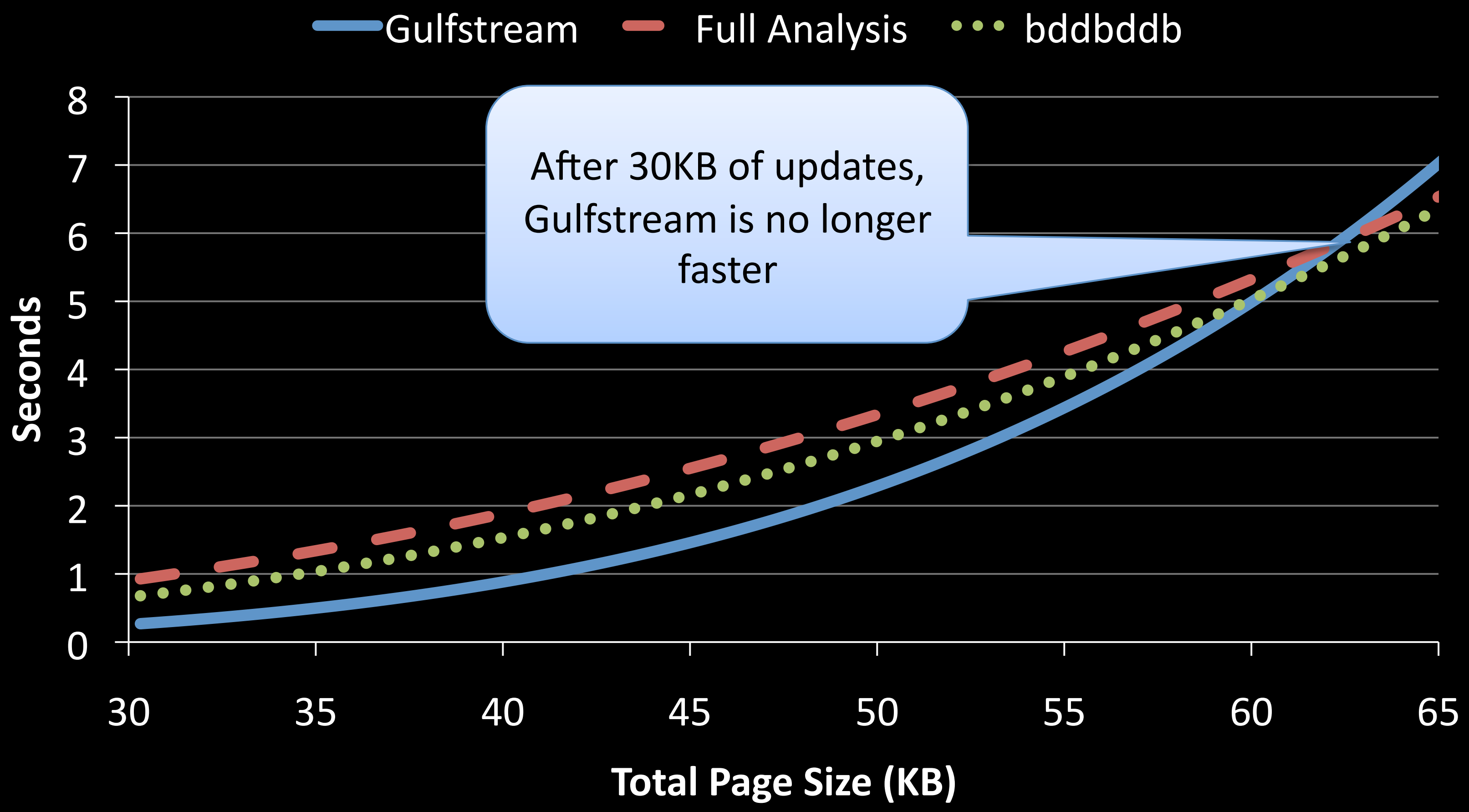
- Perform points-to analysis
  - Traverse flow graph to find all aliases
  - Follow flow through method boundaries
  - Generate points-to map for queries to use

- Queries – Use points-to data and flow graph to answer queries

Evaluation

- Question – Is Gulfstream faster than non-staged analysis
- Benchmarks
  - Synthetically generated
  - Scraped from Google code
  - Scraped from Facebook
- Simulate diverse environments
  - CPU speed and network properties
  - Cell phone, laptop, desktop, etc.

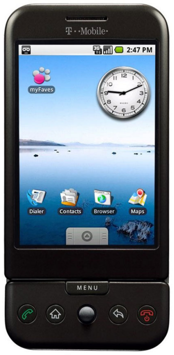
Laptop Running Time Comparison



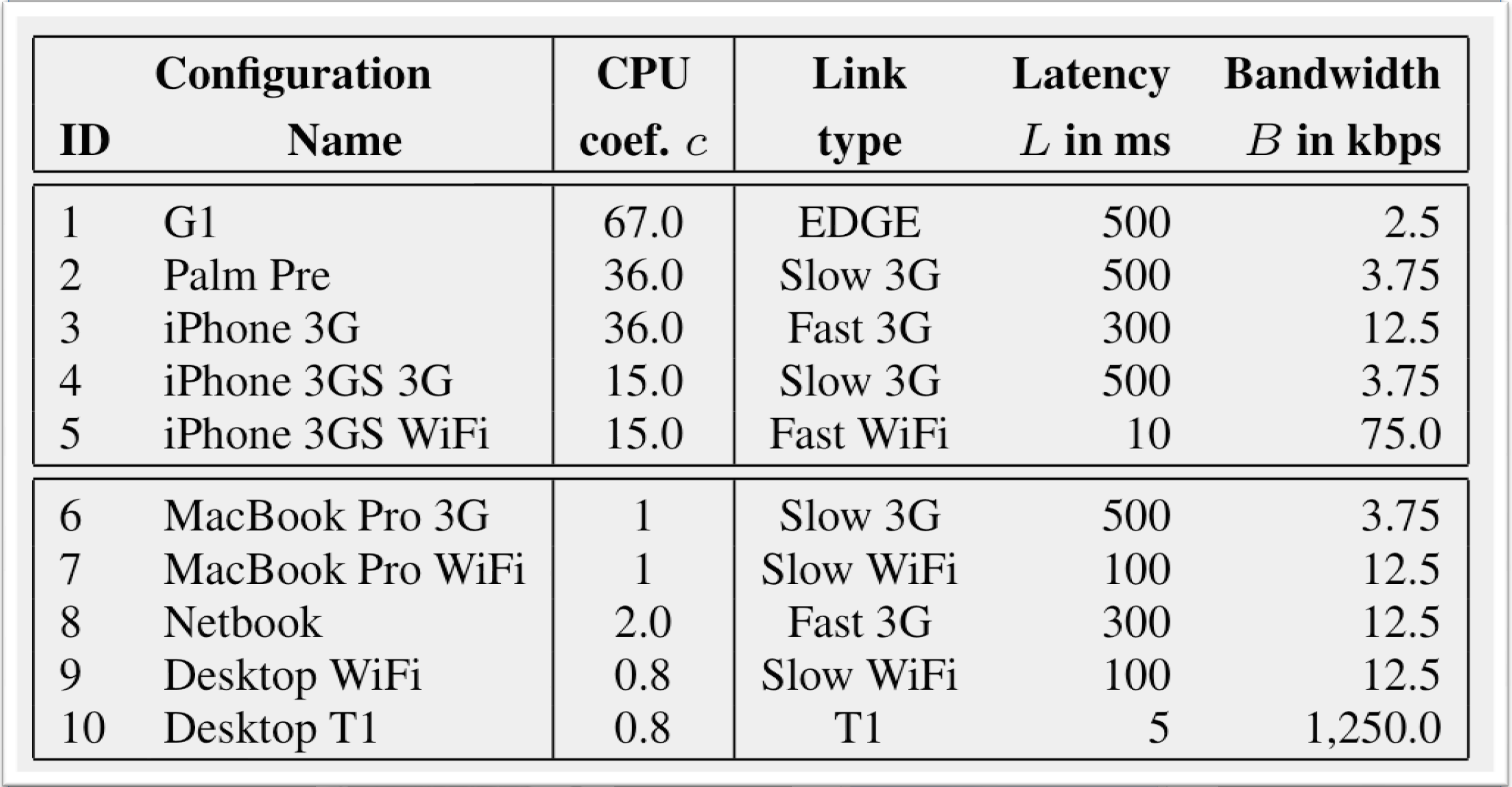
After 30KB of updates, Gulfstream is no longer faster

Simulated Devices

- Low power mobile



- High power



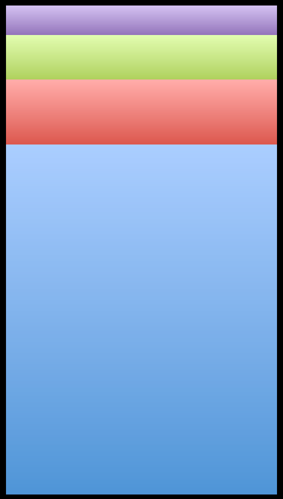
	Configuration	CPU	Link	Latency	Bandwidth
ID	**Name**	**coef.** c	**type**	L **in ms**	B **in kbps**
1	G1	67.0	EDGE	500	2.5
2	Palm Pre	36.0	Slow 3G	500	3.75
3	iPhone 3G	36.0	Fast 3G	300	12.5
4	iPhone 3GS 3G	15.0	Slow 3G	500	3.75
5	iPhone 3GS WiFi	15.0	Fast WiFi	10	75.0
6	MacBook Pro 3G	1	Slow 3G	500	3.75
7	MacBook Pro WiFi	1	Slow WiFi	100	12.5
8	Netbook	2.0	Fast 3G	300	12.5
9	Desktop WiFi	0.8	Slow WiFi	100	12.5
10	Desktop T1	0.8	T1	5	1,250.0

Lessons Learned

- **Slow devices** benefit from Gulfstream
- A **slow network** can negate the benefits of the staged analysis
- **Large page updates** don't benefit from Gulfstream

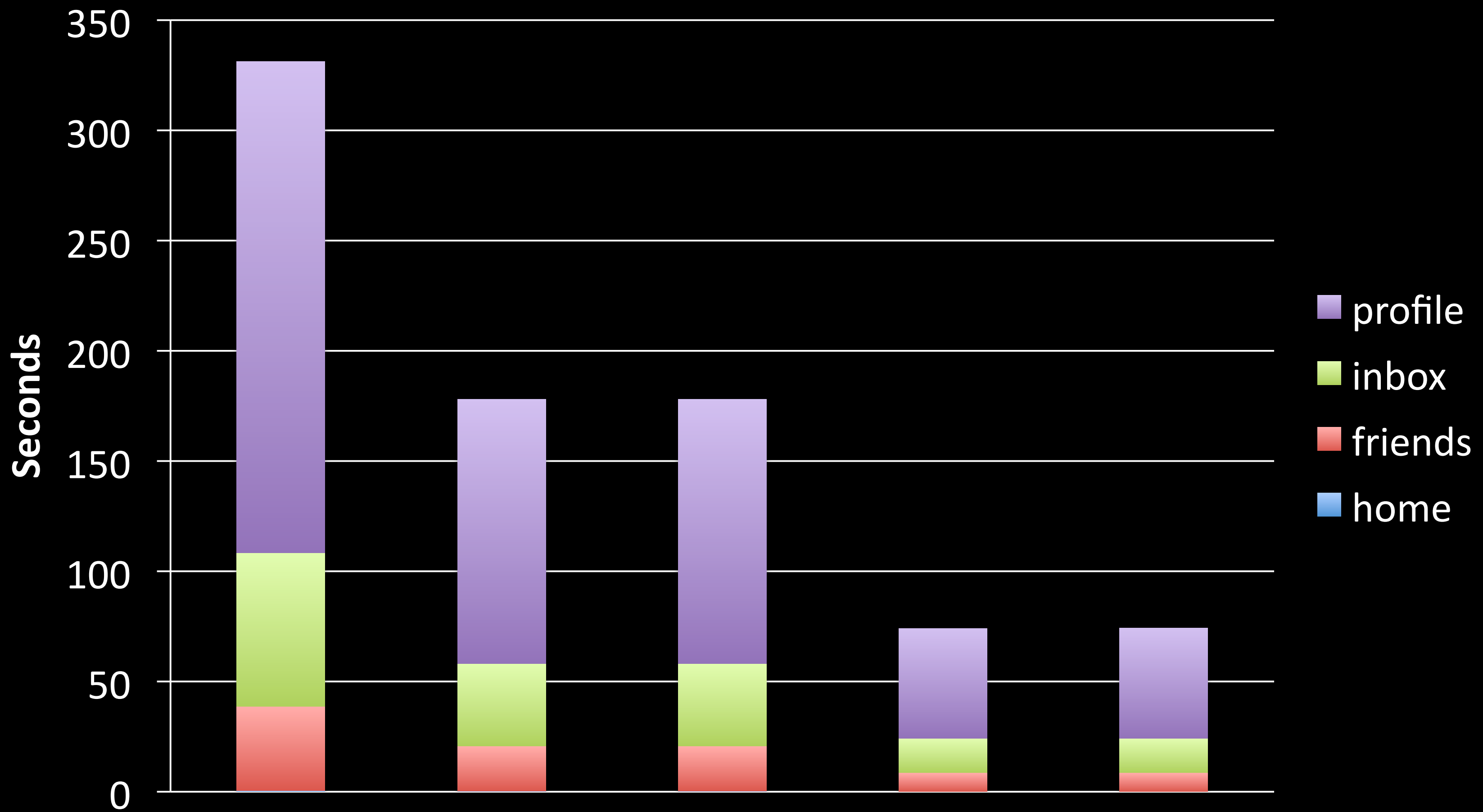
Facebook Experiment

- Visit 4 pages
  - Home
  - Friends
  - Inbox
  - Profile

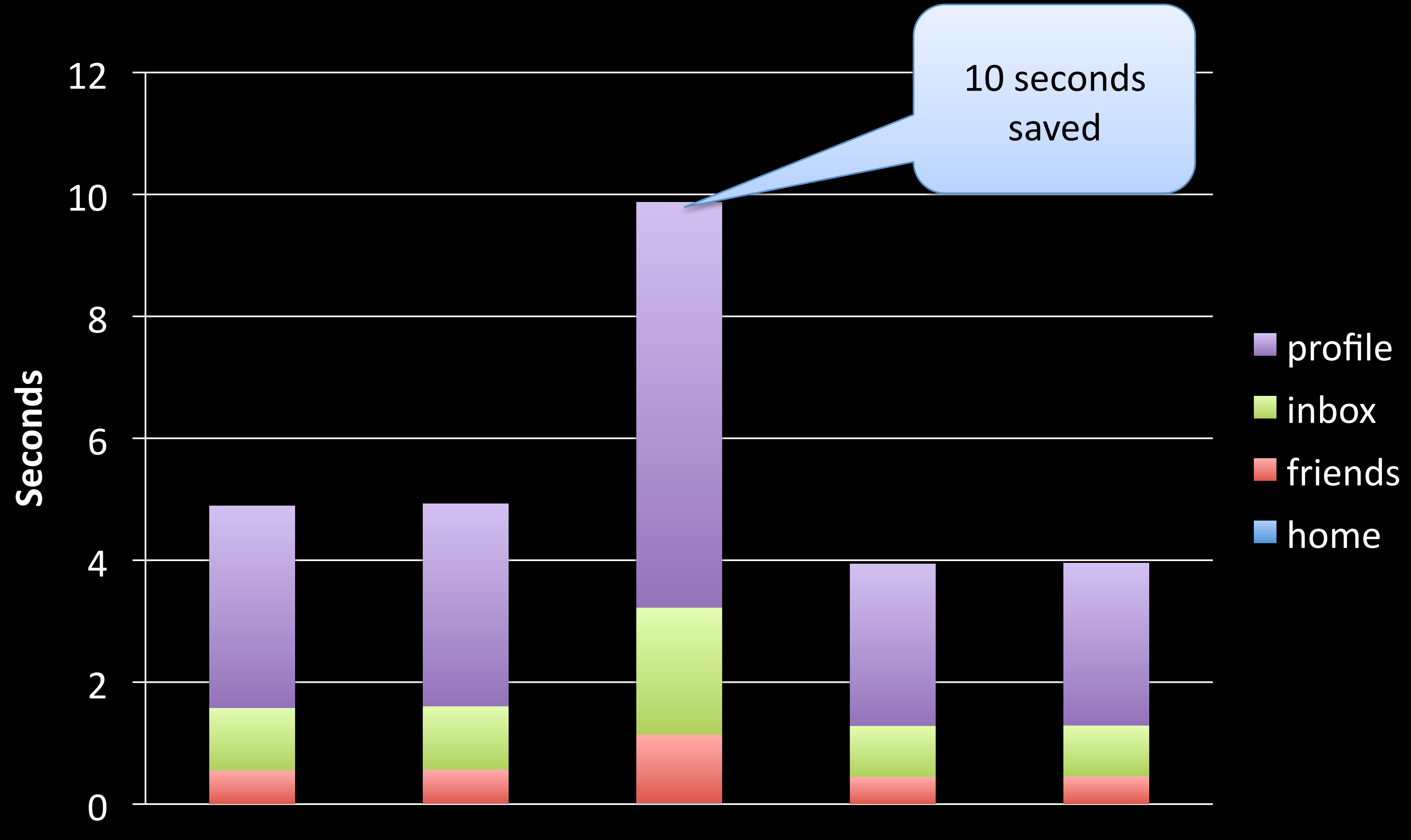


- Each page loads additional JavaScript

Gulfstream Savings: Slow Devices



Gulfstream Savings: Fast Devices



10 seconds saved

Conclusion

- Gulfstream, staged analysis for JavaScript

- Staged analysis
  - Offline on the server
  - Online in the browser

- Wide range of experiments
  - For small updates, Gulfstream is faster
  - Devices with slow CPU benefit most

The End

- Contact: salvatore.guarnieri@gmail.com