

distributed  
tracing

twitter zipkin  
google dapper  
x-trace  
tracelytics  
... more!

motivation

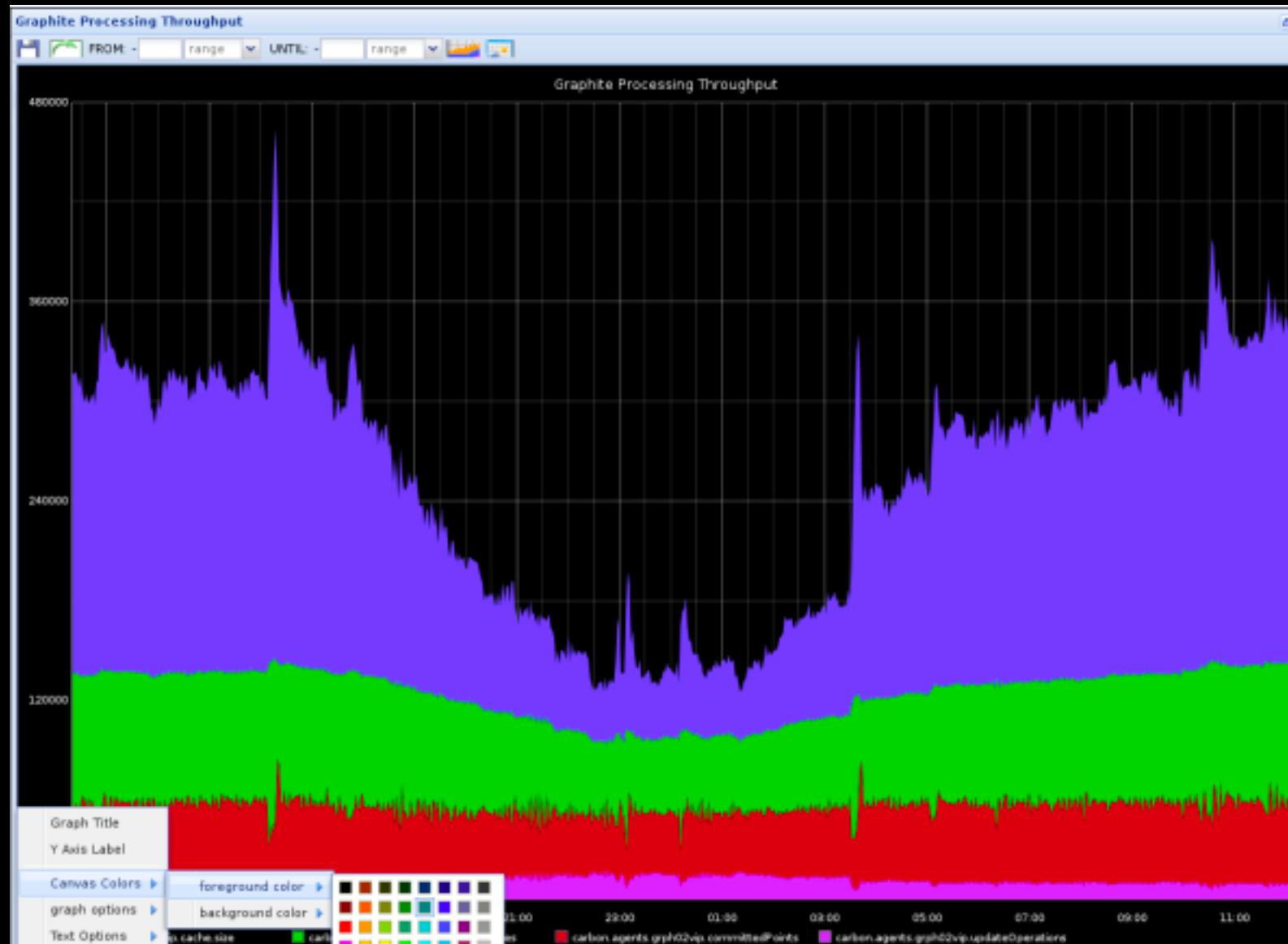


# what is slow?

```
0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:35 +0000] "GET /img/tlyguide/tlyguide_icon.png HTTP/1.1" 304 0 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:36 +0000] "GET /payments/current_plan HTTP/1.1" 200 223 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:36 +0000] "GET /notification/get HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:36 +0000] "GET /cross/layer_count?app=New_App_1&time_window=hour&time_start=1348345444949&time_end=1348347602073 HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:37 +0000] "GET /checks/summary?app=New_App_1&time_window=hour&time_start=1348345444949&time_end=1348347602073 HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:37 +0000] "GET /annotations/get?time_start=1348344810000&time_end=1348348410000&app=New_App_1 HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:37 +0000] "GET /hosts/by_app HTTP/1.1" 200 525 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:37 +0000] "GET /dashboard/layer_host_series?app=New_App_1&time_window=hour&time_start=1348344810000&time_end=1348347602073 HTTP/1.1" 200 600 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:38 +0000] "GET /dashboard/top_urls?sEcho=1&iColumns=9&sColumns=_expand%2Cdomain%2Curl%2Ccount%2Cavg_duration%2Cstddev_duration&iDisplayStart=0&iDisplayLength=10&mDataProp_0=0&mDataProp_1=1&mDataProp_2=2&mDataProp_3=3&mDataProp_4=4&mDataProp_5=5&mDataProp_6=6&mDataProp_7=7&mDataProp_8=8&mDataProp_9=9&sSearch=&bRegex=false&bSearchable_0=false&bSearchable_1=true&bSearchable_2=true&bSearchable_3=true&bSearchable_4=true&bSearchable_5=false&bSearchable_6=true&bSearchable_7=true&bSearchable_8=true&bSearchable_9=true&iSortingCols=1&iSortCol_0=5&sSortDir_0=desc&bSortable_0=false&bSortable_1=false&bSortable_2=true&bSortable_3=true&bSortable_4=true&bSortable_5=true&bSortable_6=true&bSortable_7=true&bSortable_8=true&bSortable_9=true&app=New_App_1&time_start=1348345444949&time_end=1348347602073 HTTP/1.1" 200 600 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:38 +0000] "GET /dashboard/top_controllers?sEcho=1&iColumns=9&sColumns=_expand%2Ccontroller%2Caction%2Ccount%2Cavg_duration%2Cstddev_duration&iDisplayStart=0&iDisplayLength=10&mDataProp_0=0&mDataProp_1=1&mDataProp_2=2&mDataProp_3=3&mDataProp_4=4&mDataProp_5=5&mDataProp_6=6&mDataProp_7=7&mDataProp_8=8&mDataProp_9=9&sSearch=&bRegex=false&bSearchable_0=false&bSearchable_1=true&bSearchable_2=true&bSearchable_3=true&bSearchable_4=true&bSearchable_5=false&bSearchable_6=true&bSearchable_7=true&bSearchable_8=true&bSearchable_9=true&iSortingCols=1&iSortCol_0=5&sSortDir_0=desc&bSortable_0=false&bSortable_1=false&bSortable_2=true&bSortable_3=true&bSortable_4=true&bSortable_5=true&bSortable_6=true&bSortable_7=true&bSortable_8=true&bSortable_9=true&app=New_App_1&time_start=1348345444949&time_end=1348347602073 HTTP/1.1" 200 225 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:38 +0000] "GET /dashboard/heatmap_series?app=New_App_1&time_window=hour&time_start=1348344810000&time_end=1348347602073 HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /favicon.ico HTTP/1.1" 304 0 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /dashboard/top_queries?sEcho=1&iColumns=10&sColumns=_expand%2Cfp%2Cquery_fp_id%2Clayer%2Ccount%2Cmin_duration%2Cstddev_duration&iDisplayStart=0&iDisplayLength=10&mDataProp_0=0&mDataProp_1=1&mDataProp_2=2&mDataProp_3=3&mDataProp_4=4&mDataProp_5=5&mDataProp_6=6&mDataProp_7=7&mDataProp_8=8&mDataProp_9=9&sSearch=&bRegex=false&bSearchable_0=false&bSearchable_1=true&bSearchable_2=true&bSearchable_3=true&bSearchable_4=true&bSearchable_5=false&bSearchable_6=true&bSearchable_7=true&bSearchable_8=false&bSearchable_9=true&iSortingCols=1&iSortCol_0=6&sSortDir_0=desc&bSortable_0=false&bSortable_1=false&bSortable_2=true&bSortable_3=true&bSortable_4=true&bSortable_5=true&bSortable_6=true&bSortable_7=true&bSortable_8=true&bSortable_9=true&app=New_App_1&time_start=1348345444949&time_end=1348347602073 HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /img/viz/contextgraph-left-arrow.png HTTP/1.1" 304 0 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /img/viz/contextgraph-right-arrow.png HTTP/1.1" 304 0 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /memcache/stats?app=New_App_1&time_start=1348345444949&time_end=1348347602073&op=get HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /ajax/layer/partition/?app=New_App_1&time_start=1348345444949&time_end=1348347602073&order=tott HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /img/backgrounds/loading-overlay.png HTTP/1.1" 304 0 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /hosts/hosts_series?app=New_App_1&time_window=hour&time_start=1348344810000&time_end=1348348410000 HTTP/1.1" 200 48 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:39 +0000] "GET /hosts/available HTTP/1.1" 200 473 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
n.tracelytics.com tracey [23/Sep/2012:14:26:46 +0000] "GET /hosts/by_app HTTP/1.1" 200 520 "http://internal.brian.tracelytics.com/bb" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.8; rv:15.0) Gecko/20100101 Firefox/15.0.1"
```



# what is slow?



[illegible][illegible]

# causal flow of control

[illegible][illegible]

how to



# possible approaches

# possible approaches

- **Unique identifier**

# possible approaches

- **Unique identifier**
  - propagate throughout



# possible approaches

- **Unique identifier**
  - propagate throughout
  - write instrumentation for various transports

# possible approaches

- **Unique identifier**
  - propagate throughout
  - write instrumentation for various transports

# possible approaches

- **Unique identifier**
  - propagate throughout
  - write instrumentation for various transports
- **Observe and correlate**



# possible approaches

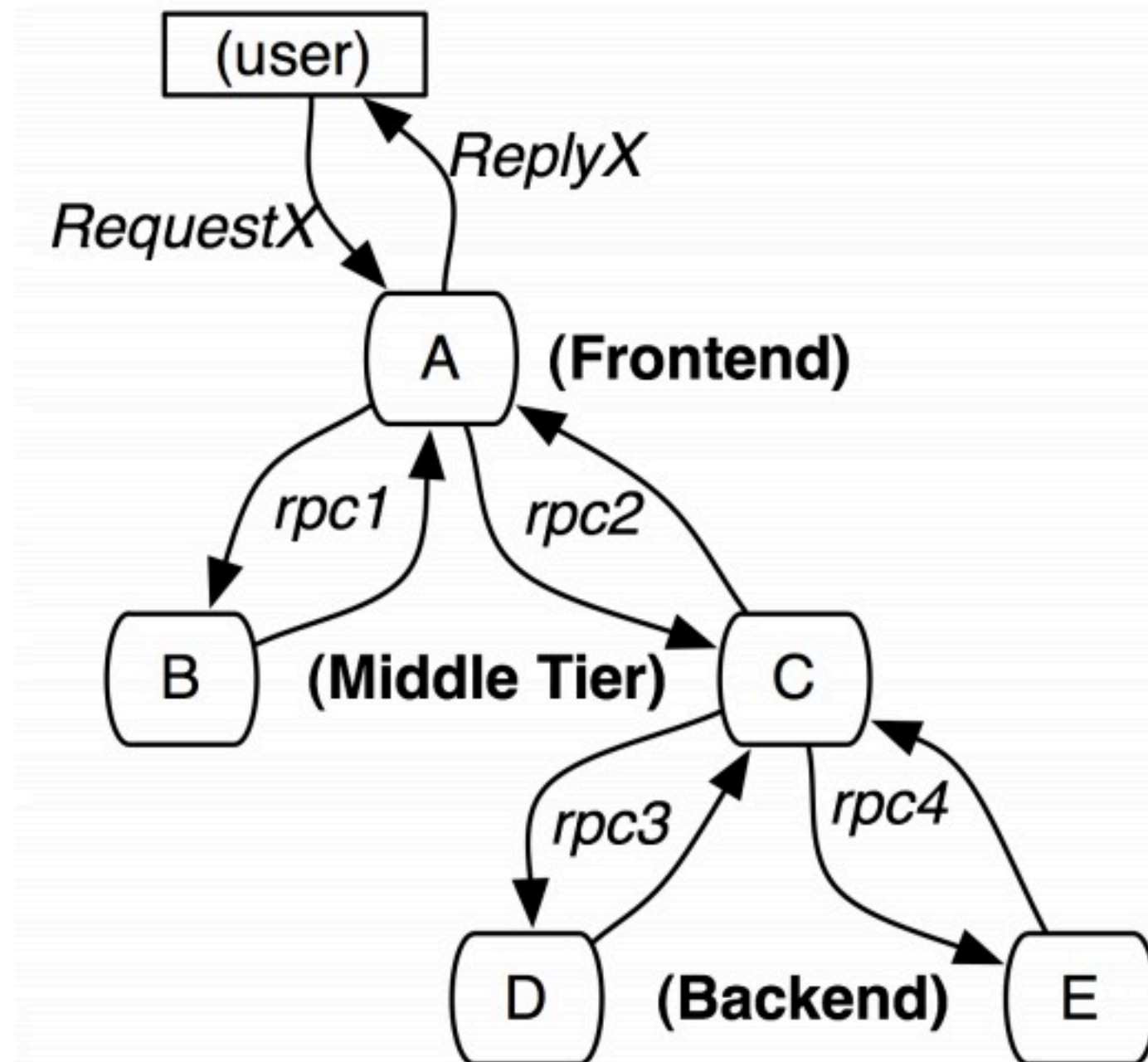
- **Unique identifier**
  - propagate throughout
  - write instrumentation for various transports
- **Observe and correlate**
  - always on the outside - black box

# possible approaches

- **Unique identifier**
  - propagate throughout
  - write instrumentation for various transports
- **Observe and correlate**
  - always on the outside - black box
  - difficult to get threaded + evented processes right

1BD57B58AE7E315BBEA  
B6795F0BDC198296357





**Figure 1: The path taken through a simple serving system on behalf of user request *X*. The letter-labeled nodes represent processes in a distributed system.**

t = start

nginx

cache

python

db

internet

the java

t = start

↓  
nginx

cache

python

db

internet

the java



t = start

nginx

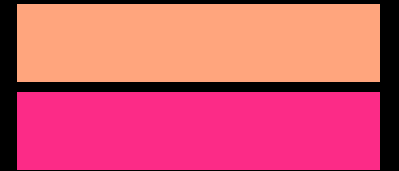
python

cache

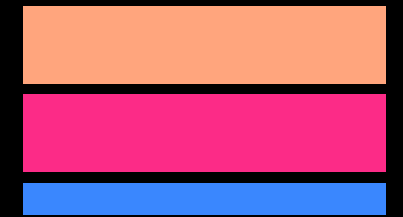
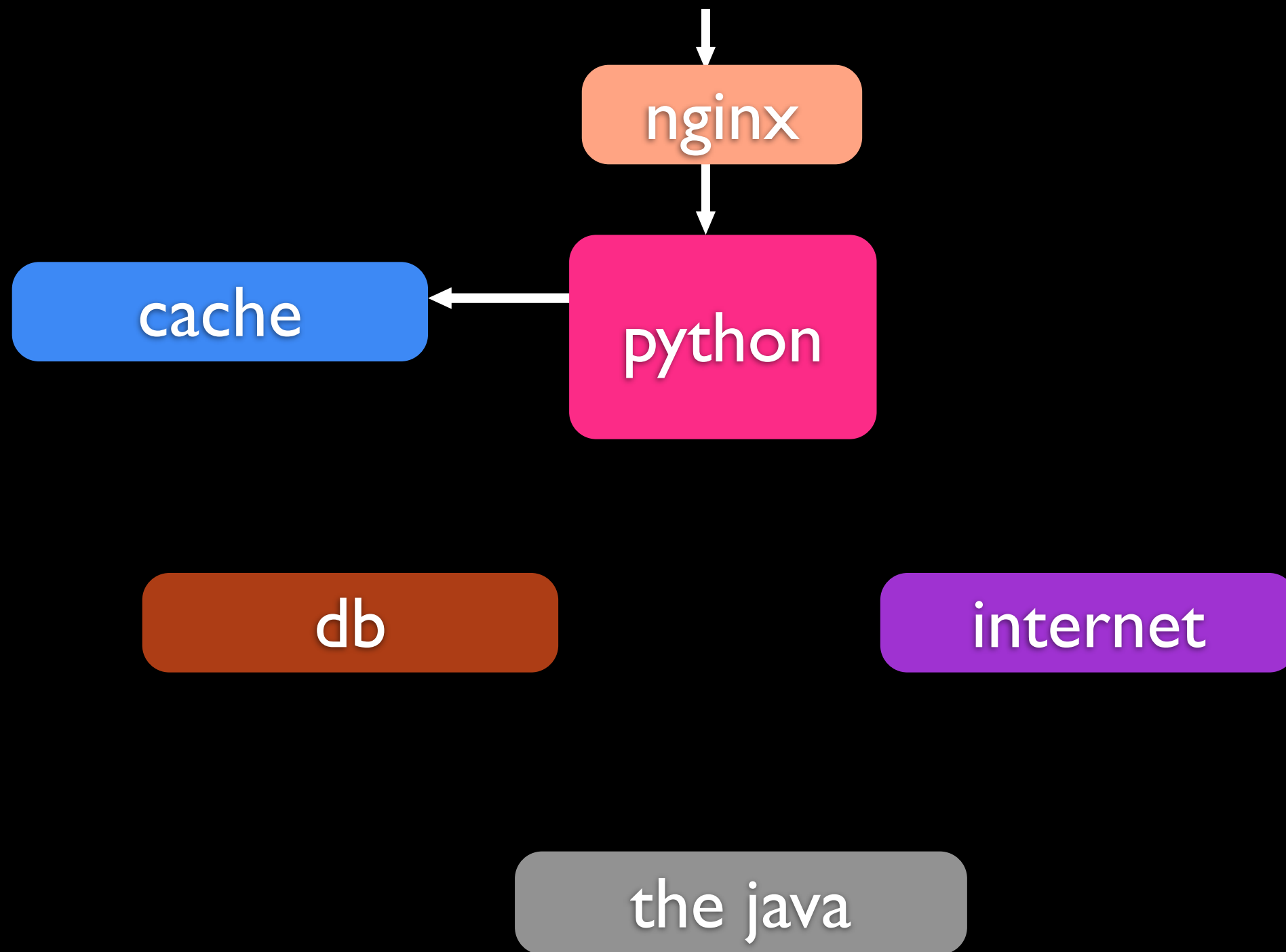
db

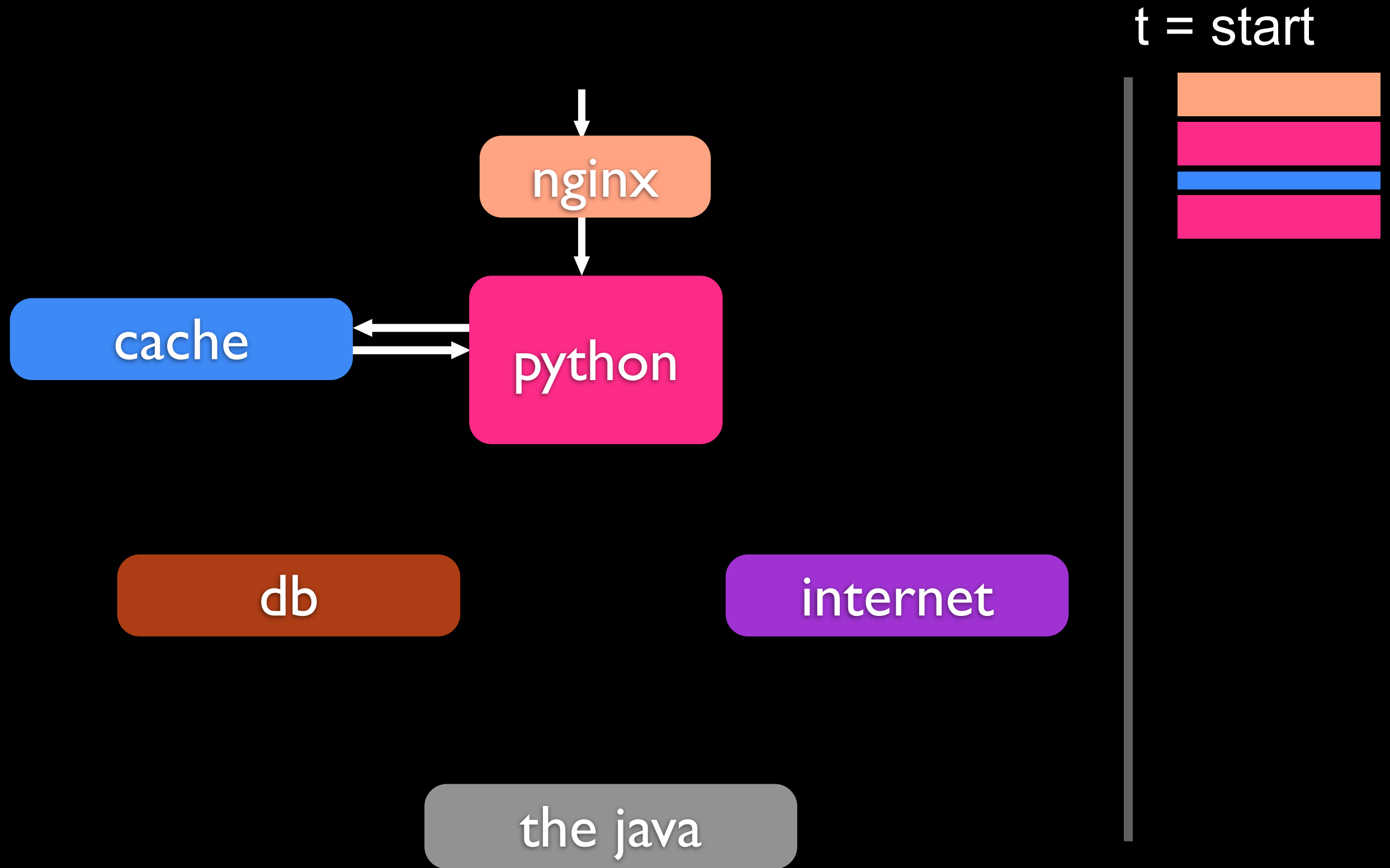
internet

the java

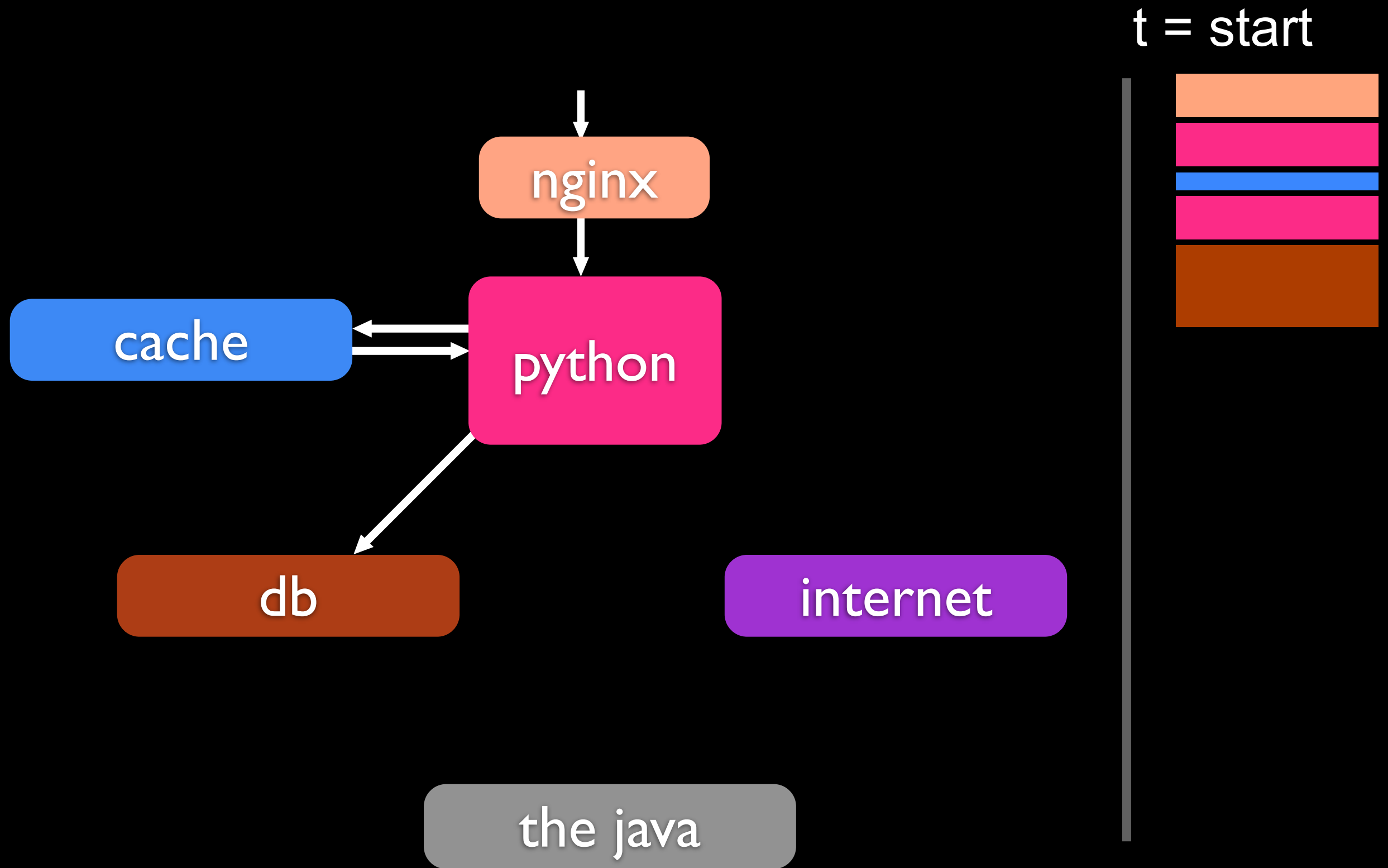


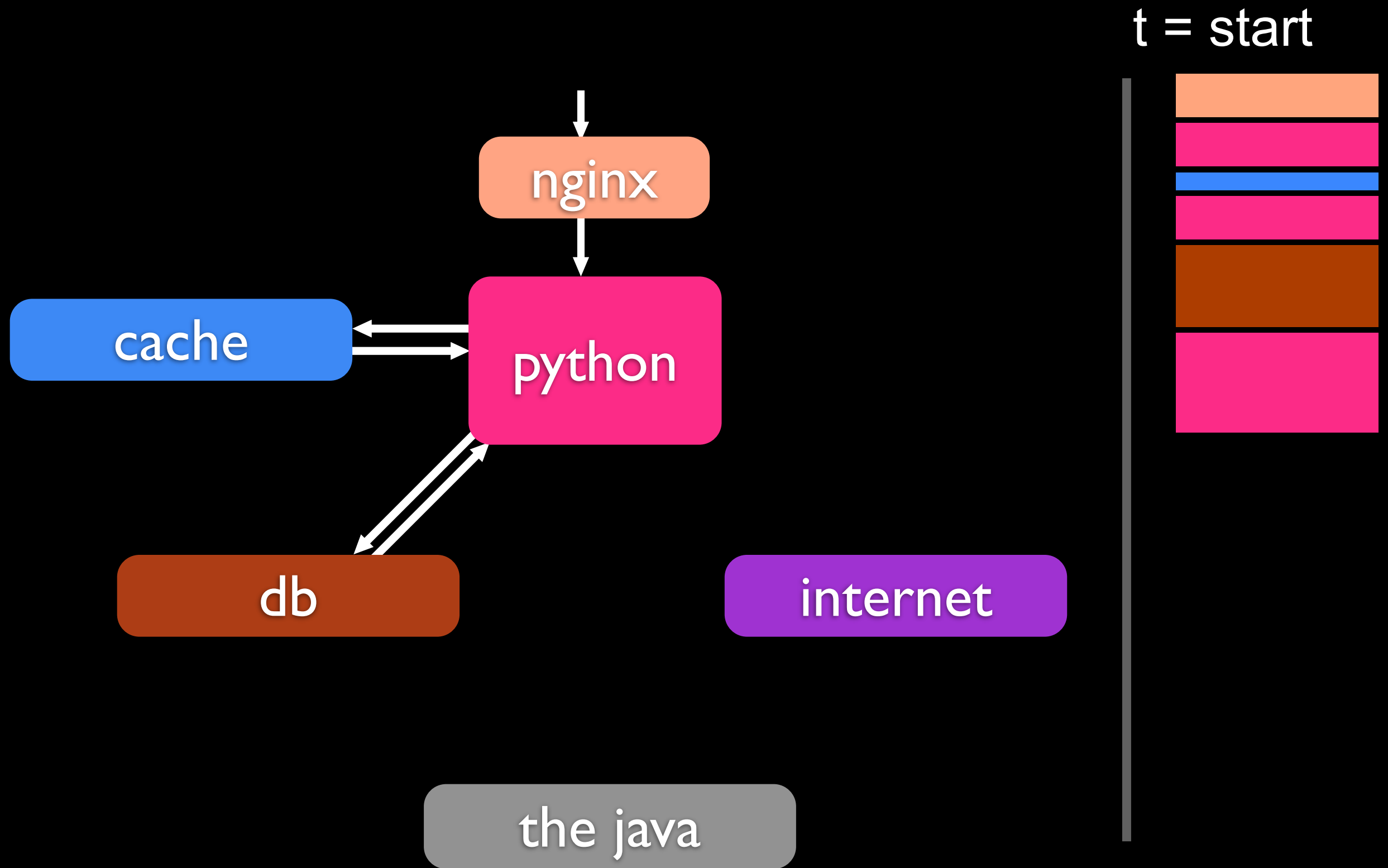
t = start

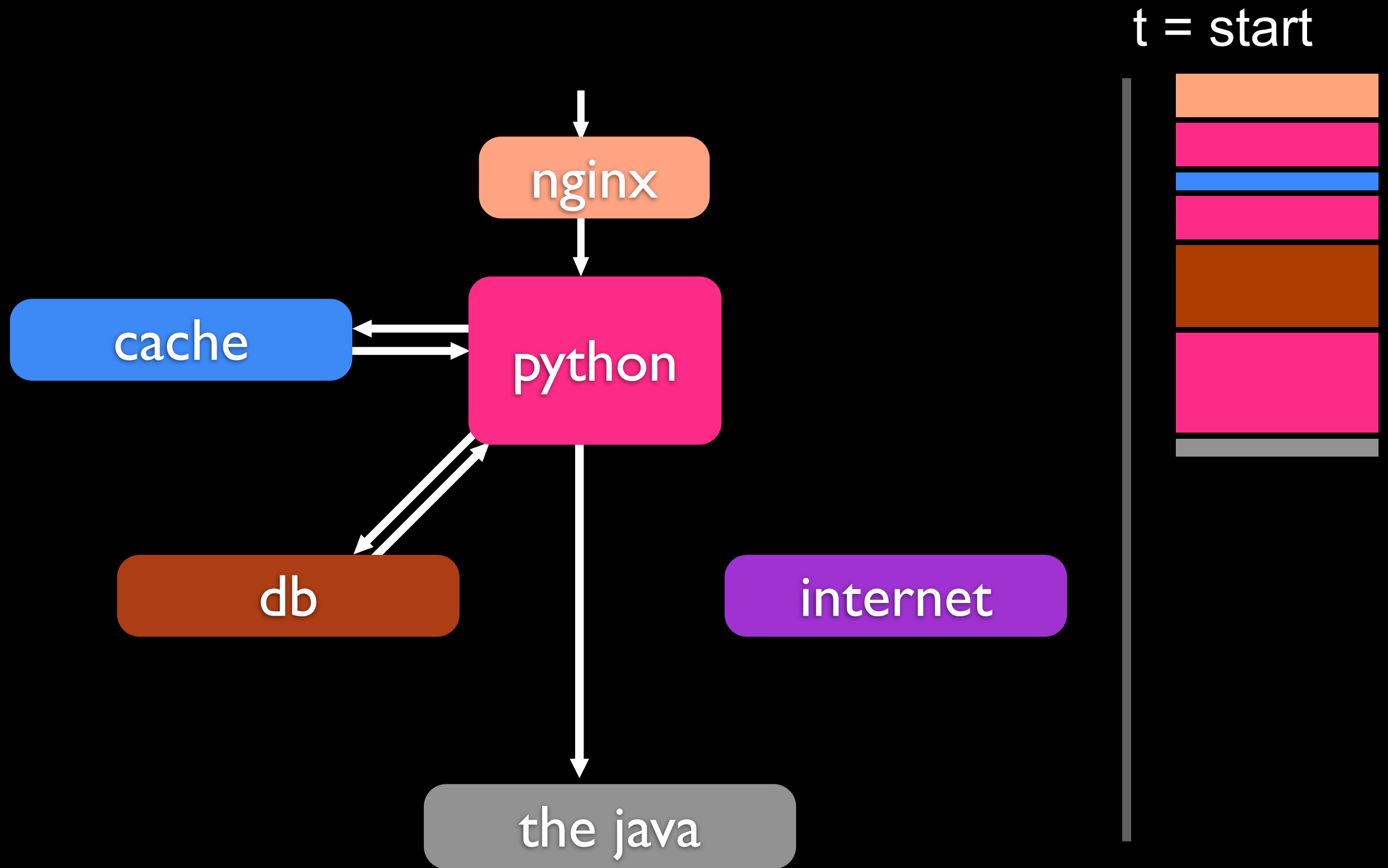


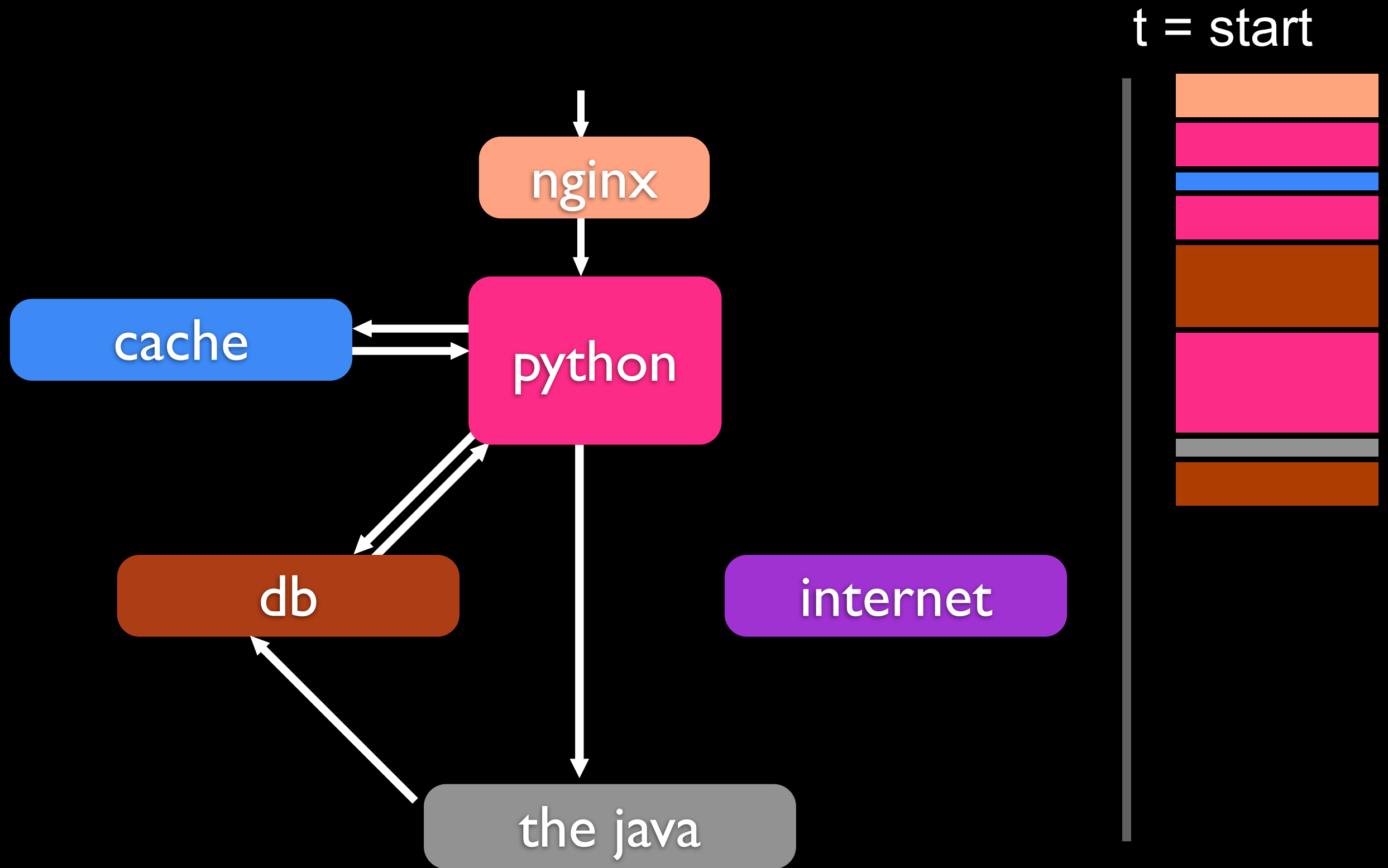


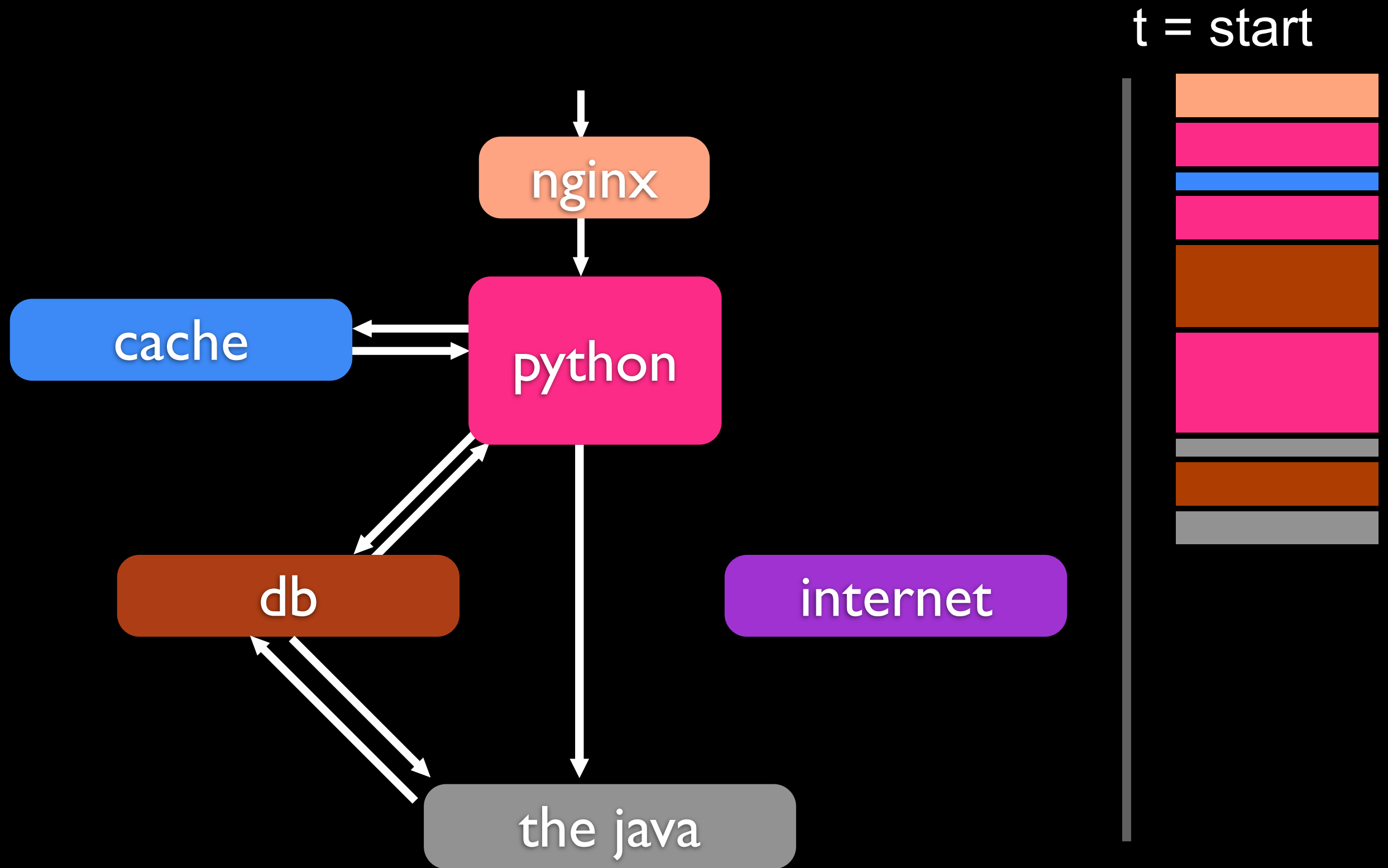


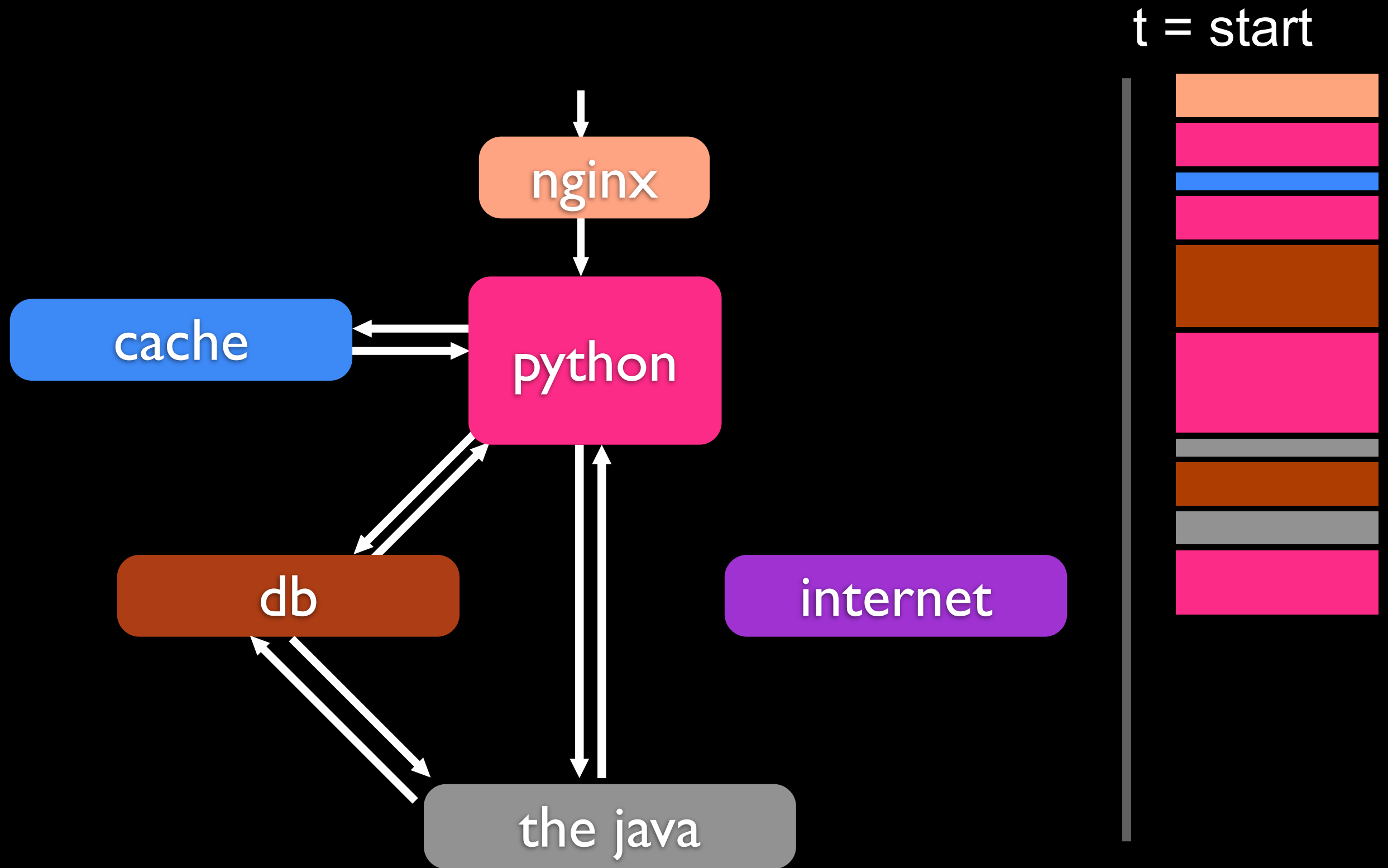




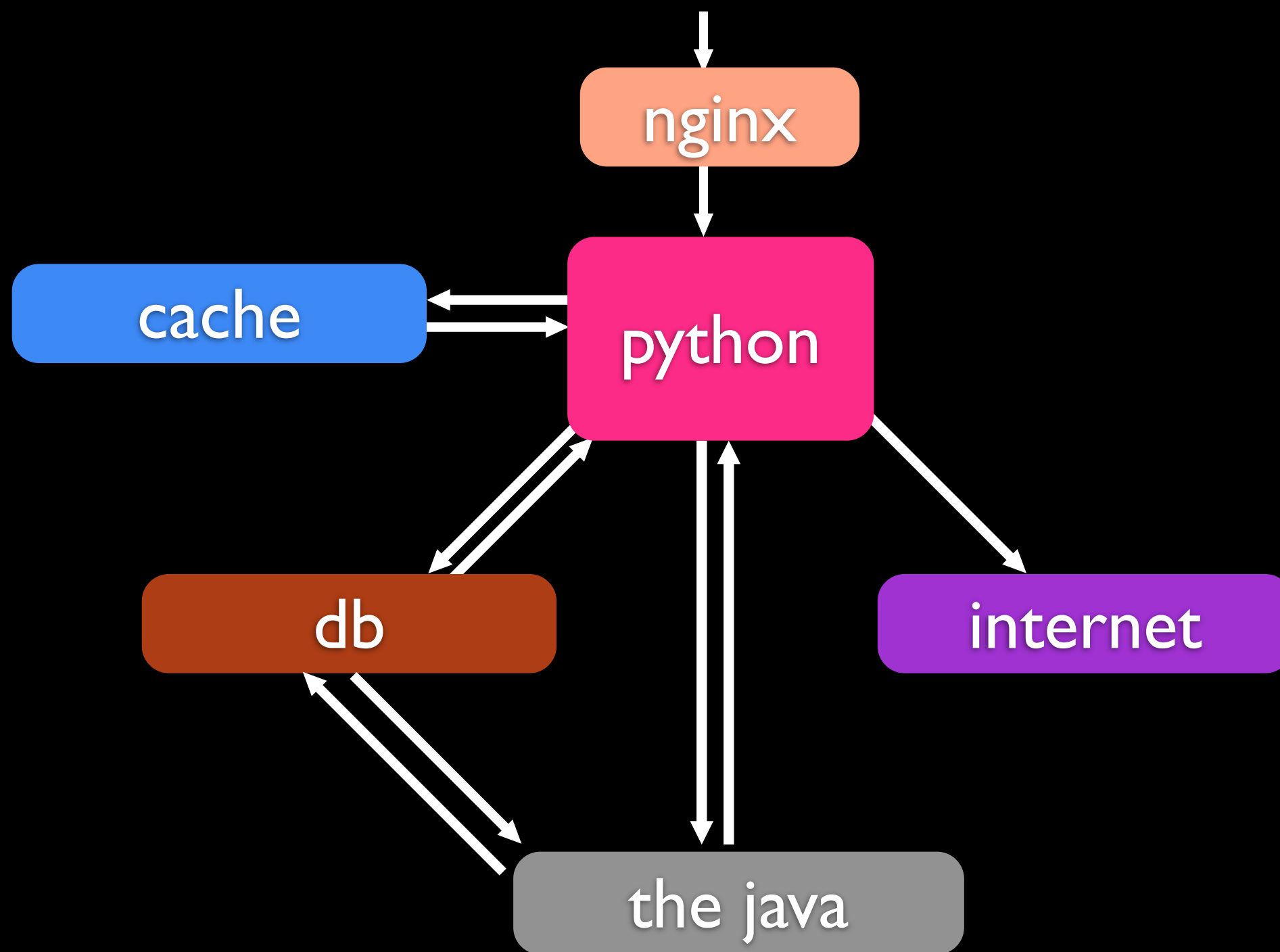




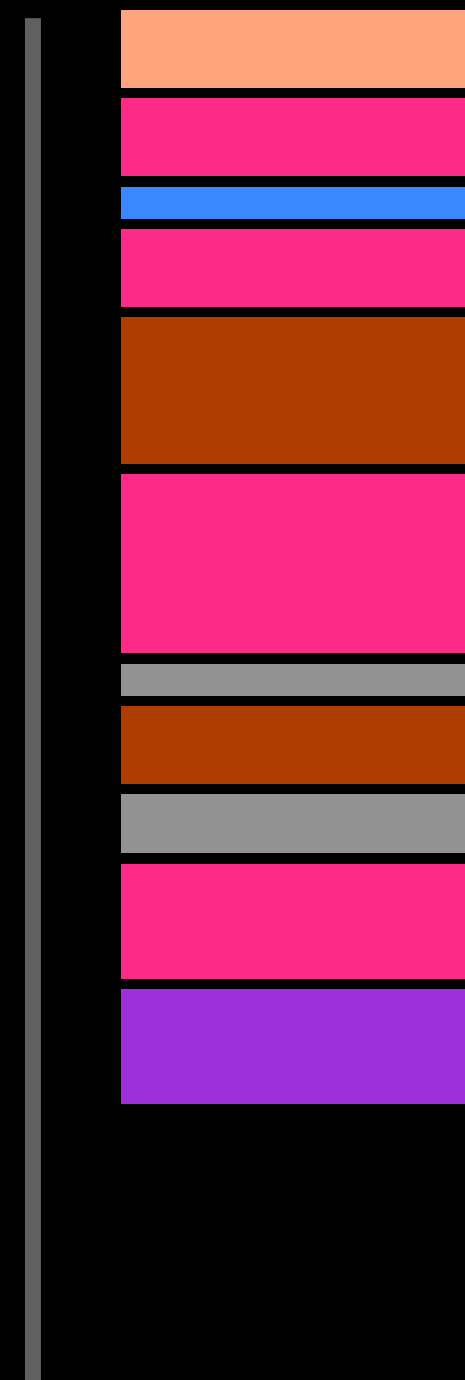


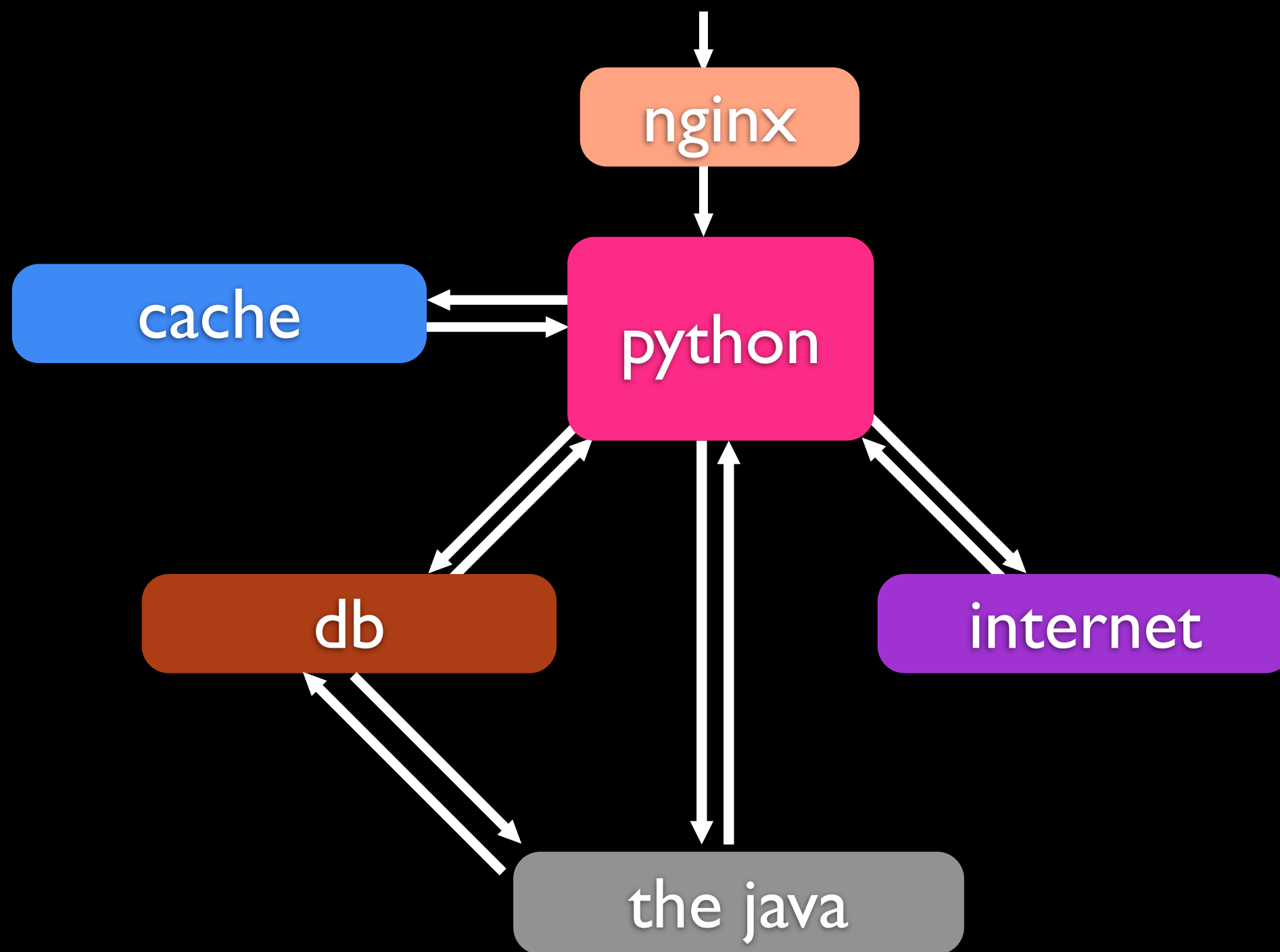




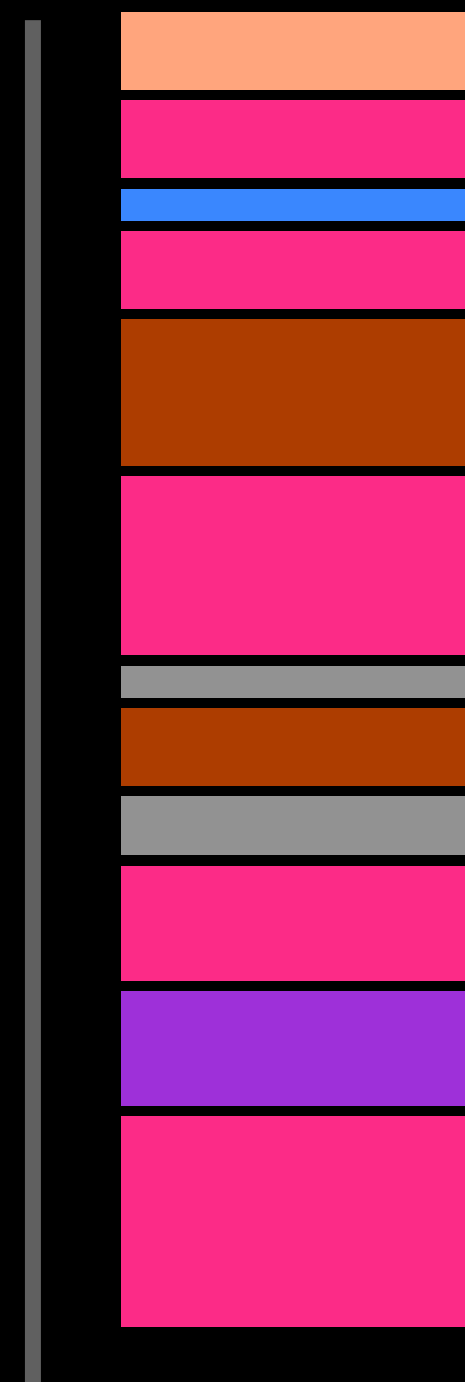


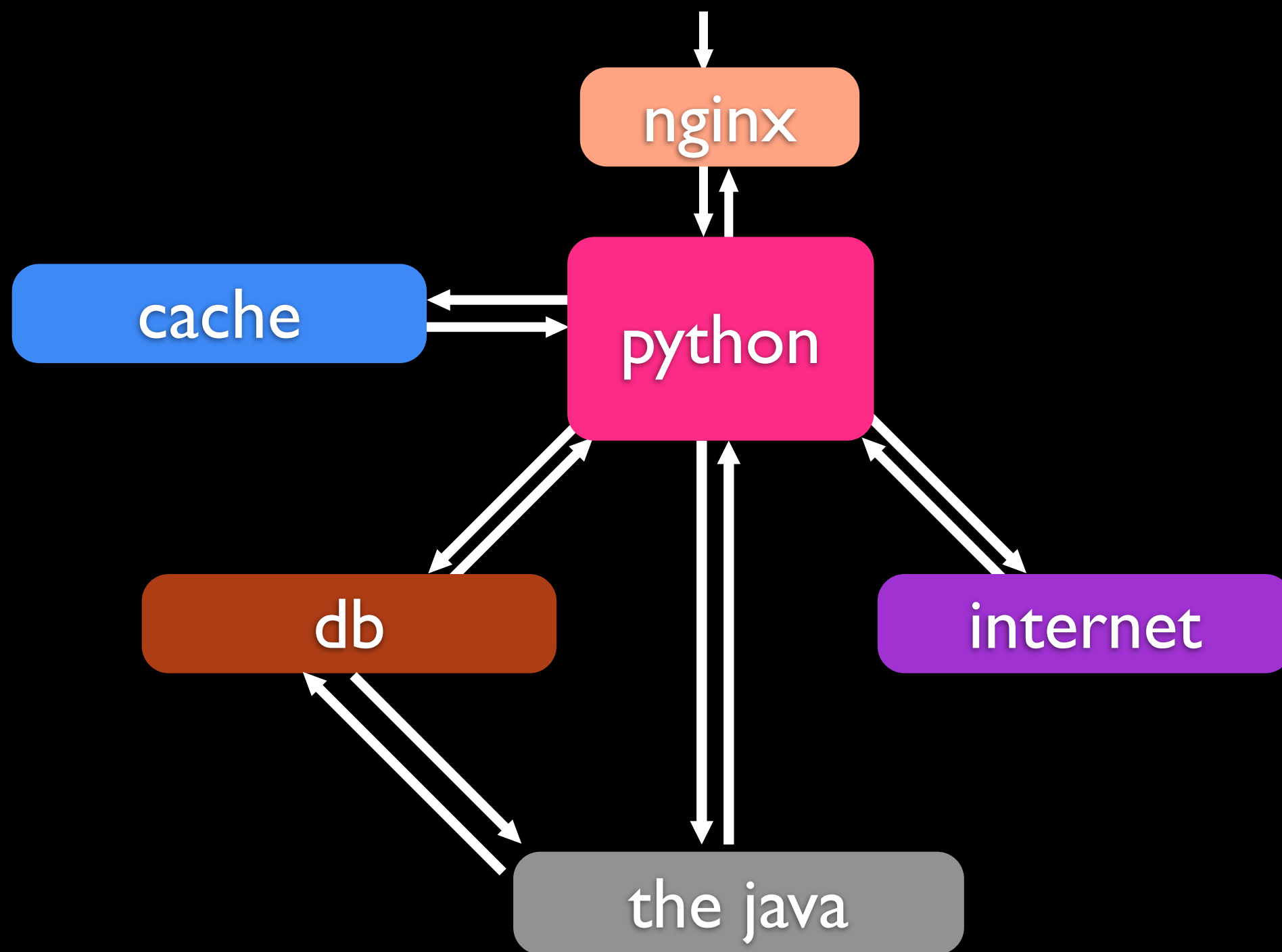
t = start



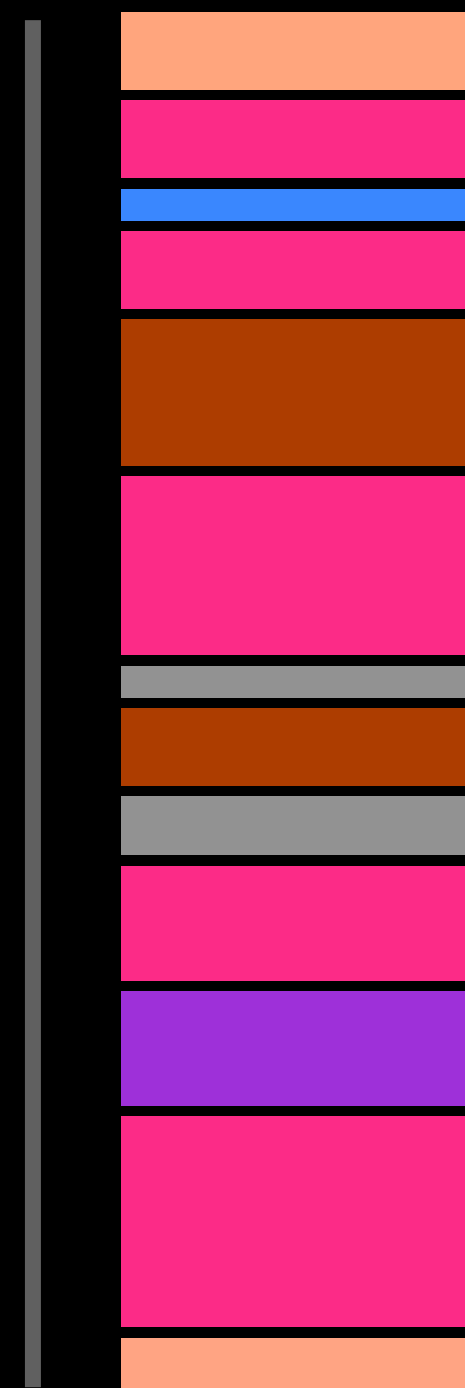


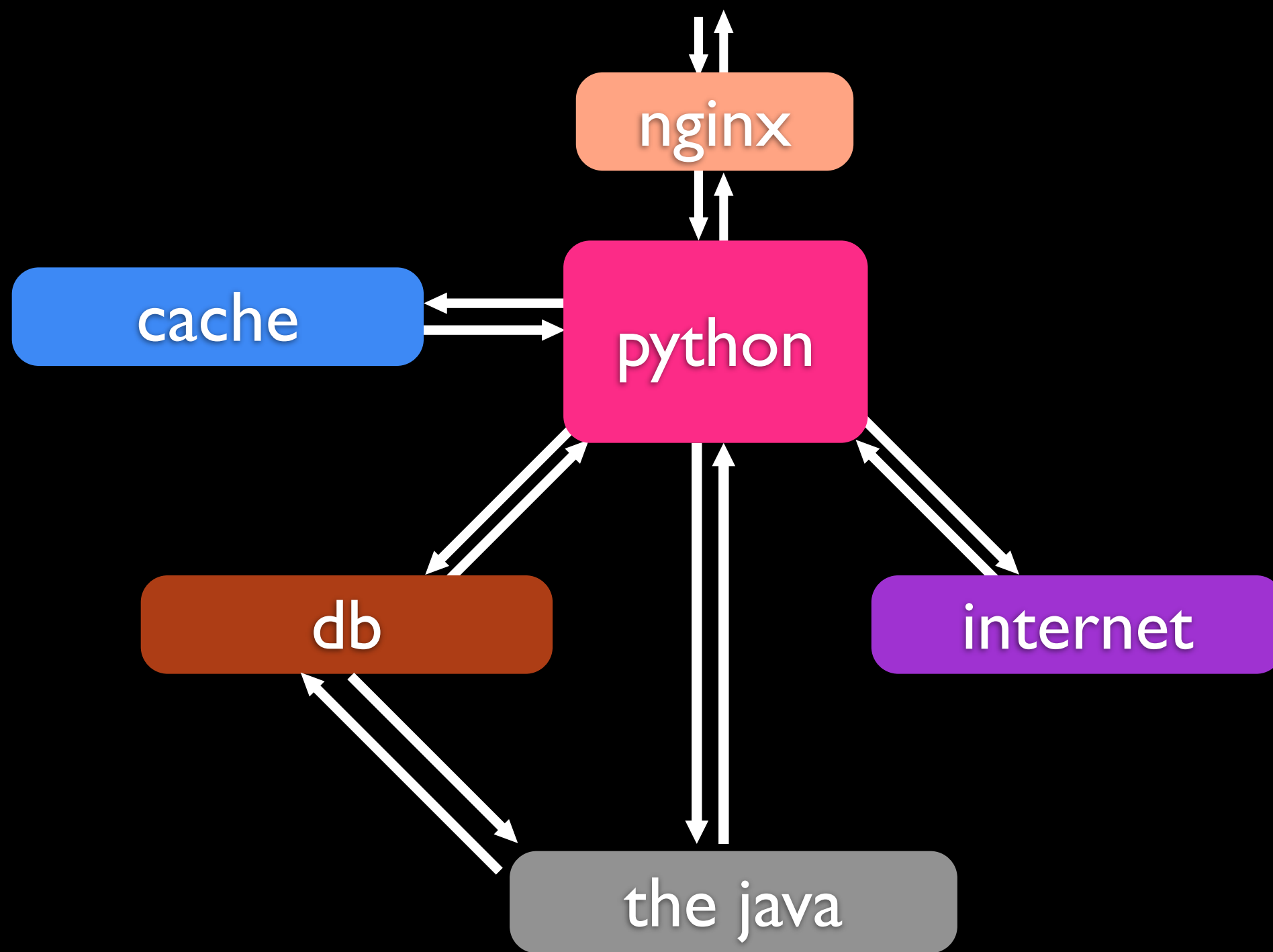
t = start



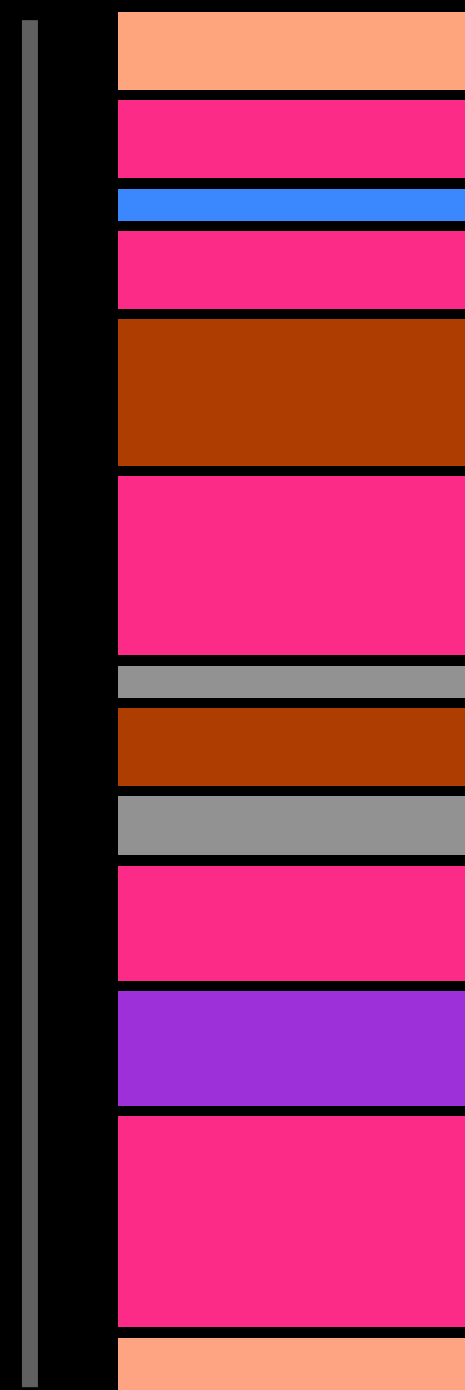


t = start





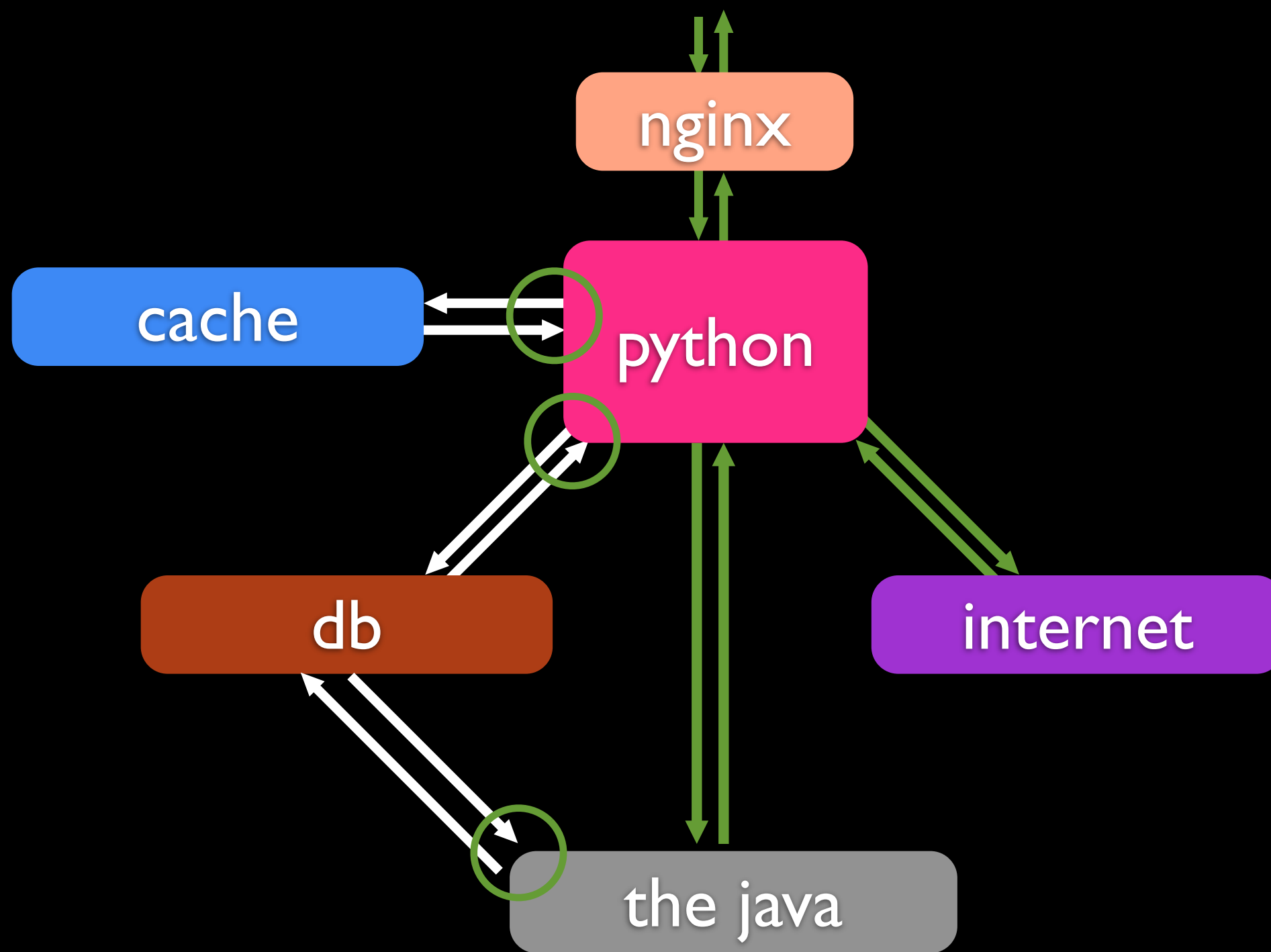
t = start



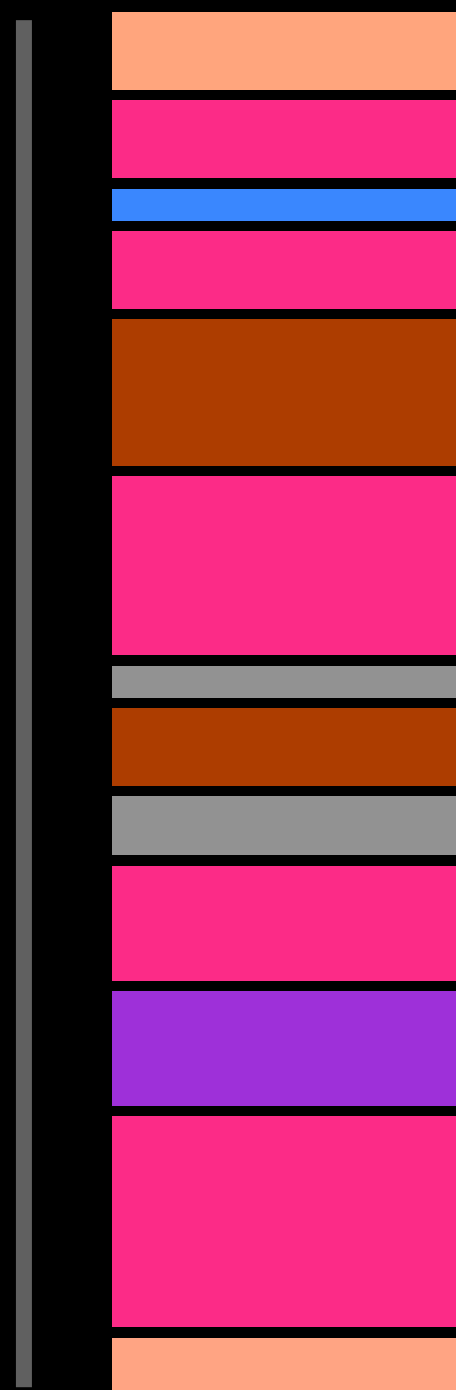
t = end

# piggyback rides

- More Doable
  - HTTP: x-headers
  - Thrift: secret argument
  - Internal RPC protocol: you're the boss
- Less Doable
  - SQL: one way ticket, also you're not percona
  - memcache: not extensible so not backwards compatible



t = start



t = end

# timing and structure

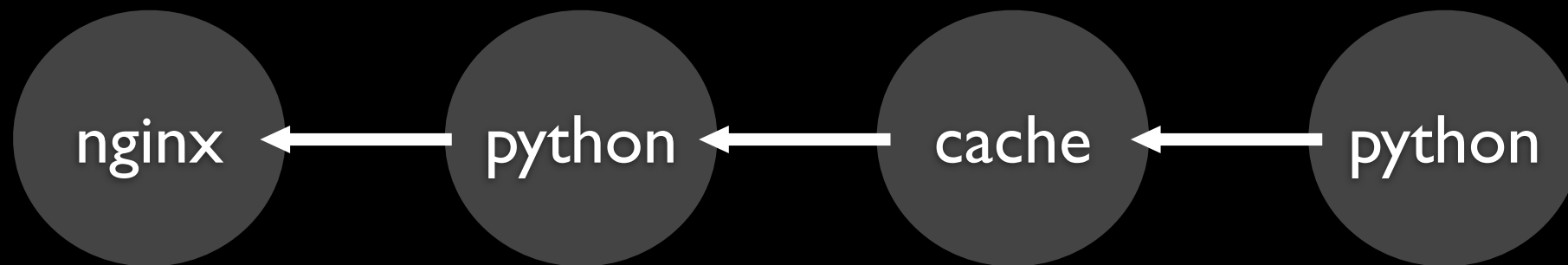
- Timing
  - distributed = clock skew
- Structure -- two approaches
  - Encode in ID
  - Encode in back-pointers



# encode in ID?

- nginx1
- nginx1python1
- nginx1python1cache1
- nginx1python1cache1python2
- nginx1python1cache1python2sql  
1
- nginx1python1cache1python2sql  
1python3
- ...

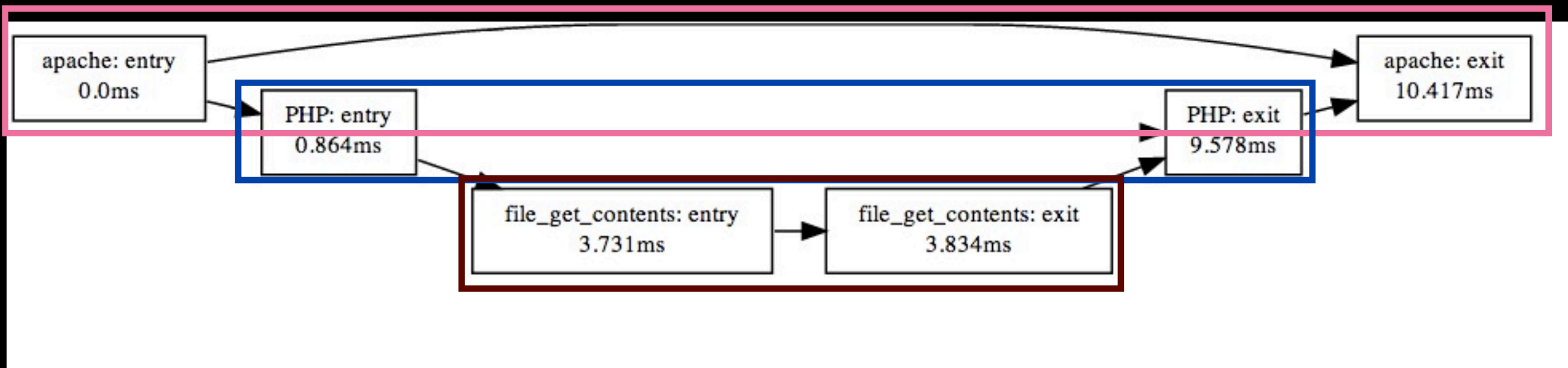
# encode in back-pointer?



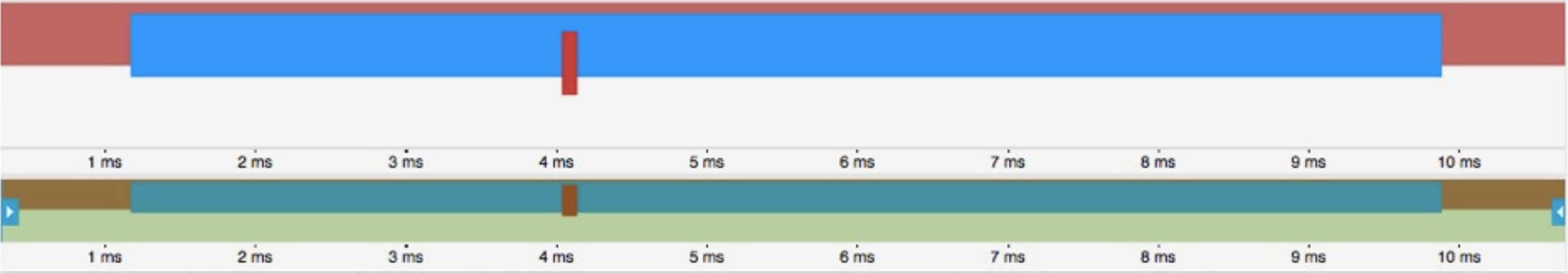
reporting

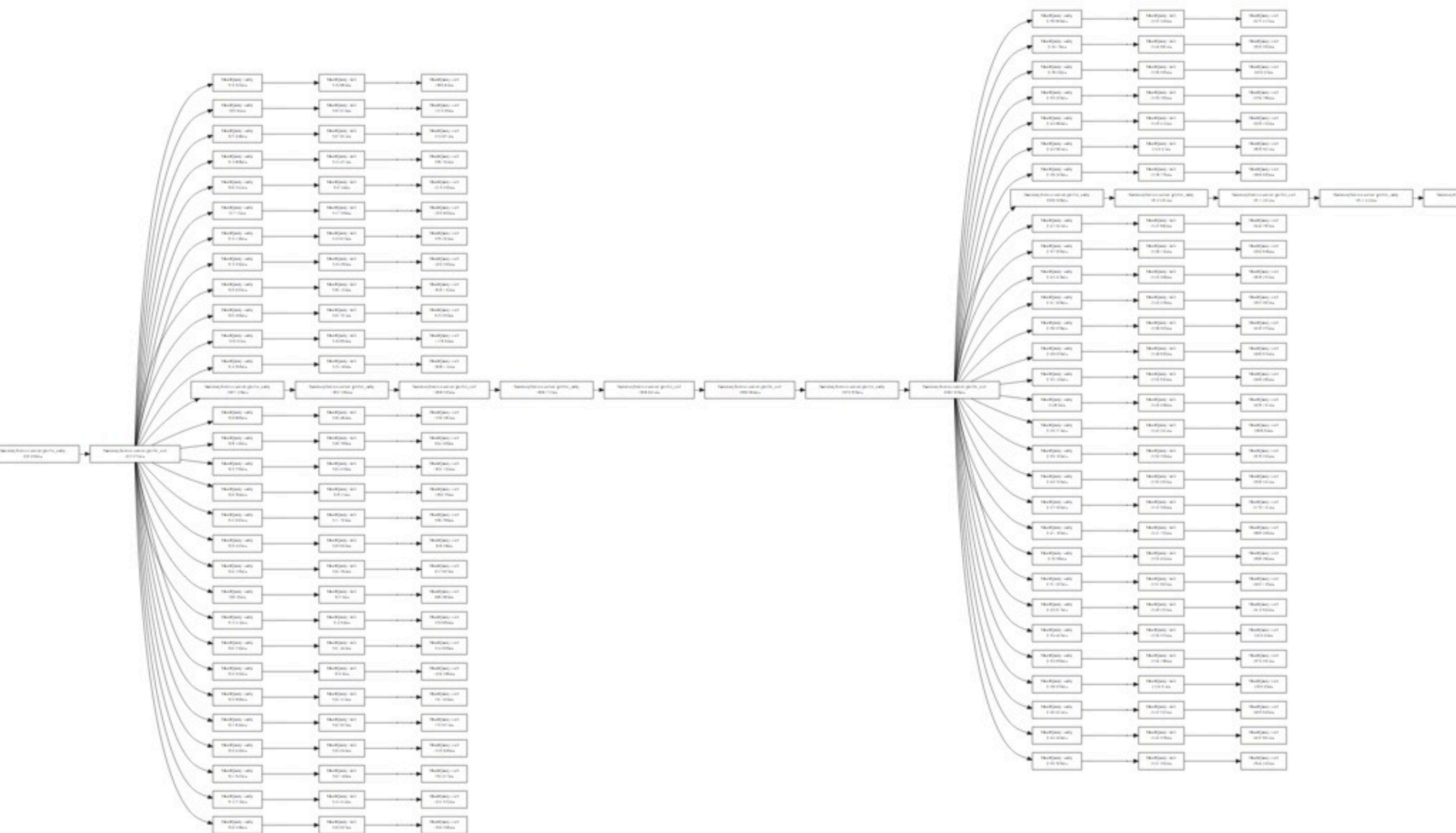
reporting





Trace timeline - 3 calls over 10 ms



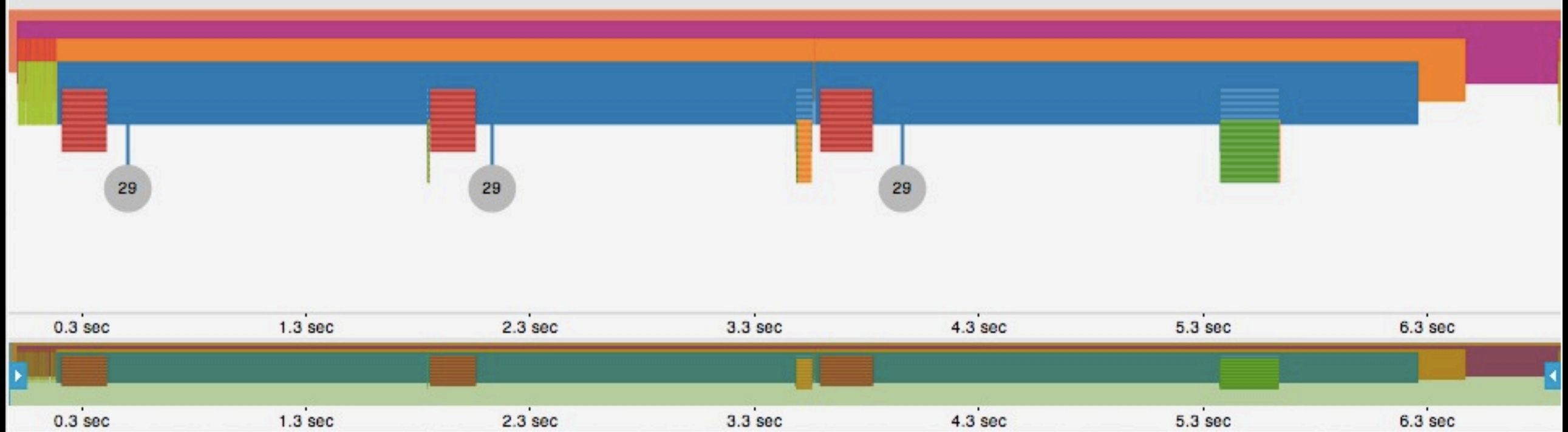




Trace timeline - 180 calls over 6.9 sec

☐ asynchronous

☒ code profile



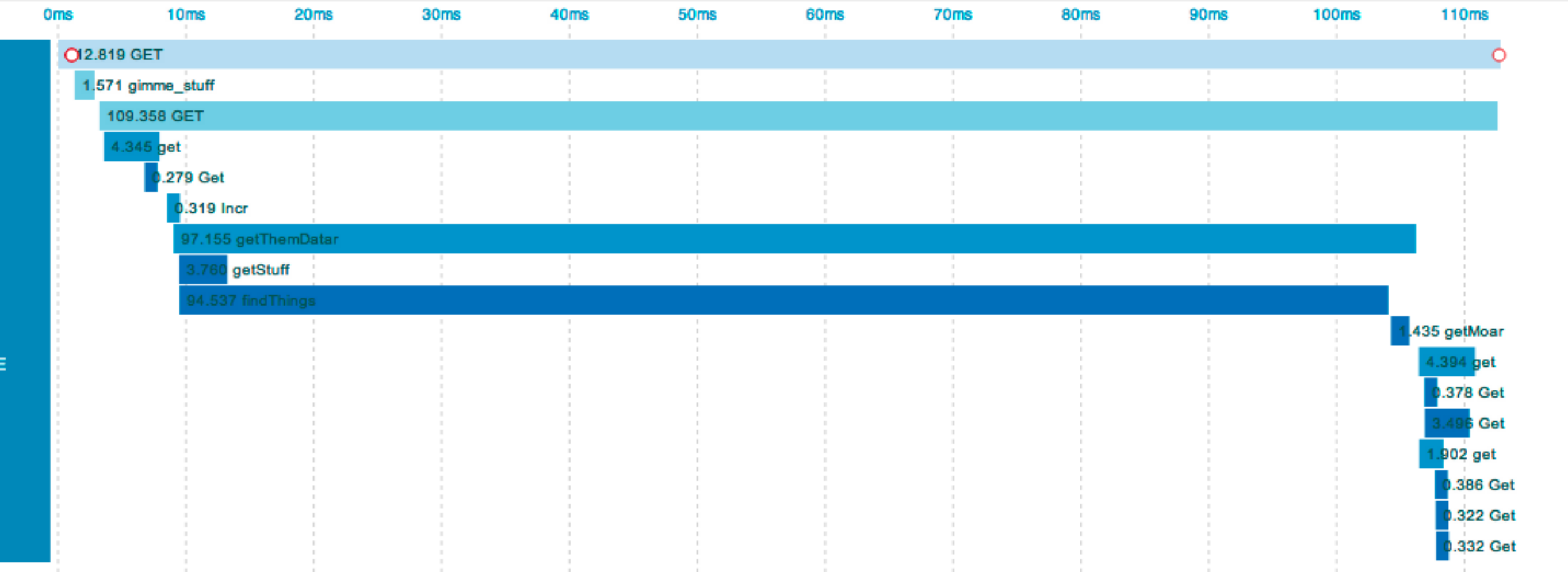
Dependencies

Search term (service)

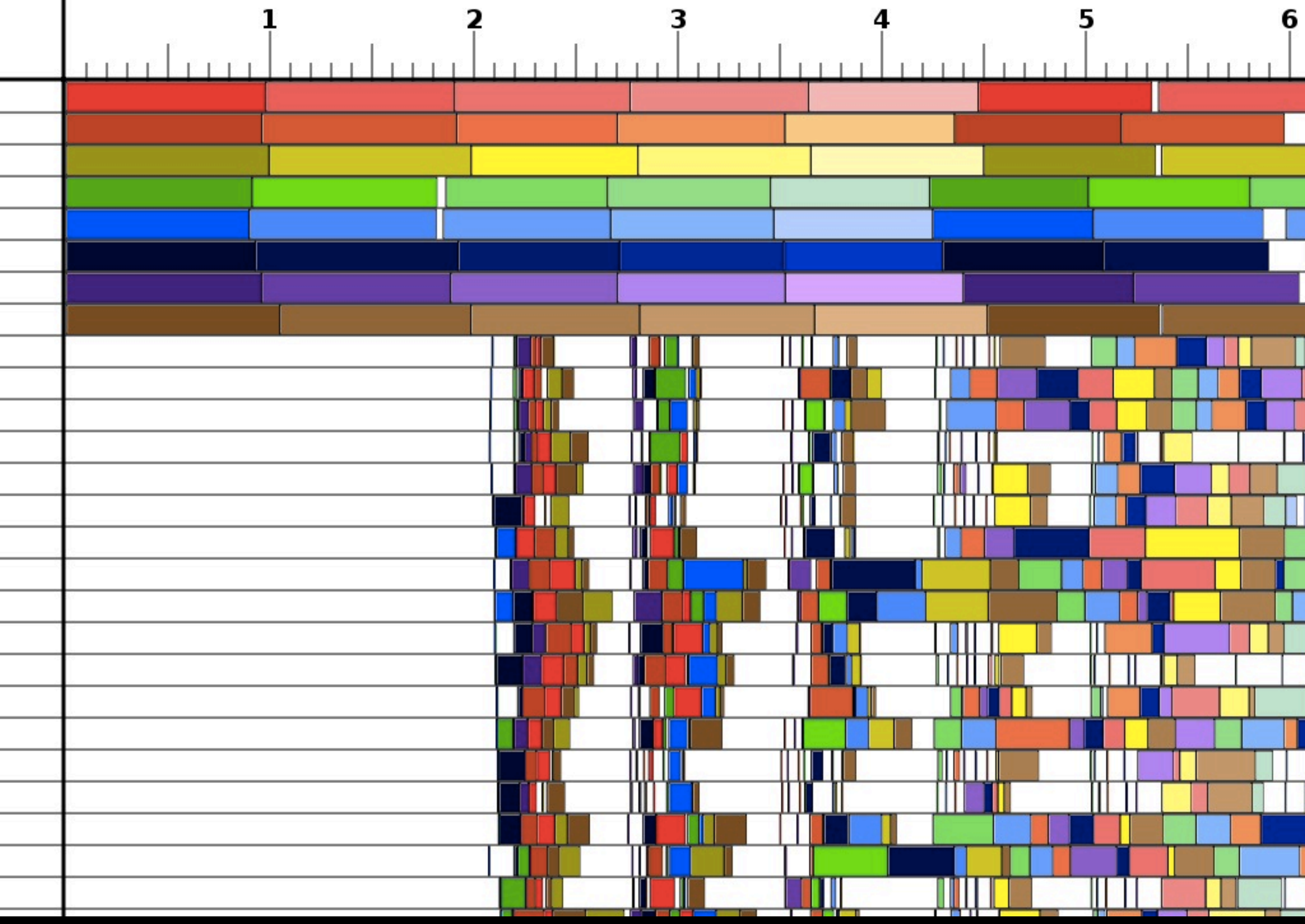


112.819 ms

16 minutes ago

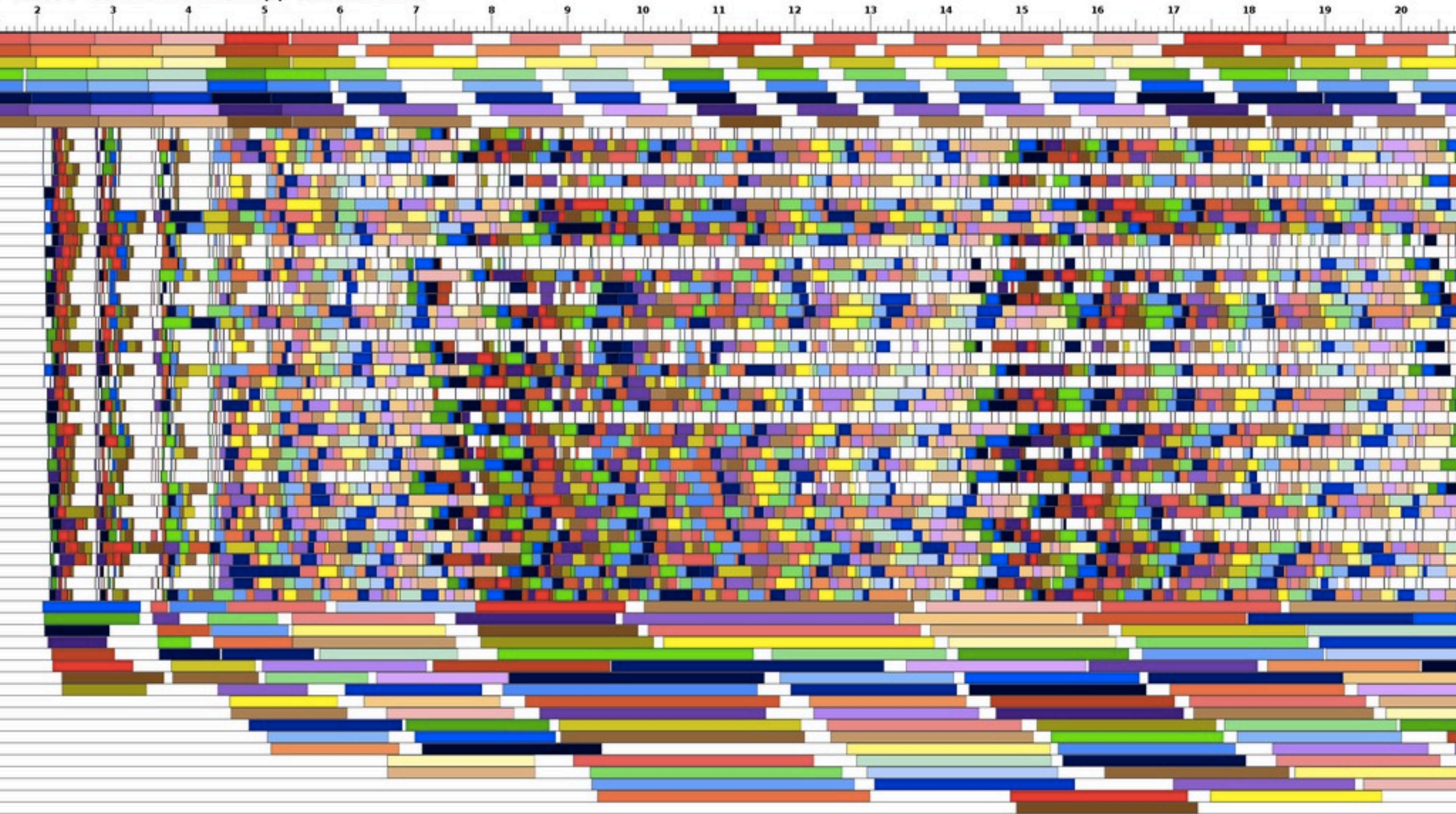


# dcswitch124 - Phase 1 - Buffer Trace Timeline(s) - All bu

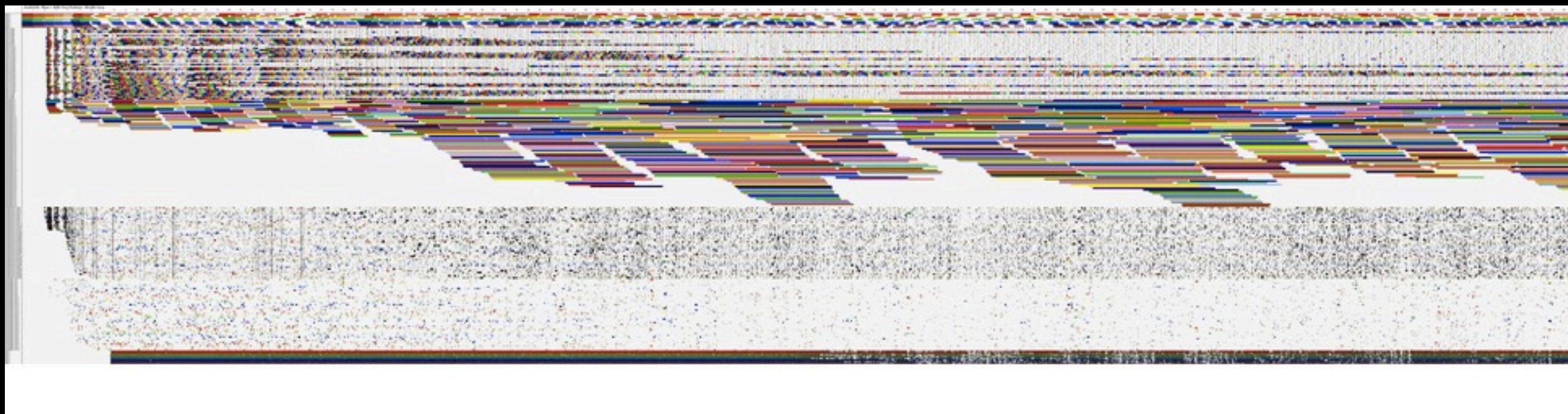




- Phase 1 - Buffer Trace Timeline(s) - All buffer traces







# other things worth figuring out

- sampling
- reporting
- aggregate analysis

# thanks!

tracelytics.com

@dankosaur

dan@tracelytics.com

# resources

- X-Trace: <http://x-trace.net>
  - <http://x-trace.net/pubs/xtr-nsdi07.pdf>
- Google Dapper: <http://research.google.com/pubs/pub36356.html>
- Twitter Zipkin: <https://github.com/twitter/zipkin>
- CMU PDL: [www.pdl.cmu.edu](http://www.pdl.cmu.edu)
  - StarDust: [http://www.pdl.cmu.edu/PDL-FTP/SelfStar/thereska\\_sigmetrics06.pdf](http://www.pdl.cmu.edu/PDL-FTP/SelfStar/thereska_sigmetrics06.pdf)
  - Trace Diff: <http://www.pdl.cmu.edu/PDL-FTP/SelfStar/NSDI11.pdf>