Copyright Policy

All content included on the Site or third-party platforms as part of the class, such as text, graphics, logos, button icons, images, audio clips, video clips, live streams, digital downloads, data compilations, and software, is the property of BitTiger or its content suppliers and protected by copyright laws.

Any attempt to redistribute or resell BitTiger content will result in the appropriate legal action being taken.

We thank you in advance for respecting our copyrighted content.For more info see https://www.bittiger.io/termsofuse and https://www.bittiger.io/termsofservice

版权声明

所有太阁官方网站以及在第三方平台课程中所产生的课程内容，如文本，图形，徽标，按钮图标，图像，音频剪辑，视频剪辑，直播流，数字下载，数据编辑和软件均属于太阁所有并受版权法保护。

对于任何尝试散播或转售BitTiger的所属资料的行为，太阁将采取适当的法律行动。

我们非常感谢您尊重我们的版权内容。

有关详情，请参阅

https://www.bittiger.io/termsofuse

https://www.bittiger.io/termsofservice

Before We Start

请扫码添加讲座小助手入群

BitTiger官方公众号

讲座负责人

后端工程师必备技能

John

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

About Me

- CS Master
- Bloomberg
  Realtime finance data, Unix, TCP, C++
- Microsoft Bing Ads
  Product Ads, Algorithm, Distributed system, Big data
- FLAG
  Big Data, Distributed system

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

Data structures & Algorithms

- Why 70% of interview questions are about data structure and algorithm
  - Data structure is a particular way of storing and organizing information in a computer
  - Backend system deal with : store data , retrieve data, process data
  - Data structures are essential ingredients of many efficient algorithms
  - Data structures make your code more readable and easy to maintain
  - Algorithm make your program run faster, solve complex problem like rank web page, make money like Ads, trading related algorithm
  - Algorithm can also ....

Break your system !!

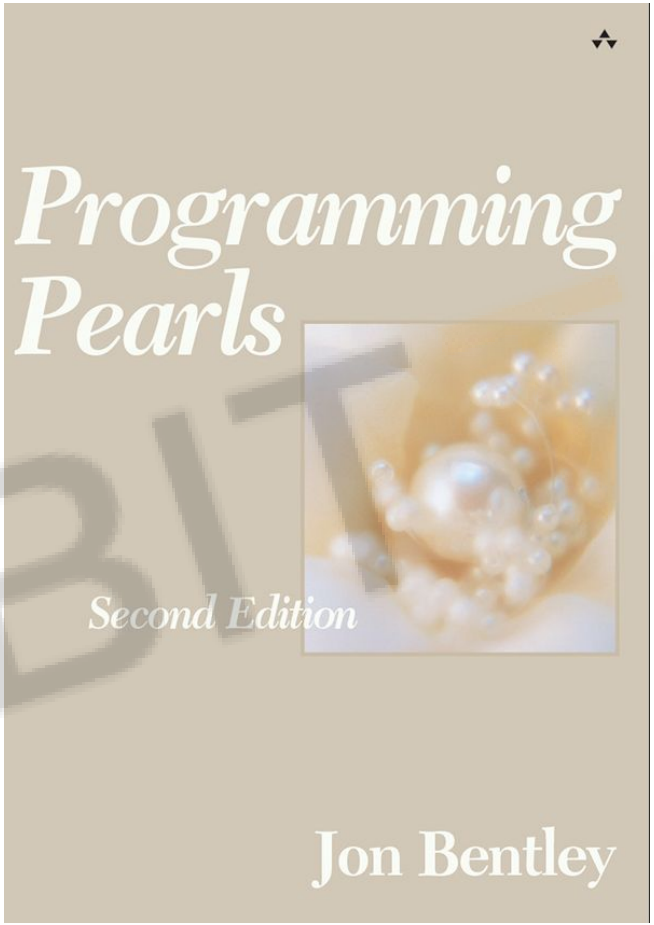
Data structures & Algorithms

- 10 most costly algorithm bug in history
  - Ariane 5 Flight 501 - $8.5 billion
  - NASA's Mars Climate Orbiter - $125 million
  - EDS Child Support System - $1 billion
  - Soviet Gas Pipeline Explosion - explosion
  - Heathrow Terminal 5 Opening - $100 million
  - The Mariner 1 Spacecraft - $18 million
  - The Morris Worm - $100 million
  - Patriot Missile Error - killed 28 American soldiers
  - Pentium FDIV bug - $475 million
  - Knight's $440 Million Error

Data structures & Algorithms


```
1:     public static int binarySearch(int[] a, int key) {
2:         int low = 0;
3:         int high = a.length - 1;
4:
5:         while (low <= high) {
6:             int mid = (low + high) / 2;
7:             int midVal = a[mid];
8:
9:             if (midVal < key)
10:                 low = mid + 1
11:             else if (midVal > key)
12:                 high = mid - 1;
13:             else
14:                 return mid; // key found
15:         }
16:         return -(low + 1);  // key not found.
17:     }
```

Data structures & Algorithms



Data structures & Algorithms

ORACLE

United States | Communities | I am a... | I want to... | Secure Search

Products and Services | Downloads | Store | Support | Education | Partners | About | Oracle Technology Network

Oracle Technology Network > Java > Java SE > Community > Bug Database

- Community
- Java Embedded
- Java Card
- Java DB
- Java EE
- JavaFX
- Java Magazine
- Java ME
- Java SE Advanced & Suite
- Java SE
- Java SE Support
- Java Advanced
- Java TV
- New to Java
- Web Tier

JDK-5045582 : (coll) binarySearch() fails for size larger than 1<<30

Details

Type: Bug
Status: Closed
Project Name: JDK
Component: core-libs
Sub-Component: java.util:collections
Priority: P2
Resolution: Fixed
Affected Versions: 5.0,6
Fixed Versions: 6 (b83)

Submit Date: 2004-05-11
Updated Date: 2012-10-08
Resolved Date: 2006-04-29
OS: generic,linux
CPU: generic,x86

Related Reports

Duplicate: JDK-6412541 - Arrays.binarySearch does not work for arrays larger than 1<<30
Relates: JDK-6437371 - (coll) TreeMap.buildFromSorted midpoint calculation
Relates: JDK-4306897 - (coll) Add java.util.Arrays.binarySearch(a, key, fromIndex, toIndex)
Relates: JDK-5050278 - BigInteger constructor gives NegativeArraySizeException

Sub Tasks

Description

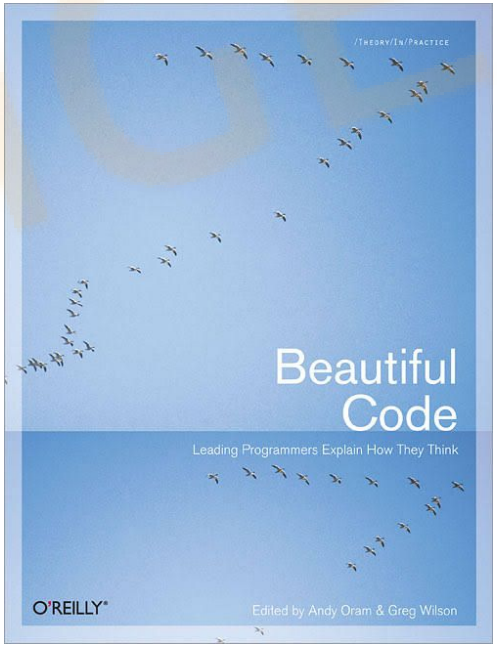
```
Name: rmT116609                Date: 05/11/2004


FULL PRODUCT VERSION :
java version "1.5.0-beta"
Java(TM) 2 Runtime Environment, Standard Edition (build 1.5.0-beta-b32c)
Java HotSpot(TM) Client VM (build 1.5.0-beta-b32c, mixed mode)


ADDITIONAL OS VERSION INFORMATION :
Linux freeway 2.4.21-4-686 #1 Sat Aug 2 23:27:25 EST 2003 i686 GNU/Linux


A DESCRIPTION OF THE PROBLEM :
java.util.Arrays.binarySearch() will throw an ArrayIndexOutOfBoundsException if
the array
is large.  This is caused by overflow in the calculation:

          int mid = (low + high) >> 1;

The correct calculation uses unsigned shift:

          int mid = (low + high) >>> 1;

There are similar problems in Collections, and TreeMap also includes the faulty
calculation
```

Data structures & Algorithms


```
6:              int mid = low + ((high - low) / 2);
```

Data structures & Algorithms

- What you can learn from shuati (刷题)
  - Apply Data structure and algorithm to real life problem
  - Code style in interview and work
  - Performance of your code



Data structures & Algorithms

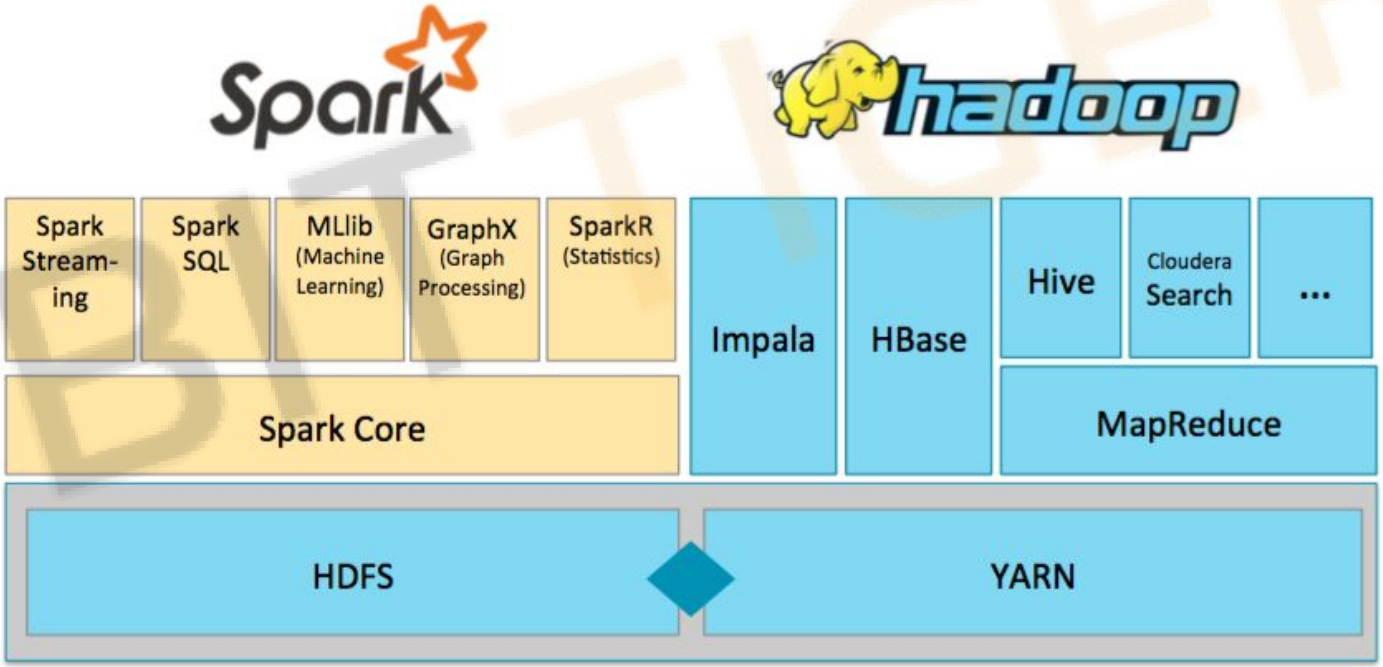
- Data structure and algorithms in C++, JAVA, Python
  - C++ std library : vector, map, unordered_map, priority_queue,algorithm
  - Java: List, HashMap, PriorityQueue, Queue, sort, search, Regular Expressions
  - Python: String, List, tuple, Dictionary,Set, Regular Expressions,Random

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

Big data

- Big data ecosystem



Big data

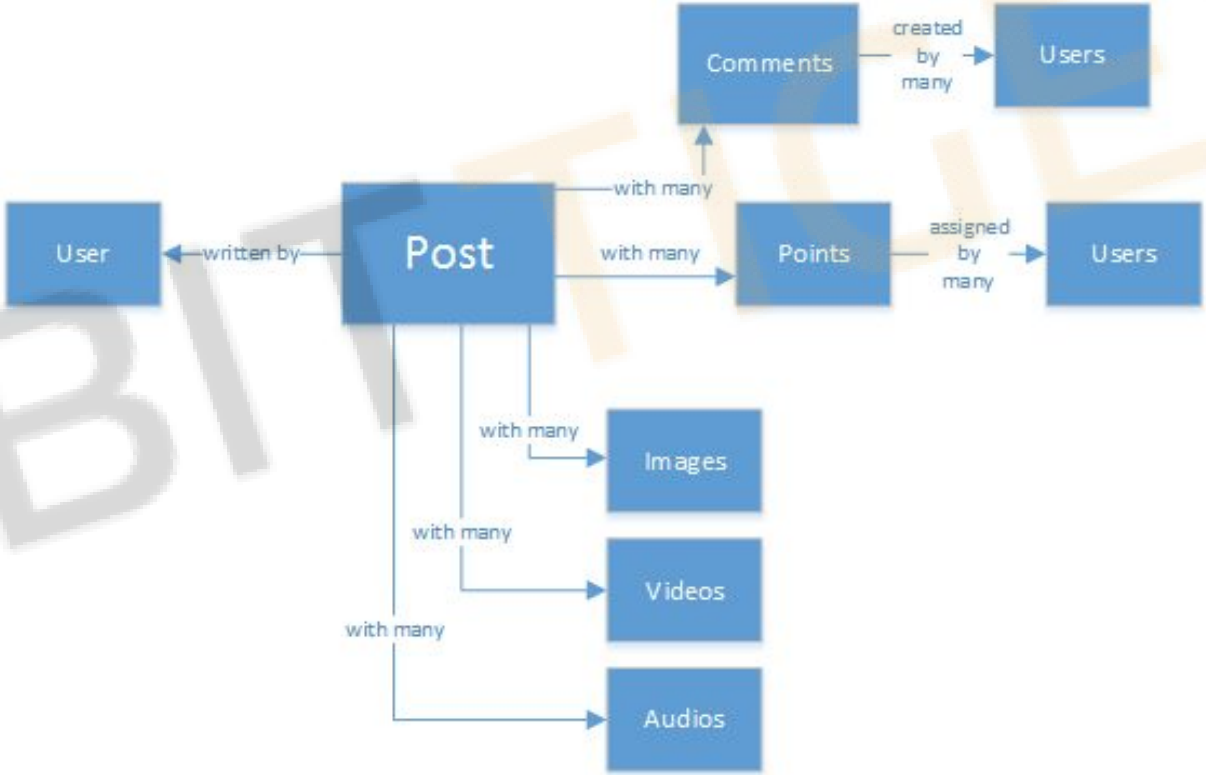
- What big data skills backend developer need
  - Run MapReduce job: Hadoop, Spark
  - Stream processing: RabbitMQ, Kafka, Storm
- Why backend developer need big data skills?
  - Offline processed data will be stored in DB or cache and used by online system

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

Relational and NoSQL databases

- When to use NoSQL
  - store facebook post in RMDB ?



Relational and NoSQL databases

- When to use NoSQL
  - Imagine dynamically load a stream of post
  - Perform query with eight table joins just to retrieve the content of one post
  - Thousands joins during loading a stream of post

Performance Disaster!

Relational and NoSQL databases

- When to use NoSQL
  Document store is a solution

```
{
   "id":"aw12-rbds2-234e-3467",
   "title":"post title",
   "date":"2017-03-01",
   "body":"this is an awesome post about bittiger",
   "createdBy": “tiger”,
   "images":["http://myfirstimage.png","http://mysecondimage.png"],
   "videos":[
      {"url":"http://myfirstvideo.mp4", "title":"The first video"},
      {"url":"http://mysecondvideo.mp4", "title":"The second video"}
   ],
   "audios":[
      {"url":"http://myfirstaudio.mp3", "title":"The first audio"},
      {"url":"http://mysecondaudio.mp3", "title":"The second audio"}
   ]
}
```

Relational and NoSQL databases

- SQL vs NoSQL

	SQL	NoSQL
Data storage	Stored in a relational model, with rows and columns. Rows contain all of the information about one specific entry/entity, and columns are all the separate data points; for example, you might have a row about a specific car, in which the columns are 'Make', 'Model', 'Colour' and so on.	The term "NoSQL" encompasses a host of databases, each with different data storage models. The main ones are: document, graph, key-value and columnar. More on the distinctions between them below.
Schemas and Flexibility	Each record conforms to fixed schema, meaning the columns must be decided and locked before data entry and each row must contain data for each column. This can be amended, but it involves altering the whole database and going offline.	Schemas are dynamic. Information can be added on the fly, and each 'row' (or equivalent) doesn't have to contain data for each 'column'.
Scalability	Scaling is vertical. In essence, more data means a bigger server, which can get very expensive. It is possible to scale an RDBMS across multiple servers, but this is a difficult and time-consuming process.	Scaling is horizontal, meaning across servers. These multiple servers can be cheap commodity hardware or cloud instances, making it alot more cost-effective than vertical scaling. Many NoSQL technologies also distribute data across servers automatically.
ACID Compliancy (Atomicity, Consistency, Isolation, Durability)	The vast majority of relational databases are ACID compliant.	Varies between technologies, but many NoSQL solutions sacrifice ACID compliancy for performance and scalability

Relational and NoSQL databases

- Document store : CouchDB and MongoDB
- Key-Value Stores:Redis, RocksDB and Dynamo
- Graph Databases : Neo4J and InfiniteGraph
- Column Databases: HBase, Cassandra

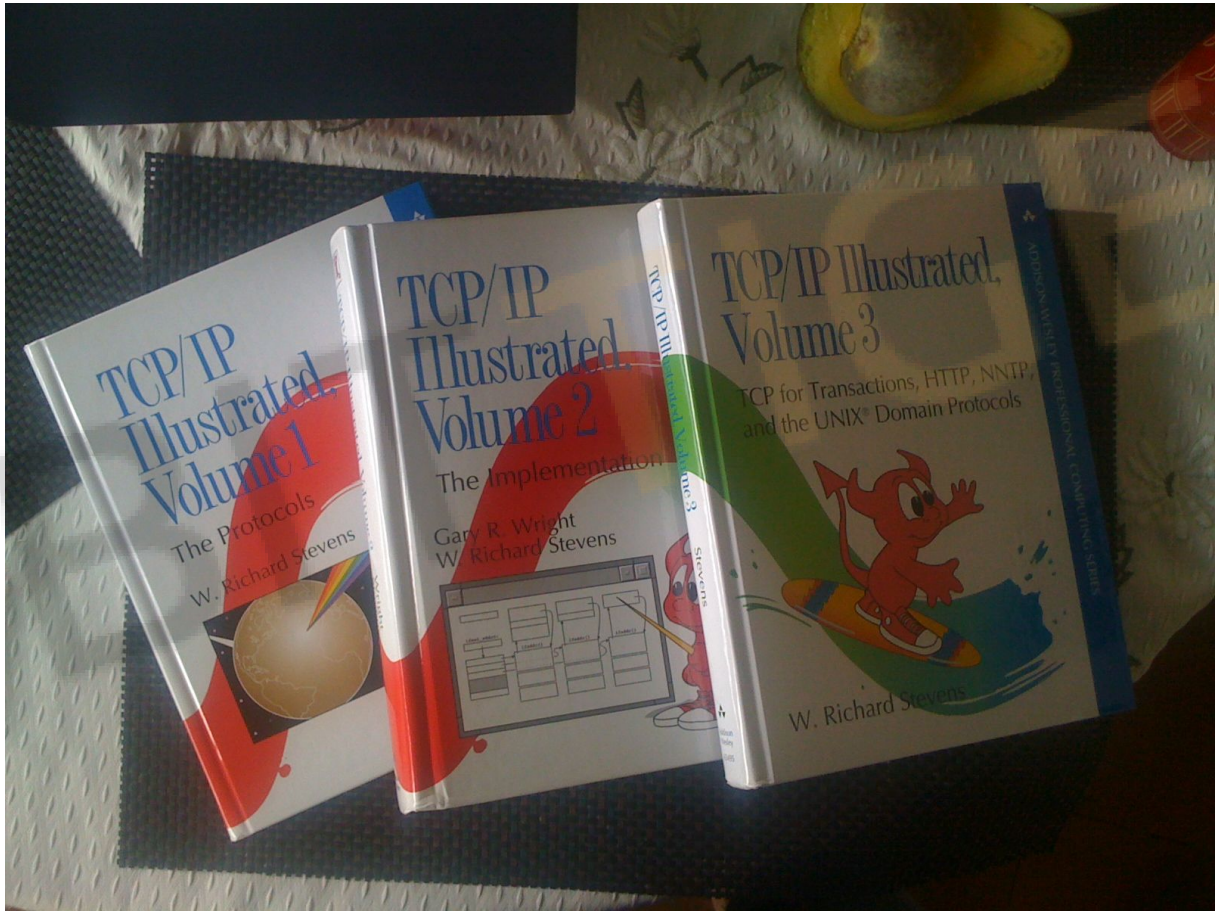
Relational and NoSQL databases

- How to get started
  - Download one of the NoSQL DB
  - Prepare sample data
  - Develop a backend system to retrieve, store data to NoSQL DB

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

Network programming



Network programming

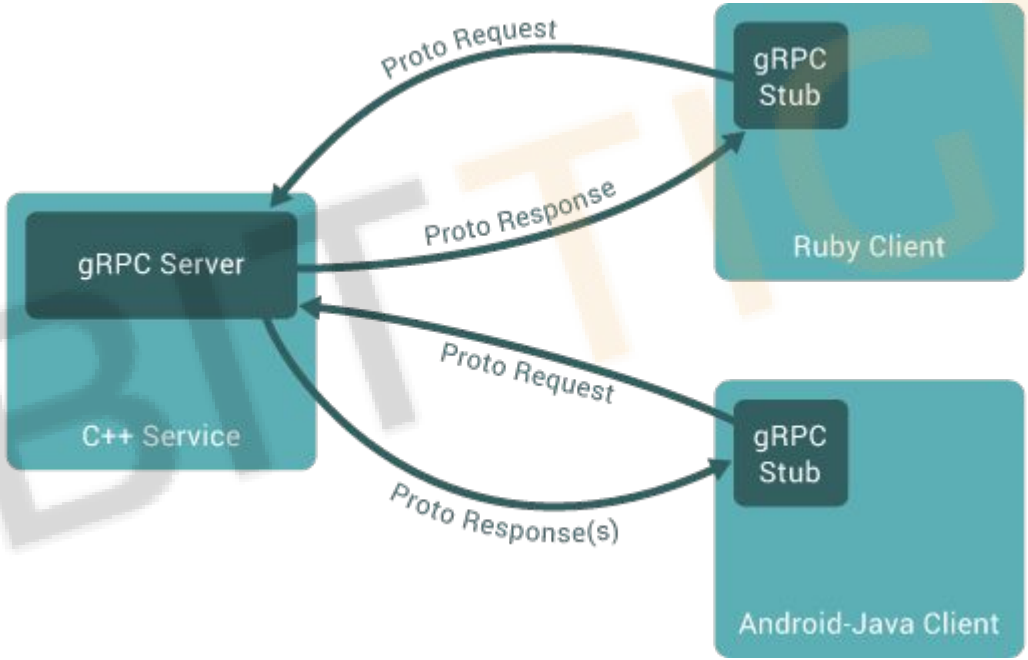
- Socket programing
  - Understand socket
  - Socket is everywhere: connect to SQL, NoSQL, rabbitMQ
  - WebSocket
  - How to get started : write a simple client - server model program with low level API like sys/socket.h
  - Playa round high level library like Netty, Twisted,Boost.Asio

Network programming

- RPC (Remote Procedure Call)
  - service interface
  - the structure of the payload messages
- example
  - Thrift : open sourced by Facebook
  - Protocol buffer: open sourced by Google
  - gRPC: based on Protocol buffer

Network programming

- RPC (Remote Procedure Call)



Network programming

- RPC (Remote Procedure Call)

Schema

```
// The greeting service definition.
service Greeter {
  // Sends a greeting
  rpc SayHello (HelloRequest) returns (HelloReply) {}
}
// The request message containing the user's name.
message HelloRequest {
  string name = 1;
}
// The response message containing the greetings
message HelloReply {
  string message = 1;
}
```

Network programming

- Why RPC (Remote Procedure Call) ?
  - Large-scale distributed systems actually composed of microservices
  - Communication between microservices are structured
  - Cross language, platform
  - Async / non-blocking io

Network programming

- Http Server
  - Linux, Apache, MySQL, PHP (LAMP)
  - Java + Jetty
  - Python + Django

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

Testing

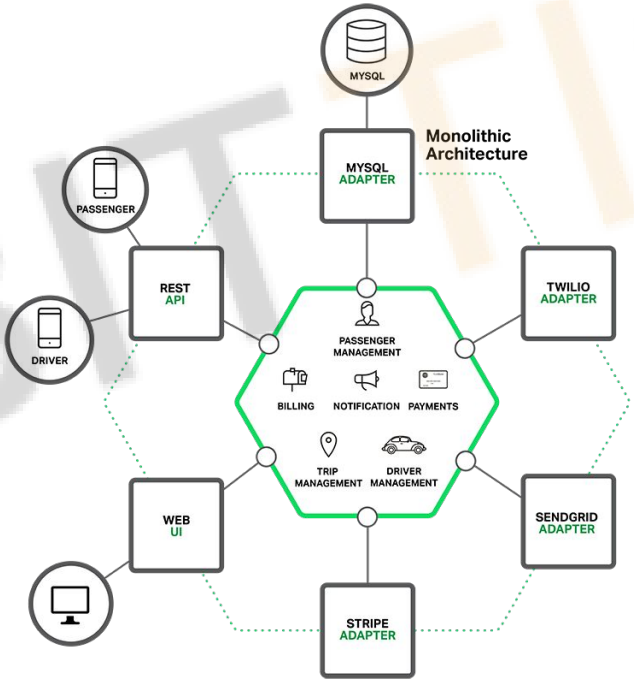
- Unit test : must have in your code review
- Integration test : test with other components
- Performance test : avoid performance regression
- A/B test : compare performance between different solution, broadly used in UI/UX, machine learning algorithm

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

Deal with complex system

- Architecture
  - Get architecture diagram in 1st week
  - If there is no architecture diagram, then draw one



Deal with complex system

- KPI monitor
  - Why monitor KPI ?

Amazon AWS S3 outage is breaking things for a lot of websites and apps

Posted Feb 28, 2017 by Darrell Etherington (@etherington)

Amazon's S3 web-based storage service is experiencing widespread issues, leading to service that's either partially or fully broken on websites, apps and devices upon which it relies. The AWS offering provides hosting for images for a lot of sites, and also hosts entire websites, and app backends including Nest.

The S3 outage is due to "high error rates with S3 in US-EAST-1," according to Amazon's AWS service health dashboard, which is where the company also says it's working on "remediating the issue," without initially revealing any further details.

Affected websites and services include Quora, newsletter provider Sailthru, Business Insider, Giphy, image hosting at a number of publisher websites, filesharing in Slack, and many more. Connected lightbulbs, thermostats and other IoT hardware is also being impacted, with many unable to control these devices as a result of the outage.

Crunchbase

Amazon Web Services

FOUNDED
2006

OVERVIEW
Amazon Web Services provides information technology infrastructure services to businesses in the form of web services. Its products and solutions include cloud computing, compute, networking, storage and content delivery, databases, analytics, application services, deployment and management, mobile services, applications, AWS marketplace software, startups, enterprises, partners, government and education, ...

LOCATION
Seattle, WA

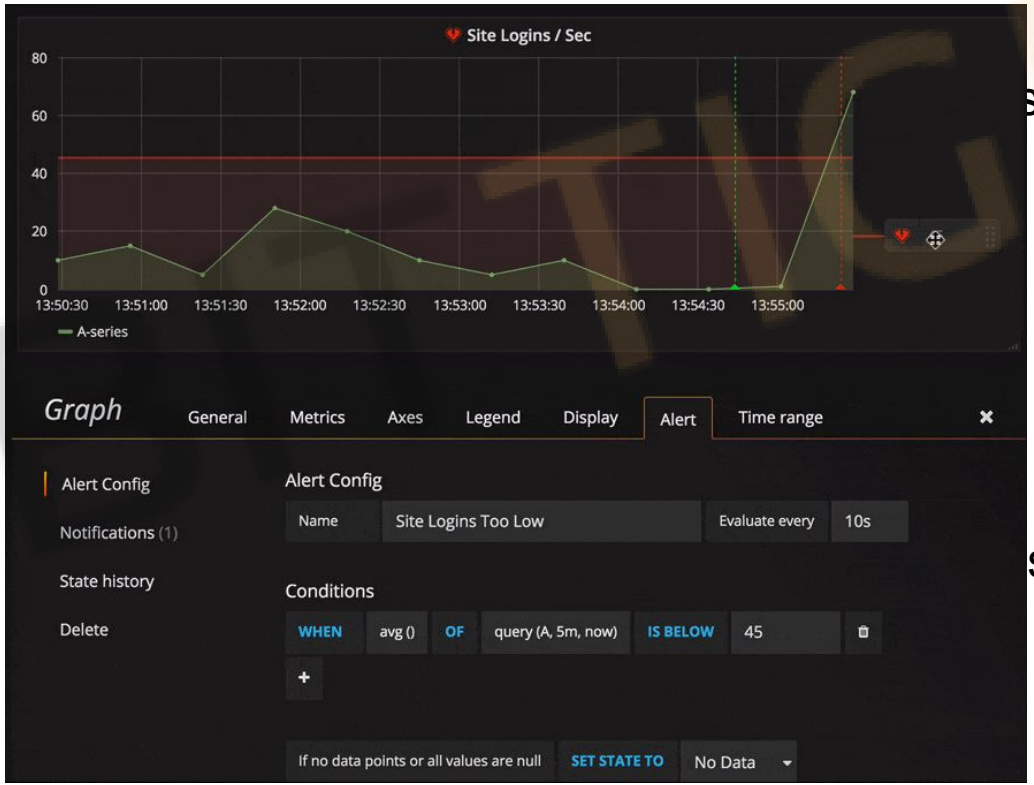
CATEGORIES
Web Development, Software, Information Technology, Information Services

Deal with complex system

- KPI monitor

Deal with complex system

- KPI monitor



siness

shold, for

Deal with complex system

- KPI monitor
  - grafana.org
  - zabbix.com

Deal with complex system

- Performance Tuning
  - Slow Response Times
    - include time taken for the different components behind your web server (application servers, DBs) to process the request
    - Response Times: The 3 Important Limits By Jakob Nielsen 1993

Response time	What user feel
0.1 second	system is reacting instantaneously
1 second	user's flow of thought to stay uninterrupted
10 second	What happened to this crappy App / Website

Deal with complex system

- Performance Tuning
  - Slow Response Times
    - include time taken for the different components behind your web server (application servers, DBs) to process the request
    - Response Times: The 3 Important Limits By Jakob Nielsen 1993
  - Steps:
    - Add latency monitor per component
    - Pressure test to find bottleneck
    - Find root cause
      - Memory Leak and Garbage Collection, java.lang.OutOfMemoryError
      - Deadlocks and Threading Issues
      - High Resource Usage (CPU/Disk/ Network)
      - Database Issues and Slow Queries
        SQL Timeout

Deal with complex system

- Diagnosis livesite issue, analyticals skill
  - Understand KPI monitor graph
  - Follow guidebook if have one
  - Login to server and read log
  - Replicate issue on dev/beta environment
  - Write analytic report

Deal with complex system

- Take Notes
  - Legacy system
  - Scripts for debug
  - Most used system monitoring linux command

    most CPU-intensive tasks : top

    processes, memory, paging, block IO : vmstat

Outline

- About me
- Data structures & Algorithms
- Big data
- Relational and NoSQL databases
- Network programming
- Testing
- Ability to deal with complex systems
- Communication Skill

Communication Skill

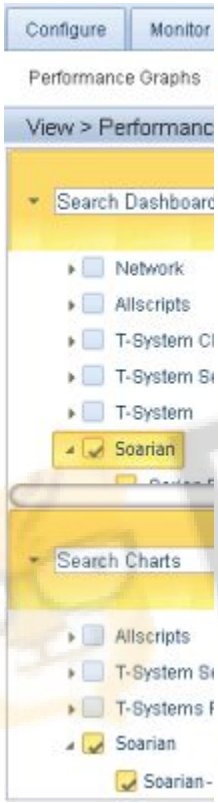
- Communication channel comparison

	Face 2 Face	Phone	Message	Email
Important and urgent	1	2	3	4
Important and not urgent	2	3	4	1
Not important and urgent	3	2	1	4
Not important and not urgent	4	3	2	1

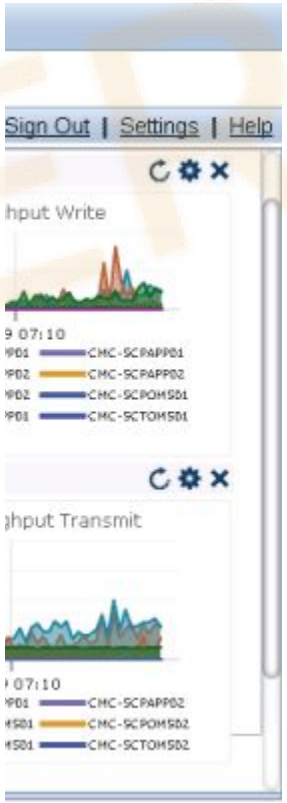
Communication Skill

- Ability to effectively articulate technical challenges and solutions
  - Describe challenges in short sentences logically
  - Describe solution with diagram, analytics, development plan, release plan

Communication Skill



Benchmark – Tech IPOs– Web Site Speed Assessment	http://www.yottaa.com/benchmarks/4e131f653911c31165000023?per_page=25	Monthly Traffic (via Compete.com)	Yottaa Score (Real User Experience)
LinkedIn	www.linkedin.com	19M	88
Zynga	www.zynga.com	11M	87
Pandora	www.pandora.com	12M	77
Trulia	www.trulia.com	6M	75
Yandex	www.Yandex.com	37k	72
DemandMedia	www.DemandMedia.com	45k	71
Zillow	www.zillow.com	8M	67
LivingSocial	www.livingsocial.com	14M	65
Compellent	www.compellent.com	2k	62
Groupon	www.groupon.com/boston	29M	62
Cavium	www.cavium.com	500	58
Renren	www.renren.com	132K	57
FusionIO	www.fusionio.com	6K	39



Question?

CS504 后端工程师直通车

第一月 → **第二月** → **第三月**

耐克跑步追踪系统强化训练
运用最新的后端开发技术，实现一个类似咕咚运动或者Nike+Running的实时位置心率模拟和追踪后台系统。

搜索广告平台项目实战
搭建电商搜索广告系统，实现广告服务器后端，健全的大数据处理pipeline，和机器学习离线训练与线上预测系统。

Capstone：工业实战项目在导师提供的公司实战级别的不同的Track之中选取一个作为主题和方向。

- Demo Day 项目成果展示
- 三大项目实战，打造最强简历
- 总结常见考点，强大面试指导
- Flag级别真实模拟面试服务
- 强大BitTiger专属内部资源库

Code Lab !

- 个性化LiveCoding
- 项目功能延展
- 即时答疑解惑

Code Lab

扫码咨询课程经理：Angela

发送简历到angela@bittiger.io，保留你的early bird席位