The Architecture Of stackoverflow

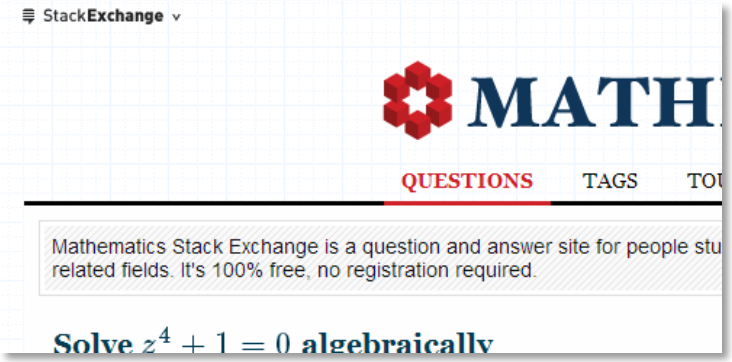
Marco Cecconi
@sklivvz
sklivvz@stackoverflow.com

expert answers to your questions

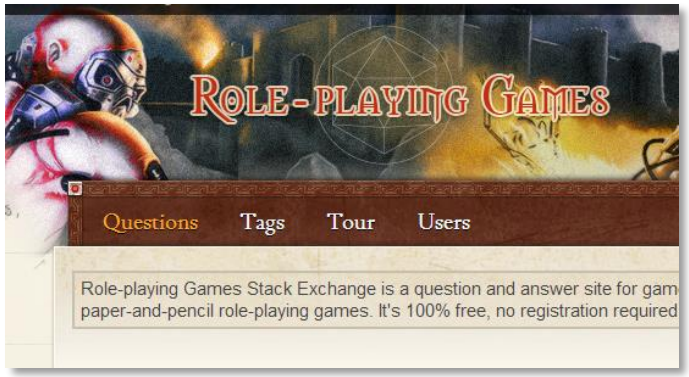
Stack Exchange is a fast-growing network of 118 question and answer sites on diverse topics from software programming to cooking to photography and gaming. We build libraries of high-quality questions and answers, focused on the most important topics in each area of expertise. From our core of Q&A, to community blogs and real-time chat, we provide experts with the tools they need to make The Internet a better place. Learn more about us...

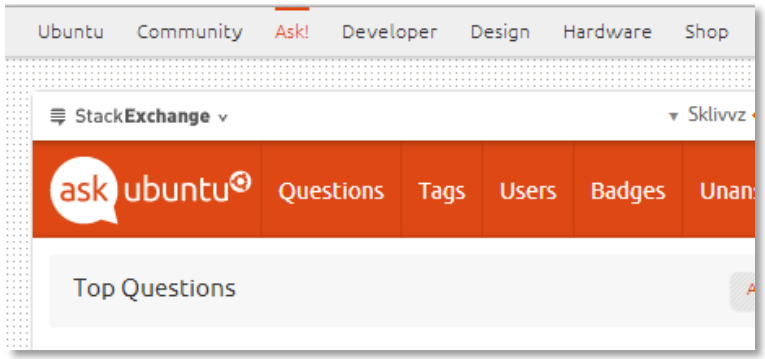
118 Q&A sites 5.2 million users 8.8 million questions 15.4 million answers

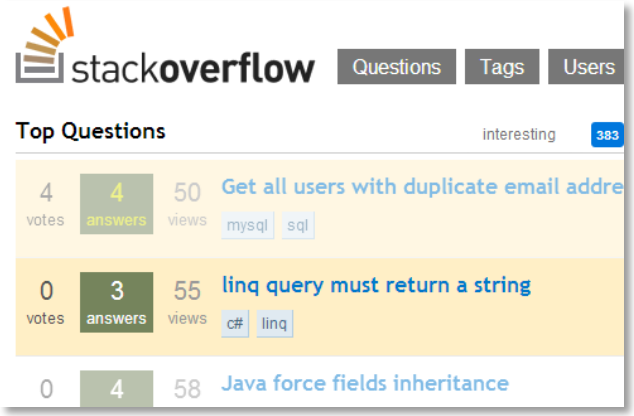
Explore our sites!

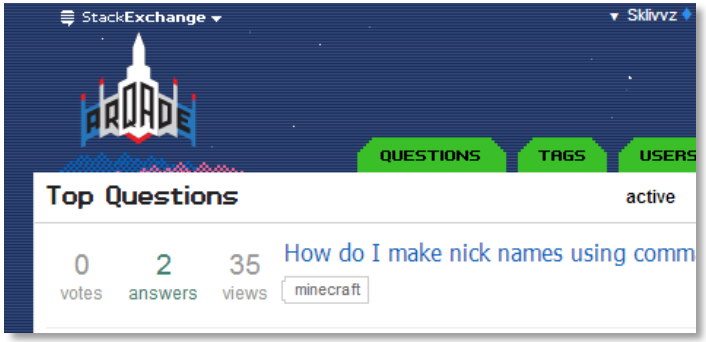

StackExchange
MATH
QUESTIONS
TAGS
Mathematics Stack Exchange is a question and answer site for people stu
related fields. It's 100% free, no registration required.
Solve $z^4 + 1 = 0$ algebraically


StackExchange
Ask Different
Answers for your Apple questions
Questions
Top Questions
5
5
2
Using an iPad for doing web deve

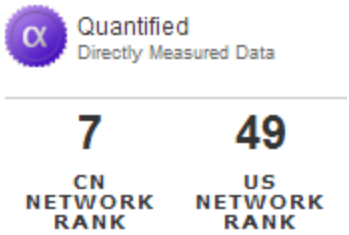

ROLE-PLAYING GAMES
Questions
Tags
Tour
Users
Role-playing Games Stack Exchange is a question and answer site for gam
paper-and-pencil role-playing games. It's 100% free, no registration required


Ubuntu
Community
Ask!
Developer
Design
Hardware
Shop
StackExchange
Sklivvz
ask ubuntu
Questions
Tags
Users
Badges
Top Questions


stackoverflow
Questions
Tags
Users
Top Questions
interesting
383
4
votes
4
answers
50
views
Get all users with duplicate email addre
mysql
sql
0
votes
3
answers
55
views
linq query must return a string
c#
linq
0
4
58
Java force fields inheritance

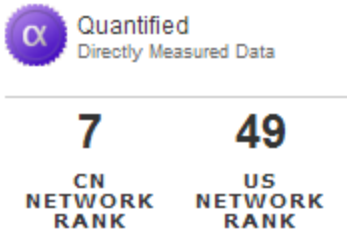

StackExchange
Sklivvz
QUESTIONS
TAGS
USERS
Top Questions
active
0
votes
2
answers
35
views
How do I make nick names using comm
minecraft

#49 network for traffic*



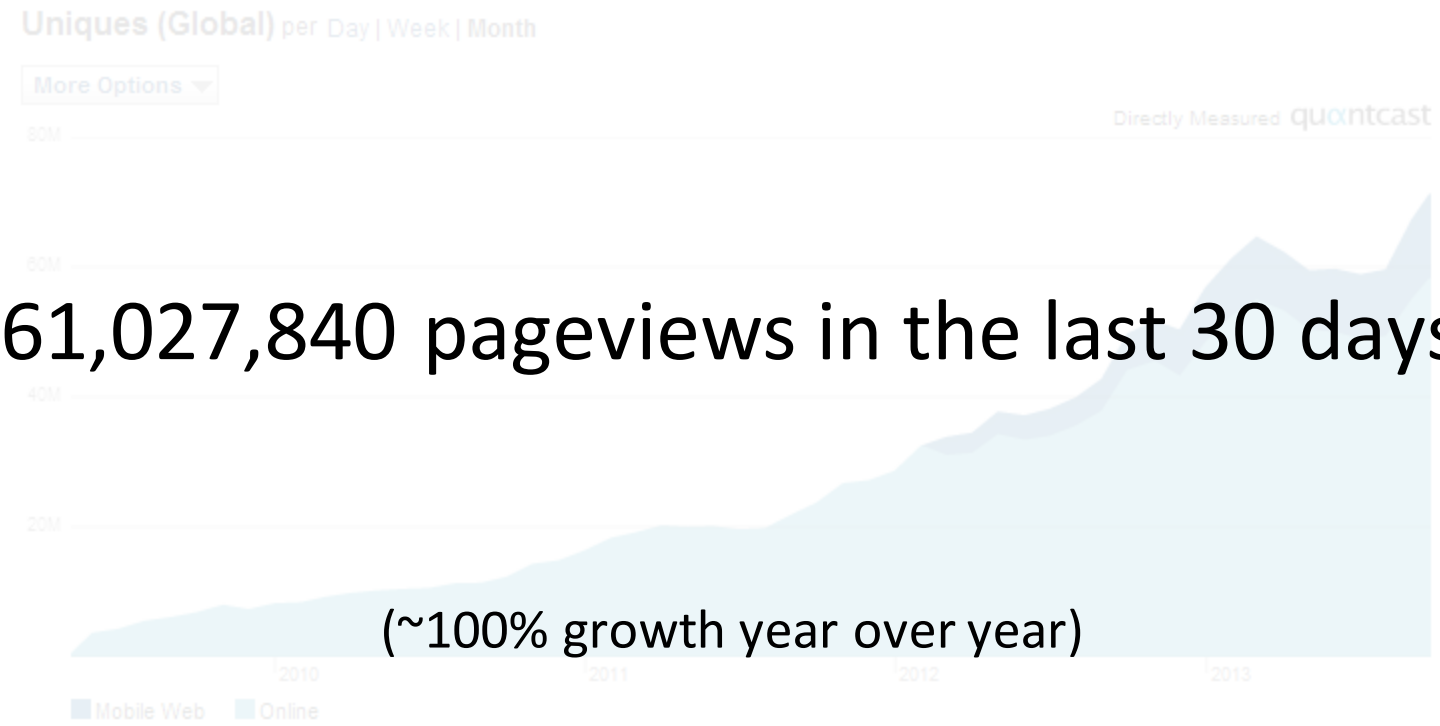
*source: Quantcast, Alexa

#49 network for traffic*



…and #7 in China!

*source: Quantcast, Alexa



561,027,840 pageviews in the last 30 days*

(~100% growth year over year)

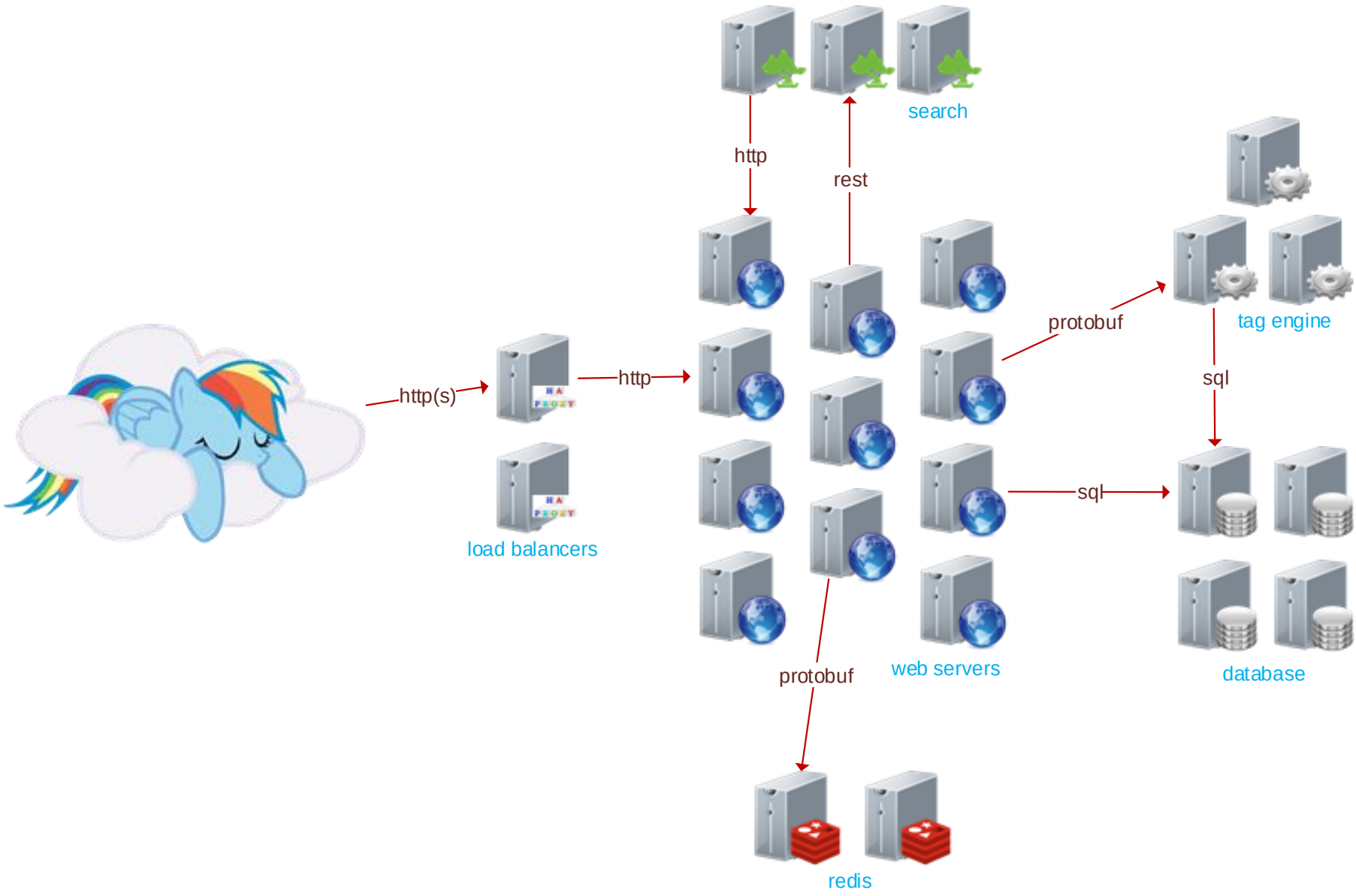
*source: Quantcast

561,027,840 pageviews in the last 30 days*

(~100% growth year over year)

*source: Quantcast


PowerEdge R610
PowerEdge R310
DO NOT PULL
PowerEdge R610
PowerEdge R610
DO NOT PULL
PowerEdge R620


search
http
rest
protobuf
tag engine
sql
http(s)
http
load balancers
sql
web servers
database
protobuf
redis

Database Servers

Node	CPU	Memory	Network
NY-SQL01	11 %	354.22 GB / 383.97 GB (92.26%)	93.25 Mb/s
NY-SQL02	27 %	354.37 GB / 383.97 GB (92.30%)	10.95 Mb/s
NY-SQL03	25 %	274.67 GB / 288.00 GB (95.37%)	59.89 Mb/s
NY-SQL04	9 %	274.54 GB / 288.00 GB (95.33%)	4.89 Mb/s
OR-SQL01	1 %	353.96 GB / 383.97 GB (92.19%)	1.01 Gb/s
OR-SQL02	2 %	5.98 GB / 288.00 GB (2.08%)	92.46 Mb/s

Web Servers

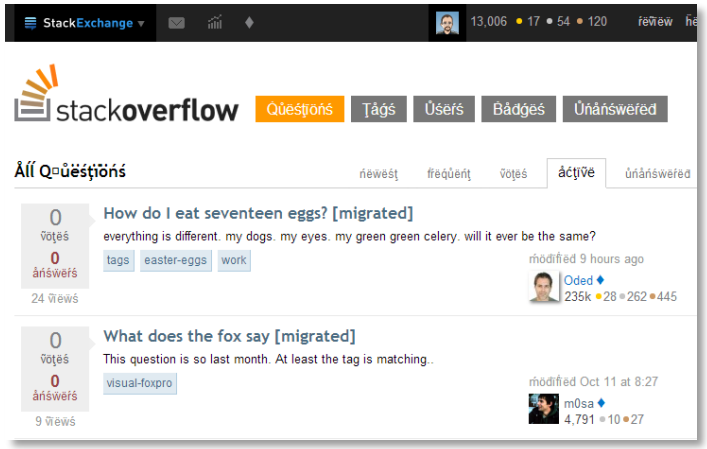
Node	CPU	Memory	Network
NY-WEB01	10 %	17.95 GB / 32.00 GB (56.13%)	25.58 Mb/s
NY-WEB02	18 %	19.40 GB / 32.00 GB (60.63%)	32.49 Mb/s
NY-WEB03	11 %	18.65 GB / 32.00 GB (58.32%)	24.42 Mb/s
NY-WEB04	12 %	19.30 GB / 32.00 GB (60.34%)	27.71 Mb/s
NY-WEB05	12 %	15.54 GB / 32.00 GB (48.57%)	41.66 Mb/s
NY-WEB06	11 %	17.23 GB / 32.00 GB (53.85%)	25.29 Mb/s
NY-WEB07	10 %	19.48 GB / 32.00 GB (60.91%)	19.12 Mb/s
NY-WEB08	19 %	20.91 GB / 32.00 GB (65.38%)	28.88 Mb/s
NY-WEB09	17 %	19.25 GB / 32.00 GB (60.19%)	28.11 Mb/s

We are still scaling up, yo!

BAT CAVE

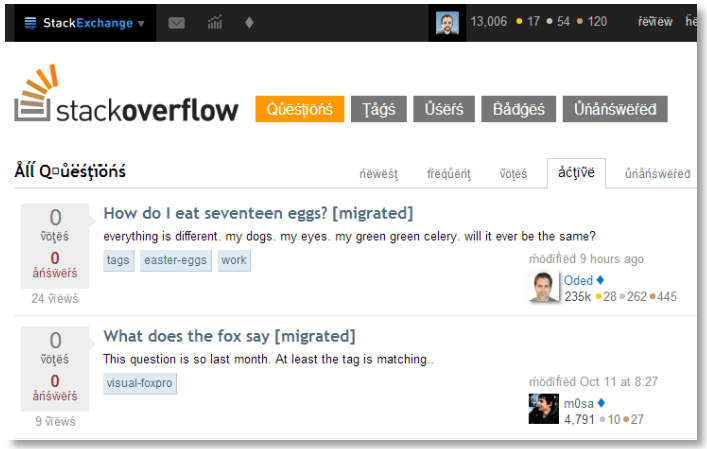
BAT CAVE

DEV.SO



BAT CAVE

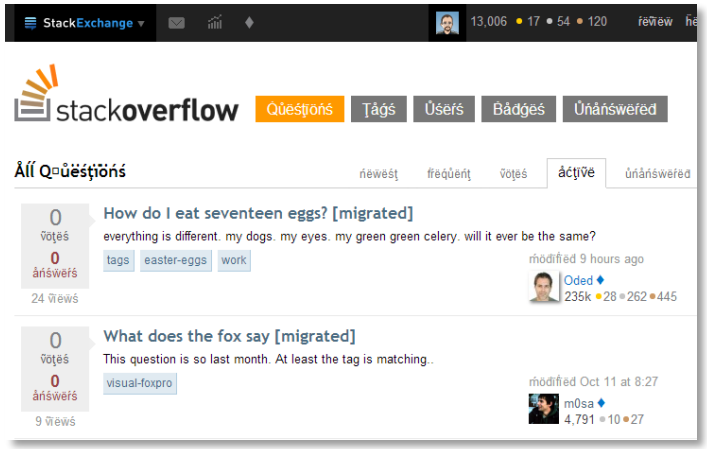
DEV.SO



META.SO

BAT CAVE

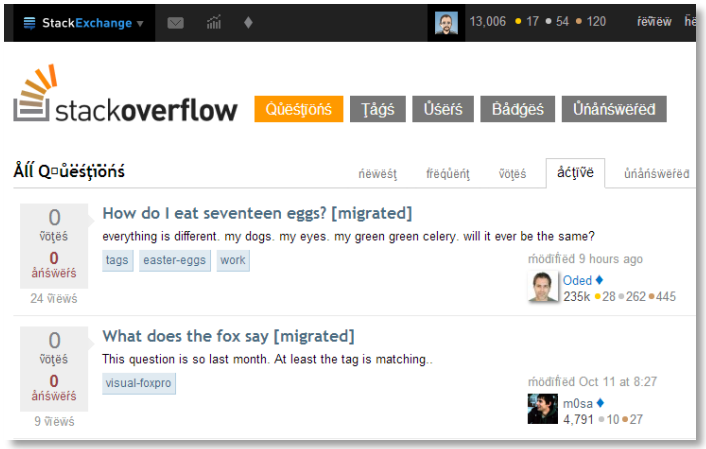
DEV.SO



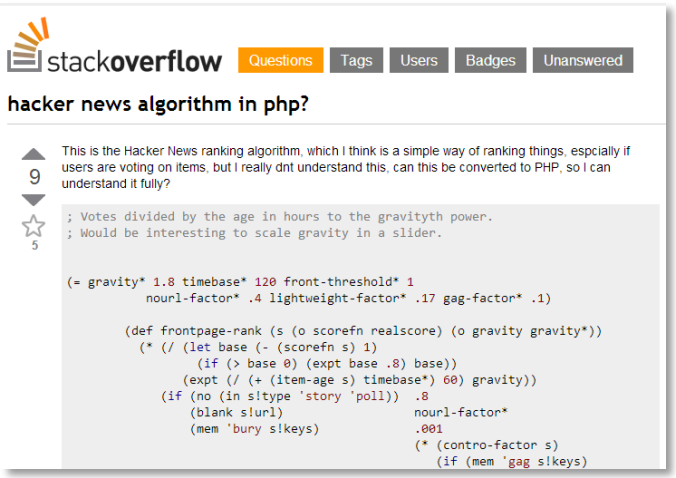
META.SO

BAT CAVE

DEV.SO



NETWORK



META.SO

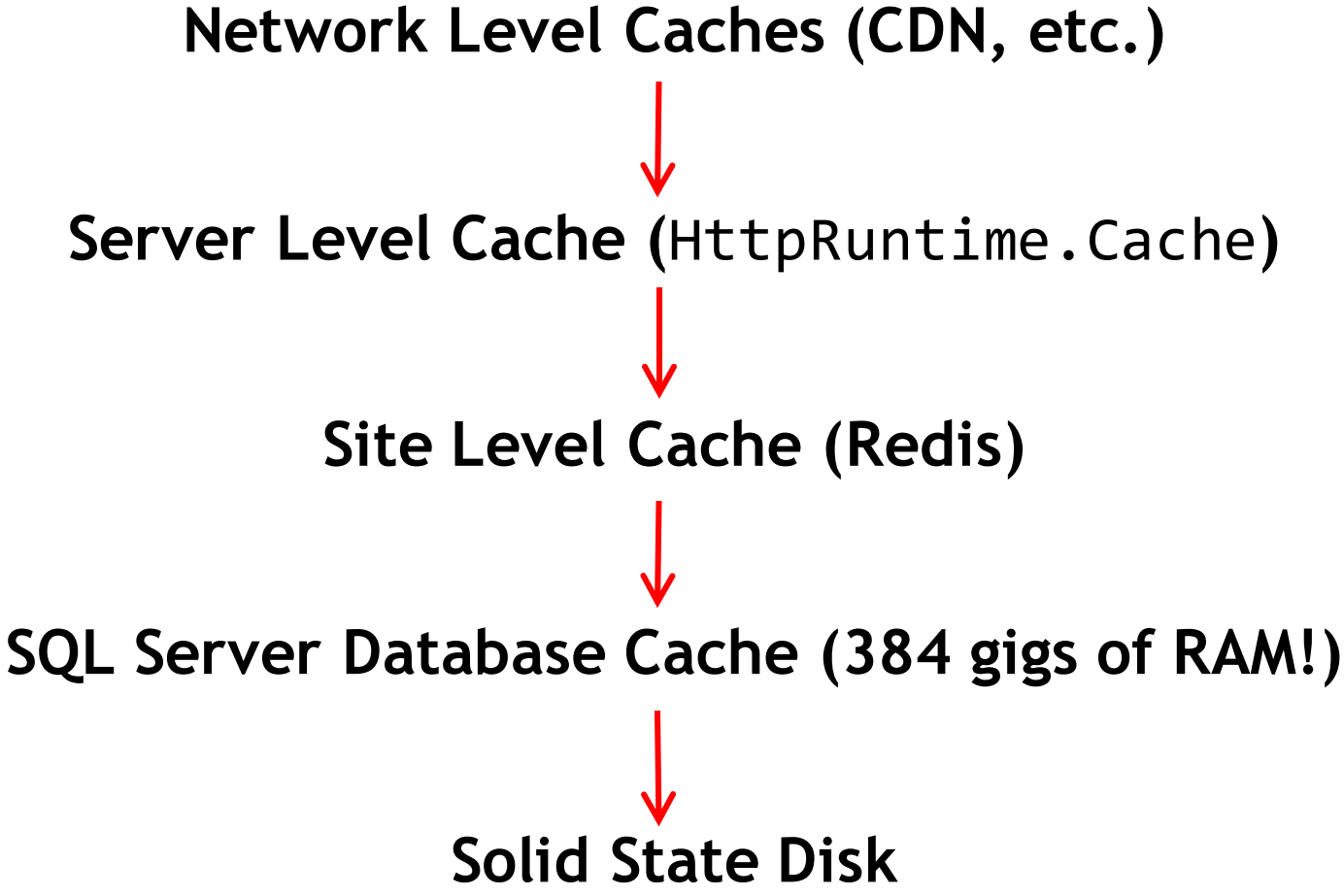
Move fast and break things

Move fast and break things*

* Not the home page or question page :-)

Move fast and break things*

* Not the home page or question page :-)


Network Level Caches (CDN, etc.)
Server Level Cache (HttpRuntime.Cache)
Site Level Cache (Redis)
SQL Server Database Cache (384 gigs of RAM!)
Solid State Disk

Help Center

New to Stack Overflow? Find out everything you need to get started by taking the tour. If you still have questions, come back and check out the pinned articles.

Take the tour

Here's how it works:

Anybody can ask a question

Anybody can answer

The best answers are voted up and rise to the top

Search help pages

search

Find out more about...

Asking

- What topics can I ask about here?
- What types of questions should I avoid asking?
- What does it mean if a question is "closed" or "on hold"?
- Why are questions no longer being accepted from my account?
- Why and how are some questions deleted?

» View more

Our model

- What kind of behavior is expected of users?
- How do I find topics I'm interested in?
- How do I search?
- What is "meta"? How does it work?
- How do I format my posts using Markdown or HTML?

» View more

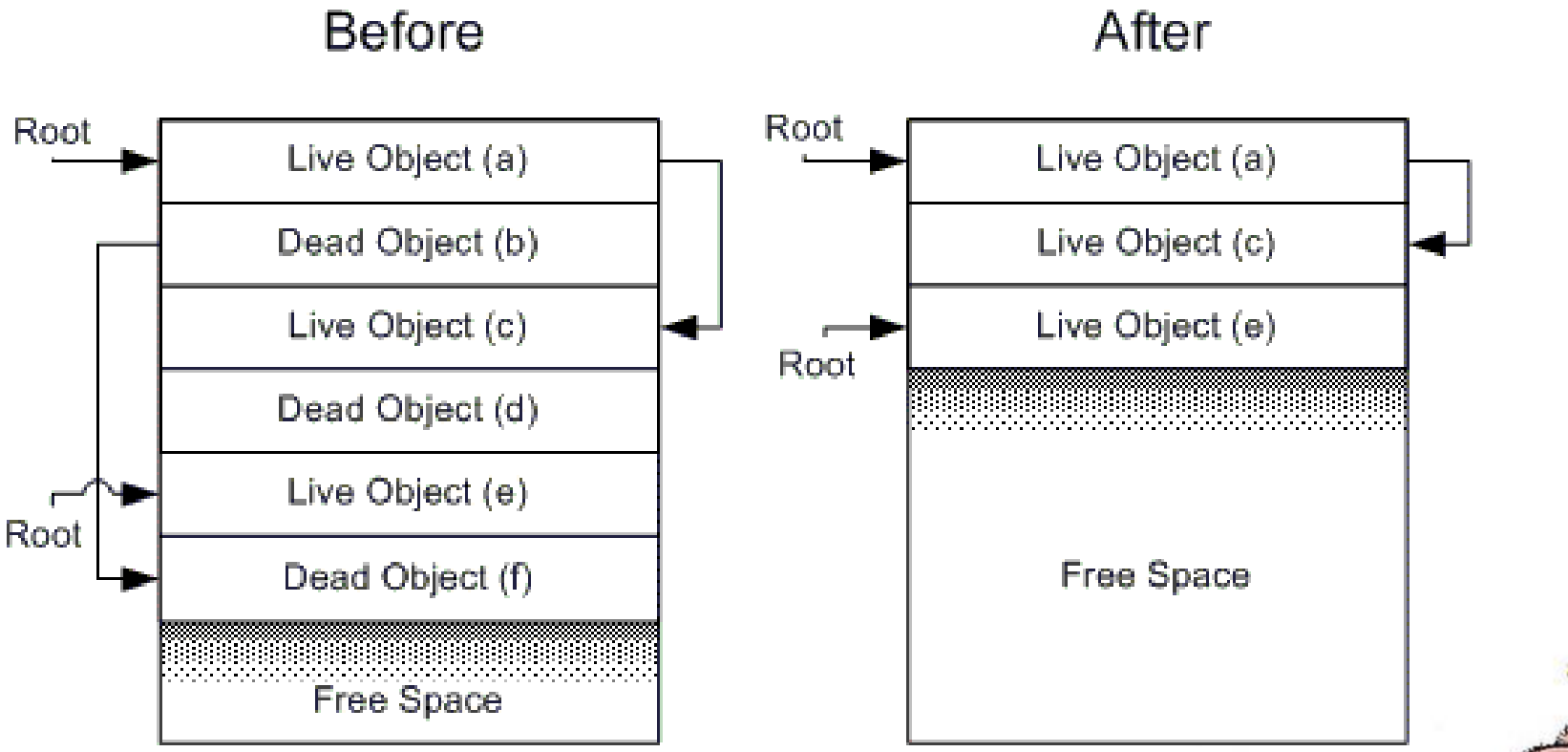
Reputation & Moderation

- What is a bounty? How can I start one?
- What is reputation? How do I earn (and lose) it?
- Who are the site moderators, and what is their role here?
- Why is voting important?

» View more

```
203
204     private static List<HelpPost> All()
205     {
206         return Current
207             .GlobalCache
208             .GetSet<List<HelpPost>>(
209                 CacheKey,
210                 (old, ctx) =>
211                 {
212                     using (SitesDBContext sitesDb = SitesDBContext.NewContext())
213                     {
214                         return sitesDb.Query<HelpPost>(@"SELECT p.Id, p.Title, p.Bc
215                     }
216                 }, 24 * 60 * 60, 24 * 60 * 60);
217     }
218
```

CACHE ALL THE THINGS !


Before
Root
Live Object (a)
Dead Object (b)
Live Object (c)
Dead Object (d)
Live Object (e)
Root
Dead Object (f)
Free Space
After
Root
Live Object (a)
Live Object (c)
Live Object (e)
Root
Free Space

Too Many Allocations

This is really the most basic thing that can go wrong.

Too Many Pointers

If you create a data structure that is a large mesh of pointers you'll have two problems. First, there will be a lot of object writes [...] and, secondly, when it comes time to collect that data structure, you will make the garbage collector follow all those pointers and if necessary change them all as things move around. [...] But if you create such a structure on a transitory basis, [...], then you will pay the cost much more often.

http://msdn.microsoft.com/en-us/library/ms973837.aspx#dotnetgcbasics_topic2

Too Many Allocations

This is really the most basic thing that can go wrong.

Too Many Pointers

If you create a data structure that is a large mesh of pointers you'll have two problems. First, there will be a lot of object writes [...] and, secondly, when it comes time to collect that data structure, you will make the garbage collector follow all those pointers and if necessary change them all as things move around. [...] But if you create such a structure on a transitory basis, [...], then you will pay the cost much more often.

http://msdn.microsoft.com/en-us/library/ms973837.aspx#dotnetgcbasics_topic2

have earned at least 10 reputation on this site.

62 Answers

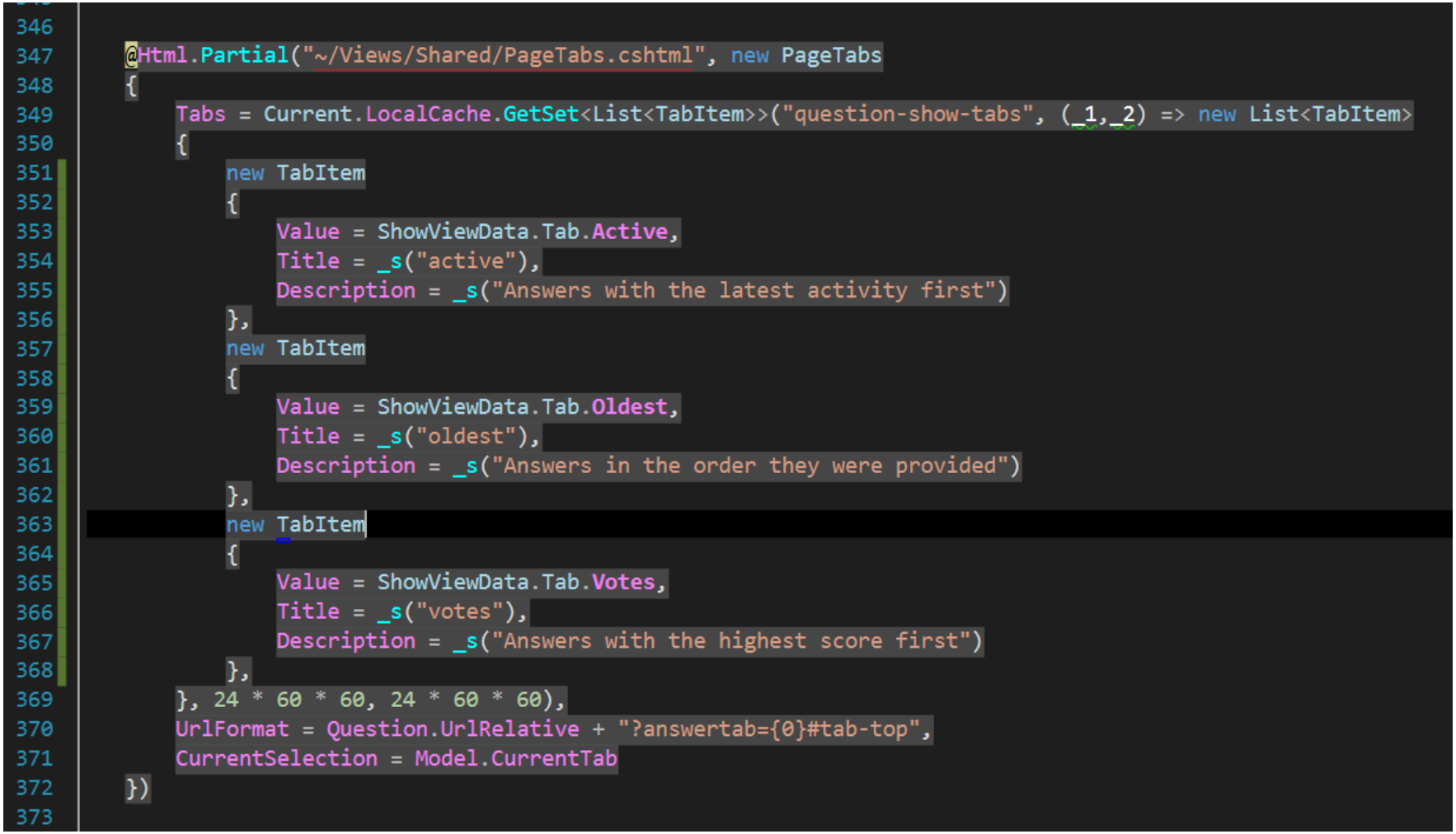
active oldest votes

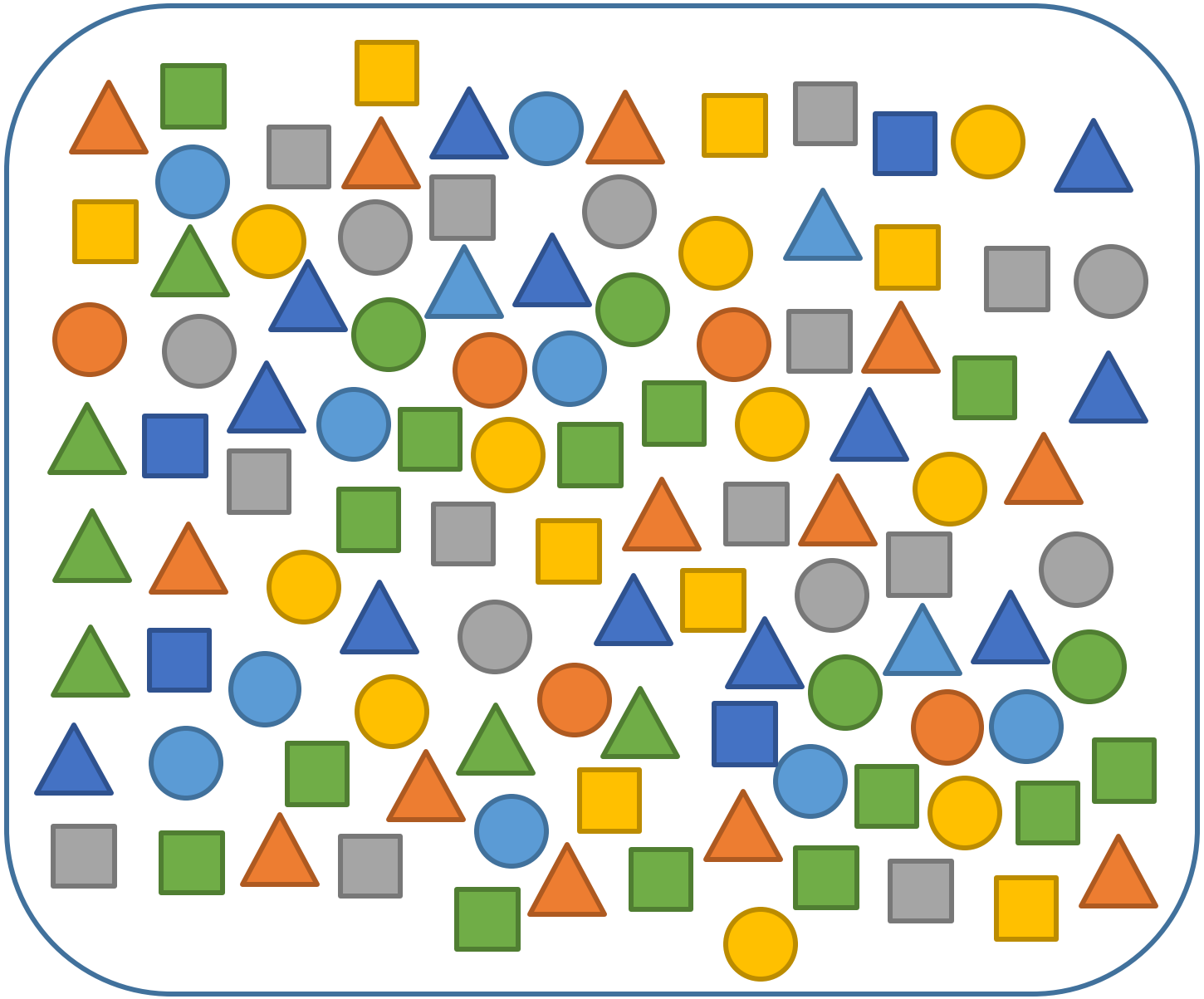
1 2 3 next

ents

4431

You can't parse [X]HTML with regex. Because HTML can't be parsed by regex. Regex is not a tool that can be used to correctly parse HTML. As I have answered in HTML-and-regex questions here so many times before, the use of regex will not allow you to consume HTML. Regular expressions are a tool that is insufficiently sophisticated to understand the constructs employed by HTML. HTML is not a regular language and hence cannot be parsed by regular expressions. Regex queries are not equipped to break down HTML into its meaningful parts. so many times but it is not getting to me. Even enhanced irregular regular expressions as used by Perl are not up to the task of parsing HTML. You will never make me crack. HTML is a language of sufficient complexity that it cannot be parsed by regular expressions. Even Jon Skeet cannot parse HTML using regular expressions. Every time you attempt to parse HTML with regular expressions, the unholy child weeps the blood of virgins, and Russian hackers pwn your webapp. Parsing HTML with regex summons tainted souls into the realm of the living. HTML and regex go together like love, marriage, and ritual infanticide. The <center> cannot hold it is too late. The force of regex and HTML together in the same conceptual space will destroy your mind like so much watery putty. If you parse HTML with regex you are giving in to Them and their blasphemous ways which doom us all to inhuman toil for the One whose Name cannot be expressed in the Basic Multilingual Plane, he comes. HTML-plus-regexp will liquify the nerves of the sentient whilst you observe, your psyche withering in the onslaught of horror. Regex-based HTML parsers are the cancer that is killing StackOverflow *it is too late it is too late we cannot be saved* the trangession of a child ensures regex will consume all living tissue (except for HTML which it cannot, as previously prophesied) *dear lord help us how can anyone survive this scourge* using regex to parse HTML has doomed humanity to an eternity of dread torture and security holes *using regex* as a tool to process HTML establishes a brea*ch between this world* and the dread realm of corrupt entities (like SGML entities, but *more corrupt) a mere glimp*se of the world of reg**ex parsers for HTML will ins**tantly transport a p*rogrammer's consciousness i*nto a world of ceaseless screaming, he comes~~, the pestilent s~~lithy regex-infection wil**l devour your HT**ML parser, application and existence for all time like Visual Basic only worse *he comes he comes do not fi*ght h**e comes, hi**s unholy radiance *destroying all enlightenment, HTML tags* **leaking from your eyes like liq***uid p*ain, the song of regular expre~~ssion parsing~~ will exti*nguish the voices of mor***tal man from the sp***here I can see it can you see it it is beautiful* the f `inal snuf` fing *of the lie***s of Man ALL IS LOST ALL IS L**OST th*e pony he comes* he c~~omes he comes~~ **the ich**or permeates all MY FACE MY FACE *ᵒh god n***o NO NOOOO N**Θ stop th*e an**̶͑̾̾ges `are` n**ot real ZALGO IS TONY THE PONY HE COMES**


@Html.Partial("~/Views/Shared/PageTabs.cshtml", new PageTabs
{
Tabs = Current.LocalCache.GetSet<List<TabItem>>("question-show-tabs", (_1,_2) => new List<TabItem>
{
new TabItem
{
Value = ShowViewData.Tab.Active,
Title = _s("active"),
Description = _s("Answers with the latest activity first")
},
new TabItem
{
Value = ShowViewData.Tab.Oldest,
Title = _s("oldest"),
Description = _s("Answers in the order they were provided")
},
new TabItem
{
Value = ShowViewData.Tab.Votes,
Title = _s("votes"),
Description = _s("Answers with the highest score first")
},
}, 24 * 60 * 60, 24 * 60 * 60),
UrlFormat = Question.UrlRelative + "?answertab={0}#tab-top",
CurrentSelection = Model.CurrentTab
})



Too Many Allocations

This is really the most basic thing that can go wrong.

Too Many Pointers

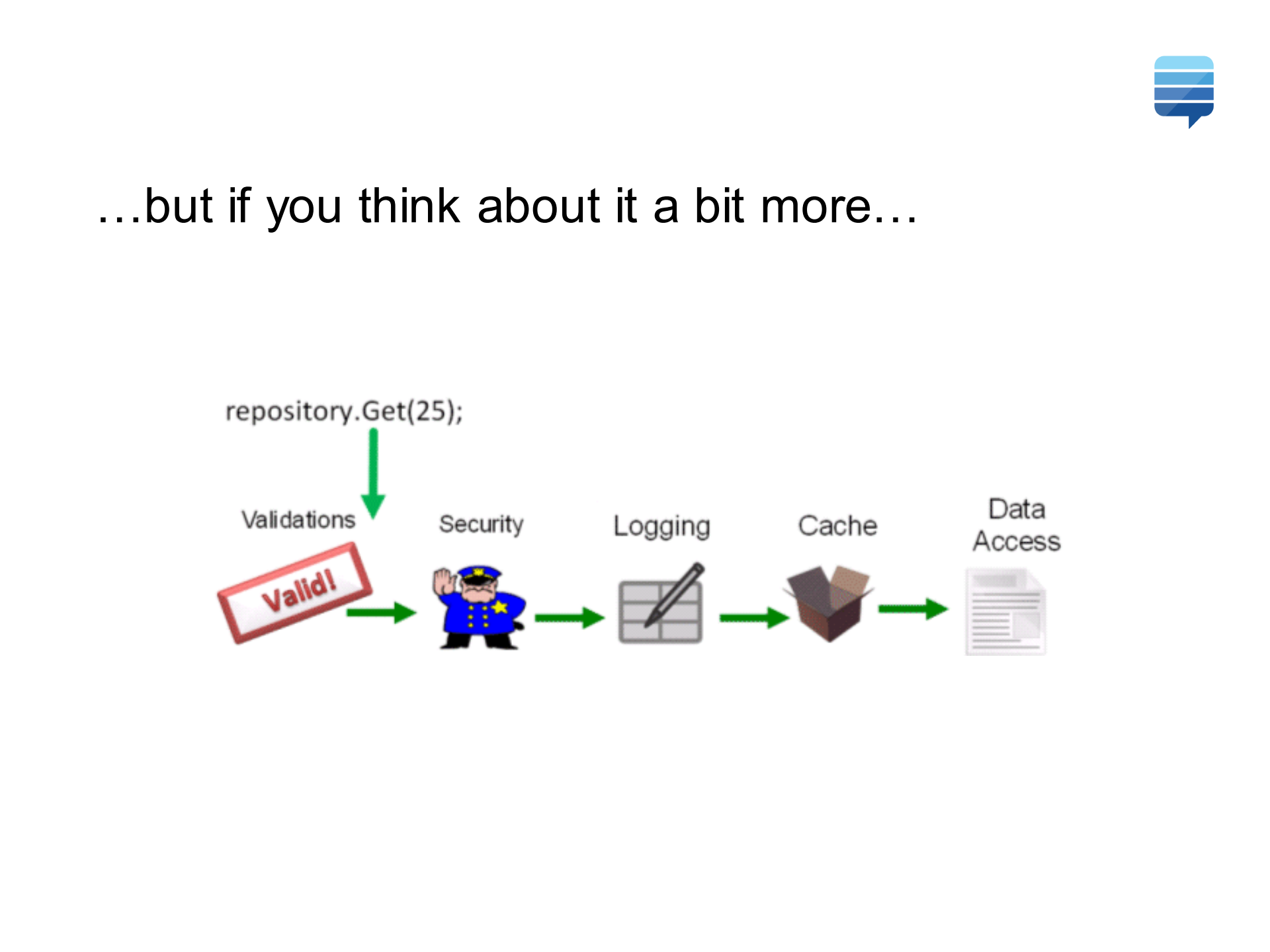
If you create a data structure that is a large mesh of pointers you'll have two problems. First, there will be a lot of object writes [...] and, secondly, when it comes time to collect that data structure, you will make the garbage collector follow all those pointers and if necessary change them all as things move around. [...] But if you create such a structure on a transitory basis, [...], then you will pay the cost much more often.

http://msdn.microsoft.com/en-us/library/ms973837.aspx#dotnetgcbasics_topic2

This is what you *think* you are doing…

```
IRepository<Order> orderRepository =
    container.Resolve<IRepository<Order>>();

Order order = orderRepository.Get(35);
```

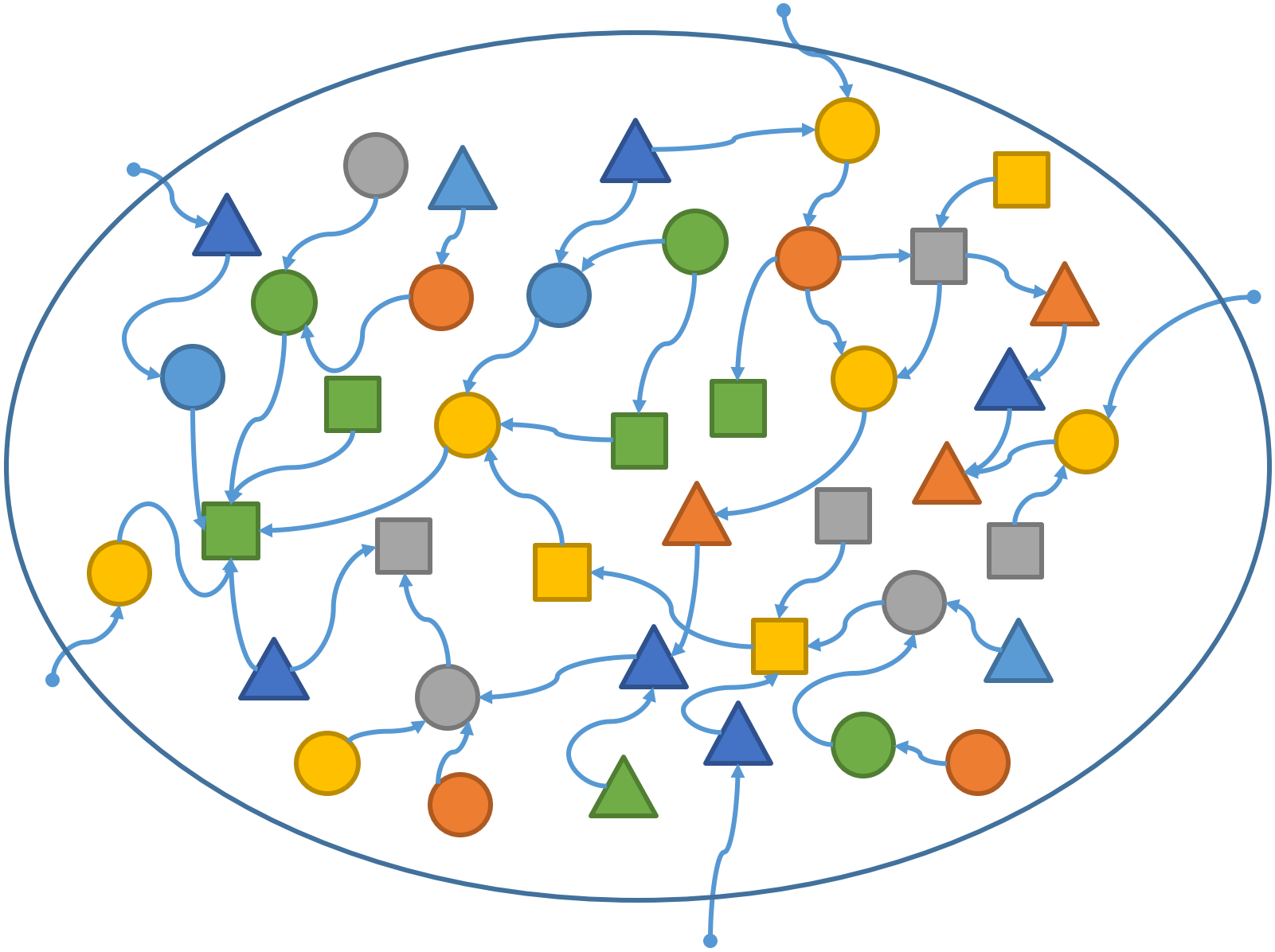


…but if you think about it a bit more…

...this is what you are *actually* doing!

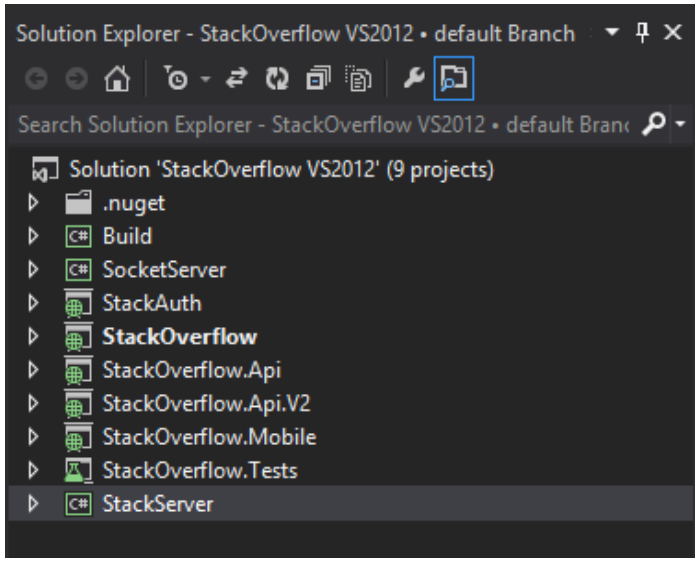
```
IRepository<Order> repository =
  new ValidatingOrderRepository (
    new SecurityRepository<Order> (
      new LoggingRepository<Order> (
        new CachingRepository<Order> (
          new NHibernateRepository<Order> ()
        )
      )
    )
  );

Order order = repository.Get(35);
```

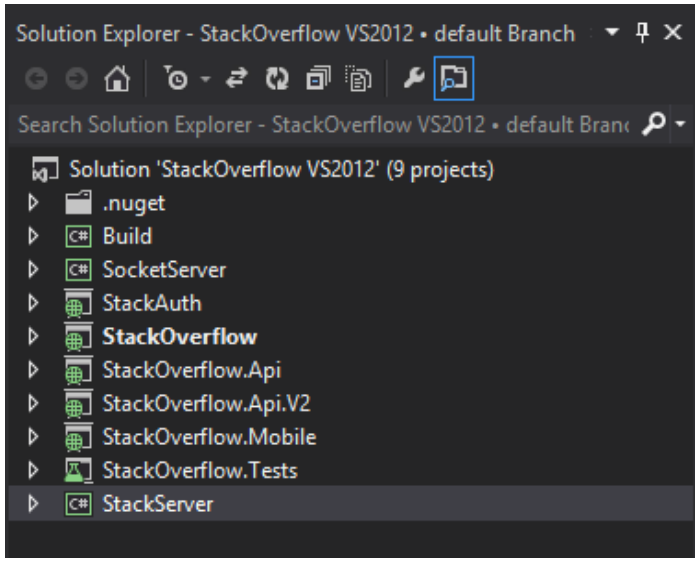


馄饨！

Few projects :-)



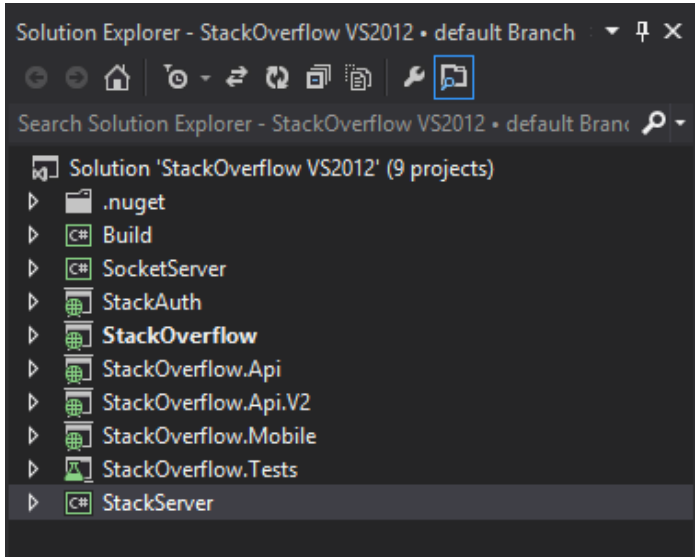
Few projects :-)



Few lines of code :-)

Project	Lines of Code
Build (Debug)	0
SocketServer (Debug)	332
StackAuth (Debug)	5,820
StackOverflow (Debug)	80,785
StackOverflow.Api (Debug)	8,015
StackOverflow.Api.V2 (Debug)	13,087
StackOverflow.Mobile (Debug)	1,718
StackOverflow.Tests (Debug)	2,989
StackServer (Debug)	669
Total	113,415

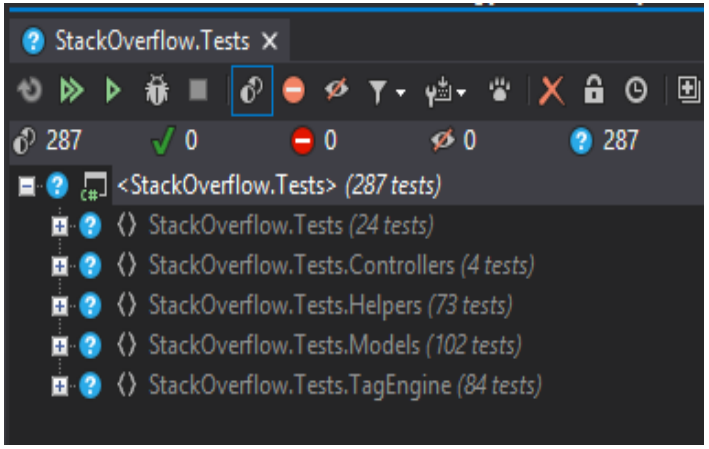
Few projects :-)



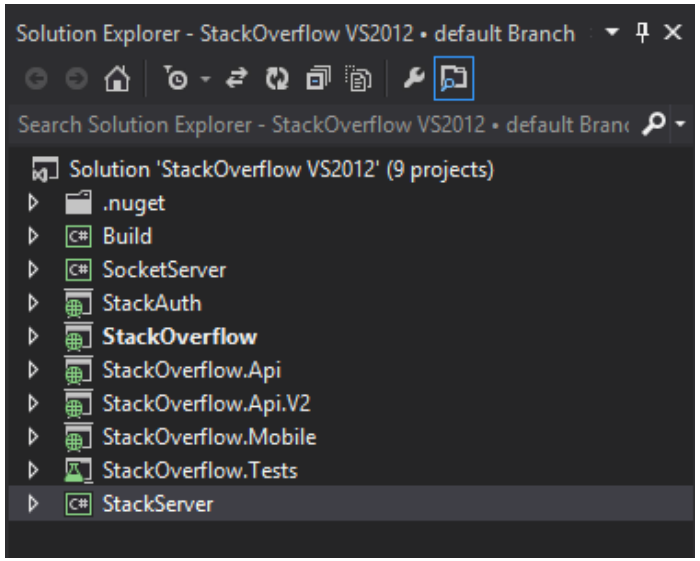
Few lines of code :-)

Project	Lines of Code
Build (Debug)	0
SocketServer (Debug)	332
StackAuth (Debug)	5,820
StackOverflow (Debug)	80,785
StackOverflow.Api (Debug)	8,015
StackOverflow.Api.V2 (Debug)	13,087
StackOverflow.Mobile (Debug)	1,718
StackOverflow.Tests (Debug)	2,989
StackServer (Debug)	669
Total	113,415

Eeek! very few tests :-S



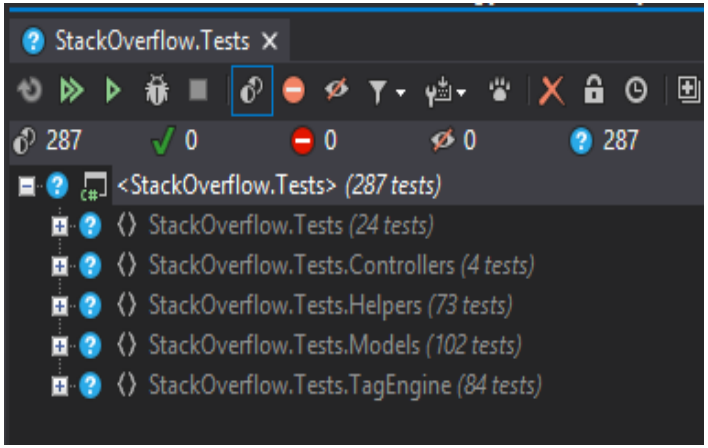
Few projects :-)



Few lines of code :-)

Project	Lines of Code
Build (Debug)	0
SocketServer (Debug)	332
StackAuth (Debug)	5,820
StackOverflow (Debug)	80,785
StackOverflow.Api (Debug)	8,015
StackOverflow.Api.V2 (Debug)	13,087
StackOverflow.Mobile (Debug)	1,718
StackOverflow.Tests (Debug)	2,989
StackServer (Debug)	669
Total	113,415

Eeek! very few tests :-S



Awesome community to help :-D

Developer -> Bugs

Tag Query: bug -status-* -careers Scope:

All Meta Sites Sort: Newest

Total In Tag Engine: 3142

1 2 3 4 5 … 105 next

Meta	Asked	S/AC	Title	Responses
Meta	12 hours ago	3/1	**Bad notification for the Mortarboard badge** bug reputation badges reputation-cap-badges	
User Experience Stack Exchange	18 hours ago	1/1	**Apple developer (HIG) links not allowed?** bug	
Meta	yesterday	17/0	**Users without full edit rights should not be able to submit edits to deleted posts** bug editing deleted-questions editing-badges	
Mathematics Stack Exchange	yesterday	1/0	**Image does not appear** bug images	
Meta	yesterday	1/0	**Chat isn't letting me move messages, and isn't showing an error either**	

YAGNI*.

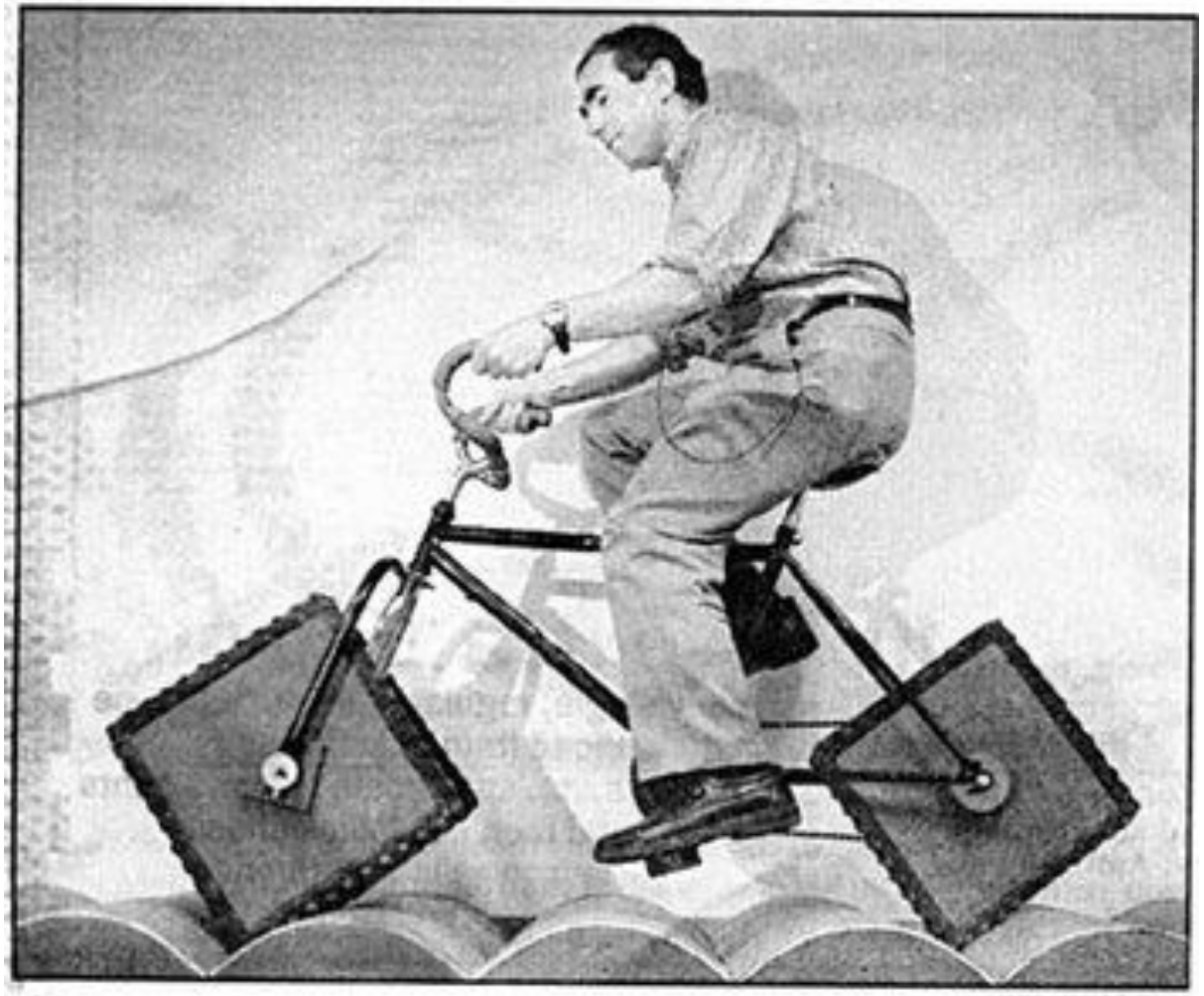
It works.

* You Ain’t Gonna Need It!

Reinventing the square wheel is the practice of unnecessarily engineering artifacts that provide functionality already provided by existing standard artifacts (reinventing the wheel) and ending up with a worse result than the standard (a square wheel). This is an anti-pattern which occurs when the engineer is unaware or contemptuous of the standard solution or does not understand the problem or the standard solution sufficiently to avoid problems overcome by the standard. It is mostly an affliction of inexperienced engineers, or the second-system effect.

* Source http://bit.ly/1eSLr8Z

BELIEVE IT OR NOT, HE'S ON A ROLL



Associated Press

Math professor Stan Wagon demonstrates his square-wheeled bicycle at Macalester College in St. Paul, Minn. In 1960, it was discovered that a square wheel would roll smoothly on a road made of catenaries (those bumpy things). Wagon said he became interested in the concept 7 years ago, did calculations and computer animations, then had the bike specially built.

pagedown

A JavaScript Markdown converter and editor

PUBLIC kevin-montrose / Sigil

dapper-dot-net

Simple SQL object mapper for ADO.NET

protobuf-net

Fast, portable, binary serialization for .NET

booksleeve

pipelined .NET bindings for redis

PUBLIC SamSaffron / MiniProfiler

stack-exchange-data-explorer

Stack Exchange Data Explorer

stackid

Stack Exchange OpenID Provider

MARKDOWNSHARP

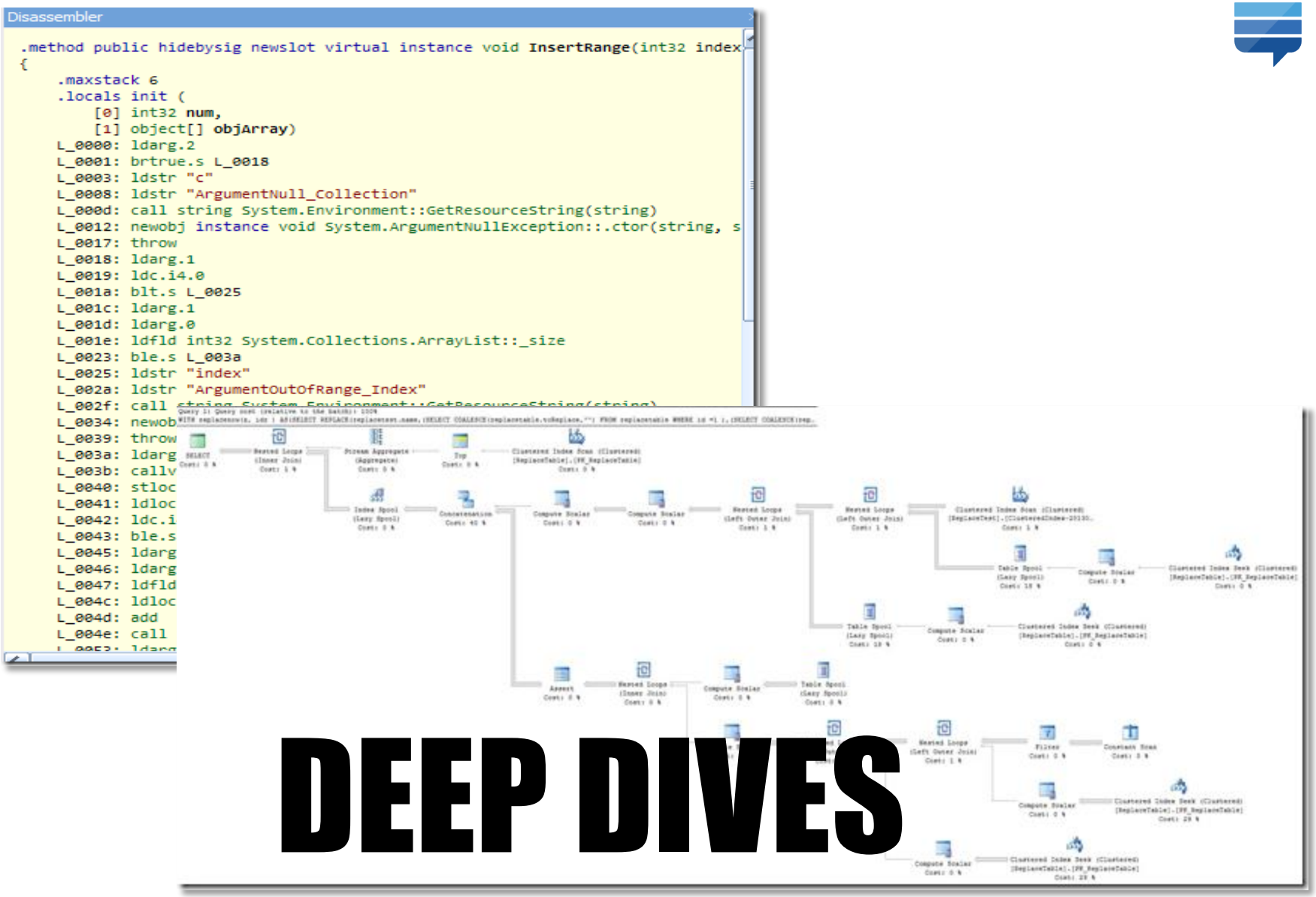
PUBLIC NickCraver / StackExchange.Exceptional

Write libraries &
open them to the world

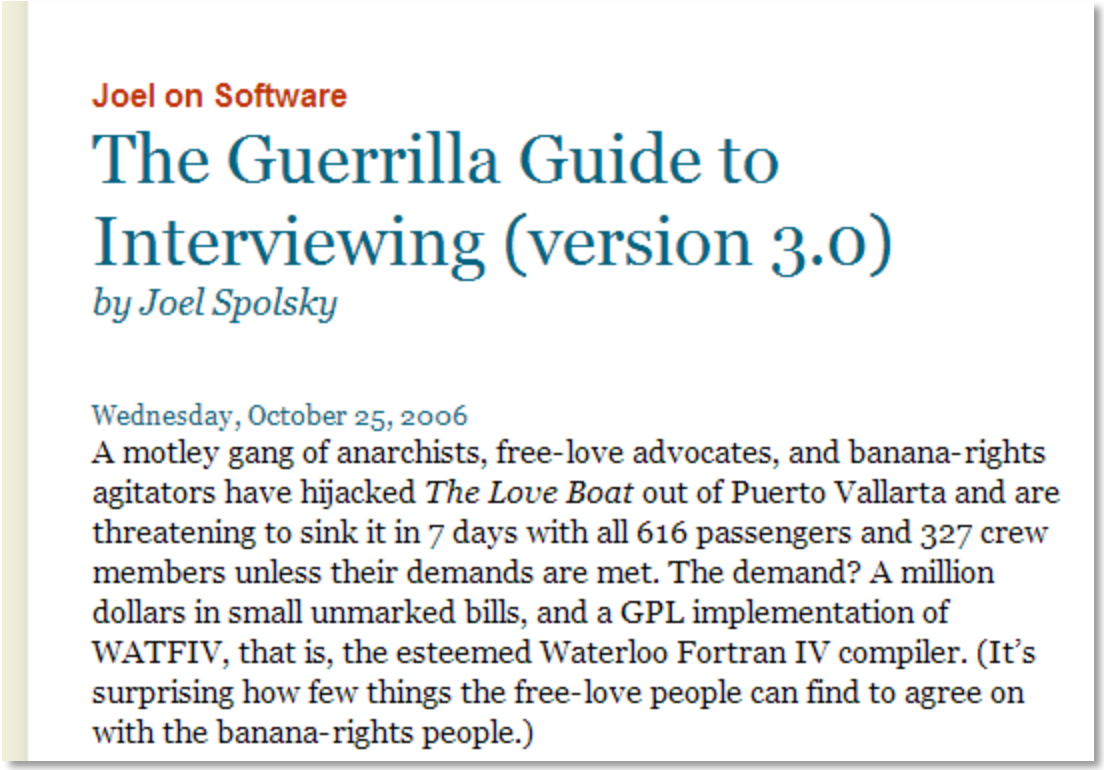

NTD
TELEVISION
SECRET SAUCE

TEAM


GEEK CULTURE


Disassembler
.method public hidebysig newslot virtual instance void InsertRange(int32 index
{
.maxstack 6
.locals init (
[0] int32 num,
[1] object[] objArray)
L_0000: ldarg.2
L_0001: brtrue.s L_0018
L_0003: ldstr "c"
L_0008: ldstr "ArgumentNull_Collection"
L_000d: call string System.Environment::GetResourceString(string)
L_0012: newobj instance void System.ArgumentNullException::.ctor(string, s
L_0017: throw
L_0018: ldarg.1
L_0019: ldc.i4.0
L_001a: blt.s L_0025
L_001c: ldarg.1
L_001d: ldarg.0
L_001e: ldfld int32 System.Collections.ArrayList::_size
L_0023: ble.s L_003a
L_0025: ldstr "index"
L_002a: ldstr "ArgumentOutOfRange_Index"
L_002f: call string System.Environment::GetResourceString(string)
L_0034: newob
L_0039: throw
L_003a: ldarg
L_003b: callv
L_0040: stloc
L_0041: ldloc
L_0042: ldc.i
L_0043: ble.s
L_0045: ldarg
L_0046: ldarg
L_0047: ldfld
L_004c: ldloc
L_004d: add
L_004e: call
DEEP DIVES

OPINIONATED HIRING



Joel on Software

The Guerrilla Guide to Interviewing (version 3.0)

by Joel Spolsky

Wednesday, October 25, 2006

A motley gang of anarchists, free-love advocates, and banana-rights agitators have hijacked *The Love Boat* out of Puerto Vallarta and are threatening to sink it in 7 days with all 616 passengers and 327 crew members unless their demands are met. The demand? A million dollars in small unmarked bills, and a GPL implementation of WATFIV, that is, the esteemed Waterloo Fortran IV compiler. (It's surprising how few things the free-love people can find to agree on with the banana-rights people.)

REMOTE WORK


IMPERIAL TIE FIGHTER
IMPERIAL TIE FIGHTER
IMPERIAL TIE FIGHTER

CONCLUSION

- Performance is a feature
- Always. Be. Shipping.
- Use your circumstances.
- Open source your libraries
- 3 obscenely big monitors.

```
$_='@mk=uf=radimdp1Z--&ewxuhhl';tr/=1m-za-l@&Z/ !a-zP@\n/&print;
```

问题？

Marco Cecconi
@sklivvz
sklivvz@stackoverflow.com