



Design & (D?)Evolution of Chartbeat's Architecture

Wes Chow / CTO / wes@chartbeat.com / @weschow

75 people on avc.com

monthly max: 95
monthly min: 0

23 new   52 returning

people reading



10

people writing



1

people idle



64

Page load time

average time



30

Recent user page load times:

[Only Ten Years Too Early](#) 10.34 second load time

[Only Ten Years Too Early](#) 30.05 second load time

[Only Ten Years Too Early](#) 18.3 second load time

[A VC](#) 9.88 second load time

[What Comes After Post Modernism?](#) 37.49 second load time

Geolocations

Page density

Only Ten Years Too Early

53

	A VC 8	Fin Mof 1	/a_vc/ 1
		/a_vc/a_vc/ /03/ve/03/in- capite-the-us though mobile on-the interne	What Comes After
		/a_vc/ /03/in- the-us mobile intern	Tran Foun Diluti
	Sometimes You Just Have To	/a_vc/ /12/bo ceo-at 1	How Much Is "Norr
Technical Innovation vs Business Model			
What Is An "Agressive Tech Agenda?"	So What Do We Think Of	/a_vc/2007	

Raw hits

[Only Ten Years Too Early](#) coming from
<http://news.ycombinator.com/>
at 5:45:58 pm

[Only Ten Years Too Early](#) coming from
<http://www.google.com/reader/view/>
at 5:45:40 pm

[Only Ten Years Too Early](#) coming from
<http://news.ycombinator.com/>
at 5:45:30 pm

[A VC](#) coming from
direct

ACTIVE VISITS

Concurrents

28,710



Recirculation

7%

Engaged Time

0:38

VISITOR FREQUENCY

<input type="checkbox"/> New	<div style="width: 100%;"></div>	8,739
<input type="checkbox"/> Returning	<div style="width: 80%;"></div>	11,883
<input type="checkbox"/> Loyal	<div style="width: 60%;"></div>	8,088

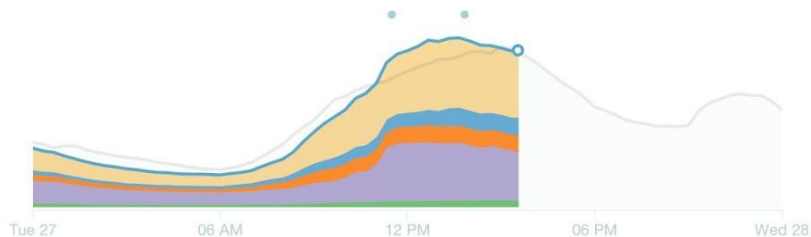
MOBILE

32% 

LOCATION

CONCURRENTS BY TRAFFIC SOURCE

Today



TOP PAGES

Sorted by Concurrents

4,321	Gizmodo - Tech By Design	0:25
3,876	Google Changed Its Logo This Weekend and You Didn't Even N...	0:30
1,530	Surface Pro 3 Review: The Greatest Laptop-Tablet Hybrid You ...	1:59
1,328	SSDs of All Sizes, the Cookw	
1,109	RIP Massimo Vignelli: The lo	

Average converting

Return rate
56%

TRAFFIC SOURCES

<input type="checkbox"/> Internal	<div style="width: 42%;"></div>	42%
<input type="checkbox"/> Social	<div style="width: 31%;"></div>	31%
<input type="checkbox"/> Links	<div style="width: 12%;"></div>	12%
<input type="checkbox"/> Direct	<div style="width: 10%;"></div>	10%
<input type="checkbox"/> Search	<div style="width: 5%;"></div>	5%

Referrers

<input type="checkbox"/> Email, apps, IM	4,333
<input type="checkbox"/> facebook.com	2,826
<input type="checkbox"/> feedly.com	935
<input type="checkbox"/> google.com	865
<input type="checkbox"/> l.facebook.com	734
<input type="checkbox"/> gawker.com	587
<input type="checkbox"/> t.co (Twitter)	510
<input type="checkbox"/> lifehacker.com	311

ACTIVE VISITS

Concurrents

540



Recirculation

17%

Engaged Time

0:24

VISITOR FREQUENCY

<input type="checkbox"/> New	<div style="width: 20%;"></div>	194
<input type="checkbox"/> Returning	<div style="width: 30%;"></div>	256
<input type="checkbox"/> Loyal	<div style="width: 10%;"></div>	90

PLATFORM

<input type="checkbox"/> Desktop	<div style="width: 62%;"></div>	62%
<input type="checkbox"/> Mobile	<div style="width: 21%;"></div>	21%

CONCURRENTS BY TRAFFIC SOURCE

30-day



TOP PAGES

Sorted by Concurrents

JESUS DIAZ x COOL x

366	Extremely rare photo of two armed F-16s escorting a ...	0:27
33	US Army helicopter drops an entire boat full of Navy ...	0:26
32	How they made the classic Old Spice "I'm on a horse..."	0:08
27	The masterful 360-degree long shot in yesterday's G...	0:21

TRAFFIC SOURCES

<input type="checkbox"/> Internal	<div style="width: 46%;"></div>	46%
<input type="checkbox"/> Social	<div style="width: 42%;"></div>	42%
<input type="checkbox"/> Links	<div style="width: 11%;"></div>	11%
<input type="checkbox"/> Search	<div style="width: 2%;"></div>	2%
<input type="checkbox"/> Direct	<div style="width: 0%;"></div>	0%

Referrers

<input type="checkbox"/> Email, apps, IM	113
<input type="checkbox"/> facebook.com	87
<input type="checkbox"/> hsr.d.yahoo.com	50
<input type="checkbox"/> l.facebook.com	16
<input type="checkbox"/> google.com	5
<input type="checkbox"/> feedly.com	5
<input type="checkbox"/> m.facebook.com	5

ping.chartbeat.net

All

Documents

Stylesheets

Images

Scripts






XHR

Fonts

WebSockets

Other

Hide data URLs

Name Path	Method	Status Text	Type	Initiator	Size Content	Time Latency	Timeline
 uuniq?u=Dn_zfNB7gtPXMbNeK&u2=Cg... ping.chartbeat.net	GET	200 OK	image/gif	chartbeat.js:17 Script	177 B 43 B	32 ms 31 ms	
 ping?h=gizmodo.com&p=%2F&u=Dn_zf... ping.chartbeat.net	GET	200 OK	image/gif	chartbeat.js:17 Script	213 B 43 B	30 ms 29 ms	
 ping?h=gizmodo.com&p=%2F&u=Dn_zf... ping.chartbeat.net	GET	200 OK	image/gif	chartbeat.js:17 Script	213 B 43 B	16 ms 15 ms	

ACTIVE VISITS

Concurrents

28,710



Recirculation

7%

Engaged Time

0:38

VISITOR FREQUENCY

<input type="checkbox"/> New	<div style="width: 100%;"></div>	8,739
<input type="checkbox"/> Returning	<div style="width: 80%;"></div>	11,883
<input type="checkbox"/> Loyal	<div style="width: 60%;"></div>	8,088

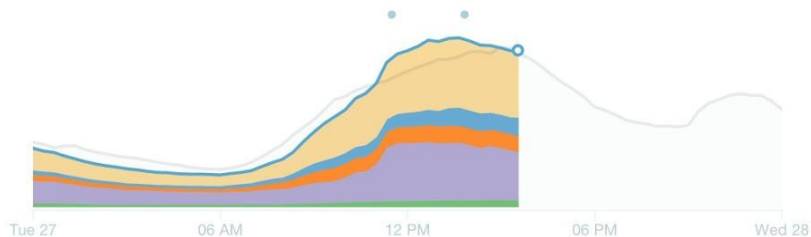
MOBILE

32%

LOCATION

CONCURRENTS BY TRAFFIC SOURCE

Today



TOP PAGES

Sorted by Concurrents

4,321	Gizmodo - Tech By Design	0:25
3,876	Google Changed Its Logo This Weekend and You Didn't Even N...	0:30
1,530	Surface Pro 3 Review: The Greatest Laptop-Tablet Hybrid You ...	1:59
1,328	SSDs of All Sizes, the Cookw	
1,109	RIP Massimo Vignelli: The lo	

Average converting

Return rate
56%

TRAFFIC SOURCES

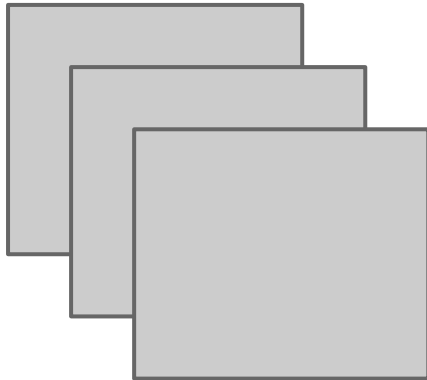
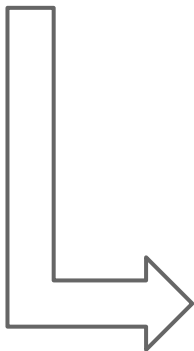
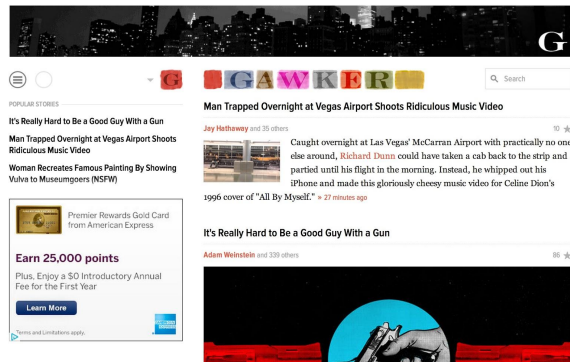
<input type="checkbox"/> Internal	<div style="width: 42%;"></div>	42%
<input type="checkbox"/> Social	<div style="width: 31%;"></div>	31%
<input type="checkbox"/> Links	<div style="width: 12%;"></div>	12%
<input type="checkbox"/> Direct	<div style="width: 10%;"></div>	10%
<input type="checkbox"/> Search	<div style="width: 5%;"></div>	5%

Referrers

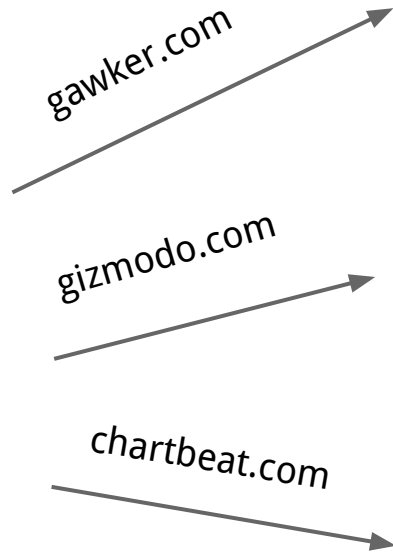
<input type="checkbox"/> Email, apps, IM	4,333
<input type="checkbox"/> facebook.com	2,826
<input type="checkbox"/> feedly.com	935
<input type="checkbox"/> google.com	865
<input type="checkbox"/> l.facebook.com	734
<input type="checkbox"/> gawker.com	587
<input type="checkbox"/> t.co (Twitter)	510
<input type="checkbox"/> lifehacker.com	311

Memoryfly

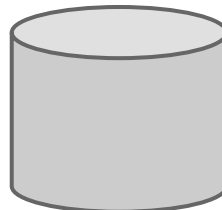
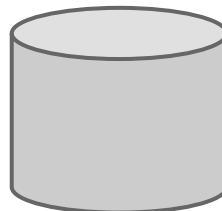
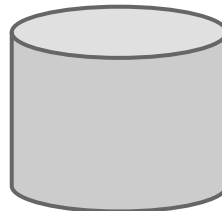
- In memory data store
- Online / streaming aggregations
- C / C++, Lua integration
- 200k pings / sec peak in average day
- 9M concurrents peak in average day
- 1B+ uniques / week



Routing Layer



Memoryfly



ACTIVE VISITS

Concurrents

28,710



Recirculation

7%

Engaged Time

0:38

VISITOR FREQUENCY

<input type="checkbox"/> New	<div style="width: 100%;"></div>	8,739
<input type="checkbox"/> Returning	<div style="width: 80%;"></div>	11,883
<input type="checkbox"/> Loyal	<div style="width: 60%;"></div>	8,088

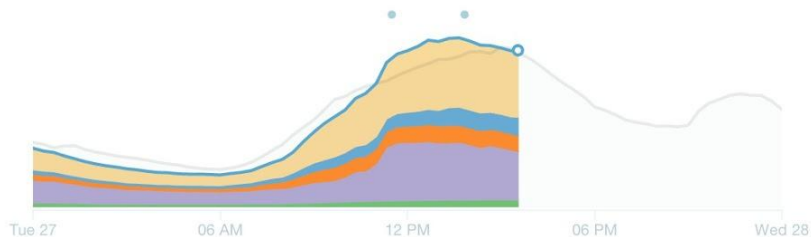
MOBILE

32%

LOCATION

CONCURRENTS BY TRAFFIC SOURCE

Today



TOP PAGES

Sorted by Concurrents

4,321	Gizmodo - Tech By Design	0:25
3,876	Google Changed Its Logo This Weekend and You Didn't Even N...	0:30
1,530	Surface Pro 3 Review: The Greatest Laptop-Tablet Hybrid You ...	1:59
1,328	SSDs of All Sizes, the Cookw...	
1,109	RIP Massimo Vignelli: The lo...	

Average converting

Return rate
56%

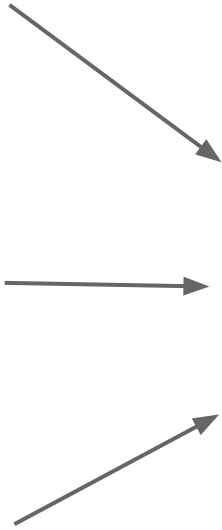
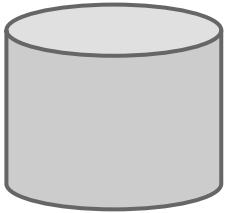
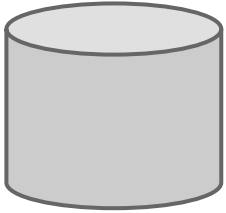
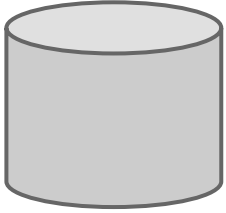
TRAFFIC SOURCES

<input type="checkbox"/> Internal	<div style="width: 42%;"></div>	42%
<input type="checkbox"/> Social	<div style="width: 31%;"></div>	31%
<input type="checkbox"/> Links	<div style="width: 12%;"></div>	12%
<input type="checkbox"/> Direct	<div style="width: 10%;"></div>	10%
<input type="checkbox"/> Search	<div style="width: 5%;"></div>	5%

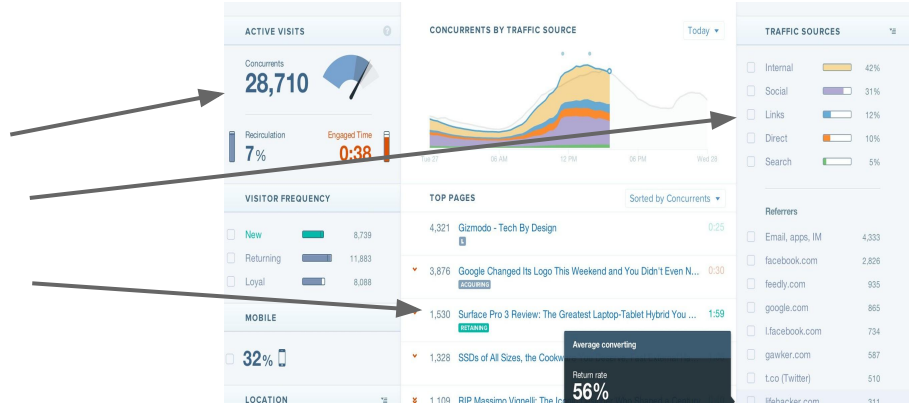
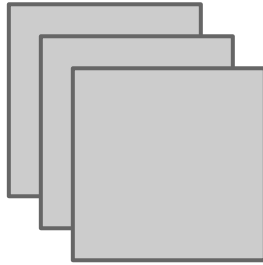
Referrers

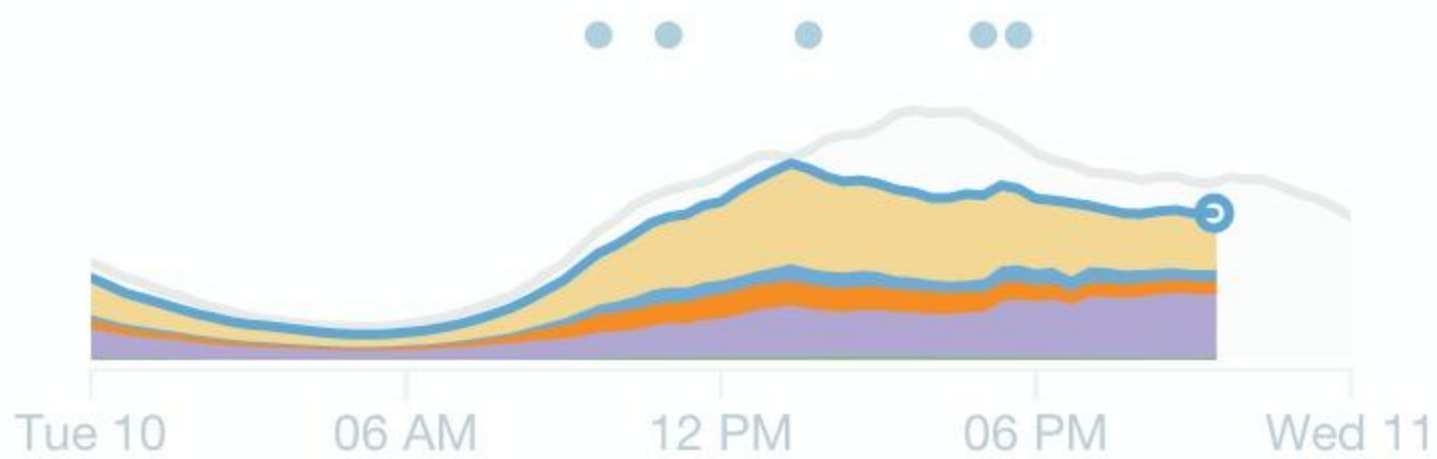
<input type="checkbox"/> Email, apps, IM	4,333
<input type="checkbox"/> facebook.com	2,826
<input type="checkbox"/> feedly.com	935
<input type="checkbox"/> google.com	865
<input type="checkbox"/> l.facebook.com	734
<input type="checkbox"/> gawker.com	587
<input type="checkbox"/> t.co (Twitter)	510
<input type="checkbox"/> lifehacker.com	311

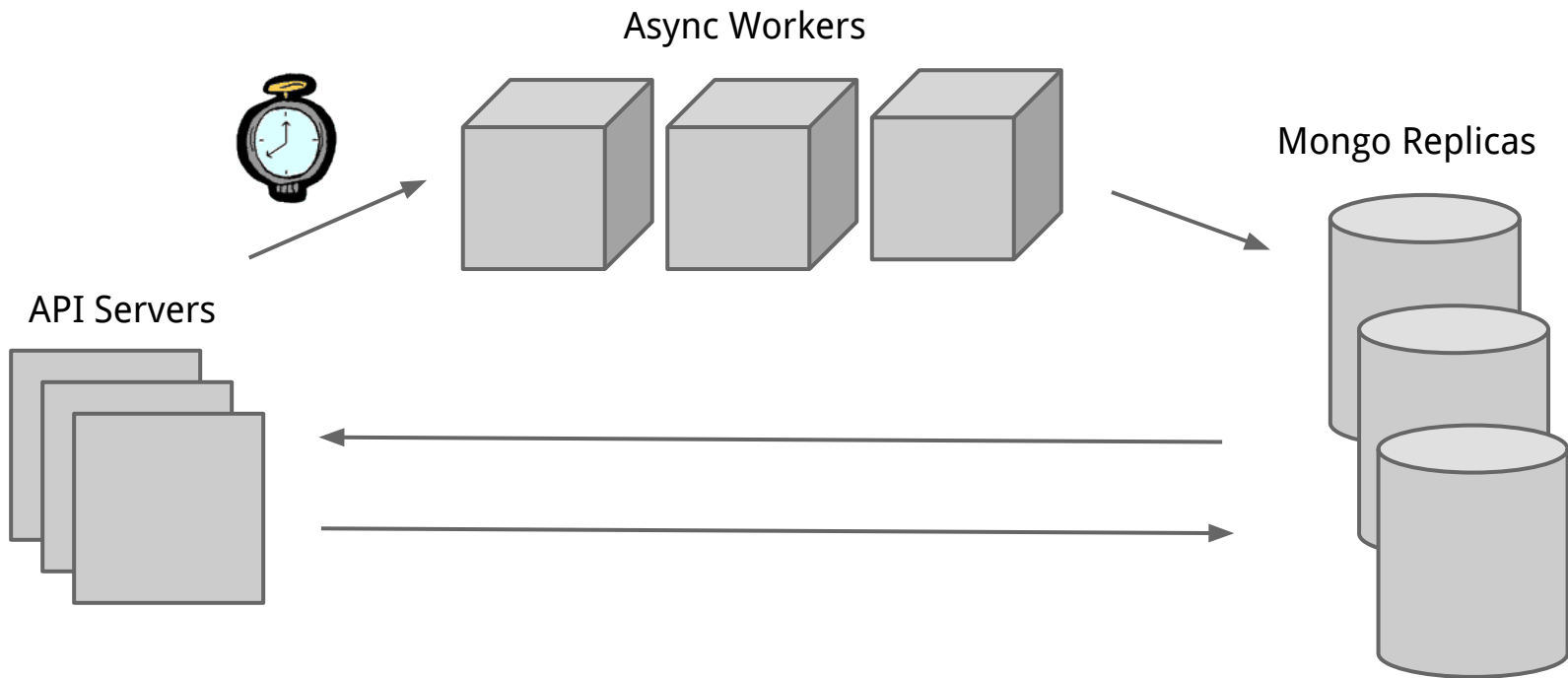
Memoryfly

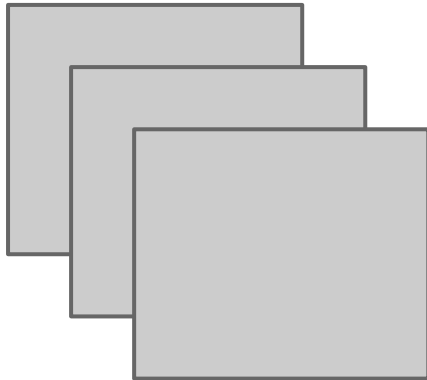
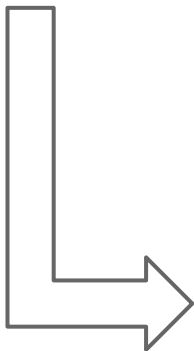
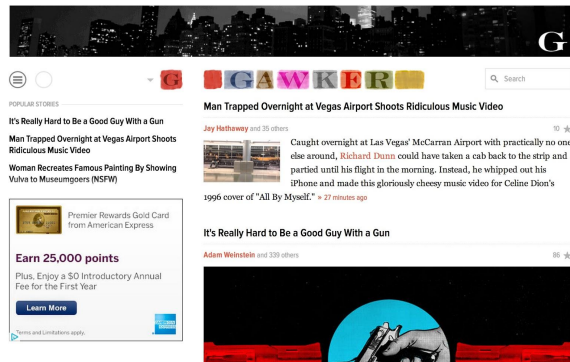


API Servers

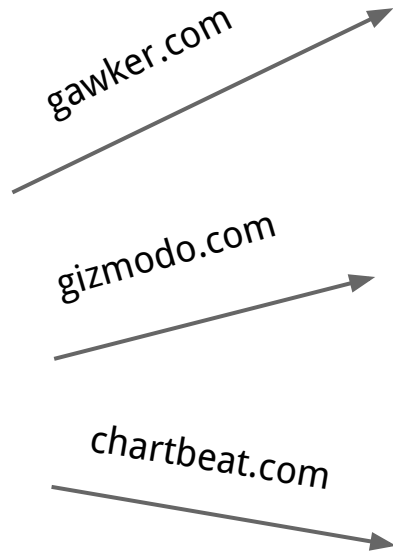




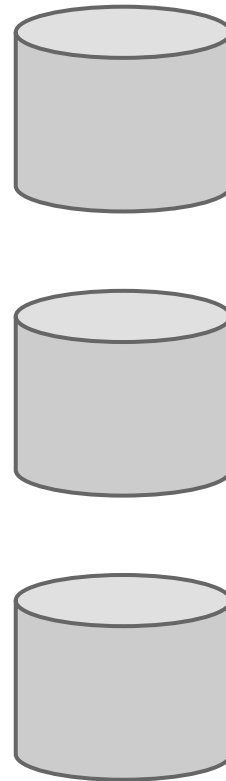




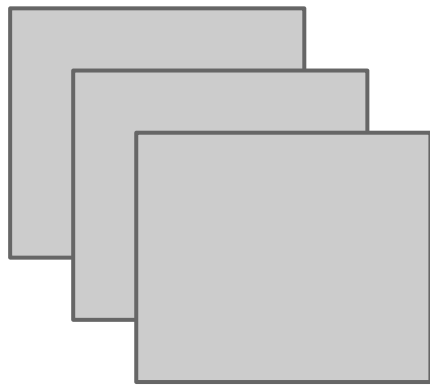
Routing Layer



Memoryfly



chartbeat.com



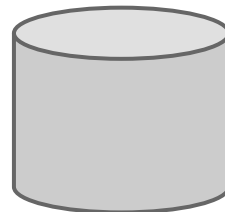
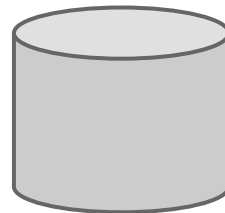
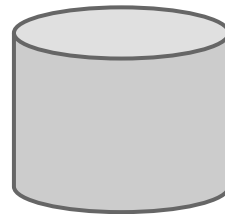
Routing Layer

33%

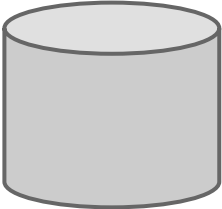
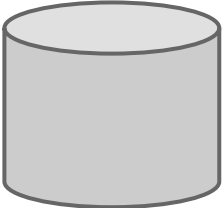
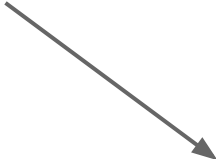
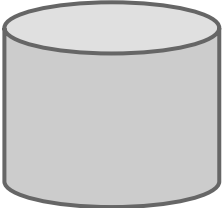
33%

33%

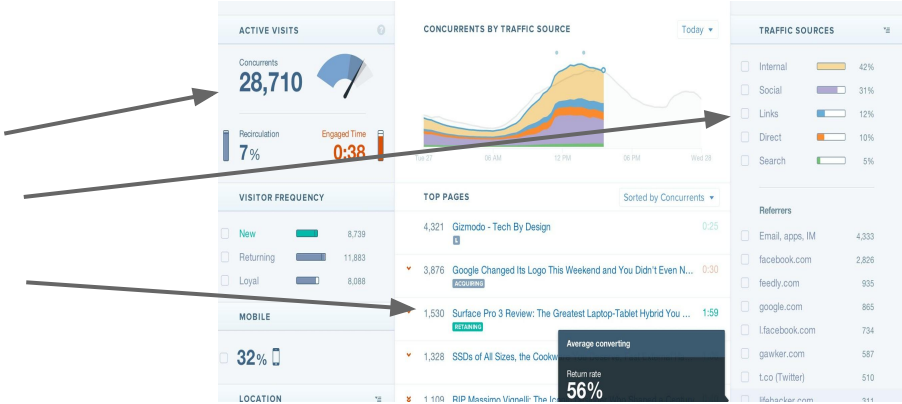
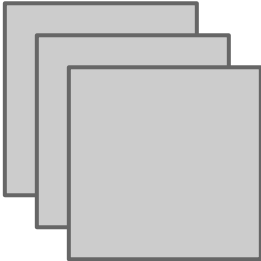
Memoryfly



Memoryfly



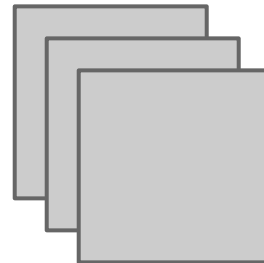
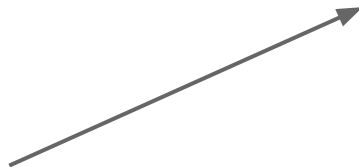
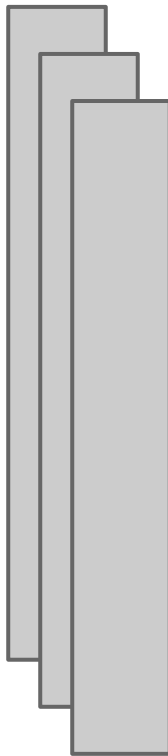
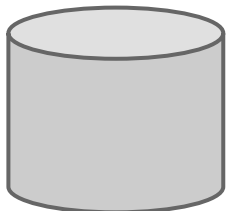
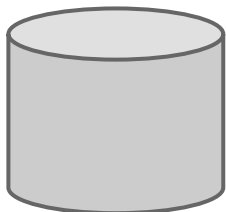
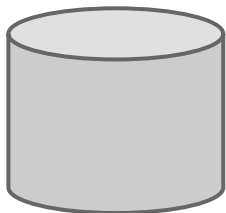
API Servers



Memoryfly

Combiners

API Servers





TOP STORIES TAKEAWAYS

The highest peak of traffic was on your #3 story

Your #7 story was your top social page

Your #19 story was your top page in search traffic, with 10% coming from search

The top pickup from any referrer came from **facebook.com** and reached a peak of **1,143** concurrents

facebook.com linked to **20** of your top pages

🕒 Engaged Min Per Person Per Week

1:17

🔄 Return Rate Per Person Per Week






21%

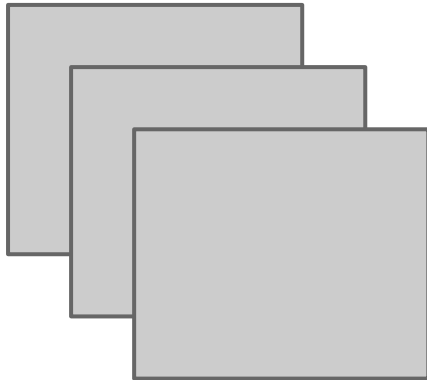
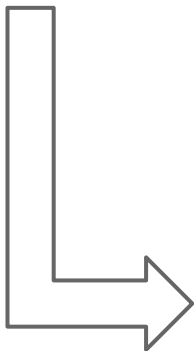
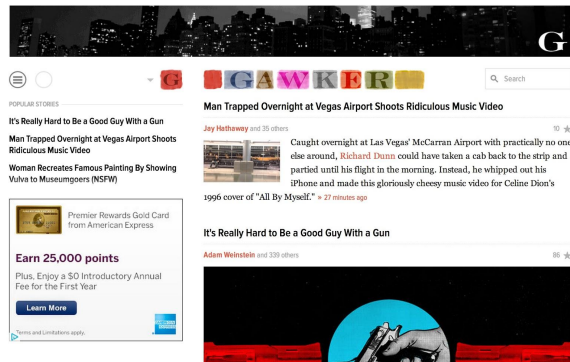
👤 Visits Per Person Per Week



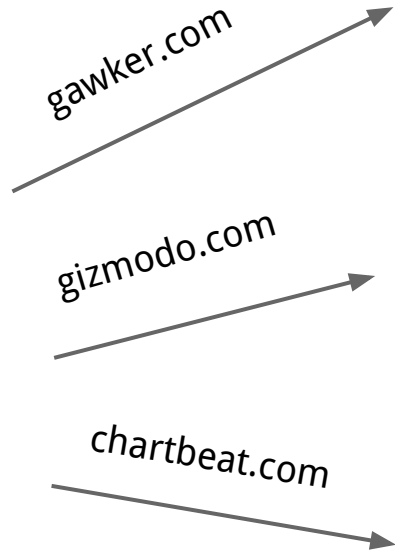
📄 Articles Read Per Person Per Week



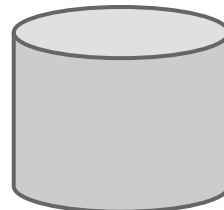
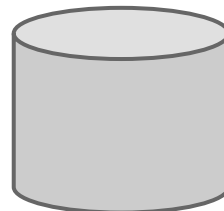
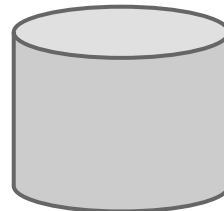
	Referrer	Number of Articles	Total Engaged Min on First Page vs. Subsequent	Return Rate	Returned Directly	Loyalty Type ?
1	m.facebook.com	629	1.0M avg 0:17  17.1K	12%	24%	LOW CONVERTING
2	facebook.com	551	642.2K avg 0:32  65.9K	30%	16%	AVERAGE CONVERTING
3	google.com	4353	514.5K avg 0:32  146.6K	29%	23%	AVERAGE CONVERTING
4	l.facebook.com	457	337.2K avg 0:33  28.1K	24%	14%	LOW CONVERTING
5	deadspin.com	90	229.7K avg 0:40  112.2K	63%	23%	HIGH CONVERTING



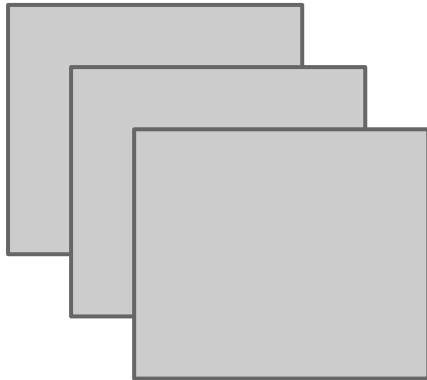
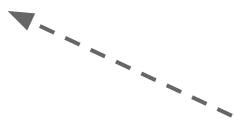
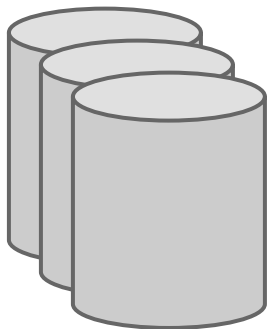
Routing Layer



Memoryfly



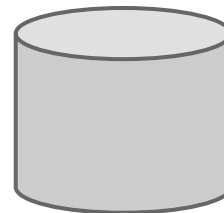
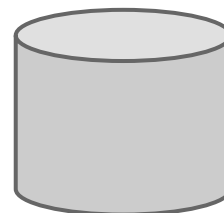
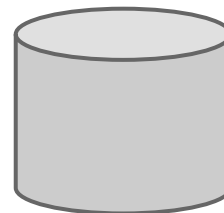
Deep Historical



Routing Layer



Memoryfly

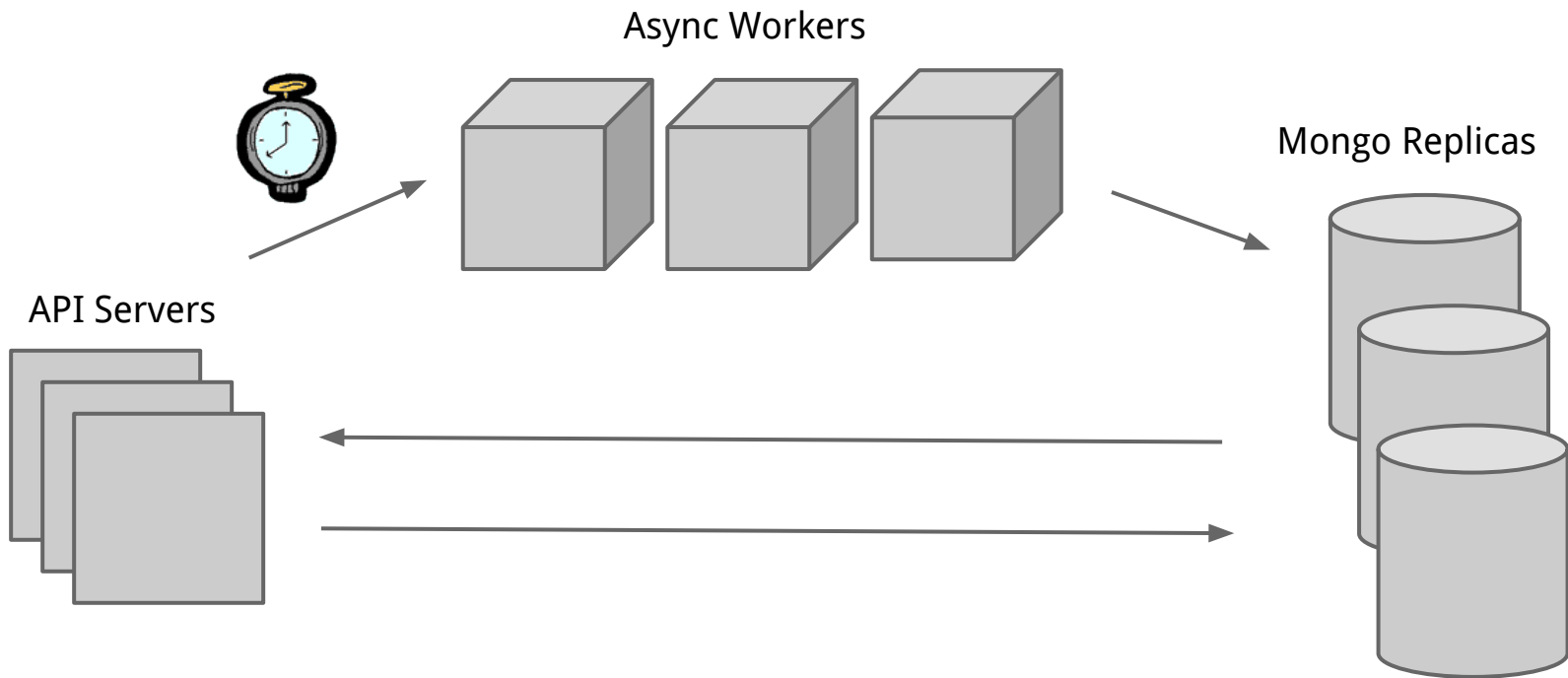


Web Tier

- Django + MySQL for user database
- Dashboards served as big Angular apps
- Data pulled in via AJAX calls

Tech Backbone

- AWS
- Languages: Python, C / C++, Lua
- RabbitMQ
- Mongo, Riak (experimental), MySQL
- Django, Tornado
- ZooKeeper, Doozer (deprecated)
- missing: Kafka, Storm



Performance Lesson

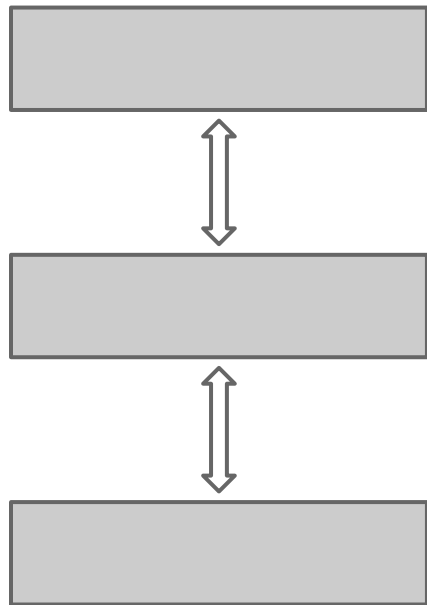
12:00	chartbeat.com	[data]	[data]	[data]
12:00	gawker.com
12:00	gizmodo.com
...
12:05	chartbeat.com
12:05	gawker.com
12:05	gizmodo.com

12:00	chartbeat.com	[data]	[data]	[data]
12:00	gawker.com
12:00	gizmodo.com
...
12:05	chartbeat.com
12:05	gawker.com
12:05	gizmodo.com

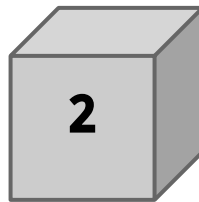
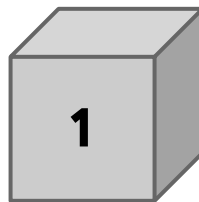
```
select where customer = 'chartbeat.com' and t in [12:00, 1:00)
```

chartbeat.com	12:00	[data]	[data]	[data]
chartbeat.com	12:05
...
gawker.com	12:00
gawker.com	12:05
...
gizmodo.com	12:00

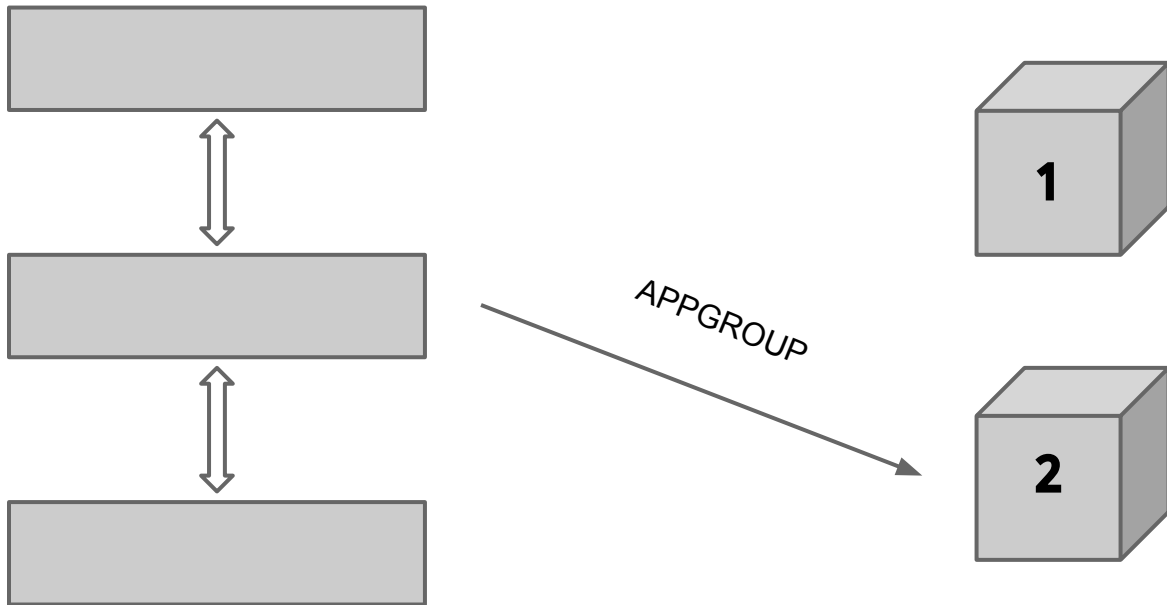
ZooKeeper



```
coord_singleton -g APPLOCK /usr/bin/app  
[args]
```



```
coord_singleton -g APPGROUP --non-block /usr/bin/app  
[args]
```



Thought Experiment

What If You're Allowed to Lose Data?

Future Work

- Current: product driven
- New: research platform
- Scale beyond what customers know
- Change how customers operate



We're hiring! (of course)

chartbeat.com/jobs / jobs@chartbeat.com

Wes Chow / CTO / wes@chartbeat.com / @weschow